

# Weighted $L^2$ -contractivity of Langevin dynamics with singular potentials

Evan Camrud<sup>1,\*</sup> , David P Herzog<sup>1</sup> , Gabriel Stoltz<sup>2</sup>   
and Maria Gordina<sup>3</sup> 

<sup>1</sup> Department of Mathematics, Iowa State University, Ames, IA, United States of America

<sup>2</sup> CERMICS, Ecole des Ponts, Marne-la-Vallée, France & MATHERIALS Team-Project, Inria Paris, France

<sup>3</sup> Department of Mathematics, University of Connecticut, Storrs, CT, United States of America

E-mail: [ecamrud@iastate.edu](mailto:ecamrud@iastate.edu), [dherzog@iastate.edu](mailto:dherzog@iastate.edu), [gabriel.stoltz@enpc.fr](mailto:gabriel.stoltz@enpc.fr) and [maria.gordina@uconn.edu](mailto:maria.gordina@uconn.edu)

Received 27 April 2021, revised 24 November 2021

Accepted for publication 8 December 2021

Published 29 December 2021



CrossMark

## Abstract

Convergence to equilibrium of underdamped Langevin dynamics is studied under general assumptions on the potential  $U$  allowing for singularities. By modifying the direct approach to convergence in  $L^2$  pioneered by Hérau and developed by Dolbeault *et al*, we show that the dynamics converges exponentially fast to equilibrium in the topologies  $L^2(d\mu)$  and  $L^2(W^* d\mu)$ , where  $\mu$  denotes the invariant probability measure and  $W^*$  is a suitable Lyapunov weight. In both norms, we make precise how the exponential convergence rate depends on the friction parameter  $\gamma$  in Langevin dynamics, by providing a lower bound scaling as  $\min(\gamma, \gamma^{-1})$ . The results hold for usual polynomial-type potentials as well as potentials with singularities such as those arising from pairwise Lennard-Jones interactions between particles.

Keywords: Langevin dynamics, Fokker–Planck equation, Convergence to equilibrium

Mathematics Subject Classification numbers: 60H10, 35Q84, 60J60 (Primary), 35B40 (Secondary).

## 1. Introduction

Langevin dynamics, and related stochastic differential equations dictating the evolution of physical systems at the atomistic scale, have long been a subject of interest in molecular dynam-

\*Author to whom any correspondence should be addressed.

Recommended by Dr Sandra Cerrai.

ics simulation [8, 18, 31, 36, 50] and more recently in machine learning [11, 37, 48, 52]. In these settings, estimates on the rate of convergence of the dynamics to equilibrium provide a rough measure for how long to run the corresponding algorithm in order to take random samples from the target Boltzmann–Gibbs distribution—as quantified by bounds on the asymptotic variance of averages over trajectories, dictated by a central limit theorem [6]. This has led to a number of works analysing the precise nature of this convergence under various assumptions on the potential  $U$ , both in the literature on statistical physics and in kinetic theory (where the typical dynamics is similar to the Fokker–Planck equation ruling the evolution of the law of the solutions of Langevin dynamics).

The first known result on the convergence of Langevin dynamics is due to Tropper [49], who proved that the dynamics mixes when the Hessian of the potential is bounded. Later, the works of Wu [53], Mattingly *et al* [40] and Talay [47] improved these results by proving convergence to equilibrium at an exponential rate for ‘polynomial-like’ potentials, in total variation. Their estimates relied on the existence of a Lyapunov function of the form  $H(q, p) + cq \cdot p$  for a small constant  $c > 0$ , where

$$H(q, p) = \frac{|p|^2}{2} + U(q) \quad (1)$$

denotes the Hamiltonian of the system. Additionally, in Talay’s work [47], exponential convergence in Sobolev spaces  $H^k(d\mu)$  for  $k \geq 1$  as well as in weighted Sobolev spaces was obtained.

These papers, as well as other seminal works such as [19, 26, 29, 42], later motivated the development of hypocoercivity by Villani [51] where, in the case of Langevin dynamics, exponential convergence to equilibrium was established in  $H^1(d\mu)$  for  $C^2(\mathbf{R}^d)$  potentials for which  $\mu$  satisfies a Poincaré inequality, and such that the global bound

$$|\nabla^2 U| \leq C(1 + |\nabla U|) \quad (2)$$

holds for some constant  $C > 0$ . There, ‘twisted gradients’ are used to obtain contractive properties in  $H^1(d\mu)$  which are ultimately transferred to  $L^2(d\mu)$  by hypoelliptic regularisation [25, 28]. Using a heuristic overdamped scaling analysis, and extending the convergence result by Hérau [27], Dolbeault *et al* [16, 17] established a convergence result similar to the one by Villani under the condition (2), but using a direct perturbative  $L^2(d\mu)$  approach—that is, instead of first obtaining a contraction in  $H^1(d\mu)$  and then transferring it to  $L^2(d\mu)$ , one can directly obtain the contraction in  $L^2(d\mu)$  by adding an appropriate perturbation to the norm.

Although the growth condition (2) is general, it does not allow for potentials with singularities. This is intuitively the case because the growth of the singular terms in  $U$  increases after taking derivatives, so that (2) cannot be satisfied. In this direction, the papers of Conrad and Grothaus [12] and Grothaus and Stilgenbauer [21] established a polynomial rate of convergence for time averages (along realisations of the stochastic dynamics) to ensemble averages (with respect to  $\mu$ ) for potentials that include Lennard–Jones interactions. Later, adopting a Lyapunov approach and denoting by  $\mathcal{O}$  the open domain on which the potential is finite, exponential convergence to equilibrium was obtained in [13, 30] in a weighted total variation distance for a class of potentials belonging to  $C^\infty(\mathcal{O})$  satisfying the general condition: for all  $\varepsilon > 0$ , there exists  $C_\varepsilon > 0$  such that

$$|\nabla^2 U| \leq \varepsilon |\nabla U|^2 + C_\varepsilon, \quad (3)$$

holds globally. In particular, it can be shown that this condition is satisfied by systems of  $N$  particles with positions  $q_1, \dots, q_N$ , interacting via pairwise singular potentials such as the

Lennard-Jones' potential

$$U(q) = \sum_{1 \leq i < j \leq N} \left( \frac{A_{ij}}{|q_i - q_j|^{12}} - \frac{B_{ij}}{|q_i - q_j|^6} \right),$$

or the Coulomb interaction in dimension  $d \geq 3$

$$U(q) = \sum_{1 \leq i < j \leq N} \frac{A_{ij}}{|q_i - q_j|^d};$$

see [4, appendix]. In particular, an explicit Lyapunov function of the form  $\exp(\delta H + \psi)$  was constructed in [13, 30], with  $\delta$  small and  $\psi$  an appropriate lower-order perturbation. See also [39] for an extension of this result to Coulomb potentials in spatial dimension  $d = 2$ . Nevertheless, explicit estimates on the convergence rate were not given, especially as they depend on key parameters of the dynamics, e.g. the friction  $\gamma > 0$  or the dimension  $d$ . This motivated the work [4] which aimed at getting an explicit dependence of the rate on the dimension  $d$ , by making use of a certain weighted  $H^1$  topology where the weight satisfies a Lyapunov-type condition. See also [10] for a related work.

The interest of working with the weighted topology  $L^2(W^* d\mu)$  is that functions in this space have a faster decay at infinity than functions in  $L^2(d\mu)$  (see remark 5 for a more precise discussion of this point). Exponential convergence in this topology then implies that the solution  $\phi$  to the Poisson equation  $-L\phi = g \in L^2(W^* d\mu)$  also belongs to  $L^2(W^* d\mu)$ , and not just to  $L^2(d\mu)$ . Since the solution  $\phi$  decays faster than typical functions in  $L^2(d\mu)$ , it is expected that spectral Galerkin discretizations such as the ones considered in [45] lead to smaller approximation errors, which turns out to be useful when solutions to Poisson equations are used in numerical methods, for instance to construct control variates [46].

In terms of the scalings in the prefactors in the convergence bounds, we obtain estimates uniform with respect to the friction, except in the weighted setting when  $\gamma \gg 1$ . This ‘break down’ makes sense intuitively because, when  $\gamma \gg 1$ , the Lyapunov function we construct is needed in a shrinking part of space as the noise dominates in a growing part of the domain. Better estimates could maybe be obtained by recent approaches which share some similarities with the direct  $L^2$  approach we use here, but do not rely on changing the underlying scalar product, either by using some space-time Poincaré inequality as in [1, 9], or by directly seeking estimates on the resolvent of the generator of the dynamics instead of looking for decay estimates of the evolution semigroup [8].

This paper is organised as follows. In section 2, we introduce the technical assumptions made throughout the paper and state the main results. Sections 3 and 4 contain the proofs of the main results assuming a key result, namely proposition 3. This proposition affords a number of formal manipulations there and is later proved in section 5 along with some important auxiliary estimates.

## 2. Mathematical setting and statement of the main result

### 2.1. Setting and basic assumptions

Throughout the paper, we study the Langevin stochastic differential equation

$$\begin{cases} dq_t = p_t dt, \\ dp_t = -\nabla U(q_t) dt - \gamma p_t dt + \sqrt{2\gamma} dB_t. \end{cases} \quad (4)$$

In the equation above, the unknowns are the vectors of positions  $q_t \in \mathbf{R}^d$  and momentum  $p_t \in \mathbf{R}^d$ . The parameter  $\gamma > 0$  is the friction constant while  $B_t$  is a standard  $d$ -dimensional Brownian motion defined on a probability space  $(\Omega, \mathcal{F}, \mathbf{P}, \mathbf{E})$ . The mapping  $U : \mathbf{R}^d \rightarrow [0, +\infty]$  is the potential energy function. Note that we allow  $U$  to take the value  $+\infty$  in the case when  $U$  is singular, as in the situation of a Lennard-Jones or Coulomb interaction force. Throughout,  $\mathcal{O}$  will denote the subset of  $\mathbf{R}^d$  of positions  $q$  with finite potential energy:

$$\mathcal{O} = \{q \in \mathbf{R}^d : U(q) < \infty\},$$

while  $\mathcal{X}$  denotes the corresponding state space

$$\mathcal{X} = \mathcal{O} \times \mathbf{R}^d$$

for the process  $x_t := (q_t, p_t)$ . Depending on the context, we will use either  $x$  or  $(q, p)$  to denote an element of  $\mathcal{X}$ .

In order guarantee the existence and uniqueness of pathwise solutions for all finite times  $t \geq 0$  and all initial conditions  $x \in \mathcal{X}$ , we make the following assumption on the potential  $U$ . Once one makes this assumption, the existence and uniqueness of pathwise solutions follows immediately by first applying the standard iteration procedure to get locally-defined solutions in time; and then extending these local solutions to global ones by using the Hamiltonian  $H$  as a basic type of Lyapunov function (see [4, section 2.2] or [34, 44] for further details).

