ANNALES
HENRI LEBESGUE

PIERRE MONMARCHÉ

WASSERSTEIN CONTRACTION AND POINCARÉ INEQUALITIES FOR ELLIPTIC DIFFUSIONS WITH HIGH DIFFUSIVITY

CONTRACTION WASSERSTEIN ET INÉGALITÉ DE POINCARÉ POUR DES DIFFUSIONS ELLIPTIQUES À FORTE DIFFUSIVITÉ.

ABSTRACT. — We consider elliptic diffusion processes on $\mathbb{R}^d$. Assuming that the drift contracts distances outside a compact set, we prove that, when the diffusion coefficient is sufficiently large, the Markov semi-group associated to the process is a contraction of the $\mathcal{W}_2$ Wasserstein distance, which implies a Poincaré inequality for its invariant measure. The result doesn't require neither reversibility nor an explicit expression of the invariant measure, and the estimates have a sharp dependency on the dimension. Some variations of the arguments are then used to study, first, the stability of the invariant measure of the process with respect to its drift and, second, systems of interacting particles, yielding a criterion for dimension-free Poincaré inequalities and quantitative long-time convergence for non-linear McKean–Vlasov type processes.

Résumé. — On s'intéresse à des processus de diffusion elliptiques sur $\mathbb{R}^d$. Sous l'hypothèse que le terme de dérive contracte les distances en-dehors d'un compact, on montre que, pour un coefficient de diffusion suffisamment grand, le semi-groupe de Markov associé au processus est une contraction de la distance de Wasserstein $\mathcal{W}_2$, ce qui implique une inégalité de Poincaré pour sa mesure invariante. Le résultat ne nécessite ni la réversibilité du processus, ni une expression explicite pour la mesure invariante, et les estimées ont une dépendance correcte en la dimension. Quelques variations du même argument sont également développées pour étudier, en premier lieu, la stabilité de la mesure invariante par rapport à sa dérive et, en second lieu, des systèmes de particules en interaction, fournissant un critère pour une inégalité de Poincaré indépendante de la dimension et une convergence en temps long quantitative pour des processus non-linéaires de type McKean–Vlasov.

1. Overview

Consider $(X_t)_{t \geqslant 0}$ a diffusion process on $\mathbb{R}^d$ solution to

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sqrt{2T}\mathrm{d}B_t \tag{1.1}$$

where $b \in \mathcal{C}^1(\mathbb{R}^d)$, $T > 0$ is a constant and $(B_t)_{t \geqslant 0}$ is a standard Brownian motion. Denote by P_t the associated semi-group, namely $P_t f(x) = \mathbb{E}_x(f(X_t))$ for all suitable f on $\mathbb{R}^d$. Let

$$k(x) := -\sup\left\{ \frac{(x-y)\cdot(b(x)-b(y))}{|x-y|^2},\ y \in \mathbb{R}^d,\ y \neq x \right\}. \tag{1.2}$$

We are interested in cases where the following holds:

Assumption 1.1. — *There exist $K, R \geqslant 0$ and $c > 0$ such that*

$$k(x) \geqslant -K \quad \forall\, x \in \mathbb{R}^d \qquad and \qquad k(x) \geqslant c \quad \forall\, x \in \mathbb{R}^d \text{ with } |x| \geqslant R. \tag{1.3}$$

Under this condition, it is standard to check that the process is non-explosive, admits a unique invariant measure μ with a positive Lebesgue density and that the law of the process converges to μ as $t \to \infty$, using e.g. Lyapunov/Doeblin conditions [BCG08, HM11, MT09]. The first main problem considered in this work is to prove that μ satisfies a Poincaré inequality, namely that there exists a constant $C_P > 0$ such that, for all $f \in \mathcal{C}^1(\mathbb{R}^d)$, writing $\mu f = \int_{\mathbb{R}^d} f \mathrm{d}\mu$,

$$\|f - \mu f\|_{L^2(\mu)}^2 \leqslant C_P \|\nabla f\|_{L^2(\mu)}^2.$$

This inequality is related to concentration inequalities for μ and to the long-time convergence in $L^2(\mu)$ of the law of the process toward μ (see e.g. [BGL14, CG08] and below). More precisely, for $\mu \in \mathcal{P}(\mathbb{R}^d)$ we write

$$C_P(\mu) := \left(\inf\{ \|\nabla f\|_{L^2(\mu)}^2 ,\ f \in \mathcal{C}^1(\mathbb{R}^d),\ \|f - \mu f\|_{L^2(\mu)} = 1 \} \right)^{-1}$$

the optimal constant in the inequality.

When $b = -\nabla U$ for some $U \in \mathcal{C}^2(\mathbb{R}^d)$, Assumption 1.1 is equivalent to say that U is convex outside a compact set, and then a Poincaré inequality is known to hold. In fact, in this case, μ has an explicit density, proportional to $\exp(-U/T)$, and moreover the process is reversible, namely its generator

$$L = T\Delta + b \cdot \nabla \tag{1.4}$$

is self-adjoint in $L^2(\mu)$. The Poincaré constant is then exactly the spectral gap of L (in non-reversible cases, the spectral gap may be larger). Many tools are available to establish Poincaré inequalities for reversible diffusions. In particular, under Assumption 1.1, a Poincaré inequality can be obtained by combining the Bakry–Emery curvature criterion and the Holley–Stroock perturbation argument (see e.g. [BGL14, Propositions 4.2.7 and Proposition 4.8.1] or Section 3.3 below) or a local inequality/Lyapunov condition as in [BCG08, CG14a], see also [BCG08, BGL14, CFG20] and references within concerning the reversible case.

Notice that different drifts b can give the same invariant measure, for instance if $b = -(I_d + J)\nabla U$ where J is any skew-symmetric matrix then $\exp(-U/T)$ is invariant for the process. Assumption 1.1 depends on b, but the Poincaré inequality depends only on μ.

Now, if b is not a gradient and if μ is not explicit, much less is known. If $k(x) \geqslant c > 0$ for all $x \in \mathbb{R}^d$ then the Bakry–Émery arguments still works [CG14b, Mon23a]. If $k(x) \geqslant c > 0$ only for $|x|$ large enough, a Poincaré inequality should be expected, but to the best of our knowledge it cannot be established by existing methods.

This issue leads to the second main question of this work, which is to prove that P_t is a contraction of the $\mathcal{W}_2$ Wasserstein distance for t large enough. Indeed, from classical arguments (see Section 3.1), this implies a Poincaré inequality for μ. Recall that, for $\alpha \in [1, \infty)$, the $\mathcal{W}_\alpha$ Wasserstein distance between $\nu, \nu' \in \mathcal{P}(\mathbb{R}^d)$ (the set of probability measures on $\mathbb{R}^d$) is defined by

$$\mathcal{W}_\alpha(\nu, \nu') = \inf_{\pi \in \Pi(\nu, \nu')} \left(\int_{\mathbb{R}^d} |x - y|^\alpha \pi(\mathrm{d}x, \mathrm{d}y) \right)^{1/\alpha},$$

where $\Pi(\nu, \nu')$ is the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals ν and ν'. Writing νP_t the law at time t of a process solving (1.1) with an initial condition X_0 distributed according to ν (so that $(\nu P_t)f = \nu(P_t f)$ for all bounded measurable f), we want to find $M, \lambda > 0$ such that

$$(1.5) \qquad \forall\, t \geqslant 0,\ \forall\, \nu, \nu' \in \mathcal{P}(\mathbb{R}^d), \qquad \mathcal{W}_\alpha(\nu P_t, \nu' P_t) \leqslant M e^{-\lambda t} \mathcal{W}_\alpha(\nu, \nu'),$$

in particular for $\alpha = 2$. Such a contraction of $\mathcal{W}_\alpha$ implies a contraction of $\mathcal{W}_\beta$ for $\beta \leqslant \alpha$, see e.g. [Mon23a]. Under Assumption 1.1, such a contraction can be proven for $\alpha = 1$ using a reflection coupling and a concave modification of the distance, see [Ebe11, Ebe16]. However, due to the convexity at 0 of $r \mapsto r^2$, for $\alpha = 2$ the method only yields estimates of the form $\mathcal{W}_2(\nu P_t, \nu' P_t) \leqslant M e^{-\lambda t} \max(\mathcal{W}_2(\nu, \nu'), \mathcal{W}_1^{1/2}(\nu, \nu'))$ (see [LW16, Wan16]), which are weaker than (1.5) with $\alpha = 2$ and are not sufficient to get a Poincaré inequality in the non-reversible case.

On the other hand, when $k(x) \geqslant c > 0$ for all $x \in \mathbb{R}^d$, the contraction (1.5) holds with $M = 1$ and $\lambda = c$ for all α, see [Mon23a]. In fact this is an equivalence, according to the Sturm–Von Renesse Theorem [S05] (which is also true in the present non-reversible case [Mon23a]): if there exists $\alpha \geqslant 1$ such that (1.5) holds with $M = 1$ and some $\lambda \in \mathbb{R}$, then necessarily it holds for all α and it implies that $k(x) \geqslant \lambda$ for all $x \in \mathbb{R}^d$. Moreover, it implies a so-called log-Sobolev inequality for the invariant measure, which is stronger than the Poincaré one (see [Mon23a] for details). Since, again, we are interested in the case where $k(x) \geqslant c > 0$ only holds for x large enough,

it means in particular that in that case (1.5) with a positive λ can only hold with $M > 1$.

Our main result is the following:

THEOREM 1.2. — *Under Assumption 1.1, suppose furthermore that*

$$(1.6) \qquad T \geqslant T_0 := (2K + c)\frac{\alpha(K + c/4)R_*^2 + 2\sup\{-x \cdot b(x), \ |x| \leqslant R_*\}}{cd}$$

for some $\alpha \geqslant 2$, *where* $R_* = R(2 + 2K/c)^{1/d}$. *Then (1.5) holds with*

$$\lambda = \frac{c}{4}, \qquad M = M_\alpha := \left(1 + \frac{\alpha(2K + c)R_*^2}{4dT}\right)^{1/\alpha}.$$

This is proven in Section 2. In contrast to the Poincaré inequality, this result is new even in the reversible case.

Let us now state some implications of a $\mathcal{W}_2$ contraction when $M > 1$, obtained from known arguments (see Section 3 for the proof of the next result). To avoid technical discussions, we assume that the coordinates of the force fields b are in $\mathcal{A}$ the set of $\mathcal{C}^\infty$ functions from $\mathbb{R}^d$ to $\mathbb{R}$ with all derivatives growing at most polynomially at infinity (with a slight abuse of notation we simply write $b \in \mathcal{A}$ in that case), and we only consider test functions in $\mathcal{A}$. Combined with Assumption 1.1 which implies a time-uniform Gaussian moment for X_t via standard Lyapunov arguments, we get that, for all $f \in \mathcal{A}$ and all $t \geqslant 0$, $L \in \mathcal{A}$, $P_t f \in \mathcal{A}$ and $\partial_t P_t f = L P_t f = P_t L f$ (see e.g. the proof of [EK86, Theorem 2.5]).

THEOREM 1.3. — *Assume that* $k(x) \geqslant -K$ *for all* $x \in \mathbb{R}^d$ *for some* $K \geqslant 0$, *that* $b \in \mathcal{A}$ *and that a Wasserstein contraction (1.5) holds with* $\alpha = 2$ *for some* $M \geqslant 1, \lambda > 0$. *Then:*

(1) *The invariant measure* μ *satisfies a Poincaré inequality with*

$$C_P(\mu) \leqslant \frac{M^2 T}{\lambda}.$$

(2) *For all* $f \in \mathcal{A}$ *and all* $t \geqslant 0$,

$$(1.7) \ \ \|P_t f - \mu f\|_{L^2(\mu)} \leqslant \min\left(e^{-tT/C_P(\mu)}, \left(1 + \frac{T\left(e^{2\lambda t} - 1\right)}{C_P(\mu)M^2\lambda}\right)^{-1/2}\right) \|f - \mu f\|_{L^2(\mu)}.$$

(3) *For all* $t > 0$ *and any probability law* ν *on* $\mathbb{R}^d$ *with finite second moment,* νP_t *has a density* h_t *with respect to* μ *and*

$$(1.8) \qquad\qquad \|\nu P_t - \mu\|_{TV}^2 \ \leqslant \ \mu(h_t \ln h_t) \ \leqslant \ \frac{J(t)}{2T}\mathcal{W}_2^2(\nu, \mu),$$

where $J(t) \leqslant K/(1 - e^{-2Kt})$ *for all* $t > 0$ *and*

$$J(t) \leqslant M^2(K + \lambda)\left(1 + K/\lambda\right)^{\lambda/K} e^{-2\lambda t} \qquad \text{for all } t \geqslant \frac{\ln(1 + K/\lambda)}{2K}$$

if $K > 0$, *and the limit of these expressions as* $K \to 0$ *if* $K = 0$.

(4) *If furthermore* $b = -\nabla U$ *for some* $U \in \mathcal{C}^2(\mathbb{R}^d)$ *then* $C_P(\mu) \leqslant T/\lambda$.

In both Theorems 1.2 and 1.3, keep in mind that μ and P_t depends on T.

As we see, we only have a partial answer to our initial questions, since the results only hold for T large enough (or, equivalently, for R small enough, which means the results hold for small perturbations of the case where $k(x) \geqslant c$ for all $x \in \mathbb{R}^d$). We didn't try to make the condition (1.6) on T as sharp as possible: it can be slightly improved, but it cannot be suppressed simply by optimizing our proof. We do not know whether, for $T = 1$, Assumption 1.1 is sufficient to get (1.5) for some M, λ or to get a Poincaré inequality for μ. However, notice that, for $T \geqslant T_0$, Theorem 1.2 gives another important information, which is that the contraction (1.5) and the Poincaré inequality hold respectively with M, λ and $C_P(\mu)/T$ which are uniform in $T \geqslant T_0$. Now, this part is clearly false if we suppress the condition that T has to be large enough, namely the statement *"Under Assumption 1.1, there exist $M, \lambda > 0$ such that, for all $T > 0$, (1.5) holds"* is clearly false, and so is *"Under Assumption 1.1, there exists $C > 0$ such that, for all $T > 0$, $C_P(\mu) \leqslant TC$"*. Indeed, the first statement would imply the second (according to Theorem 1.3), and it is well known that, if for instance $b = -\nabla U$ where U has several isolated local minima, then $C_P(\mu) \geqslant e^{a/T}$ for some $a > 0$ for T small enough [HKS89, MS14].