**Assumption 1.** The potential  $U : \mathbf{R}^d \rightarrow [0, +\infty]$  satisfies the following properties:

- (a)  $U \in C^\infty(\mathcal{O}; [0, \infty))$ ,
- (b) The set  $\mathcal{O}$  is connected. Moreover, for any  $k \in \mathbf{N}$ , the open set

$$\mathcal{O}_k = \{q \in \mathbf{R}^d : U(q) < k\}$$

has a compact closure.

- (c) The integral  $\int_{\mathcal{O}} e^{-U(q)} dq$  is finite.
- (b) For any sequence  $\{q_k\} \subset \mathcal{O}$  with  $U(q_k) \rightarrow \infty$  as  $k \rightarrow \infty$ , it holds  $|\nabla U(q_k)| \rightarrow \infty$  as  $k \rightarrow \infty$ .

Under assumption 1, the solution of (4) is a Markov process and we denote by  $\{P_t\}_{t \geq 0}$  the associated Markov semigroup with transition kernel  $P_t(x, dy)$ . Letting  $\mathcal{B}$  be the Borel  $\sigma$ -field of subsets of  $\mathcal{X}$ , the Markov semigroup acts on bounded,  $\mathcal{B}$ -measurable functions  $\phi : \mathcal{X} \rightarrow \mathbf{R}$  as

$$(P_t \phi)(x) = \mathbf{E}_x[\phi(x_t)] = \int_{\mathcal{X}} \phi(y) P_t(x, dy),$$

where the expectation  $\mathbf{E}_x$  is with respect to all realisations of the Brownian motion in (4) with initial condition  $x_0 = x$ . The Markov semigroup also acts on positive, finite  $\mathcal{B}$ -measures  $\nu$  as

$$\forall B \in \mathcal{B}, \quad (\nu P_t)(B) = \int_{\mathcal{X}} \nu(dy) P_t(y, B).$$

Throughout this work, the infinitesimal generator of  $\{P_t\}_{t \geq 0}$  is denoted by

$$L = p \cdot \nabla_q - \nabla U(q) \cdot \nabla_p - \gamma p \cdot \nabla_p + \gamma \Delta_p. \quad (5)$$

Recalling the Hamiltonian  $H(q, p)$  of the system (1), the following  $\mathcal{B}$ -measure is a probability measure:

$$\mu(\mathrm{d}q \mathrm{d}p) = \frac{1}{\mathcal{N}} e^{-H(q,p)} \mathrm{d}q \mathrm{d}p, \quad \mathcal{N} = \int_{\mathcal{X}} e^{-H(q,p)} \mathrm{d}q \mathrm{d}p < +\infty, \quad (6)$$

the finiteness of  $\mathcal{N}$  being guaranteed by assumption 1(c). Moreover, it can be shown that  $\mu$  is the unique *invariant* probability measure under the dynamics (4); that is,  $\mu$  is the unique probability measure satisfying  $\mu P_t = \mu$ . Indeed, a short calculation shows that

$$L^\dagger(e^{-H}) = 0,$$

where  $L^\dagger$  denotes the formal adjoint of  $L$  with respect to the  $L^2(\mathrm{d}x)$  inner product. Employing assumption 1, one can then establish [30, proposition 2.5] by precisely the same argument given there. This, in particular, gives the needed hypoelliptic and support properties to ensure uniqueness of  $\mu$ ; see, for example, [35, 44].

In addition to employing assumption 1, we will also make use of the following growth assumption on the Hessian of  $U$ . This assumption together with assumption 1 allows for both polynomial and singular potentials [4, 30] and, by the arguments in [51], implies that  $\mu$  satisfies a Poincaré inequality, as discussed around proposition 2 below.

**Assumption 2.** For any  $\varepsilon > 0$ , there exists a constant  $C_\varepsilon \in \mathbb{R}_+$  such that

$$\forall q \in \mathcal{O}, \quad \forall y \in \mathbf{R}^d, \quad |\nabla^2 U(q)y| \leq \varepsilon |\nabla U(q)|^2 |y| + C_\varepsilon |y|. \quad (7)$$

**Remark 1.** There is a subtle but important difference between the bound in assumption 2 and an often assumed growth condition found in the literature on Langevin dynamics (see, for example, [9, 17, 51]). A typical replacement estimate for assumption 2 often reads: there exist constants  $C, D \in \mathbb{R}_+$  such that

$$\forall q \in \mathcal{O}, \quad |\nabla^2 U(q)| \leq C |\nabla U(q)| + D. \quad (8)$$

Such conditions are usually stated in order to control the growth of the potential at infinity, having in mind that the increase in  $U$  should be polynomial in  $|q|$ . This is why there is a power one for the gradient on the right-hand side of (8) while in assumption 2 this exponent becomes two at the expense of a small constant  $\varepsilon$ . This difference in the exponent precisely affords the inclusion of potentials with singularities. As a simple illustrative example, consider

$$\mathcal{O} = (0, \infty) \quad \text{and} \quad U(q) = q^{-1} + q^2.$$

Note that this potential combines a strong repulsion at  $q = 0$  with a quadratic confinement at infinity. Now, for  $q \geq 1$ , an estimate such as (8) holds since

$$\forall q \geq 1, \quad |U''(q)| \leq 3 \leq 3 |U'(q)| + 3.$$

However, near  $q = 0$ , this estimate no longer holds. Nevertheless, for any  $\varepsilon \in (0, 1)$ , there exists  $C_\varepsilon \in \mathbb{R}_+$  such that

$$\forall q \in (0, 1], \quad |U''(q)| = 2 \left( 1 + \frac{1}{q^3} \right) \leq \varepsilon U'(q)^2 + C_\varepsilon.$$

Since  $|U'(q)| \rightarrow \infty$  as  $U(q) \rightarrow \infty$ , we finally obtain (7) by distinguishing the cases  $q \geq 1$  and  $q \leq 1$ .

**Remark 2.** Assumption 2 is ‘almost minimal’ in the sense that, if  $\varepsilon > 0$  cannot be chosen as small as wanted, there are smooth functions  $U$  on  $\mathcal{O}$  satisfying the bound (7) but such that the unnormalised measure  $\tilde{\mu}$  given by

$$\tilde{\mu}(\mathrm{d}q \, \mathrm{d}p) = e^{-H(q,p)} \mathrm{d}q \, \mathrm{d}p$$

is no longer finite, hence cannot be normalised into a probability measure. As a concrete example, consider  $\mathcal{O} = (0, \infty)$  and any  $C^\infty(\mathcal{O})$  function  $U$  such that  $U(q) = \alpha \log(q)$  for  $q \geq 2$ , with  $\alpha > 0$ . Then,

$$\forall q \geq 2, \quad |U''(q)| = \frac{\alpha}{q^2} = \frac{1}{\alpha} |U'(q)|^2,$$

Note that  $\tilde{\mu}$  is no longer finite for  $\alpha \leq 1$  since  $e^{-U(q)} = q^{-\alpha}$  for  $q \geq 2$ .

## 2.2. Statement of the main results

In order to state the main results in this paper, we now introduce some further notation and terminology used throughout this work.

*Generators.* We first introduce the  $L^2(\mathrm{d}\mu)$ -adjoint  $L^*$  of  $L$ . More precisely, for the  $L^2(\mathrm{d}\mu)$ -inner product given by

$$\langle \phi, \psi \rangle = \int_{\mathcal{X}} \phi \psi \, \mathrm{d}\mu,$$

the action of the operator  $L^*$  is defined as

$$\forall \psi, \phi \in C_c^\infty(\mathcal{X}), \quad \langle L^* \phi, \psi \rangle = \langle \phi, L\psi \rangle,$$

where  $C_c^\infty(\mathcal{X})$  is the set of real-valued,  $C^\infty(\mathcal{X})$  functions with compact support in  $\mathcal{X}$ . It will sometimes be convenient to decompose  $L$  into its symmetric and anti-symmetric parts with respect to the inner product on  $L^2(\mathrm{d}\mu)$  as

$$L = L_H + \gamma L_{OU}, \tag{9}$$

where

$$L_H = p \cdot \nabla_q - \nabla U(q) \cdot \nabla_p, \quad L_{OU} = -p \cdot \nabla_p + \Delta_p.$$

Observe that the Hamiltonian part  $L_H$  of  $L$  is anti-symmetric, while the Ornstein–Uhlenbeck part  $L_{OU}$  of  $L$  is symmetric, so that

$$L^* = -L_H + \gamma L_{OU}.$$

In fact, introducing the momentum reversal operator  $\mathcal{R}$  which is a unitary operator on  $L^2(\mathrm{d}\mu)$  acting as  $(\mathcal{R}\phi)(q, p) = \phi(q, -p)$ , a simple computation shows that

$$L^* = \mathcal{R}L\mathcal{R}. \tag{10}$$

It is also useful to introduce the generator of the *overdamped Langevin* dynamics, namely

$$L_{OD} = \Delta_q - \nabla U(q) \cdot \nabla_q = -\nabla_q^* \cdot \nabla_q.$$

Note that this operator is the infinitesimal generator of the stochastic gradient system

$$dq_t = -\nabla U(q_t) dt + \sqrt{2} dW_t, \quad (11)$$

where  $W_t$  is a standard  $d$ -dimensional Brownian motion.

**Lyapunov functions.** In this paper, we construct a distance which is contractive (at large times) for the dynamics and equivalent to the norm on  $L^2(W^* d\mu)$  where  $W^* \geq 1$  is a conveniently chosen weight function. In fact,  $W^*$  will be some Lyapunov function for  $L^*$ , which is why we write  $W^*$  instead of  $W$ . We therefore introduce the following terminology.

**Definition 1.**

- (a) We call a function  $V \in C^2(\mathcal{X}; (0, \infty))$  *strongly integrable* if there exist constants  $C > 0$  and  $\delta \in (0, 1)$  such that the following estimate holds on  $\mathcal{X}$ :

$$V + |\nabla_p V| \leq C e^{(1-\delta)H}.$$

- (b) We call a strongly integrable function  $W \in C^2(\mathcal{X}; (0, \infty))$  a *weak Lyapunov function with respect to  $L$*  with constants  $\alpha, \beta > 0$  if

$$LW \leq -\alpha W + \beta.$$

- (c) We call a strongly integrable function  $W^* \in C^2(\mathcal{X}; (0, \infty))$  a *weak Lyapunov function with respect to  $L^*$*  with constants  $\alpha, \beta > 0$  if

$$L^*W^* \leq -\alpha W^* + \beta.$$

- (d) We call a weak Lyapunov function  $V \in C^2(\mathcal{X}; (0, \infty))$  with respect to  $M \in \{L, L^*\}$  a *strong Lyapunov function with respect to  $M$*  if the sub-level sets  $\{V \leq c\}$  for  $c \in (0, +\infty)$  are compact in  $\mathcal{X}$ .