Interestingly, apart from the dependency on T, the bounds on M, λ and $C_P(\mu)$ given by Theorems 1.2 and 1.3 behave rather well with the dimension d, in contrast to what usually give the methods based on the existence of a Lyapunov function and of a local Poincaré inequality [BCG08, CG14a] or on the perturbation of a reference measure [HKS89]. For instance, consider the probability measure $\mu \propto \exp(-U)$ with $U(x) = |x|^\beta/\beta$, for $\beta > 2$. By applying Theorems 1.2 and 1.3, we get that

$$(1.9) \qquad C_P(\mu) \leqslant \frac{8(\beta - 1)\left(1 + 2^{\beta/d+1}\right)^{1-2/\beta}}{d^{1-2/\beta}},$$

see Section 3.2. By contrast, using a standard curvature+bounded perturbation argument, one cannot get better than what is given by the curvature result, which is dimension-free (see Section 3.3 or [CG14b, Remark 5.21]). Besides, here, μ is a radial log-concave probability measure, for which two-sided bounds on the Poincaré constants are known [BJM16, Bob99], in relation to the KLS conjecture [Che21, LV17]. In particular, for $U(x) = |x|^\beta/\beta$, [BJM16, Corollary 4.2] reads

$$\frac{d}{(d+2)d^{1-2/\beta}} \leqslant C_P(\mu) \leqslant \frac{d+1}{(d-1)d^{1-2/\beta}}$$

for $d \geqslant 2$. Hence, the dependency in the dimension d in (1.9) (which is based on our general result and thus does not use that μ is radial) is optimal.

As a last remark on Theorem 1.3, notice that, in (1.7), the minimum is always given by $e^{-tT/C_P(\mu)}$ in the reversible case. However, there are non-reversible cases where $\lambda > T/C_P(\mu)$, so that the second term becomes smaller for large t. For instance, in the Gaussian case $\mu \propto \exp(-U/T)$ with $U(x) = x \cdot Ax$ for some definite positive symmetric matrix A, denote by $\nu_1, \ldots, \nu_d$ the eigenvalues of A. Then it is well known that $T/C_P(\mu) = \min\{\nu_i, i \in [\![1, d]\!]\}$, while non-reversible Gaussian processes with invariant measure μ are constructed in [LNP13] with a linear drift $-Bx$ where the real parts of the eigenvalues of the matrix B are larger than $\bar{\nu} := (\nu_1 + \cdots + \nu_d)/d$,

so that a $\mathcal{W}_2$ contraction holds with $\lambda = \bar{\nu}$ (as can be seen using a synchronous coupling, see e.g. [Mon23a]).

Contrary to a simple long-time convergence at equilibrium in $\mathcal{W}_\alpha$, a contraction of $\mathcal{W}_\alpha$ can easily lead to perturbative results. For instance, consider on $\mathbb{R}^d$ a continuous process solving

$$(1.10) \qquad \mathrm{d}Y_t = \widetilde{b}(Y_t, Z_t)\mathrm{d}t + \sqrt{2T}\mathrm{d}B_t$$

where $Z = (Z_t)_{t \geqslant 0}$ is a random càdlàg process on some state space E and $\widetilde{b} : \mathbb{R}^d \times E \to \mathbb{R}^d$. Denote by $\widetilde{\nu}_t$ the law of Y_t.

PROPOSITION 1.4. — *Let b be a $\mathcal{C}^1$ vector field on $\mathbb{R}^d$ satisfying Assumption 1.1 and P_t be its associated semi-group. Let $\alpha \geqslant 2$. Assume that $T \geqslant T_0$ with T_0 given by (1.6). Then, for all $\nu \in \mathcal{P}(\mathbb{R}^d)$ and $t \geqslant 0$,*

$$(1.11) \quad \mathcal{W}_\alpha\left(\nu P_t, \tilde{\nu}_t\right)$$
$$\leqslant M_\alpha e^{-\lambda t}\mathcal{W}_\alpha(\nu, \tilde{\nu}_0) + M_\alpha^\alpha \int_0^t e^{\lambda(s-t)}\left(\mathbb{E}\left(\left|b(Y_s) - \widetilde{b}(Y_s, Z_s)\right|^\alpha\right)\right)^{1/\alpha}\mathrm{d}s.$$

where λ, M_α are as in Theorem 1.2.

This is proven in Section 3.4. The right hand side in (1.11) can be bounded given additional information on $\widetilde{b}$, for instance if $\widetilde{b}(y, z) = \widetilde{b}(y)$ and we simply assume that $\|b - \widetilde{b}\|_\infty < \infty$ as in [DEE+21, EZ19] then

$$\mathcal{W}_\alpha\left(\nu P_t, \widetilde{\nu}_0\widetilde{P_t}\right) \leqslant M_\alpha e^{-\lambda t}\mathcal{W}_\alpha(\nu, \tilde{\nu}_0) + M_\alpha^\alpha \frac{1 - e^{-\lambda t}}{\lambda}\left\|b - \widetilde{b}\right\|_\infty,$$

with $\widetilde{P_t}$ the semi-group associated to $\widetilde{b} \cdot \nabla + T\Delta$. In particular, any invariant measure $\widetilde{\mu}$ of $\widetilde{P_t}$ satisfies

$$\mathcal{W}_\alpha(\mu, \widetilde{\mu}) \leqslant \frac{M_\alpha^\alpha}{\lambda}\left\|b - \widetilde{b}\right\|_\infty.$$

Hence, under our restrictive condition (1.6), we extend the results of [EZ19], which are restricted to $\alpha = 1$.

More generally, the right hand side in (1.11) can be bounded under the assumption that $\mathbb{E}(|b(y) - \widetilde{b}(y, Z_t)|^\alpha) \leqslant Q(|y|)$ for all $y \in \mathbb{R}^d$ for some polynomial Q and then with some moment estimates on Y_t obtained by Lyapunov arguments (see e.g. the proof of Theorem 4.4 below).

A case of particular interest, in view both of the perturbation result and of the condition on the diffusion coefficient T (to be thought as a temperature parameter in statistical physics), is given by systems of interacting particles, detailed in Section 4.

As a summary, the rest of this paper is organized as follows. Section 2 is devoted to the proof of Theorem 1.2. Section 3 gathers the proofs of the other results stated in this introduction, namely Theorem 1.3, the Poincaré inequality (1.9) and Proposition 1.4, and a discussion on the reversible case. System of interacting particles are studied in Section 4. Finally, we conclude this work in Section 5 by an informal discussion on our method, related works and possible perspectives.

2. Proof of the main theorem

Assumption 1.1 is enforced in all this section, devoted to the proof of Theorem 1.2.

2.1. A probabilistic proof

2.1.1. Synchronous coupling and modified cost

For $(B_t)_{t \geqslant 0}$ a standard Brownian motion on $\mathbb{R}^d$, we consider $(X_t, Y_t)_{t \geqslant 0}$ the Markov process on $\mathbb{R}^d \times \mathbb{R}^d$ solution to

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sqrt{2T}\mathrm{d}B_t \qquad \mathrm{d}Y_t = b(Y_t)\mathrm{d}t + \sqrt{2T}\mathrm{d}B_t \,,$$

which is called the parallel or synchronous coupling of two diffusions (1.1). The generator $\mathcal{L}_s$ of this process is given by

$$\mathcal{L}_s f(x,y) = b(x) \cdot \nabla_x + b(y) \cdot \nabla_y + T\nabla \cdot (A\nabla f)(x,y) \qquad \text{where} \qquad A = \begin{pmatrix} I_d & I_d \\ I_d & I_d \end{pmatrix} .$$

Given $\alpha \geqslant 2$ and a bounded positive $\omega \in \mathcal{C}^2(\mathbb{R}^d)$, consider

$$\rho(x,y) = |x-y|^\alpha \left(2T + \alpha\omega(x) + \alpha\omega(y)\right) ,$$

which is a modification of the usual transport cost $|x-y|^\alpha$ for the Wasserstein distances. Using that $A\nabla f(x,y) = 0$ for $f(x,y) = |x-y|^2$, we get

$$\begin{aligned}
\mathcal{L}_s\rho(x,y) = & \alpha|x-y|^{\alpha-2}(x-y) \cdot (b(x) - b(y))\left(2T + \alpha\omega(x) + \alpha\omega(y)\right) \\
& + \alpha|x-y|^\alpha \left(L\omega(x) + L\omega(y)\right) \\
\leqslant & \alpha|x-y|^\alpha \left(\Psi(x) + \Psi(y)\right) ,
\end{aligned}$$

with

$$(2.1) \qquad \Psi(x) = -k(x)\left(T + \alpha\omega(x)\right) + L\omega(x) .$$

For now, assume that ω is such that there exists $\lambda > 0$ such that,

$$(2.2) \qquad \forall\, x \in \mathbb{R}^d, \qquad \Psi(x) \leqslant -\lambda\left(T + \alpha\omega(x)\right) .$$

Then, $\mathcal{L}_s\rho + \alpha\lambda\rho \leqslant 0$, i.e. $(e^{\alpha\lambda t}\rho(X_t, Y_t))_{t \geqslant 0}$ is a submartingale and for all $t \geqslant 0$,

$$\mathbb{E}\left(\rho\left(X_t, Y_t\right)\right) \leqslant e^{-\alpha\lambda t}\mathbb{E}\left(\rho(X_0, Y_0)\right) .$$

Set

$$M = \left(1 + \frac{\alpha\|\omega\|_\infty}{T}\right)^{1/\alpha} .$$

Let $\nu, \nu' \in \mathcal{P}(\mathbb{R}^d)$. For any $\pi_0 \in \Pi(\nu, \nu')$, considering an initial condition $(X_0, Y_0) \sim \pi_0$ independent from $(B_t)_{t \geqslant 0}$, we obtain

$$\mathcal{W}_\alpha\left(\nu P_t, \nu' P_t\right) \leqslant \left(\mathbb{E}\left(|X_t - Y_t|^\alpha\right)\right)^{1/\alpha} \leqslant Me^{-\lambda t}\left(\mathbb{E}_{\pi_0}\left(|X_0 - Y_0|^\alpha\right)\right)^{1/\alpha} .$$

Finally, taking the infimum over $\pi_0 \in \Pi(\nu, \nu')$ yields (1.5). The proof is thus complete if we are able to construct a bounded positive $\omega \in \mathcal{C}^2(\mathbb{R}^d)$ such that (2.2) holds for some $\lambda > 0$. This is the content of the next section.
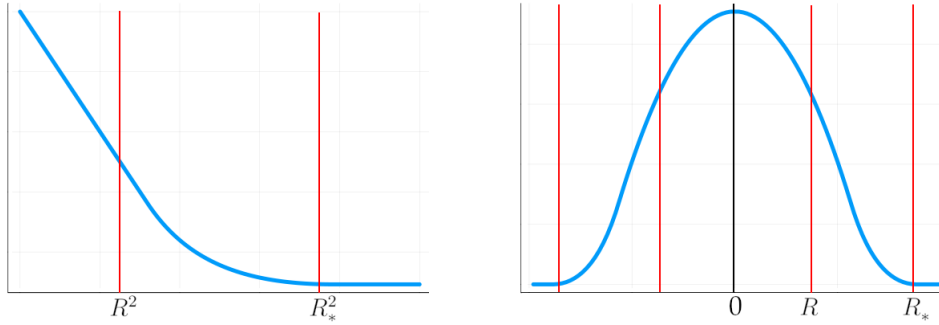
Figure 2.1. Left: $r \mapsto g(r)$. Right: $x \to \omega(x) = g(|x|^2) - \min g$ (in dimension 1). For $r \leqslant R^2$, g is affine decreasing, for $r \geqslant R_*^2$, it is constant, and in between it is convex but with g'' constrained not to be too large, which thus requires to take R_* large enough.

2.1.2. The weight function ω

Our goal is to construct a bounded positive function ω such that

$$\forall\, x \in \mathbb{R}^d, \quad \Delta\omega(x) \leqslant k(x) - \frac{c}{2} \qquad \text{and} \qquad Q := \sup_{x \in \mathbb{R}^d} b(x) \cdot \nabla\omega(x) < \infty.$$

Indeed, if this holds, taking $\lambda = c/4$, we get

$$\Psi(x) + \lambda\,(T + \alpha\omega(x)) = (\lambda - k(x))\,(T + \alpha\omega(x)) + L\omega(x)$$

$$\leqslant -\frac{c}{4}T + \alpha(K + \lambda)\|\omega\|_\infty + Q$$

and thus (2.2) holds if $T \geqslant 4\,(\alpha(K + \lambda)\|\omega\|_\infty + Q)\,/c$.

We take ω of the form $\omega(x) = g(|x|^2) - \inf g$ with $g \in \mathcal{C}^2(\mathbb{R}_+)$ to be chosen. Since

$$\nabla\omega(x) = 2xg'(|x|^2)\,, \qquad \Delta\omega = 4|x|^2 g''(|x|^2) + 2dg'(|x|^2)\,,$$

setting $K_* = K/4 + c/8$, we take g as the $\mathcal{C}^1$ solution of $g(0) = 0$, $g'(0) = -2K_*/d$ and

$$rg''(r) + \frac{d}{2}g'(r) = h(r) := \begin{cases} -K_* & \text{for } r < R^2 \\ c/8 & \text{for } r \in (R^2, R_*^2) \\ 0 & \text{for } r > R_*^2 \end{cases}$$

where $R_* > 0$ remains to be chosen so that $g'(R_*^2) = 0$ (and thus $g'(r) = g''(r) = 0$ for all $r > R_*^2$). See Figure 2.1 for a draft of the graph of g and ω. Notice that g is not $\mathcal{C}^2$, but this is easily solved by replacing this g by some $g_\varepsilon \in \mathcal{C}^2(\mathbb{R}_+)$ such that, for a small $\varepsilon > 0$, first, $g_\varepsilon''(r) = g_\varepsilon'(r) = 0$ for $r \geqslant R_*^2 + \varepsilon$ and, second, $rg_\varepsilon''(r) + (d/2)g_\varepsilon'(r)$ is always smaller than $h(r)$, equal to $-K_*$ for $r \in [0, R^2]$ and to $c/8$ for $r \in [R^2 + \varepsilon, R_*^2 - \varepsilon]$. We can conclude the proof with g_ε and finally let ε vanish in the final result. For simplicity we write the proof directly with g.

For $r \in [0, R^2]$, we simply have $g'(r) = -2K_*/d$. For $r \geqslant R^2$, since $(r^{d/2}g'(r))' = h(r)r^{d/2-1}$, we get

$$r^{d/2}g'(r) = R^d g'(R^2) + \frac{c}{8}\int_{R^2}^{r} s^{d/2-1}\mathrm{d}s = -\frac{2K_*}{d}R^d + \frac{c}{4d}\left(r^{d/2} - R^d\right)\,,$$

and thus we choose

$$R_* = R\left(1 + 8K_*/c\right)^{1/d} = R\left(2 + 2K/c\right)^{1/d} ,$$

which concludes the construction of ω. It remains to estimate Q and $\|\omega\|_\infty$. Since g is decreasing, constant on $[R_*^2, +\infty)$ and such that $\|g'\|_\infty = 2K_*/d$, we get

$$\forall\, r \geqslant 0\,, \qquad 0 \geqslant g(r) \geqslant g(R_*^2) \geqslant -\frac{2K_* R_*^2}{d}\,,$$

hence $\|\omega\|_\infty \leqslant 2K_* R_*^2/d$, and

$$Q \leqslant \frac{4K_*}{d} \sup\left\{-x \cdot b(x),\ |x| \leqslant R_*\right\} .$$

The proof is concluded by using these estimates in the definition of M and the condition on T.

2.2. Alternative proof with Bakry–Emery interpolation

In this section we give a second proof of Theorem 1.2 (see Section 5 for a discussion on the specific interest of each proof). This proof is similar to the intertwining method of Joulin, Bonnefont and their coauthors [ABJ18, BJ22, BJM16, CJ13] (who focus on reversible cases). For simplicity we only consider the case $\alpha = 2$, which is the main case of interest due to Theorem 1.3. The expressions of T_0 and M obtained along this alternative proof are slightly different than those stated in Theorem 1.2 and established in the first proof (again, we don't try to optimize the estimate on T_0). Moreover, in order to justify the time derivatives in this section, we assume that the coordinate functions of b are in $\mathcal{A}$.

The proof is organized in three steps: first we rephrase Assumption 1.1 as a local condition on the drift, namely a condition on its Jacobian. Second, similarly to Section 2.1.1, we give the proof conditionally to the existence of a suitable weight function. Third, similarly to Section 2.1.2, we construct the weight function.

2.2.1. Step 1: an infinitesimal condition

We start by an equivalent formulation of Assumption 1.1. We denote by Db the Jacobian matrix of b and

$$\widetilde{k}(x) = -\sup\left\{u \cdot Db(x)u,\ u \in \mathbb{R}^d,\ |u| = 1\right\} .$$

It is clear that, for all $x \in \mathbb{R}^d$,

$$(2.3) \qquad\qquad \widetilde{k}(x) \geqslant k(x) \geqslant \inf_{y \in \mathbb{R}^d} \int_0^1 \widetilde{k}(x + ty)\mathrm{d}t\,.$$

In particular under Assumption 1.1, $\widetilde{k}$ also satisfies (1.3), with the same K, R, c. Alternatively, assume that there exist $K, R \geqslant 0$ and $c > 0$ such that

$$(2.4) \quad \widetilde{k}(x) \geqslant -K \quad \forall\, x \in \mathbb{R}^d \qquad \text{and} \qquad \widetilde{k}(x) \geqslant c \quad \forall\, x \in \mathbb{R}^d \text{ with } |x| \geqslant R\,.$$

Then, from (2.3), $k(x) \geqslant -K$ for all $x \in \mathbb{R}^d$. Moreover, for all $x, y \in \mathbb{R}^d$, the Lebesgue measure of $\{t \in [0,1], |x+ty| < R\}$ being less than $2R/(|x|+R)$,

$$k(x) \geqslant \inf_{y \in \mathbb{R}^d} \int_0^1 \widetilde{k}(x+ty)\mathrm{d}t \geqslant -K \frac{2R}{|x|+R} + c\left(1 - \frac{2R}{|x|+R}\right) \geqslant \frac{c}{2}$$

for all $x \in \mathbb{R}^d$ with $|x| \geqslant \widetilde{R} := R\left(4(K+c)/c - 1\right)$. In other words, Assumption 1.1 is equivalent to assume that the infinitesimal condition (2.4) holds for some $K, R \geqslant 0$ and $c > 0$. This latter condition is enforced for the rest of Section 2.2.

2.2.2. Step 2: Gamma calculus

Consider a positive $a \in \mathcal{A}$, to be chosen later on. Fix $t > 0$ and $f \in \mathcal{A}$. The carré du champ of the generator L is defined as $\Gamma(f,g) = \frac{1}{2}(L(gf) - gLf - fLg)$, with the notation $\Gamma(f) = \Gamma(f,f)$. In the case of (1.4), this is simply $\Gamma(f,g) = T\nabla f \cdot \nabla g$. We write $[A, B] = AB - BA$. For a given $i \in [\![1, d]\!]$, write $g_s = \partial_{x_i} P_{t-s}f$ for $s \in [0, t]$. Then

$$\begin{aligned}
\partial_s \left(P_s(ag_s^2)\right) &= P_s\left(L(ag_s^2) - 2ag_s\partial_{x_i}LP_sf\right) \\
&= P_s\left(L(a)g_s^2 + aL(g_s^2) + 2\Gamma(a, g_s^2) - 2ag_s\left([\partial_{x_i}, L]P_{t-s}f + Lg_s\right)\right) \\
&= P_s\left(L(a)g_s^2 + 2a\Gamma(g_s) + 2\Gamma(a, g_s^2) + 2ag_s[L, \partial_{x_i}]P_{t-s}f\right).
\end{aligned}$$

Using that

$$[L, \partial_{x_i}] = [b \cdot \nabla, \partial_{x_i}] = -(\partial_{x_i}b) \cdot \nabla$$

and

$$2\Gamma(a, g_s^2) = 4g_s\Gamma(a, g_s) \geqslant -2\frac{\Gamma(a)}{a}g_s^2 - 2a\Gamma(g_s)$$

(where $\Gamma(a)/a$ is understood as 0 if $\Gamma(a) = 0$), summing over $i \in [\![1, d]\!]$, we get

$$\begin{aligned}
\partial_s &\left(P_s\left(a\left|\nabla P_{t-s}f\right|^2\right)\right) \\
&\geqslant P_s\left(\left(L(a) - 2\frac{\Gamma(a)}{a}\right)\left|\nabla P_{t-s}f\right|^2 - 2a(\nabla P_{t-s}f) \cdot Db\nabla P_{t-s}f\right) \\
&\geqslant 2P_s\left(\Phi(a)\left|\nabla P_{t-s}f\right|^2\right)
\end{aligned}$$

with

$$\Phi(a) = \frac{1}{2}L(a) - \frac{\Gamma(a)}{a} + a\widetilde{k}$$

Assume for now that a is such that there exists $\lambda > 0$ such that

(2.5) a and a^{-1} are bounded and $\Phi(a) \geqslant \lambda a$.

(This condition is similar to the ones in [ABJ18, BJ22, BJM16, CJ13], for instance [ABJ18, Theorem 3.2]; in this work, the authors do not assume that a is bounded, as they are interested in Brascamp–Lieb inequalities, namely weighted Poincaré inequalities, but in the present work we are interested in classical Poincaré inequalities and contraction of the Wasserstein distances associated to the standard

Euclidean distance, which is why we add this condition so that $a|\nabla f|^2$ is equivalent to $|\nabla f|^2$). Integrating the previous inequality yields

$$P_t\left(a\,|\nabla f|^2\right) \geqslant e^{2\lambda t}a\,|\nabla P_t f|^2\,,$$

and then

$$|\nabla P_t f|^2 \leqslant \|a\|_\infty \left\|a^{-1}\right\|_\infty e^{-2\lambda t}P_t\,|\nabla f|^2\,,$$

which concludes the proof since, thanks to the work of Kuwada [Kuw10, Kuw13], it is equivalent to the $\mathcal{W}_2$ contraction (1.5) with the same λ, $\alpha = 2$ and $M = \sqrt{\|a\|_\infty\|a^{-1}\|_\infty}$. Indeed, more precisely, in the case of a diffusion process on $\mathbb{R}^d$, as a corollary of [Kuw10, Theorem 2.2] (applied with v the Lebesgue measure), we get the following:

PROPOSITION 2.1. — *Assume that $b \in \mathcal{A}$ and that $k(x) \geqslant -K$ for all $x \in \mathbb{R}^d$ for some $K \geqslant 0$. For $\rho \in \mathbb{R}$, $M > 0$ and $t \geqslant 0$, the two following assertions are equivalent:*

- *For all $f \in \mathcal{A}$,*
$$|\nabla P_t f| \leqslant Me^{-\lambda t}\sqrt{P_t\,|\nabla f|^2}.$$
- *For all probability measures ν, μ on $\mathbb{R}^d$,*
$$\mathcal{W}_2\left(\nu P_t, \mu P_t\right) \leqslant Me^{-\lambda t}\mathcal{W}_2(\nu, \mu)\,.$$

Notice that, in [Kuw10, Theorem 2.2], the equivalence is stated for the class of functions f which are bounded and Lipschitz. For $f \in \mathcal{A}$, we can find a sequence f_n of bounded Lipschitz functions $(f_n)_{n\in\mathbb{N}}$ such that $|\nabla(f_n - f)|(x) \leqslant \varepsilon_n + \mathbb{1}_{|x|\geqslant 1/\varepsilon_n}q(x)$ for some polynomial q, where $\varepsilon_n \to 0$ as $n \to 0$. Using a synchronous coupling and that the process admits Gaussian moments, it is easily seen that $|\nabla P_t(f - f_n)| \leqslant e^{Kt}P_t|\nabla(f - f_n)| \to 0$, hence the result.

2.2.3. Step 3: the weight function

It remains to construct a weight a satisfying (2.5) for some $\lambda > 0$ under the condition (2.4). As in Section 2.1.2, we focus on the leading term for large T, setting

$$a(x) = T + 2\left(\|\omega\|_\infty - \omega(x)\right)\,,$$

where ω is a bounded positive function to be chosen. Then, using that $a \geqslant T$ and $\Gamma(\omega) = T|\nabla\omega|^2$,

$$\Phi(a) \geqslant -T\Delta\omega - b\cdot\nabla\omega - 4\,|\nabla\omega|^2 + T\widetilde{k} - 2K\|\omega\|_\infty\,.$$

Let ω be such that

$$\forall\,x \in \mathbb{R}^d,\quad \Delta\omega(x) \leqslant \widetilde{k}(x) - \frac{c}{2}$$

and

$$\widetilde{Q} := K\|\omega\|_\infty + \sup_{x\in\mathbb{R}^d}\left(b(x)\cdot\nabla\omega(x) + 4\,|\nabla\omega(x)|^2\right) < \infty\,,$$

as constructed in Section 2.1.2. Then

$$\Phi(a) \geqslant \frac{cT}{2} - \widetilde{Q} \geqslant \frac{cT}{3} \geqslant \frac{c}{4}a$$

if

$$T \geqslant \widetilde{T}_0 := 6 \max \left(\frac{\widetilde{Q}}{c}, \|\omega\|_\infty \right).$$

Moreover, for $T \geqslant \widetilde{T}_0$, we get $T \geqslant 3a/4 \geqslant 3T/4$, hence $\|a\|_\infty \|a^{-1}\|_\infty \leqslant 4/3$, i.e. (1.5) holds with $M = \sqrt{4/3}$ and $\lambda = c/4$. An explicit upper bound of $\widetilde{T}_0$ follows from the estimates on ω given in Section 2.1.2.

3. Other proofs

3.1. Proof of Theorem 1.3

Setting $V(x) = e^{c|x|^2/4}$, Assumption 1.1 classically implies that $LV \leqslant -aV + C$ for some constants $C, a > 0$, and that $V \in L^2(\mu)$. Moreover, by standard elliptic theory (see e.g. [Dyn65, Theorem 0.5 and Condition 0.24.A1]), the process (1.1) admits a continuous positive transition kernel, hence, for all compact set $\mathcal{K} \subset \mathbb{R}^d$ and all $t > 0$, there exists $\eta > 0$ such that $\inf_{x \in \mathcal{K}} P_t f(x) \geqslant \int_{\mathcal{K}} f(z) \mathrm{d}z$ for all positive f. From [HM11], we get that $P_t f(x) \to \mu f$ as $t \to \infty$ for all $f \in \mathcal{A}$ (since f/V is bounded if $f \in \mathcal{A}$). For $f \in \mathcal{A}$, $x \in \mathbb{R}^d$ and $t \geqslant 0$,

$$P_t f^2(x) - (P_t f(x))^2 = \int_0^t \partial_s \left(P_s \left(P_{t-s} f \right)^2 (x) \right) \mathrm{d}s$$

$$= 2T \int_0^t P_s \left| \nabla P_{t-s} f \right|^2 (x) \mathrm{d}s$$

$$\leqslant 2T P_t |\nabla f|^2 (x) M^2 \int_0^t e^{-2\lambda s} \mathrm{d}s$$

where we used the equivalence of Wasserstein and gradient contractions of Proposition 2.1. Letting $t \to \infty$ in the previous equality yields the Poincaré inequality (for all $f \in \mathcal{A}$, and then for all $f \in L^2(\mu)$ by density).

From the Lumer–Philips Theorem [Yos94, Chapter IX, p. 250], the Poincaré inequality is equivalent to

$$\forall f \in L^2(\mu), \forall t \geqslant 0 \qquad \|P_t f - \mu f\|_{L^2(\mu)} \leqslant e^{-tT/C_P} \|f - \mu f\|_{L^2(\mu)}.$$

Besides, for $t > 0$ and $f \in \mathcal{A}$, we can also bound

$$P_t f^2 - (P_t f)^2 = 2T \int_0^t P_s \left| \nabla P_{t-s} f \right|^2 \mathrm{d}s \geqslant 2T \left| \nabla P_t f \right|^2 M^{-2} \int_0^t e^{2\lambda s} \mathrm{d}s.$$

Integrating with respect to μ and applying this with f replaced by $f - \mu f$ yields

$$\|\nabla P_t f\|_{L^2(\mu)}^2 \leqslant \frac{M^2 \lambda}{T \left(e^{2\lambda t} - 1 \right)} \left(\|f - \mu f\|_{L^2(\mu)}^2 - \|P_t f - \mu f\|_{L^2(\mu)}^2 \right).$$

Together with the Poincaré inequality, this means that, for all $t \geqslant 0$,

$$\|P_t f - \mu f\|_{L^2(\mu)}^2 \leqslant \frac{C_P M^2 \lambda}{T \left(e^{2\lambda t} - 1 \right)} \left(\|f - \mu f\|_{L^2(\mu)}^2 - \|P_t f - \mu f\|_{L^2(\mu)}^2 \right),$$

i.e.

$$\|P_t f - \mu f\|_{L^2(\mu)}^2 \leqslant \left(1 + \frac{T(e^{2\lambda t} - 1)}{C_P M^2 \lambda}\right)^{-1} \|f - \mu f\|_{L^2(\mu)}^2 \,.$$

Now, in the reversible case where $b = -\nabla U$, since

$$-\frac{1}{t} \ln \left(1 + \frac{T\left(e^{2\lambda t} - 1\right)}{C_P M^2 \lambda}\right)^{-1} = 2\lambda \,,$$

this implies (see e.g. [CGZ13, Lemma 2.14]) that, in fact,

$$\|P_t f - \mu f\|_{L^2(\mu)}^2 \leqslant e^{-2\lambda t} \|f - \mu f\|_{L^2(\mu)}^2 \,,$$

and thus $C_P(\mu) \leqslant T/\lambda$.