**Remark 3.** Note that any positive constant function is a weak Lyapunov function with respect to both  $L$  and  $L^*$ . However, constant functions are not strong Lyapunov functions since their sub-level sets are not compact. Following the construction in [30], we show in theorem 2 that, under assumptions 1 and 2, there exists a strong Lyapunov function for  $L$ . Hence, by the structure of the generator  $L$ , there exists also a strong Lyapunov function for  $L^*$ .

Throughout, for any operator  $M \in \{L, L^*\}$ , we use the notation  $\mathcal{W}_{\alpha, \beta}(M)$  to denote the set of weak Lyapunov functions with respect to  $M$  with constants  $\alpha, \beta > 0$ , and

$$\mathcal{W}(M) = \bigcup_{\alpha, \beta > 0} \mathcal{W}_{\alpha, \beta}(M).$$

We similarly let  $\mathcal{S}_{\alpha, \beta}(M)$  denote the set of strong Lyapunov functions with respect to  $M$  with constants  $\alpha, \beta > 0$  and

$$\mathcal{S}(M) = \bigcup_{\alpha, \beta > 0} \mathcal{S}_{\alpha, \beta}(M).$$

The equality (10) implies a clear relationship between  $\mathcal{W}(L)$  and  $\mathcal{W}(L^*)$ , and hence between  $\mathcal{S}(L)$  and  $\mathcal{S}(L^*)$ , as summarised in the following proposition.

**Proposition 1.** *The function  $W$  is a weak (resp. strong) Lyapunov function for  $L$  with constants  $\alpha, \beta > 0$  if and only if  $W^*$  defined by  $W^*(q, p) = W(q, -p)$  is a weak (resp. strong) Lyapunov function with respect to  $L^*$  with constants  $\alpha, \beta > 0$ .*

**Remark 4.** The existence of a Lyapunov function is often assumed in the literature to establish geometric ergodicity for SDEs in a (weighted) total variation distance; see, for example, [24, 41, 44]. Using the one-to-one correspondence between  $\mathcal{W}(L)$  and  $\mathcal{W}(L^*)$  in our context, we will see here that contraction for the dynamics (4) in a weighted  $L^2(d\mu)$  sense also follows from the existence of a weak Lyapunov function.

**Remark 5.** One can always take our weak Lyapunov function to be a positive constant. However, for potentials satisfying assumptions 1 and 2 (see [4, 30, 39]), the set  $\mathcal{S}(L)$  is non-empty, hence so is  $\mathcal{S}(L^*)$  by this one-to-one correspondence. Consequently, the weight  $W^* \in \mathcal{S}(L^*)$  in the weighted norm  $L^2(W^* d\mu)$  actually diverges at large values of the Hamiltonian  $H$ , so the topology induced by  $L^2(W^* d\mu)$  is finer than the one induced by the usual (unweighted)  $L^2(d\mu)$  norm.

*Poincaré inequalities.* Because of the product structure of the invariant measure  $\mu$ , the  $q$ -marginal  $\mu_{\text{OD}}$  of  $\mu$  given by

$$\mu_{\text{OD}}(dq) = \frac{1}{\mathcal{N}_{\text{OD}}} e^{-U(q)} dq, \quad \mathcal{N}_{\text{OD}} = \int_{\mathcal{O}} e^{-U(q)} dq,$$

satisfies a Poincaré inequality if and only if  $\mu$  does, in view of tensorization results for Poincaré inequalities. We rely here on the fact that the other term in the product, the  $p$ -marginal of  $\mu$  denoted by  $\mu_{\text{OU}}$ , is a Gaussian measure with identity covariance, so that, for any  $\phi \in H^1(d\mu_{\text{OU}})$  with  $\int_{\mathbb{R}^d} \phi d\mu_{\text{OU}} = 0$ , it holds

$$\|\phi\|_{L^2(d\mu_{\text{OU}})}^2 \leq \|\nabla_p \phi\|_{L^2(d\mu_{\text{OU}})}^2. \quad (12)$$

Thus, in order to analyse constants later, we express the Poincaré constant for  $\mu$  in terms of the Poincaré inequality for  $\mu_{\text{OD}}$ . The following estimate is a consequence of the results stated in [3] for instance.

**Proposition 2.** Suppose that  $U$  satisfies assumptions 1 and 2. Then there exists a constant  $\rho > 0$  such that the following bound holds for all  $\phi \in H^1(d\mu_{\text{OD}})$  with  $\int_{\mathcal{O}} \phi d\mu_{\text{OD}} = 0$ :

$$\rho \int_{\mathcal{O}} \phi^2 d\mu_{\text{OD}} \leq \int_{\mathcal{O}} |\nabla \phi|^2 d\mu_{\text{OD}}.$$

When assumptions 1 and 2 hold,  $\mu$  therefore also satisfies a Poincaré inequality: for any  $\varphi \in H^1(d\mu)$  with  $\int_{\mathcal{X}} \varphi d\mu = 0$ ,

$$\min(1, \rho) \|\varphi\|_{L^2(d\mu)}^2 \leq \|\nabla_q \varphi\|_{L^2(d\mu)}^2 + \|\nabla_p \varphi\|_{L^2(d\mu)}^2. \quad (13)$$

*Modified norms.* For convenience in the arguments in later sections of the paper, we make use of the following notation:

$$\|\phi\|_T^2 = \int_{\mathcal{X}} [T\phi] \phi d\mu, \quad (14)$$

where  $T$  is a positive operator. When the argument  $T$  in the above is a nonnegative function, one should understand this as being a multiplication operator. We can further define an inner product  $\langle \cdot, \cdot \rangle_T$  associated to (14) by polarisation. When no argument is indicated in the norm,  $\|\phi\|$  denotes the usual norm on  $L^2(d\mu)$ .

*Exponential convergence in  $L^2(d\mu)$ .* Our first result establishes exponential convergence to equilibrium in the unweighted setting of  $L^2(d\mu)$  when the potential satisfies assumptions 1 and



2. In particular, the result applies in the setting of a singular Lennard-Jones interaction as in [30]. Note that the Lyapunov structure is not directly employed here.

**Theorem 1.** Suppose that the potential  $U$  satisfies assumptions 1 and 2. Then,

(a) For any  $\gamma > 0$ , there exist an explicit constant  $\lambda > 0$  such that, for any  $\phi \in L^2(d\mu)$  with  $\int_{\mathcal{X}} \phi d\mu = 0$ ,

$$\forall t \geq 0, \quad \|P_t \phi\| \leq 3 \exp(-\lambda t) \|\phi\|.$$

(b) There exists  $\bar{\lambda}$  such that the following lower bounds on the exponential decay rate holds:

$$\forall \gamma \in (0, +\infty), \quad \lambda \geq \bar{\lambda} \min\{\gamma, \gamma^{-1}\}.$$

*Exponential convergence in weighted  $L^2$  spaces.* Our next result shows that assumptions 1 and 2 imply explicit exponential convergence to equilibrium in a weighted topology constructed from the appropriate Lyapunov functional. The existence of a weak Lyapunov function follows by remark 3, but, as remarked earlier, we also have the existence of an explicit strong Lyapunov function  $W^*$  under assumptions 1 and 2. In fact, we construct explicit Lyapunov functions, which parametrically depend on the friction  $\gamma > 0$ , in order to obtain the correct scaling for the convergence rate.

**Theorem 2.** Suppose that the potential  $U$  satisfies assumptions 1 and 2.

(a) Consider  $W^* \in \mathcal{W}_{\alpha, \beta}(L^*)$  for some constants  $\alpha, \beta > 0$ . Let  $\lambda > 0$  be as in the conclusion of theorem 1 and set  $m = 5\eta\lambda/\beta$  for some  $\eta \in (0, 1)$ . Then, for any  $\phi \in L^2(W^* d\mu)$  with  $\int_{\mathcal{X}} \phi d\mu = 0$ ,

$$\forall t \geq 0, \quad \|P_t \phi\|_{mW^*+1} \leq 3 \exp\left(-\min\left\{\lambda(1-\eta), \frac{\alpha}{2}\right\} t\right) \|\phi\|_{mW^*+1}. \quad (15)$$

(b) For any  $\gamma > 0$  and  $\eta \in (0, 1)$ , there exist  $\alpha, \beta > 0$  and  $W^* \in \mathcal{S}_{\alpha, \beta}(L^*)$  of the form

$$W^* = \exp(\eta(H + \psi_\gamma)),$$

where  $\psi_\gamma = o(H)$  as  $H \rightarrow \infty$ . Moreover, there exist  $c > 0$  and  $C, D \in \mathbb{R}_+$  (which all depend on  $\eta$  but not on  $\gamma$ ) such that

$$\forall \gamma \leq 1, \quad \alpha \geq c\gamma, \quad \beta \leq C\gamma, \quad (16)$$

$$\forall \gamma \geq 1, \quad \alpha \geq c\gamma, \quad \beta \leq C\gamma^3 \mathcal{M}_\gamma, \quad \mathcal{M}_\gamma = \max_{|p| \leq D, |\nabla U| \leq D\gamma} W^*. \quad (17)$$

Note that the decay rate in (15) is slightly smaller than the one in (1). Somehow, the larger the Lyapunov function is (i.e. the larger  $m$  is), the smaller the decay rate is. Note also that if  $W^* \in \mathcal{W}(L^*) \setminus \mathcal{S}(L^*)$ , then theorem 2 does not provide additional information when compared with theorem 1. Essentially, one can think of  $W^* \in \mathcal{W}(L^*) \setminus \mathcal{S}(L^*)$  as being a positive constant, in which case the norms  $\|\cdot\|_{mW^*+1}$  and  $\|\cdot\|$  are equivalent. However, if  $W^* \in \mathcal{S}(L^*)$ , then  $W^*$  has compact sub-level sets  $\{W^* \leq c\}$  for all  $c > 0$ . This means that  $W^* \rightarrow \infty$  as  $H \rightarrow \infty$ , in which case the norm  $\|\cdot\|_{mW^*+1}$  dominates the norm  $\|\cdot\|$  (see also remark 5).

Before proceeding further, let us discuss more precisely the convergence result provided by theorem 2 through some remarks.

**Remark 6.** Item (b) of theorem 2 implies that the effective exponential decay rate  $\min\{\lambda(1-\varepsilon), \alpha/2\}$  in (15) is still of order  $\min\{\gamma, \gamma^{-1}\}$ , as in theorem 1. When  $\gamma > 0$  is small, we also note that  $m$  is of order 1, so that the norm  $\|\cdot\|_{mW^*+1}$  is genuinely stronger than the

standard norm on  $L^2(d\mu)$ . On the other hand, when  $\gamma > 0$  is large, the factor  $m$  is small in  $\gamma$ , of order  $\gamma^{-4} \mathcal{M}_\gamma^{-1}$ . To estimate  $\mathcal{M}_\gamma$ , one needs more knowledge than assumptions 1 and 2 in order to compare  $|\nabla U|$  and  $U$ . One should however typically think of  $\mathcal{M}_\gamma$  as being exponentially large in some power of  $\gamma$  as  $\gamma \rightarrow \infty$ . The poor scaling of this constant with respect to  $\gamma$  is not surprising given that the construction of the Lyapunov function in the previous result is based on the analysis of the dynamics at large energies, outside the region  $\{|p| \leq D\} \cap \{|\nabla U| \leq D\gamma\}$ . Note that this region grows as  $\gamma \rightarrow \infty$  so that the analysis works in a shrinking part of the phase space.