Finally, the first inequality in (1.8) is simply the Pinsker inequality, and the entropy/$\mathcal{W}_2$ regularisation is proven in [RW10]. It is assumed in the latter work that b is Lipschitz, but it is not used in this part of the proof. We briefly recall the proof for completeness. Remark that if Assumption 1.1 holds with $K = 0$ then it also holds with all $K > 0$ and thus it is sufficient to treat the case $K > 0$. From [Wan12, Theorem 3.3], under Assumption 1.1, for all $t > 0$ all positive f and all $x, y \in \mathbb{R}^d$,

$$P_t \ln f(x) \leqslant \ln P_t f(y) + \frac{K|x - y|^2}{2T\left(1 - e^{-2Kt}\right)} \,.$$

Denoting by P_t^* the dual of P_t in $L^2(\mu)$, we apply the previous inequality with f replaced by $P_t^* f$ for some positive f with $\mu f = 1$ and integrate with a coupling measure $\pi \in \Pi(f\mu, \mu)$ to get

$$\int_{\mathbb{R}^d} P_t^* f(x) \ln P_t^* f(x) \mu(\mathrm{d}x)$$

$$= \int_{\mathbb{R}^d} P_t \left(\ln P_t^* f(x)\right) f(x) \mu(\mathrm{d}x)$$

$$\leqslant \int_{\mathbb{R}^d} \ln P_t P_t^* f(y) \mu(\mathrm{d}y) + \frac{K}{2T\left(1 - e^{-2Kt}\right)} \int_{\mathbb{R}^{2d}} |x - y|^2 \pi(\mathrm{d}x, \mathrm{d}y) \,.$$

Using that μ is invariant by $P_t P_t^*$ and Jensen's inequality,

$$\mu \left(\ln P_t P_t^* f\right) \leqslant \ln \mu \left(P_t P_t^* f\right) = \ln \mu f = 0 \,,$$

and taking the infimum over π concludes the proof of (1.8) with $J(t) = K/(1 - e^{-2Kt})$. Then, using the $\mathcal{W}_2$ contraction, for all $s \in [0, t)$,

$$\nu P_t \left(\ln \frac{\nu P_t}{\mu}\right) \leqslant \frac{K}{2T(1 - e^{-2K(t-s)})} M^2 e^{-2\lambda s} \mathcal{W}_2(\nu, \mu) \,.$$

The minimum of $s \mapsto e^{-2\lambda s}/(1 - e^{-2K(t-s)})$ for $s \leqslant t$ for a fixed $t > 0$ is attained at $s = s_* := t - \ln\left(1 + K/\lambda\right)/(2K)$. When $s_* \geqslant 0$, the proof is concluded by taking $s = s_*$ in the previous bound, since

$$\frac{K}{2T(1 - e^{-2K(t-s_*)})}M^2 e^{-2\lambda s_*} = \frac{M^2 K \left(1 + K/\lambda\right)^{\lambda/K} e^{-2\lambda t}}{2T \left(1 - (1 + K/\lambda)^{-1}\right)}$$

$$= \frac{M^2}{2T}(K + \lambda)\left(1 + K/\lambda\right)^{\lambda/K} e^{-2\lambda t}.$$

3.2. Degenerate convex potential

Let $U(x) = |x|^\beta / \beta$ for some $\beta > 2$. For $T > 0$, let μ_T be the probability law on $\mathbb{R}^d$ with density proportional to $\exp(-U/T)$. If X is a random variable with law μ_1, then $T^{1/\beta}X$ is distributed according to μ_T. The scaling properties of the Poincaré inequality imply that $C_P(\mu_T) = T^{2/\beta}C_P(\mu_1)$. For $b = -\nabla U$, $Db(0) = 0$ and, for $x \neq 0$,

$$u \cdot Db(x)u = -u \cdot \nabla^2 U(x)u = -|x|^{\beta-2}\left[\frac{\beta - 2}{\beta}\frac{|x \cdot u|^2}{|x|^2} + |u|^2\right] \leqslant -|x|^{\beta-2}|u|^2.$$

Then, for all $x \neq 0$ and all $y \neq x$,

$$\frac{(x - y) \cdot (b(x) - b(y))}{|x - y|^2} = -\int_0^1 \frac{x - y}{|x - y|} \cdot \nabla^2 U(x + t(x - y))\frac{x - y}{|x - y|}\mathrm{d}t$$

$$\leqslant -\int_0^1 |x + t(x - y)|^{\beta-2}\mathrm{d}t$$

$$\leqslant -|x|^{\beta-2}\inf_{r \in \mathbb{R}}\int_0^1 |1 + tr|^{\beta-2}\mathrm{d}t$$

$$= -|x|^{\beta-2}\inf_{r \geqslant 0}\int_0^1 |1 - tr|^{\beta-2}\mathrm{d}t$$

$$= -|x|^{\beta-2}\inf_{r \geqslant 0}\frac{1}{r}\int_{1-r}^1 |s|^{\beta-2}\mathrm{d}s.$$

The term in the inf is non-increasing for $r \leqslant 1$, non-decreasing for $r \geqslant 2$ and, for $r \in [1, 2]$,

$$\frac{1}{r}\int_{1-r}^1 |s|^{\beta-2}\mathrm{d}s \geqslant \frac{1}{2}\int_0^1 |s|^{\beta-2}\mathrm{d}s = \frac{1}{2(\beta - 1)}.$$

As a conclusion, for all $x \in \mathbb{R}^d$

$$k(x) \geqslant \frac{1}{2(\beta - 1)}|x|^{\beta-2},$$

which means that, for all $r > 0$, Assumption 1.1 holds with $K = 0$, $R = r$ and $c = c(r) := r^{\beta-2}/(2\beta - 2)$. Besides, $-x \cdot b(x) = |x|^\beta$ for all $x \in \mathbb{R}^d$ and thus Theorem 1.2 applied with $\alpha = 2$ and

$$T = T(r) := \frac{2^{2/d-1}r^2 c(r) + 2(2^{1/d}r)^\beta}{d} = \frac{r^\beta}{d}\left(\frac{1}{2^{2-2/d}(\beta - 1)} + 2^{\beta/d+1}\right)$$

$$\leqslant \frac{r^\beta}{d}\left(1 + 2^{\beta/d+1}\right),$$

yields (according to Theorem 1.3 since we are in the reversible case)

$$C_P(\mu_{T(r)}) \leqslant T(r)\frac{4}{c(r)}$$

hence

$$C_P(\mu_1) \leqslant (T(r))^{1-2/\beta}\,\frac{4}{c(r)} \leqslant \frac{8(\beta-1)\left(1+2^{\beta/d+1}\right)^{1-2/\beta}}{d^{1-2/\beta}}\,.$$

Notice that, as expected due to the homogeneity of the problem, the powers of r have disappeared.

3.3. The reversible case

If $b = -\nabla U$ and Assumption 1.1 holds, let us check what explicit estimate can be obtained by the Holley–Stroock perturbation argument together with the Bakry–Emery curvature one. As in Section 2.2.1, we use that Assumption 1.1 implies that $\nabla^2 U(x) \geqslant c\mathbb{1}_{|x|\geqslant R} - K\mathbb{1}_{|x|<R}$. Consider $V(x) = -g(|x|^2)$ for some $g \in \mathcal{C}^2$ to be chosen. Assume that g is bounded and that $\nabla^2(U+V)(x) \geqslant c/2$ for all $x \in \mathbb{R}^d$. Then the Bakry–Emery criterion states that $\widetilde{\mu} \propto \exp(-(U+V)/T)$ satisfies a Poincaré inequality with constant $2T/c$, and by bounded perturbation we get that $\mu \propto \exp(-U/T)$ satisfies a Poincaré inequality with constant $2Te^{(\max g - \min g)/T}/c$.

It remains to choose g. We choose g to be non-increasing and convex, so that we bound

$$\nabla^2 V(x) = -4xx^T g''(|x|^2) - 2I_d g'(|x|^2) \geqslant -4|x|^2 g''(|x|^2) - 2g'(|x|^2)\,.$$

As in Section 2.1.2, we simply take g' as (a $\mathcal{C}^1$ non-decreasing approximation of) the continuous solution of $g'(0) = -K/2 - c/4$ and

$$rg''(r) + \frac{1}{2}g'(r) = \begin{cases} -(K/4+c/8) & \text{for } r < R^2 \\ c/8 & \text{for } r \in (R^2, R_*^2) \\ 0 & \text{for } r > R_*^2 \end{cases}$$

with $R_* = 2R(1+K/c)$. In other words, g is exactly such as constructed in Section 2.1.2 with $d = 1$. We end up with

$$\max g - \min g \leqslant (2K+c)R^2(1+K/c)^2\,.$$

As a consequence,

$$C_P(\mu) \leqslant \frac{2T}{c}\exp\left(\frac{(2K+c)R^2(1+K/c)^2}{T}\right)\,.$$

3.4. Perturbation of the drift

This section is devoted to the proof of Proposition 1.4.

Proof. — Define ρ as

$$\rho(x,y) = |x-y|^\alpha \left(T + \alpha\omega(x)\right), \qquad x, y \in \mathbb{R}^d$$

for some $\alpha \geqslant 2$, where ω is a positive bounded function to be chosen. Let $(X_t)_{t \geqslant 0}$ and $(Y_t)_{t \geqslant 0}$ be the solution respectively of (1.1) and (1.10) driven by $(B_t)_{t \geqslant 0}$ with some initial condition ν. Then

$$\partial_t \mathbb{E}\left(\rho(X_t, Y_t)\right)$$
$$= \alpha \mathbb{E}\left(|X_t - Y_t|^{\alpha-2}(X_t - Y_t)\cdot\left(b(X_t) - \widetilde{b}(Y_t, Z_t)\right)(T + \alpha\omega(X_t)) + |X_t - Y_t|^\alpha L\omega(X_t)\right)$$
$$\leqslant \alpha\mathbb{E}\left(|X_t - Y_t|^\alpha \Psi(X_t) + |X_t - Y_t|^{\alpha-1}(T + \alpha\|\omega\|_\infty)\left|b(Y_t) - \widetilde{b}(Y_t, Z_t)\right|\right)$$

with Ψ defined in (2.1). Let ω be as defined in Section 2.1.2, and $\lambda = c/4$. Then, writing $m(t) = \mathbb{E}\left(\rho(X_t, Y_t)\right)$ and using the Hölder inequality,

$$m'(t)$$
$$\leqslant -\lambda\alpha m(t) + \alpha(T + \alpha\|\omega\|_\infty)\left(\mathbb{E}\left(|X_t - Y_t|^\alpha\right)\right)^{1-1/\alpha}\left(\mathbb{E}\left(\left|b(Y_t) - \widetilde{b}(Y_t, Z_t)\right|^\alpha\right)\right)^{1/\alpha}$$
$$\leqslant -\lambda\alpha m(t) + \alpha\widetilde{C}(t)m(t)^{1-1/\alpha}$$

with

$$\widetilde{C}(t) = \frac{T + \alpha\|\omega\|_\infty}{T^{1-1/\alpha}}\left(\mathbb{E}\left(\left|b(Y_t) - \widetilde{b}(Y_t, Z_t)\right|^\alpha\right)\right)^{1/\alpha}.$$

Then

$$\left(m^{1/\alpha}\right)'(t) \leqslant -\lambda m^{1/\alpha}(t) + \widetilde{C}(t)$$

and thus

$$m^{1/\alpha}(t) \leqslant e^{-\lambda t}m^{1/\alpha}(0) + \int_0^t e^{\lambda(s-t)}\widetilde{C}(s)\mathrm{d}s.$$

Using the equivalence between ρ and $|x-y|^\alpha$ and taking the infimum over all coupling of the initial conditions yields

$$\mathcal{W}_\alpha\left(\nu P_t, \widetilde{\nu}_t\right) \leqslant M_\alpha e^{-\lambda t}\mathcal{W}_\alpha(\nu, \widetilde{\nu}_0) + T^{-1/\alpha}\int_0^t e^{\lambda(s-t)}\widetilde{C}(s)\mathrm{d}s,$$

which concludes the proof of Proposition 1.4. $\square$

4. Interacting particles at high temperature

4.1. General framework

Let $\mathbf{X}_t = (X_{1,t}, \ldots, X_{N,t})$ be a process on $(\mathbb{R}^d)^N$ solving

$$\forall\, i \in [\![1, N]\!], \qquad \mathrm{d}X_{i,t} = F(X_{i,t})\mathrm{d}t + G_i(\mathbf{X}_t)\mathrm{d}t + \sqrt{2T}\mathrm{d}B_{i,t}$$

where $(B_{1,t}, \ldots, B_{N,t})_{t \geqslant 0}$ are N independent standard d-dimensional Brownian motions, $F \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$, $G_i \in \mathcal{C}^1(\mathbb{R}^{dN}, \mathbb{R}^d)$. In other words, $\mathbf{X}$ solves (1.1) with a drift b whose i^{th} d-dimensional component is $b_i(\mathbf{x}) = F(x_i) + G_i(\mathbf{x})$ for all $i \in [\![1, N]\!]$. Write $G(\mathbf{x}) = (G_1(\mathbf{x}), \ldots, G_N(\mathbf{x})) \in (\mathbb{R}^d)^N$.

THEOREM 4.1. — *Assume that there exist $c > 0$, $a < c$ and $C_F, C_G, R, M_G \geqslant 0$ such that*

$$(4.1) \quad (x - y) \cdot (F(x) - F(y)) \leqslant \begin{cases} C_F |x - y|^2 & \text{for all } x, y \in \mathbb{R}^d \\ -c|x - y|^2 & \text{for all } x, y \in \mathbb{R}^d \text{ with } |x| \geqslant R \end{cases},$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{dN}$

$$(4.2) \sum_{i=1}^{N} |x_i - y_i||G_i(\mathbf{x}) - G_i(\mathbf{y})| \leqslant C_G |\mathbf{x} - \mathbf{y}|^2, \qquad (\mathbf{x} - \mathbf{y}) \cdot (G(\mathbf{x}) - G(\mathbf{y})) \leqslant a|\mathbf{x} - \mathbf{y}|^2$$

and $|G_i(\mathbf{x})| \mathbb{1}_{|x_i| \leqslant R_*} \leqslant M_G$ for all $i \in [\![1, N]\!]$ with $R_* = R \left(2 + 2(C_F + a)/(c - a)\right)^{1/(2d)}$. Writing $K_* = C_F + (c + a)/2$, assume furthermore that

$$T \geqslant T_0 = \frac{K_*}{d(c - a)} \left(R_*^2 \left(C_F + C_G\right) + 2 \sup \left\{ -F(x) \cdot x, |x| \leqslant R_* \right\} + 4 M_G R_* \right).$$

Then the semi-group P_t associated to the process $\mathbf{X}$ satisfies the $\mathcal{W}_2$ contraction (1.5) (with $\alpha = 2$) with

$$M = \sqrt{1 + 2K_* R_*^2/(Td)} \qquad \lambda = \frac{c - a}{4 + 8K_* R_*^2/(Td)}.$$

Notice that the first part of (4.2) implies the second one with $a = C_G$, but in many cases we can have $a < C_G$ (possibly $a \leqslant 0$, see next section) and the result is much more sensible to the value of a (in particular with the condition $a < c$) than to the value of C_G.

Proof. — Consider

$$\rho(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{N} |x_i - y_i|^2 \left(T + \omega(x_i) + \omega(y_i)\right)$$

where ω is a positive function to be chosen. As in Section 2.1.2, writing $\mathcal{L}_s$ the generator on $\mathbb{R}^{dN} \times \mathbb{R}^{dN}$ of a parallel coupling of two processes, we consider separately the leading terms with respect to T and the rest in

$$\mathcal{L}_s \rho(\mathbf{x}, \mathbf{y}) = TA + B$$

with

$$A = \sum_{i=1}^{N} \left[2(x_i - y_i) \left(F(x_i) + G_i(\mathbf{x}) - F(y_i) - G_i(\mathbf{y})\right) + |x_i - y_i|^2 \left(\Delta\omega(x_i) + \Delta\omega(y_i)\right) \right]$$

$$\leqslant \sum_{i=1}^{N} |x_i - y_i|^2 \left(2a - k_F(x_i) - k_F(y_i) + \Delta\omega(x_i) + \Delta\omega(y_i)\right)$$

where $k_F(x_i) = c$ if $|x_i| \geqslant R$ and $k_F(x_i) = -C_F$ otherwise and, assuming that ω is constant outside the ball $\{|x| \leqslant R_*\}$,

$$B = \sum_{i=1}^{N} \Big[2(x_i - y_i) \cdot (b_i(\mathbf{x}) - b_i(\mathbf{y})) \left(\omega(x_i) + \omega(y_i)\right)$$

$$+ |x_i - y_i|^2 \left(b_i(\mathbf{x}) \cdot \nabla\omega(x_i) + b_i(\mathbf{y}) \cdot \nabla\omega(y_i)\right) \Big]$$

$$\leqslant \Big(4\|\omega\|_\infty \left(C_F + C_G\right) + 2\sup\{F \cdot \nabla\omega\} + 2M_G\|\nabla\omega\|_\infty \Big)|\mathbf{x} - \mathbf{y}|^2.$$

We take ω as in Section 2.1.2 but with K, c replaced respectively by $C_F + a$ and $c - a$ (in particular ω is indeed constant outside the ball $\{|x| \leqslant R_*\}$). The following holds:

$$\forall \, x \in \mathbb{R}^d, \quad a - k_F(x_i) + \Delta\omega(x_i) \leqslant -\frac{c-a}{2}$$

$$\|\omega\|_\infty \leqslant \frac{K_* R_*^2}{d}$$

$$\sup\{F \cdot \nabla\omega\} \leqslant \frac{K_*}{d} \sup\{-F(x) \cdot x, |x| \leqslant R_*\}$$

$$\|\nabla\omega\|_\infty \leqslant 2R_* \frac{K_*}{d} \,.$$

Hence, provided $T \geqslant T_0$, the previous bounds yield

$$\mathcal{L}_s \rho(\mathbf{x}, \mathbf{y}) \leqslant -\frac{c-a}{2} T |\mathbf{x} - \mathbf{y}|^2 \leqslant -\frac{c-a}{2 + 4\|\omega\|_\infty/T} \rho(\mathbf{x}, \mathbf{y}) \,.$$

The conclusion is now similar to the end of the proof of Theorem 1.2, using that

$$|\mathbf{x} - \mathbf{y}|^2 \leqslant \frac{1}{T} \rho(\mathbf{x}, \mathbf{y}) \leqslant \left(1 + \frac{2\|\omega\|_\infty}{T}\right) \|\mathbf{x} - \mathbf{y}\|^2 \,. \qquad \square$$

The point of Theorem 4.1 is that if all the constants in the assumption are independent from N, then so are T_0, λ and M. In particular, from Theorem 1.3, we get for the invariant measure of the process a Poincaré inequality independent from N (for T large enough). The restriction to a sufficiently high temperature is very natural for interacting particles systems where phase transitions are expected in the behaviour of the Poincaré inequality at low temperature [Tug14].

A Poincaré constant uniform in N for T large enough is established in [GLWZ19] in a reversible framework with an explicit invariant measure. Although Theorem 4.1 does not require reversibility, on the other hand it needs the interaction force G to be bounded, which is not the case in [GLWZ19] and is a restrictive condition. It is however satisfied in many cases of interest, for instance in adaptive algorithms such as studied in [CdRDM+21], a typical choice is

$$G_i(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N \nabla W \left(x_i - x_j\right) , \qquad W(x) = e^{-|x|^2} ,$$

which induces a local repulsion of particles, enhancing the exploration of the state space.

More generally, assume that there exist a graph on $[\![1, N]\!]$ of degree D and a bounded and Lipschitz function $H \in \mathcal{C}^1(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ such that

$$G_i(\mathbf{x}) = \frac{1}{D} \sum_{j \sim i} H \left(x_i, x_j\right)$$

where $i \sim j$ means that (i, j) is an edge of the graph. This is the case for mean field interaction (with the complete graph and $D = N$) or for interaction with closest neighbors in $(\mathbb{Z}/n\mathbb{Z})^k$ or $[\![1, n]\!]^k$ (with $i \sim j$ if $|i - j| = 1$, $D = 2k$). Then G satisfies the assumptions of Theorem 4.1 with M_G, C_G, a which only depend on H (and thus not on the number of particles). For instance, in the particular (reversible) case

where the forces are the gradients of some potentials, we get the following corollary of Theorem 4.1.

PROPOSITION 4.2. — *Let $V \in \mathcal{C}^1(\mathbb{R}^d)$, $W \in \mathcal{C}^1(\mathbb{R}^d \times \mathbb{R}^d)$. Assume that ∇W is bounded and Lipschitz and that there exists $c > 0$ and $a < c$ such that $\nabla^2 V \geqslant c > 0$ outside a compact set and $\nabla^2 W \geqslant -a/2$. Then there exist $T_0, C > 0$ such that the following holds. For all $T \geqslant T_0$, all $N \in \mathbb{N}$, all graph on $[\![1, N]\!]$, denoting by D the degree of the graph and considering on $\mathbb{R}^{dN}$ the potential*

$$U(\mathbf{x}) = \sum_{i=1}^N V(x_i) + \frac{1}{D} \sum_{i=1}^N \sum_{j \sim i} W(x_i - x_j) ,$$

then $e^{-U/T}$ is integrable and the probability measure proportional to this density satisfies a Poincaré inequality with constant $C_P \leqslant CT$.

This is a result in the spirit of [GLWZ19, Theorem 1].

4.2. The mean-field case and propagation of chaos

Let us now focus on the case of mean field interactions. More precisely, we work under the following condition:

ASSUMPTION 4.3. — *The drift b on $\mathbb{R}^{dN}$ is of the form*

$$b_i(\mathbf{x}) = F(x_i) + \frac{1}{N} \sum_{i=1}^N H(x_i, x_j)$$

where $F \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ and $H \in \mathcal{C}^1(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$. Moreover, there exist $c > 0$, $a < c$ and $C_F, R, C_G, M_G \geqslant 0$ such that F satisfies (4.1), H is $2C_H/3$-Lipschitz continuous and for all $x, y, x', y' \in \mathbb{R}^d$, $\mathbb{1}_{|x| \leqslant R} H(x, y) \leqslant M_H$ and

$$(4.3) \quad (x - y) \cdot (H(x, x') - H(y, y')) + (x' - y') \cdot (H(x', x) - H(y', y))$$
$$\leqslant a \left(|x - y|^2 + |x' - y'|^2 \right) .$$

Finally, $T \geqslant T_0$, where T_0 is given in Theorem 4.1.

It is straightforward to check that this condition implies the assumptions of Theorem 4.1. As soon as H is $2C_H/3$-Lipschitz, (4.3) holds with $a = C_H$, and the condition $a < c$ is then satisfied if the interaction is sufficiently small. However, in some cases, a may be smaller than C_H, in particular, in the usual case where $H(x, y) = \widetilde{H}(x - y) = -\widetilde{H}(y - x)$ for some $\widetilde{H}$, the condition (4.3) reads

$$(x - y - x' + y') \cdot \left(\widetilde{H}(x - x') - \widetilde{H}(y - y') \right) \leqslant a \left(|x - y|^2 + |x' - y'|^2 \right) ,$$

and this holds with $a = 0$ if $x \cdot \widetilde{H}(x) \leqslant 0$ for all $x \in \mathbb{R}^d$. This is the case for instance for $\widetilde{H}(x) = -\nabla W(x)$ with $W(x) = \gamma \sqrt{1 + |x|^2}$, for any $\gamma \geqslant 0$. Since ∇W is bounded, in this case, Theorem 4.1 applies whatever the value of γ, i.e. even if the interaction force is not small with respect to the confining force (however, as γ increases, so does the temperature T_0).

As $N \to +\infty$, according to the propagation of chaos phenomenon, it is well-known that two given particles of the system behave like independent McKean-Vlasov processes solving

$$(4.4) \quad \mathrm{d}\bar{X}_t = F(\bar{X}_t)\mathrm{d}t + \int_{\mathbb{R}^d} H\left(\bar{X}_t, z\right) \rho_t(\mathrm{d}z)\mathrm{d}t + \sqrt{2T}\mathrm{d}B_t\,, \qquad \bar{\nu}_t = \mathcal{L}aw\left(\bar{X}_t\right)\,.$$

In other words, $\bar{\nu}_t$ solves the non-linear equation

$$(4.5) \qquad\qquad \partial_t\bar{\nu}_t = \nabla \cdot \left(T\nabla\bar{\nu}_t - \left(F + (H * \bar{\nu}_t)\right)\bar{\nu}_t\right)\,,$$

where $H*\bar{\nu}(x) = \int_{\mathbb{R}^d} H(x,y)\bar{\nu}(\mathrm{d}y)$. The existence, uniqueness of the process (4.4) and of the solution of the equation (4.5), together with time-dependent propagation of chaos estimates, follow from standard arguments [Mél96, Szn91] for initial conditions $\bar{\nu}_0$ in $\mathcal{P}_2(\mathbb{R}^d)$ the set of probability measures on $\mathbb{R}^d$ with finite second moment.

With a $\mathcal{W}_2$ Wasserstein contraction such as given by Theorem 4.1 at hand, it is straightforward to obtain time-uniform propagation of chaos and a Wasserstein contraction for the limit equation.

THEOREM 4.4. — *Under Assumption 4.3, there exist (explicit) constants $\alpha, \beta > 0$ such that the following holds. For $N \in \mathbb{N}$, let P_t^N be the semi-group associated to $L = b \cdot \nabla + T\Delta$ on $\mathbb{R}^{dN}$ and let $\bar{\nu}_t$ be a solution of (4.5). Then, for all $t \geqslant 0$,*

$$\mathcal{W}_2\left(\nu P_t^N, \bar{\nu}_t^{\otimes N}\right) \leqslant Me^{-\lambda t}\mathcal{W}_2\left(\nu, \bar{\nu}_0^{\otimes N}\right) + \alpha e^{-\frac{c-a}{2}t}\sqrt{\int_{\mathbb{R}^d} |y|^2\bar{\nu}_0(\mathrm{d}y)} + \beta\,,$$

where M and λ are given in Theorem 4.1.

Proof. — The proof is essentially the same as the proof of Proposition 1.4, except that we consider a cost ρ as in the proof of Theorem 4.1, namely

$$\rho(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{N} |x_i - y_i|^2 \left(T + 2\omega(x_i)\right)\,.$$

Let $(\mathbf{X}_t)_{t \geqslant 0}$ be a system of particles with drift b and initial condition ν and $\mathbf{Y} = (Y_1, \ldots, Y_N)$ be solutions of (4.4) (with the same Brownian motions as $\mathbf{X}$) with initial condition $\nu_0^{\otimes N}$. In particular, $\mathbf{Y}_t \sim \bar{\nu}_t^{\otimes N}$ for all $t \geqslant 0$. As in the proof of Proposition 1.4, writing $m(t) = \mathbb{E}(\rho(\mathbf{X}_t, \mathbf{Y}_t))$, we get

$$(4.6) \qquad m'(t) \leqslant -2\lambda m(t) + 2(T + 2\|\omega\|_\infty) \sum_{i=1}^{N} \sqrt{\mathbb{E}\left(|X_{i,t} - Y_{i,t}|^2\right)\widehat{C}(t)}$$

where λ is given in Theorem 4.1 and, using that the Y_j's all have the same law,

$$\widehat{C}(t) = \mathbb{E}\left(\left|\frac{1}{N}\sum_{j=1}^{N} H\left(Y_{1,t}, Y_{j,t}\right) - H * \bar{\nu}_t(Y_{1,t})\right|^2\right)\,.$$

Developing the square and using that the variables $A_j := H(Y_{1,t}, Y_{j,t}) - H * \bar{\nu}_t(Y_{1,t})$ for $j \neq 1$ are independent and centered, we get

$$\widehat{C}(t) \leqslant \frac{1}{N^2} \mathbb{E} \left(\sum_{j=1}^{N} |A_j|^2 + A_1 \cdot \sum_{j \neq 1} A_j \right)$$

$$\leqslant \frac{1}{N} \mathbb{E} \left(|A_1|^2 + |A_2|^2 \right) .$$

Then we bound

$$\mathbb{E}\left(|A_j|^2\right) = \mathbb{E}\left(\left| \int_{\mathbb{R}^d} : \left(H\left(Y_{1,t}, Y_{j,t}\right) - H\left(Y_{1,t}, y\right) \right) \bar{\nu}_t(\mathrm{d}y) \right|^2 \right)$$

$$\leqslant L_H^2 \mathbb{E}\left(\int_{\mathbb{R}^d} |Y_{j,t} - y|^2 \, \bar{\nu}_t(\mathrm{d}y) \right) = 2L_H^2 \left(\int_{\mathbb{R}^d} |y|^2 \bar{\nu}_t(\mathrm{d}y) - \left| \int_{\mathbb{R}^d} y \bar{\nu}_t(\mathrm{d}y) \right|^2 \right)$$

and thus,

$$\widehat{C}(t) \leqslant \frac{4L_H^2}{N} \int_{\mathbb{R}^d} |y|^2 \bar{\nu}_t(\mathrm{d}y) .$$