**Remark 7.** Following the calculations in [4], it is possible to use the methods of this paper to marginally improve the dimensionality dependence in  $\lambda$  of theorem 1. In particular, it is possible to replace assumption 2 with the condition: there exists  $\varepsilon > 0$  small enough but independent of the dimension such that condition (7) is satisfied. Indeed, note from the proof of proposition 2.40 in the appendix of [4] that it is possible to choose  $\varepsilon$  in assumption 2 independent of the dimension for Lennard-Jones-like potentials. This can be done by carefully choosing the dimension dependence of the constants  $C_1, C_2$  in the latter work as functions of  $d$ . However, the resulting constant  $C_\varepsilon$  in (2) will diverge as  $d^a$  (the exponent  $a$  being determined by the choices of  $C_1, C_2$ ). It is possible to remedy to this issue by rescaling variables as  $q \mapsto d^{-a/2}q$ , so that  $|d^a \nabla^2 U(q)y| \leq |d^{a/2} \nabla U(q)|^2 |y| + C_\varepsilon$ , which leads to  $|\nabla^2 U(q)y| \leq |\nabla U(q)||y| + d^{-a}C_\varepsilon$ . Hence, under this spatial rescaling, both  $\varepsilon$  and  $C_\varepsilon$  can be chosen independently of the dimension.

Although the scaling of  $\varepsilon$  and  $C_\varepsilon$  in assumption 2 can be made precise and controlled, the dependence of the Poincaré constant  $\rho$  with respect to the dimension is less clear. Dimension-free Poincaré constants are obtained only in unrealistic situations, for instance particles not interacting with each other (in which case the potentials is separable, namely  $U(q) = \sum_{i=1}^d u(q_i)$ , and the Poincaré constant is simply the minimum of the Poincaré constants associated with the potentials  $u_i$ ), or uniformly convex potentials and perturbations thereof. This makes it difficult to determine the dimension dependence of the convergence rate  $\lambda$ , which is based on the constants in (32).

**Remark 8.** Note that  $\alpha$  and  $\lambda$  have different scalings with respect to  $\gamma > 0$  in part (b) when  $\gamma > 0$  is large. In particular,  $\alpha = \mathcal{O}(\gamma)$  while  $\lambda = \mathcal{O}(\gamma^{-1})$  as  $\gamma \rightarrow \infty$ . To see heuristically why one should expect this discrepancy, note that in the region

$$\{|p| \geq D\} \cup \{|\nabla U| \geq D\gamma\}$$

for  $D > 0$  large, but independent of  $\gamma$ , the effects of the noise in the system (4) can heuristically be considered to be negligible. Furthermore, suppose that at high energy levels the system (4) moves approximately to leading order in time according to the deterministic Hamiltonian dynamics

$$\begin{cases} \dot{Q} = P, \\ \dot{P} = -\nabla U(Q). \end{cases} \quad (18)$$

This is heuristically the case if, at higher and higher energy levels, the periods of the ‘orbits’ of (18) (should such orbits even exist) are short compared to  $\gamma^{-1}$ . Introduce

$$\langle P^2 \rangle_h = \frac{1}{T_h} \int_0^{T_h} |P_s|^2 ds,$$

where  $T_h$  denotes the time spent during one complete Hamiltonian orbit for the dynamics (18) on  $\{(Q, P) : H(Q, P) = h\}$ . If we believe for large  $h$  that

$$\langle P^2 \rangle_h \approx ch$$

for some  $c > 0$ , then we obtain for  $H(q, p) = h$  large that

$$\begin{aligned} \frac{1}{T_h} \mathbf{E}_{(q,p)} [H(q_{T_h}, p_{T_h})] &= \frac{1}{T_h} H(q, p) - \frac{\gamma}{T_h} \mathbf{E}_{(q,p)} \left[ \int_0^{T_h} |p_s|^2 ds \right] + \gamma d \\ &\approx \frac{1}{T_h} H(q, p) - \frac{c\gamma}{T_h} \int_0^{T_h} H(q_s, p_s) ds + \gamma d. \end{aligned}$$

Since intuitively the time  $T_h$  is small for  $H = h$  large, we expect some Lyapunov-like condition on  $H$ , which suggests that  $\alpha = \gamma c$  for some  $c > 0$  independent of  $\gamma$  when  $\gamma$  is large.

**Remark 9.** We motivate here the scaling  $\min\{\alpha, \lambda\}$  for the convergence rate in (15). Consider  $\eta > 0$  and suppose that  $W^* \in \mathcal{S}_{(1+\eta)\alpha, \beta}(L^*)$  for some  $\alpha, \beta > 0$  with  $W^* \geq 1$ . Then there exists  $K^* \subseteq \mathcal{X}$  compact such that

$$\forall (q, p) \in (K^*)^c, \quad (L^* W^*)(q, p) \leq -\alpha W^*(q, p).$$

In view of (10), the function  $W$  defined by  $W(q, p) := (\mathcal{R}W^*)(q, p) = W^*(q, -p)$  satisfies an analogous estimate for  $L$  on the set  $K = \{(q, p) : (q, -p) \in K^*\}$ , namely

$$\forall (q, p) \in K^c, \quad (LW)(q, p) \leq -\alpha W(q, p).$$

Let  $\tau = \inf\{t \geq 0 : (q_t, p_t) \in K\}$  be the first time the process  $(q_t, p_t)$  solving (4) enters  $K$ . Since  $W^* \geq 1$  and hence  $W \geq 1$  as well, the following equality follows from Itô's formula applied to  $\psi(t, q_t, p_t) = e^{\alpha t} W(q_t, p_t)$  started in  $K^c$  and stopped at time  $\tau$ :

$$\forall (q, p) \in K^c, \quad \mathbf{E}_{(q,p)} [\exp(\alpha\tau)] \leq W(q, p) < \infty. \quad (19)$$

One way to interpret the above exponential moment estimate on  $\tau$  is that the speed at which the process  $(q_t, p_t)$  returns to  $K$  is governed by the parameter  $\alpha$ . Outside of this compact set, the dynamics returns to  $K$  exponentially fast on average as dictated by the value  $\alpha > 0$  for which the Laplace transform of  $\tau$  in (19) is finite. Once the process enters  $K$ , mixing occurs according to the local topology given by the norm  $\|\cdot\|$ . This suggests in particular the scaling  $\min\{\alpha, \lambda\}$  in the exponential convergence rate of theorem 2.

**Remark 10.** Some convergence results for Fokker–Planck operators associated with Langevin dynamics can be extended to other types of generators, in particular generators associated with piecewise deterministic Markov processes [2, 17, 38]. The generator of the linear Boltzmann dynamics corresponds to replacing the differential operator  $L_{OU}$  in (9) by the integral operator  $\Pi - 1$ . An inspection of the proof of theorem 1 shows that this result will still hold, at least formally. Some estimates are unchanged, such as most of the bounds in sections 5.1 and 5.2, which involve either functions of the  $q$  variable only or functions in the image of the generator of the Hamiltonian part of the dynamics. Some work would however be required in the proof of proposition 3 in section 5.3 to make all estimates rigorous since hypoellipticity is lost, and some manipulations based on truncations and carré-du-champ formulas would have to be adapted. In contrast, it is not clear whether the results of theorem 2 hold. The issue there is to find a Lyapunov function. In the framework of section 4.2.1, this is done in [7, lemma 3.2] for a Lyapunov function quadratic in  $p$  (although the scaling

with respect to the equivalent of the parameter  $\gamma$  would have to be made explicit). It is not obvious how to extend the approach to Lyapunov functions such as (43) since our algebraic manipulations in section 4.2.2 rely on the fact that we work with second order differential operators.

### 3. Proof of the main general results

We prove in this section the main general results of this paper, namely theorem 1 part (a) and theorem 2 part (a). These results are established by assuming some technical estimates, whose proofs are postponed to section 5. The analysis of the claimed scalings with respect to the friction parameter  $\gamma$ , as outlined in the statements of theorem 1 part (b) and theorem 2 part (b), are studied in a second stage, in section 4.

We first motivate and discuss the main ideas behind the proofs in sections 3.1 and 3.2. In these two motivating subsections, we make a number of formal manipulations in order to simplify the presentation, without justification, using generic ‘nice’ functions  $\phi \in L^2(d\mu)$  with mean zero with respect to  $\mu$ . The needed manipulations, as well as what is meant by a ‘nice’  $\phi \in L^2(d\mu)$  with mean zero, are made precise in the proofs of the main results in section 3.3.

#### 3.1. Idea of proof in the unweighted setting (theorem 1)

The first observation we make is that working in the topology  $L^2(d\mu)$  is a good start in order to see some elements of a contraction. Indeed, recalling that

$$2\phi L\phi = L(\phi^2) - 2\gamma|\nabla_p\phi|^2, \quad (20)$$

we have, by invariance of  $\mu$ , that

$$\frac{d}{dt}\|P_t\phi\|^2 = 2\langle LP_t\phi, P_t\phi \rangle = \langle L(P_t\phi)^2, 1 \rangle - 2\gamma\|\nabla_p P_t\phi\|^2 = -2\gamma\|\nabla_p P_t\phi\|^2. \quad (21)$$

The Poincaré inequality (13) satisfied by the measure  $\mu$  cannot be used at this stage since only  $p$ -derivatives appear in (21) through  $-2\gamma\|\nabla_p P_t\phi\|^2$  above. We need to uncover the ‘missing’  $q$ -derivatives, namely  $-\|\nabla_q P_t\phi\|^2$ . This is done by adding a small perturbation to the norm in  $L^2(d\mu)$  in order to couple the  $q$  and  $p$  degrees of freedom, and spread the dissipation from  $p$  to  $q$ .

This perturbation is encoded in practice by some operator  $A$ , following the approach of [17, 27]. We introduce

$$\begin{aligned} A &= -(1 + (L_H\Pi)^*(L_H\Pi))^{-1}(L_H\Pi)^* = (1 - \Pi L_H^2\Pi)^{-1}\Pi L_H \\ &= (1 - L_{OD}\Pi)^{-1}\Pi L_H, \end{aligned} \quad (22)$$

where  $\Pi$  is the projection operator on  $L^2(d\mu)$  given by

$$(\Pi\phi)(q) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbf{R}^d} \phi(q, p) e^{-\frac{|p|^2}{2}} dp. \quad (23)$$

Effectively, the perturbation  $A$  plays a role similar to the twisted gradient as in [42, 47, 51], but it has been renormalised to be an operator on  $L^2(d\mu)$  as opposed to  $H^1(d\mu)$ . Thus, when we uncover the missing  $q$ -derivatives, they appear in renormalised form.

**Remark 11.** Strictly speaking, the operator  $A$  is defined on a dense domain in  $L^2(d\mu)$ . Once we prove that it is bounded operator in  $L^2(d\mu)$  when restricted to this dense domain, we can extend it to an operator on all of  $L^2(d\mu)$  with the same norm; see section 5, in particular lemma 3.

Recalling the notation (14), we next consider the modified ‘norm’  $\|\cdot\|_{1+\delta A}$  with  $\delta > 0$  a parameter to be determined—chosen in particular so that  $\|\cdot\|_{1+\delta A}$  is indeed a norm, equivalent to the standard norm  $\|\cdot\|$ . Then, using (21),

$$\begin{aligned} \frac{d}{dt} \|P_t \phi\|_{1+\delta A}^2 &= \frac{d}{dt} \|P_t \phi\|^2 + \delta \frac{d}{dt} \langle A P_t \phi, P_t \phi \rangle \\ &= -2\gamma \|\nabla_p P_t \phi\|^2 + \delta \langle A L P_t \phi, P_t \phi \rangle + \delta \langle A P_t \phi, L P_t \phi \rangle \\ &= -2\gamma \|\nabla_p P_t \phi\|^2 + \delta \langle A L_H \Pi P_t \phi, P_t \phi \rangle + \delta \langle A L_H (1 - \Pi) P_t \phi, P_t \phi \rangle \\ &\quad + \delta \gamma \langle A L_{OU} P_t \phi, P_t \phi \rangle + \delta \langle L^* A P_t \phi, P_t \phi \rangle \\ &=: -2\gamma \|\nabla_p P_t \phi\|^2 + \delta [T_1(P_t \phi) + T_2(P_t \phi) + T_3(P_t \phi) + T_4(P_t \phi)]. \quad (24) \end{aligned}$$

We now highlight that the term  $T_1(P_t \phi)$  provides the missing dissipation in  $q$ , while we will see later on, in section 3.3, that  $T_i(P_t \phi)$  for  $i = 2, 3, 4$  are ‘lower-order’ order terms in a sense to be made precise.

Note first that, for  $\phi \in L^2(d\mu)$  with  $\int_{\mathcal{X}} \phi d\mu = 0$ , one has  $\int_{\mathcal{X}} P_t \phi d\mu = \int_{\mathcal{O}} \Pi P_t \phi d\mu_{OD} = 0$ . We can therefore use the Poincaré inequality for the overdamped measure  $\mu_{OD}$  (see proposition 2) to obtain

$$\begin{aligned} T_1(P_t \phi) &= \langle A L_H \Pi P_t \phi, P_t \phi \rangle = \langle (1 - L_{OD} \Pi)^{-1} L_{OD} \Pi P_t \phi, \Pi P_t \phi \rangle \\ &= - \langle (1 + \nabla_q^* \cdot \nabla_q \Pi)^{-1} \nabla_q^* \cdot \nabla_q \Pi P_t \phi, \Pi P_t \phi \rangle \\ &\leq - \frac{\rho}{\rho + 1} \|\Pi P_t \phi\|^2. \quad (25) \end{aligned}$$

Thus, because the operator  $\Pi$  is playing the role of a renormalised gradient in  $q$ , the term on the last line above can be combined with  $-2\gamma \|\nabla_p P_t \phi\|^2 \leq -2\gamma \|(1 - \Pi) P_t \phi\|^2$ , where the last inequality follows from the Poincaré inequality (12) in the  $p$ -marginal  $\mu_{OU}$ . This leads to the sought after global dissipation and allows to apply a Gronwall lemma.

For the approach to be effective, we need to carefully choose  $\delta > 0$  so that (a)  $\|\cdot\|_{1+\delta A}$  is actually a norm, equivalent to the standard one on  $L^2(d\mu)$ , (b) the remainder terms  $T_i(P_t \phi)$  for  $i = 2, 3, 4$  can be controlled, and (c) the decay rate has the correct scaling in  $\gamma$  as claimed in theorem 1(b). The originality of our results compared to related results for Langevin dynamics [15, 22, 33] is that we consider weaker conditions on  $U$  than in previous works, allowing in particular the possibility of potentials with singularities.

### 3.2. Idea of the proof for a weighted norm (theorem 2)

Building off the heuristics in the previous section, we now outline some of the differences from the above when we switch to the weighted topology  $L^2(W^* d\mu)$  where  $W^* \in \mathcal{W}_{\alpha, \beta}(L^*)$  for some  $\alpha, \beta > 0$ . In fact, if one makes good choices in the construction of the norm and uses the mathematical setup of section 2, then the result in the weighted setting, in particular

theorem 2, readily follows from the previous analysis. Again, we proceed here formally using ‘nice’ functions with mean zero.

Recalling the notation (14), as well as (20), we obtain for  $W^* \in \mathcal{W}_{\alpha,\beta}(L^*)$  that

$$\frac{d}{dt} \|P_t \phi\|_{W^*}^2 = 2 \langle LP_t \phi, P_t \phi \rangle_{W^*} = \langle (P_t \phi)^2, L^* W^* \rangle - 2\gamma \|\nabla_p P_t \phi\|_{W^*}^2.$$

Using the fact that  $L^* W^* \leq -\alpha W^* + \beta$ , we thus arrive at the inequality

$$\frac{d}{dt} \|P_t \phi\|_{W^*}^2 \leq -2\gamma \|\nabla_p P_t \phi\|_{W^*}^2 - \alpha \|P_t \phi\|_{W^*}^2 + \beta \|P_t \phi\|^2. \quad (26)$$

Somewhat similar to the unweighted setting, we are faced with the problem that, although we have a globally dissipative factor, namely  $-\alpha \|P_t \phi\|_{W^*}^2$ , we have introduced a growth factor, namely  $\beta \|P_t \phi\|^2$ . Thus we need to somehow add a perturbation to the norm that compensates for this term.

In what follows, we perturb the norm  $\|\cdot\|_{W^*}$  by  $1 + \delta A$ ; that is, we consider  $\|\cdot\|_{W^*+1+\delta A}$  where  $A$  is as in (22) and, provided it exists following the analysis in section 3.1,  $\delta > 0$  satisfies the following two properties:

- (a)  $\|\cdot\|_{1+\delta A}$  and  $\|\cdot\|$  are equivalent norms;
- (b) There exists  $\lambda > 0$  for which the following estimate holds for all  $\phi$  in a ‘nice’ dense family in  $L^2(d\mu)$  with  $\int_{\mathcal{X}} \phi \, d\mu = 0$ :

$$\forall t \geq 0, \quad \frac{d}{dt} \|P_t \phi\|_{1+\delta A}^2 \leq -2\lambda \|P_t \phi\|_{1+\delta A}^2.$$

If such a  $\delta > 0$  exists, we then find that

$$\begin{aligned} \frac{d}{dt} \|P_t \phi\|_{W^*+1+\delta A}^2 &= \frac{d}{dt} \|P_t \phi\|_{W^*}^2 + \frac{d}{dt} \|P_t \phi\|_{1+\delta A}^2 \\ &\leq -\alpha \|P_t \phi\|_{W^*}^2 + \beta \|P_t \phi\|^2 - 2\lambda \|P_t \phi\|_{1+\delta A}^2. \end{aligned}$$

In order to make sure that  $\lambda > 0$  is large enough to subsume the  $\beta$  term, we rescale the weak Lyapunov function by a factor  $m > 0$  to be determined, so that  $mW^* \in \mathcal{W}_{\alpha,m\beta}(L^*)$ . Then, repeating the above produces

$$\frac{d}{dt} \|P_t \phi\|_{mW^*+1+\delta A}^2 \leq -\alpha \|P_t \phi\|_{mW^*}^2 + m\beta \|P_t \phi\|^2 - 2\lambda \|P_t \phi\|_{1+\delta A}^2.$$

Picking  $m > 0$  small enough depending on  $\lambda, \delta > 0$  allows to conclude the argument.

### 3.3. Proof of parts (i) of theorems 1 and 2

In order to prove the main general results, we first make precise the class of sufficiently smooth test functions  $\phi : \mathcal{X} \rightarrow \mathbf{R}$  we consider. This allows us to more easily manipulate expressions to arrive at the desired inequalities, and then apply density arguments to see that the inequalities are satisfied for a larger class of functions. More precisely, we define

$$C_{c,0}^\infty(\mathcal{X}) = \left\{ \phi \in C^\infty(\mathcal{X}) \mid \phi = \psi - \int_{\mathcal{X}} \psi \, d\mu \quad \text{for some } \psi \in C_c^\infty(\mathcal{X}) \right\}.$$

We note that a function in  $C_{c,0}^\infty(\mathcal{X})$  does not have compact support in general, since it is generally constant and non-zero outside of a compact set. We observe that, by assumption 1 and Hörmander's theorem [32], hypoellipticity implies that the function  $P_t\phi$  belongs to  $C_{b,0}^\infty(\mathcal{X})$  for  $t > 0$  whenever  $\phi \in C_{c,0}^\infty(\mathcal{X})$ , where

$$C_{b,0}^\infty(\mathcal{X}) = \left\{ \phi \in C^\infty(\mathcal{X}) \mid \sup_{x \in \mathcal{X}} |\phi(x)| < +\infty, \int_{\mathcal{X}} \phi \, d\mu = 0 \right\}.$$

In order to conclude theorems 1(a) and 2(a), we assume that proposition 3 below holds true. This result provides some integrated form of dissipation. The proof of this proposition is postponed to section 5 since it relies on a number of technical estimates and commutations of operators. The approach for obtaining it, however, is exactly the same as in [17], but there are some key differences in our setting, especially in the proof of the ‘elliptic regularity estimate’, which make the estimates slightly more involved as we are allowing a weaker growth condition on  $U$  than in [17].

In order to state the proposition, observe that assumption 2 implies that there exist constants  $c_1 \in (0, 1)$  and  $C_2 > 0$  independent of  $\gamma$  such that

$$\Delta U(q) \leq c_1 |\nabla U(q)|^2 + C_2. \quad (27)$$

Consider  $\varepsilon > 0$  such that

$$0 < \varepsilon < \frac{(1 - c_1)^2}{4}, \quad (28)$$

and let  $C_\varepsilon$  be the corresponding constant as provided by assumption 2. Finally, define the constant  $\eta_\varepsilon > 0$  as

$$\eta_\varepsilon = \sqrt{2 \left( 1 - \frac{4\varepsilon}{(1 - c_1)^2} \right)^{-1} \max \left\{ 1, \frac{1}{2} \left( C_\varepsilon + \frac{2C_2\varepsilon}{1 - c_1} \right) \right\}}. \quad (29)$$

**Proposition 3.** Suppose that  $U$  satisfies assumptions 1 and 2, and consider  $W^* \in \mathcal{W}_{\alpha,\beta}(L^*)$  for some  $\alpha, \beta > 0$ . Then, the following properties hold.