Under Assumption 4.3, for all $y, y' \in \mathbb{R}^d$,

$$y \cdot F(y) \leqslant |y||F(0)| - c|y|^2 + (C_F + c)R^2$$

$$y \cdot H(y, y') + y' \cdot H(y', y) \leqslant a\left(|y|^2 + |y'|^2\right) + (|y| + |y'|) |H(0,0)|$$

from which

$$\partial_t \int_{\mathbb{R}^d} |y|^2 \bar{\nu}_t(\mathrm{d}y) = 2 \int_{\mathbb{R}^d \times \mathbb{R}^d} \left[y \cdot (F(y) + H(y, y')) + Td \right] \bar{\nu}_t(\mathrm{d}y) \bar{\nu}_t(\mathrm{d}y')$$

$$\leqslant -(c - a) \int_{\mathbb{R}^d} |y|^2 \bar{\nu}_t(\mathrm{d}y) + Q$$

with

$$Q = Td + 2(C_F + c)R^2 + \frac{1}{c - a} \left(|H(0,0)| + |F(0)| \right)^2 .$$

Integrating in time,

$$\int_{\mathbb{R}^d} |y|^2 \bar{\nu}_t(\mathrm{d}y) \leqslant e^{-(c-a)t} \int_{\mathbb{R}^d} |y|^2 \bar{\nu}_0(\mathrm{d}y) + Q$$

Going back to (4.6),

$$m'(t) \leqslant -2\lambda m(t) + 2(T + 2\|\omega\|_\infty) \sqrt{N \sum_{i=1}^{N} \mathbb{E}\left(|X_{i,t} - Y_{i,t}|^2\right) \widehat{C}(t)}$$

$$\leqslant -2\lambda m(t) + 4\frac{L_H(T + 2\|\omega\|_\infty)}{\sqrt{T}} \sqrt{m(t) \left(e^{-(c-a)t} \int_{\mathbb{R}^d} |y|^2 \bar{\nu}_0(\mathrm{d}y) + Q \right)}$$

$$:= -2\lambda m(t) + 2\sqrt{m(t) \left(e^{-(c-a)t} A^2 + B^2 \right)},$$

hence

$$\left(\sqrt{m}\right)'(t) \leqslant -\lambda \sqrt{m(t)} + \sqrt{e^{-(c-a)t} A^2 + B^2} \leqslant -\lambda \sqrt{m(t)} + e^{-(c-a)t/2} A + B .$$

Integrating in time (and noticing that $\lambda \leqslant (c-a)/4$),

$$\sqrt{m(t)} \leqslant e^{-\lambda t}\sqrt{m(0)} + \int_0^t e^{\lambda(s-t)}\left(e^{-(c-a)s/2}A + B\right)\mathrm{d}s$$

$$\leqslant e^{-\lambda t}\sqrt{m(0)} + \frac{e^{-(c-a)t/2}}{(c-a)/2 - \lambda}A + \frac{1}{\lambda}B,$$

and the proof of Theorem 4.4 is concluded by using the equivalence between ρ and the Euclidean norm and taking the infimum over the couplings of the initial distributions. $\qquad\square$

Theorem 4.4 has the following consequences.

COROLLARY 4.5. — *Under Assumption 4.3, considering $M, \lambda, \alpha, \beta$ as in Theorem 4.4, then, for all $N \in \mathbb{N}$, $k \in [\![1, N]\!]$ and all $\bar{\nu}_0, \bar{\mu}_0 \in \mathcal{P}_2(\mathbb{R}^d)$, the following holds. Let $(\mathbf{X}_t)_{t \geqslant 0}$ be a system of N interacting particles on $\mathbb{R}^{dN}$ with drift b and with initial condition $\bar{\nu}_0^{\otimes N}$ and denote by $\nu_t^{k,N}$ the law of $(X_{1,t}, \ldots, X_{k,t})$. Let $\bar{\nu}_t, \bar{\mu}_t$ be the solutions of (4.5) with respective initial conditions $\bar{\nu}_0, \bar{\mu}_0$. Then, for all $t \geqslant 0$,*

$$\mathcal{W}_2\left(\nu_t^{k,N}, \bar{\nu}_t^{\otimes k}\right) \leqslant \sqrt{\frac{k}{N}}\left(\alpha e^{-\frac{c-a}{2}t}\sqrt{\int_{\mathbb{R}^d}|y|^2\bar{\nu}_0(\mathrm{d}y)} + \beta\right)$$

$$\mathcal{W}_2\left(\bar{\mu}_t, \bar{\nu}_t\right) \leqslant Me^{-\lambda t}\mathcal{W}_2\left(\bar{\mu}_0, \bar{\nu}_0\right)$$

and there exists a unique stationary solution to (4.5) in $\mathcal{P}_2(\mathbb{R}^d)$.

Moreover, if $\int_{\mathbb{R}^d}|y|^5\bar{\nu}_0(\mathrm{d}y) < +\infty$ then for all $t \geqslant 0$,

$$\mathbb{E}\left(\mathcal{W}_2^2\left(\frac{1}{N}\sum_{i=1}^N\delta_{X_{i,t}}, \bar{\nu}_t\right)\right) \leqslant C'\alpha(N)\left(e^{-(c-a)t}\int_{\mathbb{R}^d}|y|^5\bar{\nu}_0(\mathrm{d}y) + 1\right)^{2/5}.$$

where C' is a constant which depends only on d and on the parameters of Assumption 4.3, and

$$\alpha(N) = \begin{cases} N^{-1/2} & \text{if } d < 4 \\ N^{-1/2}\ln(1+N) & \text{if } d = 4 \\ N^{-2/d} & \text{if } d > 4. \end{cases}$$

In the last claim we assumed a finite 5th moment to get a simple statement but, as can be seen in the proof and from the results of [FG15], a similar result would hold assuming only a qth finite moment for any $q > 2$. If only a second moment is available, we still get a similar result if $\mathcal{W}_2^2$ is replaced by $\mathcal{W}_p^p$ for any $p < 2$.

Proof. — Using the interchangeability of particles and that any coupling of νP_t and $\bar{\nu}_t^{\otimes N}$ gives a coupling of the k first particles immediately yield

$$\mathcal{W}_2\left(\nu_t^{k,N}, \bar{\nu}_t^{\otimes k}\right) \leqslant \sqrt{\frac{k}{N}}\mathcal{W}_2\left(\bar{\nu}_0^{\otimes N}P_t^N, \bar{\nu}_t^{\otimes N}\right).$$

The first claim then follows from Theorem 4.4. By the same argument, denoting by $\mu_t^{1,N}$ the first d-dimensional marginal of $\bar{\mu}_0^{\otimes N}P_t^N$, we get

$$(4.7) \quad \mathcal{W}_2\left(\nu_t^{1,N}, \mu_t^{1,N}\right) \leqslant \frac{1}{\sqrt{N}}\mathcal{W}_2\left(\bar{\nu}_0^{\otimes N}P_t^N, \bar{\mu}_0^{\otimes N}P_t^N\right) \leqslant \frac{1}{\sqrt{N}}Me^{-\lambda t}\mathcal{W}_2\left(\bar{\nu}_0^{\otimes N}, \bar{\mu}_0^{\otimes N}\right),$$

using Theorem 4.1. Considering a coupling of $\bar{\nu}_0^{\otimes N}$ and $\bar{\mu}_0^{\otimes N}$ of the form $\pi_0^{\otimes N}$ where $\pi_0 \in \Pi(\bar{\nu}_0, \bar{\mu}_0)$ and then taking the infimum over π_0 yields

$$\frac{1}{\sqrt{N}} \mathcal{W}_2\left(\bar{\nu}_0^{\otimes N}, \bar{\mu}_0^{\otimes N}\right) \leqslant \mathcal{W}_2\left(\bar{\nu}_0, \bar{\mu}_0\right),$$

and the second claim is thus obtained by letting N go to infinity in (4.7).

As a consequence, for t large enough, the function $\Phi_t : \bar{\nu}_0 \mapsto \bar{\nu}_t$, where $(\bar{\nu}_t)_{t \geqslant 0}$ is the solution of (4.5) with initial condition $\bar{\nu}_0$, is a contraction of $\mathcal{P}_2(\mathbb{R}^d)$ endowed with the $\mathcal{W}_2$ distance, which is complete. Hence, Φ_t admits a unique fixed point for t large enough, and using that $\Phi_t \Phi_s = \Phi_s \Phi_t$ for all $s \geqslant 0$ and the uniqueness of the fixed point we get that the fixed point of Φ_t is in fact a fixed point of Φ_s for all $s \geqslant 0$, i.e. is a stationary solution of (4.5).

For the last claim, let $(\mathbf{X}_t, \mathbf{Y}_t) \sim \pi_t \in \Pi(\bar{\nu}_0^{\otimes N} P_t, \bar{\nu}_t^{\otimes N})$. We bound

$$\mathbb{E}\left(\mathcal{W}_2^2\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_{i,t}}, \bar{\nu}_t\right)\right)$$

$$\leqslant 2\mathbb{E}\left(\mathcal{W}_2^2\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_{i,t}}, \frac{1}{N}\sum_{i=1}^N \delta_{Y_{i,t}}\right)\right) + 2\mathbb{E}\left(\mathcal{W}_2^2\left(\frac{1}{N}\sum_{i=1}^N \delta_{Y_{i,t}}, \bar{\nu}_t\right)\right)$$

$$\leqslant \frac{2}{N}\mathbb{E}\left(|\mathbf{X}_t - \mathbf{Y}_t|^2\right) + c_d \alpha(N)\left(\int_{\mathbb{R}^d} |y|^5 \bar{\nu}_t(\mathrm{d}y)\right)^{2/5}$$

with c_d a constant that depends only on d, where we used the coupling $(X_{J,t}, Y_{J,t})$ with J a random variable uniformly distributed over $[\![1, N]\!]$ independent from $(\mathbf{X}_t, \mathbf{Y}_t)$ to bound the first term, and [FG15, Theorem 1] for the second one. Then, reasoning exactly as in the proof of Theorem 4.4, we get that

$$\int_{\mathbb{R}^d} |y|^5 \bar{\nu}_t(\mathrm{d}y) \leqslant e^{-(c-a)t} \int_{\mathbb{R}^d} |y|^5 \bar{\nu}_0(\mathrm{d}y) + Q'$$

for some $Q' > 0$ which depend on the parameters of Assumption 1.1. Taking the infimum over $\pi_t \in \Pi\left(\bar{\nu}_0^{\otimes N} P_t, \bar{\nu}_t^{\otimes N}\right)$ we end up with

$$\mathbb{E}\left(\mathcal{W}_2^2\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_{i,t}}, \bar{\nu}_t\right)\right) \leqslant \frac{2}{N}\mathcal{W}_2^2\left(\bar{\nu}_0^{\otimes N} P_t, \bar{\nu}_t^{\otimes N}\right)$$

$$+ c_d \alpha(N)\left(e^{-(c-a)t}\int_{\mathbb{R}^d}|y|^5 \bar{\nu}_0(\mathrm{d}y) + Q'\right)^{2/5},$$

and Theorem 4.4 concludes the proof. $\qquad\square$

5. Discussion

On the two proofs

First, let us notice that the main ingredient of the two proofs of Theorem 1.2 is the very simple construction of a weighted distance. Indeed, the weighted gradient of the second proof is

$$a(x)|\nabla f(x)|^2$$
$$= \lim_{r \to 0} \sup_{|y-x| \leqslant r} \frac{|f(x) - f(y)|^2}{|x-y|^2 \left(a^{-1}(x) + a^{-1}(y)\right)/2} := \lim_{r \to 0} \sup_{|y-x| \leqslant r} \frac{|f(x) - f(y)|^2}{r^2(x,y)},$$

in other words $\sqrt{a}\nabla$ is the gradient associated to the weighted distance r. Weighted distances are a very standard tool, in particular for the study of the long-time convergence of Markov processes. However, it seems to us that the main originality of our work is that the role of the weight is exactly the converse of the usual one. Indeed, under Assumption 1.1, typically (see e.g. [EGZ19, EM19, HM08, HM11, Mon23b]), one considers costs of the form $\rho(x,y) = \mathrm{d}(x,y)(1+V(x)+V(y))$ where $\mathrm{d}(x,y) = \mathbb{1}_{x \neq y}$ or $\mathrm{d}(x,y) = f(|x-y|)$ for some f is the initial distance we are interested in and V is a Lyapunov function (say, $V(x) = |x|^2$), which satisfies $LV \leqslant -cV$ outside a compact ball, thanks to the deterministic drift part $b \cdot \nabla$. This Lyapunov condition is then combined with some local information (local Poincaré inequality, local Doeblin condition, local coupling condition...) which is available on compact sets. In other words, the weight is used to obtain a decay outside a given compact set. On the contrary, in our case, we take a weight of the form $V(x) = C - |x|^2$ in a compact ball, so that $LV \leqslant -cV$ in this ball thanks to the dissipative part $T\Delta$.

One of the interest of the first proof is that it gives more information than simply a contraction of the Wasserstein distance at time $t > 0$: it shows that

$$\forall\, t \geqslant 0, \qquad \mathbb{E}\left(|X_t - Y_t|^\alpha\right) \leqslant M_\alpha^\alpha e^{-\alpha \lambda t} \mathbb{E}\left(|X_0 - Y_0|^\alpha\right)$$

where $(X_t, Y_t)_{t \geqslant 0}$ is the synchronous coupling of two diffusions. In particular, this is a Markovian coupling (i.e. (X,Y) is a Markov process), realized with a single process for all times.

On the other hand, one of the interest of the second proof is that it can be adapted to deal with quantities integrated with respect to μ, which could be interesting in the perspective of proving a Poincaré inequality for non-explicit invariant measure of non-reversible diffusion processes without any restriction on T but with the assumption that μ satisfies local Poincaré inequalities (which straightforwardly follow from elliptic lower and upper bounds on the transition density). Indeed, once integrated with respect to μ, the computations of Section 2.2 reads

$$\partial_t \mu\left(a|\nabla P_t|^2\right) \leqslant -2\lambda \mu\left(a|\nabla P_t|^2\right)$$

and thus

$$\mu\left(a|\nabla P_t|^2\right) \leqslant \|a\|_\infty \left\|a^{-1}\right\|_\infty e^{-2\lambda t} \mu|\nabla f|^2,$$

which doesn't give a $\mathcal{W}_2$-contraction but is sufficient to get a Poincaré inequality (see Section 3.1). An argument somehow in this spirit is given in [BGH21], although in a completely different framework, i.e. with a non-elliptic hypoelliptic diffusion, a singular drift and weighted Poincaré inequalities. On the other hand, although the process is non-reversible in this case, its invariant measure is known, and thus the proof relies on the knowledge of L^* the adjoint of L in $L^2(\mu)$, which is unavailable if μ is unknown. Besides, in the elliptic case, if L^* is known, it should be possible to

adapt the arguments of [BCG08, CG14a] to get a Poincaré inequality by considering a Lyapunov function with respect to L^* rather than L, as in [BGH21].