(a) The operator  $A$  is bounded, and satisfies the estimate:

$$\forall \phi \in L^2(d\mu), \quad \|A\phi\| \leq \frac{1}{2} \|(1 - \Pi)\phi\|.$$

(b) For any  $\phi \in C_{c,0}^\infty(\mathcal{X})$  and  $0 \leq s \leq t$ ,

$$\begin{aligned} \langle AP_t\phi, P_t\phi \rangle - \langle AP_s\phi, P_s\phi \rangle + \int_s^t \frac{\rho}{\rho + 1} \|\Pi P_u\phi\|^2 \, du \\ \leq \int_s^t \left[ \left( \eta_\varepsilon + \frac{\gamma}{2} \right) \|(1 - \Pi)P_u\phi\| \|\Pi P_u\phi\| + \|(1 - \Pi)P_u\phi\|^2 \right] \, du, \end{aligned}$$

where the constant  $\eta_\varepsilon > 0$  is defined in (29).

(c) For any  $\phi \in C_{c,0}^\infty(\mathcal{X})$  and  $0 \leq s \leq t$ ,

$$\|P_t\phi\|_{W^*}^2 \leq \|P_s\phi\|_{W^*}^2 - \int_s^t [\alpha \|P_u\phi\|_{W^*}^2 - \beta \|P_u\phi\|^2 + 2\gamma \|\nabla_p(P_u\phi)\|_{W^*}^2] \, du.$$

Note that, if  $W^* \equiv 1$ , it is possible to take  $\alpha = \beta$  in the estimate of part (c), which then corresponds to the integral version of the simple decay estimate (21).

With proposition 3 at hand, we can now complete the proofs of theorem 1(i) and 2(i).

**Proof of theorem 1(i)** Let  $\phi \in C_{c,0}^\infty(\mathcal{X})$  and  $\delta > 0$ . Applying propositions 3(b) and (c) (with  $W^* \equiv 1$  and  $\alpha = \beta$  in part (c)), we have, for  $0 \leq s \leq t$ ,

$$\begin{aligned} \|P_t \phi\|_{1+\delta A}^2 &\leq \|P_s \phi\|_{1+\delta A}^2 - \left( \int_s^t 2\gamma \|\nabla_p(P_u \phi)\|^2 + \frac{\delta \rho}{\rho+1} \|\Pi P_u \phi\|^2 du \right) \\ &\quad + \delta \int_s^t \left[ \left( \eta_\varepsilon + \frac{\gamma}{2} \right) \|(1-\Pi)P_u \phi\| \|\Pi P_u \phi\| + \|(1-\Pi)P_u \phi\|^2 \right] du \\ &=: \|P_s \phi\|_{1+\delta A}^2 - \int_s^t S_1(P_u \phi) du + \delta \int_0^t S_2(P_u \phi) du. \end{aligned}$$

Next, observe that, for any  $\phi \in C_{b,0}^\infty(\mathcal{X})$ , the Poincaré inequality (12) gives

$$S_1(\phi) = 2\gamma \|\nabla_p \phi\|^2 + \frac{\delta \rho}{\rho+1} \|\Pi \phi\|^2 \geq 2\gamma \|(1-\Pi)\phi\|^2 + \frac{\delta \rho}{\rho+1} \|\Pi \phi\|^2.$$

Combining this estimate with the other term  $S_2(\phi)$  produces

$$S_1(\phi) - \delta S_2(\phi) \geq X^T \mathbf{B} X, \quad (30)$$

with

$$X = \begin{pmatrix} \|(1-\Pi)\phi\| \\ \|\Pi \phi\| \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2\gamma - \delta & -\frac{\delta}{2} \left( \eta_\varepsilon + \frac{\gamma}{2} \right) \\ -\frac{\delta}{2} \left( \eta_\varepsilon + \frac{\gamma}{2} \right) & \frac{\delta \rho}{1+\rho} \end{pmatrix}.$$

Therefore,

$$S_1(\phi) - S_2(\phi) \geq \min\{\Lambda_+, \Lambda_-\} \|\phi\|^2,$$

where  $\Lambda_+, \Lambda_-$  are the eigenvalues of the symmetric matrix  $\mathbf{B}$ . To analyse the eigenvalues, we denote the trace and determinant of  $\mathbf{B}$  by

$$\mathbf{T} := \text{Tr}(\mathbf{B}) = 2\gamma - \frac{\delta}{1+\rho}, \quad \mathbf{D} := \det(\mathbf{B}) = \frac{\delta \rho}{1+\rho} (2\gamma - \delta) - \frac{\delta^2}{4} \left( \eta_\varepsilon + \frac{\gamma}{2} \right)^2. \quad (31)$$

We can then express the eigenvalues of  $\mathbf{B}$  as

$$\Lambda_\pm = \frac{\mathbf{T}}{2} \pm \sqrt{\frac{\mathbf{T}^2}{4} - \mathbf{D}} = \frac{\mathbf{D}}{\frac{\mathbf{T}}{2} \mp \sqrt{\frac{\mathbf{T}^2}{4} - \mathbf{D}}}, \quad (32)$$

provided the denominator in the last equality is non-zero. We wish to choose  $\delta > 0$  such that both eigenvalues are indeed positive. Since  $\mathbf{B}_{22} > 0$ , this is the case when  $\mathbf{D} > 0$ , i.e.

$$\delta < \frac{2\gamma}{1 + \frac{1+\rho}{4\rho} \left( \eta_\varepsilon + \frac{\gamma}{2} \right)^2}. \quad (33)$$



On the other hand, we need to make sure we are picking  $\delta > 0$  so that  $\|\cdot\|$  and  $\|\cdot\|_{1+\delta A}$  are equivalent norms. Proposition 3(a) implies that, for  $0 < \delta < 2$ ,

$$\left(1 - \frac{\delta}{2}\right) \|\phi\| \leq \|\phi\|_{1+\delta A} \leq \left(1 + \frac{\delta}{2}\right) \|\phi\|. \quad (34)$$

Therefore, for any

$$\delta < \min \left\{ 1, \frac{2\gamma}{1 + \frac{1+\rho}{4\rho} \left(\eta_\varepsilon + \frac{\gamma}{2}\right)^2} \right\}, \quad (35)$$

it holds that  $\|\phi\|_{1+\delta A} \leq \frac{3}{2} \|\phi\|$  and

$$\begin{aligned} \|P_t \phi\|_{1+\delta A}^2 &\leq \|P_s \phi\|_{1+\delta A}^2 - 5\lambda \int_s^t \|P_u \phi\|^2 du \\ &\leq \|P_s \phi\|_{1+\delta A}^2 - 2\lambda \int_s^t \|P_u \phi\|_{1+\delta A}^2 du, \end{aligned} \quad (36)$$

where

$$\lambda = \frac{1}{5} \min\{\Lambda_+, \Lambda_-\} = \frac{\Lambda_-}{5}.$$

The inequality (36) implies by a Gronwall lemma that  $\|P_t \phi\|_{1+\delta A}^2 \leq \exp(-2\lambda t) \|\phi\|_{1+\delta A}^2$ . Hence, by the equivalence of norms (34) and the choice of  $\delta$ , we have

$$\|P_t \phi\|^2 \leq 9 \exp(-2\lambda t) \|\phi\|^2. \quad (37)$$

Since the above inequality is satisfied for any  $\phi \in C_{c,0}^\infty(\mathcal{X})$ , it follows by density that it is also satisfied for any  $\phi \in L^2(d\mu)$  with  $\int_{\mathcal{X}} \phi d\mu = 0$ , which allows to conclude the proof.  $\square$

With the proof of theorem 1(i) at hand, we can next turn to the proof of theorem 2(i).

**Proof of theorem 2(i).** Suppose that  $W^* \in \mathcal{W}_{\alpha,\beta}(L^*)$  for some  $\alpha, \beta > 0$  and let  $\delta > 0$  be as in (35). We first observe that  $mW^* \in \mathcal{W}_{\alpha,m\beta}(L^*)$  for any  $m > 0$ . Applying proposition 3(c) and (36), we have, for any  $\phi \in C_{c,0}^\infty(\mathcal{X})$ ,

$$\begin{aligned} \|P_t \phi\|_{mW^*+1+\delta A}^2 &= \|P_t \phi\|_{mW^*}^2 + \|P_t \phi\|_{1+\delta A}^2 \\ &\leq \|P_s \phi\|_{mW^*+1+\delta A}^2 - \int_s^t [\alpha \|P_u \phi\|_{mW^*}^2 - m\beta \|P_u \phi\|^2 + 5\lambda \|P_u \phi\|^2] du, \end{aligned} \quad (38)$$

where  $\lambda$  is as in (35). Observe that, by choosing

$$m = \frac{5\eta\lambda}{\beta} \quad (39)$$

for some  $\eta \in (0, 1)$ , we find that

$$-\alpha \|P_u \phi\|_{mW^*}^2 + m\beta \|P_u \phi\|^2 - 5\lambda \|P_u \phi\|^2 \leq -\min\{\alpha, 2\lambda(1-\eta)\} \|P_u \phi\|_{mW^*+1+\delta A}^2,$$

where we used the equivalence of norms (34) and the choice of  $\delta \in [0, 1]$  as in (35). In particular, it follows from a Gronwall lemma, an approximation argument and the equivalence of norms (34), that, for any  $\phi \in L^2(W^* d\mu)$  with  $\int_{\mathcal{X}} \phi d\mu = 0$ ,

$$\forall t \geq 0, \quad \|P_t \phi\|_{mW^*+1}^2 \leq 9 \exp(-\min\{2\lambda(1-\varepsilon), \alpha\}t) \|\phi\|_{mW^*+1}^2.$$

This gives the claimed result.  $\square$

#### 4. Scalings and consequences of Lyapunov structure

In this section, we analyse the rate of convergence to equilibrium in both the unweighted and weighted settings with respect to the friction parameter  $\gamma > 0$ , ultimately proving theorems 1(b) and 2(b). In the unweighted setting (section 4.1), we study in particular the behaviour of the parameter  $\lambda$  as  $\gamma \rightarrow 0$  or  $\gamma \rightarrow \infty$ . In the weighted setting (section 4.2), we need of course the previous analysis of  $\lambda$ , but we also have to construct a Lyapunov function  $W^* \in \mathcal{S}_{\alpha,\beta}(L^*)$  for some  $\alpha, \beta > 0$ , with an explicit dependence of  $\alpha, \beta$  on  $\gamma > 0$ .

##### 4.1. The unweighted setting and $\lambda$

Here we recall that

$$\lambda = \frac{1}{5} \max\{\Lambda_+, \Lambda_-\} = \frac{1}{5} \Lambda_-,$$

where  $\Lambda_{\pm}$  are the eigenvalues introduced in (32). We also recall that  $\delta > 0$  needs to be chosen so that the restriction (35) is satisfied, and that  $\eta_{\varepsilon}, \rho > 0$  are constants which do not depend on  $\gamma$ .