Flat torus

As emphasized in the previous paragraph, the proof relies on a synchronous coupling. Now, consider the very simple case of the Brownian motion on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, i.e. $L = \Delta$. Considering a synchronous coupling of two such processes leads to $|X_t - Y_t| = |X_0 - Y_0|$ (where we write $|x - y|$ the distance on $\mathbb{T}$). Hence, any $\rho : \mathbb{T} \times \mathbb{T} \to \mathbb{R}_+$ such that $c|x - y|^\alpha \leqslant \rho(x, y) \leqslant C|x - y|^\alpha$ for some $c, C, \alpha > 0$ necessarily satisfies

$$x\rho(X_t, Y_t) \geqslant \frac{c}{C}\rho(X_0, Y_0)\,.$$

Hence, the first proof of Theorem 1.2 cannot apply in this case.

Notice that, for compact manifolds, Poincaré inequalities can be obtained by lower and upper bounds on the transition density and then perturbation of the Lebesgue measure. See also [Wan97, Wan14] in the reversible case. Besides, notice that a weighted distance is used in [Wan14], but with the motivation of handling boundaries.

Exponential tail

Consider on $\mathbb{R}$ the drift $b(x) = -U'(x)$ with $U(x) = \sqrt{1 + x^2}$. It is well-known that $\mu \propto \exp(-U)$ satisfies a Poincaré inequality. However, using that $\|U'\|_\infty = 1$, and using that $\partial_t \mathbb{E}(X_t) = \mathbb{E}(b(X_t))$, we get that

$$|\mathbb{E}(X_t) - \mathbb{E}(X_0)| \leqslant t\,.$$

This clearly forbids a Wasserstein contraction (1.5), for any $\alpha \geqslant 1$. Indeed,

$$\mathcal{W}_\alpha(\delta_x P_t, \delta_y P_t) \geqslant |\mathbb{E}_x(X_t) - \mathbb{E}_y(X_t)| \geqslant |x - y| - 2t = \mathcal{W}_\alpha(\delta_x, \delta_y) - 2t\,,$$

and thus

$$\sup_{x,y \in \mathbb{R}^d, x \neq y} \frac{\mathcal{W}_\alpha(\delta_x P_t, \delta_y P_t)}{\mathcal{W}_\alpha(\delta_x, \delta_y)} \geqslant 1$$

for all $t \geqslant 0$.

Wasserstein contraction versus Wasserstein convergence

In the case where $\mu \propto \exp(-U)$ and U is convex at infinity, μ is known to satisfy a so-called log-Sobolev inequality, see [BGL14], which is stronger than the Poincaré inequality and implies (combining the exponential decay of the entropy, the $\mathcal{T}_2$ Talagrand inequality implied by the log-Sobolev one and the $\mathcal{W}_2$/entropy regularization of [RW10], see e.g. the proof of [GM21, Theorem 2]) that there exist $C, \lambda > 0$ such that for all $t \geqslant 0$ and any probability law ν on $\mathbb{R}^d$,

$$\mathcal{W}_2(\nu P_t, \mu) \leqslant Ce^{-\lambda t}\mathcal{W}_2(\nu, \mu)\,.$$

However this convergence towards equilibrium in $\mathcal{W}_2$ is weaker than the contraction (1.5) for $\alpha = 2$. Besides, we are interested in cases where the Poincaré inequality does not follow from standard arguments, and thus neither does the log-Sobolev inequality.

Besides, in [Wan14, Corollary 1.3], the log-Sobolev inequality is proven from the exponential decay of weighted gradients along the semi-group, but in contrast to the present work it concerns reversible processes, more precisely the perturbation argument used at the end of [Wan14, Corollary 1.3] requires an explicit invariant measure. In order to use a perturbation argument when the invariant measure of the semi-group P_t is unknown, we can still say that the measure $\phi\mu$, where ϕ is bounded above and below by positive constants, is invariant by the semi-group $P_t^\phi f = \phi^{-1} P_t(\phi f)$, but it is unclear whether this could be used to adapt the proof of [Wan14] to non-reversible cases with unknown μ.

<div align="center">Non-reversible sampling</div>

In the context of Markov Chain Monte Carlo algorithms, a question is the following: given a known target probability measure $\mu \propto \exp(-U)$, one would like to find, among all the drifts b such that μ is invariant for $L = b \cdot \nabla + \Delta$, the one for which convergence toward equilibrium is the fastest (here for simplicity we fix the diffusion matrix to be the identity), see e.g. [HHMS93, HHMS05, LNP13, GM16]. If the convergence toward equilibrium is quantified in terms of the $L^2(\mu)$-norm, then

$$\sup_{f \in L^2(\mu)} \frac{\|P_t f - \mu f\|_{L^2(\mu)}}{\|f - \mu f\|_{L^2(\mu)}} \leqslant e^{-t/C_P(\mu)} \, ,$$

see e.g. Theorem 1.3, and moreover this is an equality in the reversible case. As a consequence, adding a non-reversible part to the generator $-\nabla U \cdot \nabla + \Delta$ can only improve the L^2 convergence rate. For instance, as already mentioned in Section 1, in the Gaussian case where $U(x) = x \cdot Ax$ for some definite positive symmetric matrix A, $C_P(\mu)$ is the minimum of the eigenvalues of A while the optimal rate obtained with non-reversible elliptic diffusions is the mean of the eigenvalues of A, see [LNP13].

Alternatively, the efficiency of the process can be measured in terms of $\mathcal{W}_2$-contraction (and this leads to the same conclusion for Gaussian processes). Let $\mathcal{B}$ be the set of drifts b which are K-Lipschitz and such that μ is invariant for $b \cdot \nabla + \Delta$ (in practice, the Lipschitz constant impacts the stability of the numerical schemes used to discretize (1.1), and thus the non-reversible part should not be too large). Then, instead of seeing Theorem 1.3 as a way to obtain a Poincaré inequality from a Wasserstein contraction, here we can use this result to obtain a constraint on M and λ such that (1.5) holds in terms of μ, uniformly over $\mathcal{B}$. Another way to see this is the following: for a fixed $t > 0$ (corresponding to a fixed computational budget), what is the smallest $\gamma(t)$ one can obtain such that there exist a drift $b \in \mathcal{B}$ such that

$$(5.1) \qquad \forall \, \nu, \nu' \in \mathbb{R}^d \, , \qquad \mathcal{W}_2(\nu P_t, \nu' P_t) \leqslant \gamma(t) \mathcal{W}_2(\nu, \nu') \, ,$$

where P_t is the semi-group associated to b? More generally, for a given b and for $s \geqslant 0$, let

$$\gamma(s) = \sup_{\nu \neq \nu'} \frac{\mathcal{W}_2(\nu P_s, \nu' P_s)}{\mathcal{W}_2(\nu, \nu')} .$$

Since b is K-Lipschitz, we immediately obtain from a synchronous coupling that $\gamma(s) \leqslant e^{Ks}$ and, following the proof of Theorem 1.3, (5.1) implies that

$$(5.2) \qquad C_P(\mu) \leqslant \int_0^\infty \gamma(s) \mathrm{d}s \leqslant \frac{e^{Kt} - 1}{K} \sum_{k=0}^\infty \gamma(t)^k = \frac{e^{Kt} - 1}{K(1 - \gamma(t))} .$$

Hence the lower bound on the contraction rate

$$\gamma(t) \geqslant 1 - \frac{e^{Kt} - 1}{C_P(\mu) K}$$

uniformly over $\mathcal{B}$. Of course this is not very informative for large t.

In fact, for sampling algorithm, one is more interested in ergodic averages rather than marginal laws at given times. The bias of an MCMC estimator for a Lipschitz test function is typically bounded as

$$\left| \mathbb{E}_\nu \left(\frac{1}{t} \int_0^t f(X_s) \mathrm{d}s \right) - \mu f \right| \leqslant \|\nabla f\|_\infty \frac{1}{t} \int_0^t \mathcal{W}_2(\nu P_s, \mu) \mathrm{d}s$$

$$\leqslant \|\nabla f\|_\infty \mathcal{W}_2(\nu, \mu) \frac{1}{t} \int_0^t \gamma(s) .$$

Trying to minimize the right hand side by a suitable choice of drift in $\mathcal{B}$, in any cases it is not possible to get better than $C_P(\mu)/t + o(1/t)$ due to the constraint (5.2). Having a better contraction rate in large times is thus only useful if the estimator is $1/(t - t_0) \int_{t_0}^t f(X_s) \mathrm{d}s$ for a suitable warm-up time t_0.

Finally, notice that, in Theorem 1.2, $M \leqslant \sqrt{2}$, and thus $\lambda \leqslant 2T/C_P(\mu)$, which means that, for instance, the improvement of the L^2 decay rate from $T/C_P(\mu)$ to λ obtained by combining Theorems 1.2 and 1.3 is small.

Link with a Feynman-Kac eigenvalue problem

This section has to be credited to the anonymous referee who made the following remark: in Section 2.1, writing $u(x) = T + \alpha \omega(x)$, the proof works as soon as we find $\lambda > 0$ and a positive function u such that

$$(5.3) \qquad Lu(x) - \alpha k(x)u(x) \leqslant -\lambda u(x) \qquad \forall\, x \in \mathbb{R}^d \,,$$

which is (2.2), but in fact looking for u, λ such that equality holds in (5.3) is an eigenvalue problem for the operator $Lu - \alpha ku$, and the Krein-Rutman theorem (which holds under Assumption 1.1, see e.g. [CV23, Corollary 4.2]) states that in fact such an eigenpair u, λ always exist, with $u > 0$.

The remaining question is whether $\lambda > 0$. Since $f_t(x) = e^{-\lambda t} u(x)$ solves

$$\partial_t f_t = L f_t - \alpha k u \,,$$

we have for u a Feynman-Kac representation

$$u(x) = e^{\lambda t} \mathbb{E}_x \left(e^{-\alpha \int_0^t k(X_s) \mathrm{d}s} u(X_t) \right)$$

for any $t \geqslant 0$. Normalizing u to have $\mu(u) = 1$, using that u grows at most polynomially (again thanks to [CV23, Corollary 4.2] since under Assumption 1.1, $x \mapsto |x|^2$ is a Lyapunov function for L) and that μ has all its polynomial moment finite, we get by integrating the previous inequality with respect to μ and using the Hölder inequality that

$$1 \leqslant e^{-\lambda t} \left(\mathbb{E}_\mu \left(e^{-a\alpha \int_0^t k(X_s)\mathrm{d}s} \right) \right)^{1/a} \left(\mu(u^b) \right)^{1/b}$$

with $1/b + 1/a = 1$ for all $a > 1$, $t \geqslant 0$. This implies that

$$\lambda \geqslant -\limsup_{t \geqslant 0} \frac{1}{t} \ln \left(\mathbb{E}_\mu \left(e^{-a\alpha \int_0^t k(X_s)\mathrm{d}s} \right) \right)^{1/a},$$

for all $a > 1$, and thus for $a = 1$, namely

$$\lambda \geqslant \bar{\lambda} := -\limsup_{t \geqslant 0} \frac{1}{t} \ln \mathbb{E}_\mu \left(e^{-\alpha \int_0^t k(X_s)\mathrm{d}s} \right).$$

This quantity naturally appears when intertwining the semi-group P_t with the gradient, as in [ABJ18, CFG20, CJ13] (however, notice that in our case, we are just interested in the asymptotic exponential rate, and the expectation is with respect to the invariant measure μ instead of taking the supremum over all initial conditions $x \in \mathbb{R}^d$).

It remains to see whether $\bar{\lambda} > 0$. Notice that, by Jensen's inequality,

$$\bar{\lambda} \leqslant -\limsup_{t \geqslant 0} \frac{1}{t} \mathbb{E}_\mu \left(-\alpha \int_0^t k(X_s)\mathrm{d}s \right) = a\mu(k),$$

which means that having a positive mean curvature $\mu(k)$ is necessary to proceed with our proof based on the synchronous coupling.

Under Assumption 1.1, it is classically seen that $V(x) = e^{\delta|x|^2}$ is a Lyapunov function for L (in the sense that $LV \leqslant -rV$ outside some compact for some $r > 0$) as soon as $\delta < c/(2T)$. According to [DGW04, Theorem 2.3.], μ thus satisfies a T_1 Talagrand transport inequality (for the Euclidean distance on $\mathbb{R}^d$) with constant θT for some $\theta > 0$ independent from T (we refer to [DGW04] for definitions and details). Then, we can use [GLWY09, Corollary 2.4] (although it is written for reversible processes, this assumption is only used in its second part; the first part is a direct corollary of [GLWY09, Theorem 2.2] where reversibility is not assumed) to bound, for any $v > 0$,

$$\mathbb{P}_\mu \left(-\frac{\alpha}{t} \int_0^t k(X_{s/T})\mathrm{d}s \geqslant -\alpha\mu(k) + v \right) \leqslant \exp \left(-\frac{v^2 t}{4\theta T \alpha^2 \|k\|_{lip}^2} \right)$$

(where, for consistency with [GLWY09], we rescaled the process in time so that the corresponding carré du champ is $\Gamma(f) = |\nabla f|^2$, corresponding to the standard distance, instead of $T|\nabla f|^2$). Here we have to assume that k is a Lipschitz function, which is not a problem since our proof in Section 2.1 works if k is replaced by a lower bound of k.

Then, we follow the proof of [CFG20, Corollary 4.4]. Assuming that $\mu(k) > 0$, we introduce the event

$$\mathcal{A} = \left\{ -\frac{\alpha}{t} \int_0^t k(X_s) \mathrm{d}s \geqslant -\alpha\mu(k) + v \right\} = \left\{ -\frac{\alpha}{Tt} \int_0^{Tt} k(X_{s/T}) \mathrm{d}s \geqslant -\alpha\mu(k) + v \right\}$$

for some $v > 0$ to be chosen later on. Since k is lower bounded by $-K$,

$$\mathbb{E}_\mu \left(e^{-\alpha \int_0^t k(X_s) \mathrm{d}s} \right) \leqslant e^{\alpha K t} \mathbb{P}(\mathcal{A}) + e^{-\alpha\mu(k)t + vt}$$

$$\leqslant \exp \left(\alpha K t - \frac{v^2 t}{4\theta\alpha^2 \|k\|_{lip}^2} \right) + e^{-\alpha\mu(k)t + vt}.$$

Taking e.g. $v = \alpha\mu(k)/2$, we end up with

$$\lambda \geqslant \bar{\lambda} > 0$$

provided

$$\alpha K < \frac{(\mu(k))^2}{16\theta\|k\|_{lip}^2}.$$

Finally, as already mentioned, since we only need k to be a lower bound of the curvature, we can take it with a Lipschitz constant arbitrary small. However, this reduces $\mu(k)$, which can become negative. But then by assuming T is large enough we can make $\mu(k)$ positive again. More precisely: first, fix a lower bound $\widetilde{k}$ of the real curvature k defined by (1.2) (with $\inf \widetilde{k} = -K$ and $\widetilde{k}(x) = c$ for x large enough) with a Lipschitz constant $\|\widetilde{k}\|_{lip}$ sufficiently small so that

$$\alpha K < \frac{c^2}{64\theta\|\widetilde{k}\|_{lip}^2}.$$

(θ being related to μ, it is not affected by choice of $\widetilde{k}$). Then, since $\mu(\{x \in \mathbb{R}^d, |x| \geqslant R\}) \to 1$ for all $R > 0$ as $T \to \infty$, there exists $T_0 > 0$ such that for $T \geqslant T_0$, $\mu(\widetilde{k}) \geqslant c/2$, and thus $\lambda > 0$.

At the conclusion of this sketch of proof, we have thus obtained the following: under Assumption 1.1, for any $\alpha > 0$, there exists $T_0 > 0$ such that the operator $L - \alpha\widetilde{k}$ has an eigenpair (u, λ) with $u > 0$ and $\lambda > 0$ if $T \geqslant T_0$, where $\widetilde{k}$ is some lower bound of the curvature (which implies in particular that $Lu - \alpha ku \leqslant -\lambda u$).

Now, in order to conclude with a result similar to Theorem 1.2, it remains to do a bit of work on u, which we will not discuss here. Notice that, with an argument which starts with a non-explicit existence of the positive eigenfunction u, it is not necessarily easy to end up with explicit estimates as in Theorem 1.2 (in particular for the constant M). However this approach can be applied in a much more general framework, for instance for non-elliptic hypoelliptic diffusion processes. This question will be the topic of a future work. Moreover, this discussion suggests that the high-temperature regime is in fact a necessary condition for the synchronous coupling to contract distances for sufficiently large times under Assumption 1.1 (which is consistent with the remarks after Theorem 1.3).

BIBLIOGRAPHY

[ABJ18]   Marc Arnaudon, Michel Bonnefont, and Aldéric Joulin, *Intertwinings and generalized Brascamp–Lieb inequalities*, Rev. Mat. Iberoam. **34** (2018), no. 3, 1021–1054. ↑949, 950, 968

[BCG08]   Dominique Bakry, Patrick Cattiaux, and Arnaud Guillin, *Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré*, J. Funct. Anal. **254** (2008), no. 3, 727–759. ↑942, 943, 945, 965

[BGH21]   Fabrice Baudoin, Maria Gordina, and David P. Herzog, *Gamma Calculus Beyond Villani and Explicit Convergence Estimates for Langevin Dynamics with Singular Potentials*, Arch. Ration. Mech. Anal. **241** (2021), no. 2, 765–804. ↑964, 965

[BGL14]   Dominique Bakry, Ivan Gentil, and Michel Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften, vol. 348, Springer, 2014. ↑942, 943, 965

[BJ22]   Michel Bonnefont and Aldéric Joulin, *A note on eigenvalues estimates for one-dimensional diffusion operators*, Bernoulli **28** (2022), no. 1, 64–86. ↑949, 950

[BJM16]   Michel Bonnefont, Aldéric Joulin, and Yutao Ma, *Spectral gap for spherically symmetric log-concave probability measures, and beyond*, J. Funct. Anal. **270** (2016), no. 7, 2456–2482. ↑945, 949, 950

[Bob99]   Sergey G. Bobkov, *Isoperimetric and Analytic Inequalities for Log-Concave Probability Measures*, Ann. Probab. **27** (1999), no. 4, 1903–1921. ↑945

[CdRDM+21] Paul-Eric Chaudru de Raynal, Manh Hong Duong, Pierre Monmarché, Milica Tomašević, and Julian Tugaut, *Reducing exit-times of diffusions with repulsive interactions*, 2021, `https://arxiv.org/abs/2110.13230`. ↑958

[CFG20]   Patrick Cattiaux, Max Fathi, and Arnaud Guillin, *Self-improvement of the Bakry–Emery criterion for Poincaré inequalities and Wasserstein contraction using variable curvature bounds*, 2020, `https://arxiv.org/abs/2002.09221`. ↑943, 968, 969

[CG08]   Patrick Cattiaux and Arnaud Guillin, *Deviation bounds for additive functionals of Markov processes*, ESAIM, Probab. Stat. **12** (2008), 12–29. ↑942

[CG14a]   _____, *Functional inequalities via Lyapunov conditions*, Optimal Transport: Theory and Applications (Yann Ollivier, Hervé Pajot, and CedricEditors Villani, eds.), London Mathematical Society Lecture Note Series, vol. 413, Cambridge University Press, 2014, pp. 274—-287. ↑943, 945, 965

[CG14b]   _____, *Semi Log-Concave Markov Diffusions*, Séminaire de Probabilités XLVI (Catherine Donati-Martin, Antoine Lejay, and Alain Rouault, eds.), Lecture Notes in Mathematics, vol. 2123, Springer, 2014, pp. 231–292]. ↑943, 945

[CGZ13]   Patrick Cattiaux, Arnaud Guillin, and Pierre-André Zitt, *Poincaré inequalities and hitting times*, Ann. Inst. Henri Poincaré, Probab. Stat. **49** (2013), no. 1, 95–118. ↑953

[Che21]   Yuansi Chen, *An Almost Constant Lower Bound of the Isoperimetric Coefficient in the KLS Conjecture*, Geom. Funct. Anal. **31** (2021), 34–61. ↑945

[CJ13]   Djalil Chafaï and Aldéric Joulin, *Intertwining and commutation relations for birth–death processes*, Bernoulli **19** (2013), no. 5A, 1855–1879. ↑949, 950, 968

[CV23]   Nicolas Champagnat and Denis Villemonais, *General criteria for the study of quasi-stationarity*, Electron. J. Probab. **28** (2023), article no. 22. ↑967, 968

[DEE+21]   Alain Durmus, Andreas Eberle, Aurélien Enfroy, Arnaud Guillin, and Pierre Monmarché, *Discrete sticky couplings of functional autoregressive processes*, 2021, `https://arxiv.org/abs/2104.06771`. ↑946

[DGW04]   Hacène Djellout, Arnaud Guillin, and Liming Wu, *Transportation cost-information inequalities and applications to random dynamical systems and diffusions*, Ann. Probab. **32** (2004), no. 3B, 2702–2732. ↑968

[Dyn65]   Evgenij B. Dynkin, *Markov Processes. Volume II*, Springer, 1965. ↑952

[Ebe11]   Andreas Eberle, *Reflection coupling and Wasserstein contractivity without convexity*, C. R. Math. Acad. Sci. Paris **349** (2011), no. 19, 1101–1104. ↑943

[Ebe16]   _____, *Reflection couplings and contraction rates for diffusions*, Probab. Theory Relat. Fields **166** (2016), 851–886. ↑943

[EGZ19]   Andreas Eberle, Arnaud Guillin, and Raphael Zimmer, *Couplings and quantitative contraction rates for Langevin dynamics*, Ann. Probab. **47** (2019), no. 4, 1982–2010. ↑964

[EK86]   Stewart N. Ethier and Thomas G. Kurtz, *Markov processes. Characterization and convergence*, Wiley Series in Probability and Mathematical Statistics: Probability and mathematical statistics, John Wiley & Sons, 1986. ↑944

[EM19]   Andreas Eberle and Mateusz B. Majka, *Quantitative contraction rates for Markov chains on general state spaces*, Electron. J. Probab. **24** (2019), article no. 26. ↑964

[EZ19]   Andreas Eberle and Raphael Zimmer, *Sticky couplings of multidimensional diffusions with different drifts*, Ann. Inst. Henri Poincaré, Probab. Stat. **55** (2019), no. 4, 2370–2394. ↑946

[FG15]   Nicolas Fournier and Arnaud Guillin, *On the rate of convergence in Wasserstein distance of the empirical measure*, Probab. Theory Relat. Fields **162** (2015), no. 3-4, 707–738. ↑962, 963

[GLWY09]   Arnaud Guillin, Christian Léonard, Liming Wu, and Nian Yao, *Transportation-information inequalities for Markov processes*, Probab. Theory Relat. Fields **144** (2009), 669–695. ↑968

[GLWZ19]   Arnaud Guillin, Wei Liu, Liming Wu, and Chaoen Zhang, *Uniform Poincaré and logarithmic Sobolev inequalities for mean field particles systems*, 2019, `https://arxiv.org/abs/1909.07051`. ↑958, 959

[GM16]   Arnaud Guillin and Pierre Monmarché, *Optimal linear drift for the speed of convergence of an hypoelliptic diffusion*, Electron. Commun. Probab. **21** (2016), article no. 74. ↑966

[GM21]   _____, *Uniform Long-Time and Propagation of Chaos Estimates for Mean Field Kinetic Particles in Non-convex Landscapes*, J. Stat. Phys. **185** (2021), no. 2, article no. 15. ↑965

[HHMS93]   Chii-Ruey Hwang, Shu-Yin Hwang-Ma, and Shuenn-Jyi Sheu, *Accelerating Gaussian diffusions*, Ann. Appl. Probab. **3** (1993), 897–913. ↑966

[HHMS05]   _____, *Accelerating diffusions*, Ann. Appl. Probab. **15** (2005), no. 2, 1433–1444. ↑966

[HKS89]   Richard A. Holley, Shigeo Kusuoka, and Daniel W. Stroock, *Asymptotics of the spectral gap with applications to the theory of simulated annealing*, J. Funct. Anal. **83** (1989), no. 2, 333–347. ↑945

[HM08]   Martin Hairer and Jonathan C. Mattingly, *Spectral Gaps in Wasserstein Distances and the 2D Stochastic Navier–Stokes Equations*, Ann. Probab. **36** (2008), no. 6, 2050–2091. ↑964

[HM11]   _____, *Yet another look at Harris' ergodic theorem for Markov chains*, Seminar on Stochastic Analysis, Random Fields and Applications VI, Progress in Probability, vol. 63, Birkhäuser/Springer, 2011, pp. 109–117. ↑942, 952, 964

[Kuw10]   Kazumasa Kuwada, *Duality on gradient estimates and Wasserstein controls*, J. Funct. Anal. **258** (2010), no. 11, 3758–3774. ↑951

[Kuw13] _____, *Gradient estimate for Markov kernels, Wasserstein control and Hopf–Lax formula*, Potential theory and its related fields, RIMS Kôkyûroku Bessatsu, B43, Res. Inst. Math. Sci. (RIMS), Kyoto, 2013, pp. 61–80. ↑951

[LNP13] Tony Lelièvre, Francis Nier, and Grigorios A. Pavliotis, *Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion*, J. Stat. Phys. **152** (2013), no. 2, 237–274. ↑945, 966

[LV17] Yin Tat Lee and Santosh S. Vempala, *The Kannan–Lovász–Simonovits conjecture*, Papers based on selected lectures given at the current development mathematics conference, Harvard University, Cambridge, MA, USA, November 2017, Current Developments in Mathematics, International Press, 2017, pp. 1–36. ↑945

[LW16] Dejun Luo and Jian Wang, *Exponential convergence in L^p-Wasserstein distance for diffusion processes without uniformly dissipative drift*, Math. Nachr. **289** (2016), no. 14-15, 1909–1926. ↑943

[Mél96] Sylvie Méléard, *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, Probabilistic models for nonlinear partial differential equations, Springer, 1996, pp. 42–95. ↑960

[Mon23a] Pierre Monmarché, *Almost sure contraction for diffusions on Rd. Application to generalized Langevin diffusions*, Stochastic Processes Appl. **161** (2023), 316–349. ↑943, 946

[Mon23b] _____, *Elementary coupling approach for non-linear perturbation of Markov processes with mean-field jump mechanisms and related problems*, ESAIM, Probab. Stat. **27** (2023), 278–323. ↑964

[MS14] Georg Menz and André Schlichting, *Poincaré and logarithmic Sobolev inequalities by decomposition of the energy landscape*, Ann. Probab. **42** (2014), no. 5, 1809–1884. ↑945

[MT09] Sean Meyn and Robert L. Tweedie, *Markov chains and stochastic stability*, second ed., Cambridge Mathematical Library, Cambridge University Press, 2009, with a prologue by Peter W. Glynn. ↑942

[S05] Max-K. von Renesse and Karl-Theodor Sturm, *Transport inequalities, gradient estimates, entropy and Ricci curvature*, Commun. Pure Appl. Math. **58** (2005), no. 7, 923–940. ↑943

[RW10] Michael Röckner and Feng-Yu Wang, *Log-Harnack inequality for stochastic differential Equations in Hilbert spaces and its consequences*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **13** (2010), no. 1, 27–37. ↑953, 965

[Szn91] Alain-Sol Sznitman, *Topics in propagation of chaos*, Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989 (Paul-Louis Hennequin, ed.), Lecture Notes in Mathematics, Springer, 1991, pp. 165–251. ↑960

[Tug14] Julian Tugaut, *Phase transitions of McKean–Vlasov processes in double-wells landscape*, Stochastics **86** (2014), no. 2, 257–284. ↑958

[Wan97] Feng-Yu Wang, *Logarithmic Sobolev inequalities on noncompact Riemannian manifolds*, Probab. Theory Relat. Fields **109** (1997), 417–424. ↑965

[Wan12] _____, *Coupling and Applications*, Stochastic Analysis and Applications to Finance, Interdisciplinary Mathematical Sciences, vol. 13, World Scientific, 2012, pp. 411–424. ↑953

[Wan14] _____, *Modified Curvatures on Manifolds with Boundary and Applications*, Potential Anal. **41** (2014), 699–714. ↑965, 966

[Wan16] Jian Wang, *L^p-Wasserstein distance for stochastic differential equations driven by Lévy processes*, Bernoulli **22** (2016), no. 3, 1598–1616. ↑943

[Yos94]   Kosaku Yosida, *Functional analysis*, Classics in Mathematics, Springer, 1994, Reprint of the sixth (1980) edition. ↑952

Pierre MONMARCHÉ
Laboratoire Jacques-Louis Lions,
Sorbonne Université,
4 place Jussieu, 75 005 Paris, France
pierre.monmarche@sorbonne-universite.fr