**Proof of theorem 1(ii).** We briefly recall the approach of [33, section 5.3] for instance. Consider first the case when  $\gamma \geq 1$ . We choose  $\delta = c\gamma^{-1}$  with  $c > 0$  small enough (independent of  $\gamma$ ) so that (35) is satisfied. Then, by simplifying the expression of  $\Lambda_-$ , we find that  $\mathbf{T} \sim 2\gamma$  and  $\mathbf{D} \sim \frac{2c\rho}{1+\rho} - \frac{c^2}{16}$  as  $\gamma \rightarrow \infty$ , so that

$$\Lambda_- = \frac{c}{\gamma} \left( \frac{\rho}{\rho+1} - \frac{c}{32} \right) + \mathcal{O}\left(\frac{1}{\gamma^2}\right).$$

The prefactor of the dominant term  $\gamma^{-1}$  is positive for  $c > 0$  sufficiently small, hence giving the correct scaling for  $\lambda$  as  $\gamma \rightarrow \infty$ .

We next consider the situation when  $\gamma \leq 1$ , for which we choose  $\delta = c\gamma$  with  $c > 0$  sufficiently small so that (35) holds. A simple computation shows that, as  $\gamma \rightarrow 0$ ,

$$\Lambda_- = \gamma \left[ 1 - \frac{c}{2(1+\rho)} - \sqrt{\left(1 - \frac{c}{2(1+\rho)}\right)^2 - \frac{c\rho}{1+\rho}(2-c) + \frac{c^2}{4}\eta_{\varepsilon}^2} \right] + \mathcal{O}(\gamma^2).$$

The prefactor of the dominant term  $\gamma$ , equal to  $c\rho/(1+\rho) + \mathcal{O}(c^2)$ , is positive for  $c > 0$  sufficiently small, hence giving the correct scaling for  $\lambda$  as  $\gamma \rightarrow 0$ . This concludes the proof of the result.  $\square$

#### 4.2. Explicit Lyapunov functions and convergence rates

Our goal in this section is to construct an explicit Lyapunov function  $W^* \in \mathcal{S}_{\alpha,\beta}(L^*)$  and analyse the value  $\alpha > 0$  we obtain from the construction. Since the convergence rate parameter in the weighted setting is  $\min\{\alpha/2, \lambda(1 - \varepsilon)\}$  where  $\lambda > 0$  is the parameter analysed in section 4.1 and  $\varepsilon \in (0, 1)$ , our ultimate goal here is to check whether the parameter  $\alpha > 0$  also scales as  $\min(\gamma, \gamma^{-1})$  as  $\gamma \rightarrow 0$  or  $\gamma \rightarrow +\infty$ .

As a simple first example, we consider in section 4.2.1 a commonly employed condition on  $U$  in the literature (see [40, 47, 53]), relevant for potentials which have a polynomial-like growth at infinity. In this setting, we follow various previous works [40, 47, 53] and construct our Lyapunov function by adding a term  $p \cdot q$  to the Hamiltonian. We make sure that the scaling of  $\alpha$  with respect to  $\gamma$  is indeed  $\min(\gamma, \gamma^{-1})$ . We then consider in section 4.2.2 the general situation of a potential satisfying assumptions 1 and 2, for which we construct a different Lyapunov function allowing us to write the proof of theorem 2(ii).

**4.2.1. Polynomial-like potentials.** We suppose here that the potential  $U$  satisfies assumptions 1 and 2 and the following additional growth assumption.

**Assumption 3.** The potential  $U$  is such that  $U \in C^\infty(\mathbf{R}^d)$  and there exist constants  $c_3, C_4, c_5 > 0$  for which

$$\forall q \in \mathbf{R}^d, \quad \nabla U(q) \cdot q \geq c_3 U(q) - C_4, \quad U(q) \geq c_5 |q|^2.$$

Under these assumptions, we now build a Lyapunov function  $W \in \mathcal{S}_{\alpha,\beta}(L)$  in the same spirit as in [40, 47, 53].

**Lemma 1.** Suppose that  $U$  satisfies assumptions 1–3, and introduce

$$W_\kappa(q, p) = H(q, p) + \kappa q \cdot p. \quad (40)$$

Then there exist  $\bar{\kappa} > 0$  such that

$$\begin{aligned} \forall \gamma \in (0, 1], \quad LW_{\bar{\kappa}\gamma}(q, p) &\leq -\frac{\bar{\kappa}c_3}{2}\gamma W_{\bar{\kappa}\gamma} + \gamma d + C_4\bar{\kappa}\gamma, \\ \forall \gamma \in [1, \infty), \quad LW_{\bar{\kappa}/\gamma}(q, p) &\leq -\frac{\bar{\kappa}c_3}{2\gamma}W_{\bar{\kappa}/\gamma} + \gamma d + \frac{C_4\bar{\kappa}}{\gamma}. \end{aligned}$$

Note that the scaling we obtain on the decay rate  $\alpha$  for  $\gamma \geq 1$  is not of order  $\gamma$  as in theorem 2(b), but of order  $1/\gamma$ . This is not an issue since the associated exponential decay rate in (15) will still scale as  $\min(\gamma, \gamma^{-1})$ . The difference with the scaling obtained in theorem 2(b) comes from the fact that the Lyapunov function is not of exponential type here. The smaller value of  $\alpha$  is compensated by a much smaller and more explicit value of  $\beta$ .

**Proof.** In order for  $W_\kappa > 0$  and  $W_\kappa \rightarrow \infty$  as  $H \rightarrow \infty$ , it is sufficient by assumption 3 and a Cauchy–Schwarz inequality that

$$\kappa < \sqrt{2c_5}. \quad (41)$$

In view of assumptions 1 and 3,

$$\begin{aligned} LW_\kappa(q, p) &= -(\gamma - \kappa)|p|^2 - \kappa \nabla U(q) \cdot q - \kappa \gamma p \cdot q + \gamma d \\ &\leq -(\gamma - \kappa)|p|^2 - c_3 \kappa U(q) - \kappa \gamma p \cdot q + \gamma d + C_4 \kappa. \end{aligned} \quad (42)$$

When  $\delta \leq c_3 \kappa$ , it holds

$$(\gamma - \kappa)|p|^2 + c_3 \kappa U(q) + \kappa \gamma p \cdot q - \delta W_\kappa(q, p) \geq \begin{pmatrix} q \\ p \end{pmatrix}^T \begin{pmatrix} (c_3 \kappa - \delta) c_5 & \kappa(\gamma - \delta)/2 \\ \kappa(\gamma - \delta)/2 & \gamma - \kappa - \delta/2 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}.$$

The aim is to prove that the matrix appearing on the right-hand side of the previous inequality is nonnegative.

We first consider the case  $\gamma \leq 1$ . Setting  $\kappa = \bar{\kappa} \gamma$  and  $\delta = c_3 \kappa/2$ , we obtain

$$\begin{pmatrix} (c_3 \kappa - \delta) c_5 & \kappa(\gamma - \delta)/2 \\ \kappa(\gamma - \delta)/2 & \gamma - \kappa - \delta/2 \end{pmatrix} = \gamma \begin{pmatrix} \frac{c_3 c_5 \bar{\kappa}}{2} & \frac{\bar{\kappa}}{2} \gamma \left(1 - \frac{c_3 \bar{\kappa}}{2}\right) \\ \frac{\bar{\kappa}}{2} \gamma \left(1 - \frac{c_3 \bar{\kappa}}{2}\right) & 1 - \bar{\kappa} \left(1 + \frac{c_3}{4}\right) \end{pmatrix}.$$

We finally choose  $\bar{\kappa} > 0$  sufficiently small so that the latter matrix is positive for all  $\gamma \leq 1$  (which is possible since the determinant is of order  $c_3 c_5 \bar{\kappa}/2 + O(\bar{\kappa}^2)$  uniformly in  $\gamma \leq 1$ ) and (41) is satisfied for  $\gamma = 1$ , which provides the claimed inequality for  $\gamma \leq 1$ .

For  $\gamma \geq 1$ , we still set  $\delta = c_3 \kappa/2$  but consider now  $\kappa = \bar{\kappa}/\gamma$ . Then,

$$\begin{pmatrix} (c_3 \kappa - \delta) c_5 & \kappa(\gamma - \delta)/2 \\ \kappa(\gamma - \delta)/2 & \gamma - \kappa - \delta/2 \end{pmatrix} = \begin{pmatrix} \frac{c_3 c_5 \bar{\kappa}}{2 \gamma} & \frac{\bar{\kappa}}{2} \left(1 - \frac{c_3 \bar{\kappa}}{2 \gamma^2}\right) \\ \frac{\bar{\kappa}}{2} \left(1 - \frac{c_3 \bar{\kappa}}{2 \gamma^2}\right) & \gamma - \frac{\bar{\kappa}}{\gamma} \left(1 + \frac{c_3}{4}\right) \end{pmatrix}.$$

When  $\bar{\kappa} > 0$  is sufficiently small, the determinant of the matrix on the right-hand side is positive, of order  $c_3 c_5 \bar{\kappa}/2 + O(\bar{\kappa}^2)$  uniformly in  $\gamma \geq 1$ . Upon further reducing the value  $\bar{\kappa}$  found for  $\gamma \leq 1$ , we finally obtain the claimed inequalities.  $\square$

**4.2.2. The general case.** If we remove assumption 3, then this limits the types of known Lyapunov functions one can consider. In fact, the function in (40) will not satisfy (42) as a Lyapunov function if we merely consider assumptions 1 and 2. In order to deal with these issues, we will slightly modify the form of the function as in [4, 30]. Note that a similar form was used in [14, 39].

**Proof of theorem 2(ii).** Our goal here is to construct a Lyapunov function for  $L$  of the form

$$W(q, p) = e^{\eta \tilde{H}(q, p)}, \quad (43)$$

where

$$\tilde{H}(q, p) = H(q, p) + \kappa \frac{p \cdot \nabla U(q)}{|\nabla U(q)|^2 + \sigma} =: H(q, p) + \kappa \psi(q, p),$$

for some constants  $\kappa, \sigma > 0$  and  $\eta \in (0, 1)$  to be determined later. The Lyapunov function for  $L^*$  is then obtained by proposition 1.

First observe that  $|\psi(q, p)| \leq \sigma^{-1/2}|p|$ , so that, for any choice of  $\kappa, \sigma, \eta > 0$ , it holds  $W > 0$  and  $W \rightarrow +\infty$  as  $H \rightarrow +\infty$ . Furthermore, for any  $\eta \in (0, 1)$  and  $\kappa, \sigma > 0$ , it is easy to check that  $W$  is strongly integrable (see definition 1). Next, to help compute  $LW$ , note that

$$\begin{aligned} L\tilde{H}(q, p) &= LH + \kappa L\psi = -\gamma|p|^2 + \gamma d - \kappa\gamma \frac{p \cdot \nabla U(q)}{|\nabla U(q)|^2 + \sigma} - \kappa \frac{|\nabla U(q)|^2}{|\nabla U(q)|^2 + \sigma} \\ &\quad + \kappa \frac{p \cdot \nabla^2 U(q)p}{|\nabla U(q)|^2 + \sigma} - 2\kappa \frac{p \cdot \nabla U(q)}{|\nabla U(q)|^2 + \sigma} \frac{\nabla U(q) \cdot \nabla^2 U(q)p}{|\nabla U(q)|^2 + \sigma}. \end{aligned}$$

Using assumption 2 on the two terms of the second line gives

$$\begin{aligned} L\tilde{H}(q, p) &\leq \gamma d + \left[ -\gamma + \kappa \frac{\varepsilon |\nabla U(q)|^2 + C_\varepsilon}{|\nabla U(q)|^2 + \sigma} \left( 1 + 2 \frac{|\nabla U(q)|^2}{|\nabla U(q)|^2 + \sigma} \right) \right] |p|^2 \\ &\quad - \kappa \frac{|\nabla U(q)|^2}{|\nabla U(q)|^2 + \sigma} - \kappa\gamma \frac{p \cdot \nabla U(q)}{|\nabla U(q)|^2 + \sigma}. \end{aligned}$$

We choose  $\sigma > C_\varepsilon/\varepsilon$ , so that the coefficient of  $|p|^2$  achieves a maximum as  $|\nabla U(q)| \rightarrow +\infty$ ; the corresponding value being smaller than  $-\gamma + 3\kappa\varepsilon$ . Therefore,

$$L\tilde{H}(q, p) \leq \gamma d + (-\gamma + 3\kappa\varepsilon)|p|^2 - \kappa \frac{|\nabla U(q)|^2}{|\nabla U(q)|^2 + \sigma} - \kappa\gamma \frac{p \cdot \nabla U(q)}{|\nabla U(q)|^2 + \sigma}.$$

Using this inequality along with the fact that for all  $V \in C^2(\mathcal{O})$  and  $\eta > 0$ ,

$$L(e^{\eta V}) = \eta(LV + \eta\gamma|\nabla_p V|^2)e^{\eta V},$$

the function  $W = e^{\eta\tilde{H}}$  in (43) satisfies

$$\begin{aligned} \frac{LW(q, p)}{\eta W(q, p)} &= L\tilde{H}(q, p) + \eta\gamma|\nabla_p \tilde{H}(q, p)|^2 \\ &\leq \gamma d + (-\gamma + 3\kappa\varepsilon)|p|^2 - \kappa \frac{|\nabla U(q)|^2}{|\nabla U(q)|^2 + \sigma} - \kappa\gamma \frac{p \cdot \nabla U(q)}{|\nabla U(q)|^2 + \sigma} \\ &\quad + \eta\gamma \left( |p|^2 + 2\kappa \frac{p \cdot \nabla U(q)}{|\nabla U(q)|^2 + \sigma} + \kappa^2 \frac{|\nabla U(q)|^2}{(|\nabla U(q)|^2 + \sigma)^2} \right) \\ &\leq \gamma d - ((1 - \eta)\gamma - 3\kappa\varepsilon)|p|^2 - \kappa \frac{|\nabla U(q)|^2}{|\nabla U(q)|^2 + \sigma} + \frac{\kappa^2\gamma}{|\nabla U(q)|^2 + \sigma} \\ &\quad + |2\eta - 1|\kappa\gamma \frac{|p \cdot \nabla U(q)|}{|\nabla U(q)|^2 + \sigma}, \end{aligned}$$

where we used that  $\eta \in [0, 1]$  (so that  $|2\eta - 1| \leq 1$  in particular). Using Young's inequality for the last term, hereby introducing a constant  $C_\eta > 0$  which can be made as large as wanted depending on the value of  $\eta$ , we finally obtain

$$\begin{aligned} \frac{LW(q, p)}{\eta W(q, p)} &\leq \gamma d - [(1 - \eta - C_\eta^{-1})\gamma - 3\kappa\varepsilon]|p|^2 - \kappa \frac{|\nabla U(q)|^2}{|\nabla U(q)|^2 + \sigma} \\ &\quad + \left( \frac{C_\eta}{4} + 1 \right) \frac{\kappa^2\gamma}{|\nabla U(q)|^2 + \sigma}. \end{aligned} \tag{44}$$

We next choose the parameters for the function  $W$  in (43). These parameters are the same for the two limiting regimes we consider, namely  $\gamma \ll 1$  and  $\gamma \gg 1$ , but the estimate for  $\beta$  changes depending on which case is considered. Let us first fix some value  $\eta \in (0, 1)$ , independently of  $\gamma$ . The regime where  $U$  is large but  $|p|$  is small forces us to choose  $\kappa$  to be order  $\gamma$ . This is because, in particular, when  $|p|$  is small and  $U(q)$  is sufficiently large, we need

$$\gamma d - \kappa \frac{|\nabla U(q)|^2}{|\nabla U(q)|^2 + \sigma} < 0.$$

Specifically, we set  $\kappa = 2\gamma d$  and consider  $\varepsilon, C_\eta > 0$  independent of  $\gamma$  and such that

$$6d\varepsilon + C_\eta^{-1} \leq \frac{1-\eta}{2},$$

so that  $(1 - \eta - C_\eta^{-1})\gamma - 3\kappa\varepsilon \geq (1 - \eta)\gamma/2$ .

With these choices, consider first the case  $\gamma \leq 1$ . Then, on a region of the form

$$\mathcal{R} = \{|p| > D\} \cup \{|\nabla U| > D\},$$

where  $D > 0$  is large but independent of  $\gamma$ , we have that

$$LW \leq -c\gamma W$$

where  $c > 0$  is independent of  $\gamma$ . Now the complementary region

$$\mathcal{R}^c = \{|p| \leq D\} \cap \{|\nabla U| \leq D\}$$

is compact. Moreover,  $W$  is in this case bounded on  $\mathcal{R}^c$  independently of  $\gamma \leq 1$  since  $0 \leq \kappa \leq 2d$ . Since the right-hand side (44) is of order  $\gamma$  on  $\mathcal{R}$ , this in turn implies the estimate

$$LW \leq -c\gamma W + C\gamma$$

for some constant  $C > 0$  independent of  $\gamma$ , which indeed gives (16). Note that this agrees with the estimate in lemma 1 obtained when  $U$  satisfies assumption 3 in addition to assumption 2, but the function  $W$  constructed here grows much faster for  $H$  large than the one in (40).

We next consider the case  $\gamma \geq 1$ . Again, we maintain the same choices of  $\kappa, \varepsilon, C_\eta$  as before, but this time we need to enlarge the region depending on the size of  $\gamma$  in order to control the remainder term  $(1 + C_\eta/4)\kappa^2\gamma/(|\nabla U|^2 + \sigma)$ . More precisely, in the region

$$\mathcal{R}_\gamma = \{|p| > D\} \cup \{|\nabla U| > D\gamma\}$$

where  $D > 0$  is large but independent of  $\gamma$ , it holds

$$\forall (q, p) \in \mathcal{R}_\gamma, \quad (LW)(q, p) \leq -c\gamma W(q, p)$$

for some  $c > 0$  independent of  $\gamma$ . On the complementary region

$$\mathcal{R}_\gamma^c = \{|p| \leq D\} \cap \{|\nabla U| \leq D\gamma\},$$

the function  $W$  is bounded by  $\mathcal{M}_\gamma := \max_{\mathcal{R}_\gamma^c} W$ . We therefore have the global estimate

$$LW \leq -c\gamma W + C\gamma^3 \mathcal{M}_\gamma$$

for some constant  $C > 0$  independent of  $\gamma$ , which is indeed (17).  $\square$

### 5. Auxiliary estimates and the proof of proposition 3

In this section, we make precise the definition of the operator  $A$  introduced in (22) and deduce a number of estimates that will be crucial in the proof of proposition 3. At times, we will also need to make use of a sequence of cutoff functions indexed by  $n \in \mathbb{N}$ ,

$$\chi_n \in C^\infty([0, \infty); [0, 1]), \quad (45)$$

satisfying

$$\chi_n(t) = \begin{cases} 1 & \text{if } t \leq n, \\ 0 & \text{if } t \geq n+1, \end{cases} \quad \chi'_n \leq 0, \quad \sup_{n \geq 1} \{\|\chi'_n\|_\infty + \|\chi''_n\|_\infty\} < +\infty, \quad (46)$$

where  $\|\varphi\|_\infty := \sup_{x \in \mathcal{X}} |\varphi(x)|$  for  $\varphi$  bounded measurable. Depending on the context, we use cutoff functions of the form  $\chi_n(U(q))$  or  $\chi_n(H(q, p))$ .

In what follows, we will also utilise the gradient dynamics on  $\mathcal{O}$  given by (11). Under assumptions 1 and 2, it follows that the solution  $q_t$  of (11) is non-explosive and has a unique invariant probability measure given by the  $q$ -marginal of  $\mu$ , namely  $\mu_{\text{OD}}$ . Indeed, invariance of  $\mu_{\text{OD}}$  follows by assumption 1(a) and a direct calculation showing that

$$L_{\text{OD}}^\dagger(e^{-U}) = 0,$$

where  $L_{\text{OD}}^\dagger$  denotes the formal  $L^2(dq)$  adjoint of  $L_{\text{OD}}$ . Furthermore, uniqueness of  $\mu_{\text{OD}}$  follows by assumptions 1(a) and (b) since  $q_t$  is a uniformly non-degenerate diffusion [35, 44].

#### 5.1. Regularity estimates and moments of $|\nabla U|$

In this section, we establish some basic facts about solutions  $\psi$  of the equation

$$(1 - L_{\text{OD}})\psi = \phi, \quad (47)$$

for a given  $\phi \in C_b^\infty(\mathcal{O})$ . We also show that the gradient of  $|\nabla U|$  has finite moments with respect to  $\mu_{\text{OD}}$ .

**Proposition 4.** *Suppose that  $U$  satisfies assumptions 1 and 2, and let  $\phi \in C_b^\infty(\mathcal{O})$ . Then there exists a unique classical solution  $\psi \in C_b^\infty(\mathcal{O})$  of the equation (47), which can moreover be expressed by the Green's formula*

$$\psi(q) = \mathbf{E}_q \left[ \int_0^\infty e^{-s} \phi(q_s) ds \right], \quad (48)$$

where  $q_t$  satisfies (11).

**Proof.** We first prove the existence of solutions belonging to the class  $C_b^\infty(\mathcal{O})$ . In fact, we show that  $\psi$  defined by the formula (48) belongs to this class of functions and satisfies the desired equation.

Consider the bounded open domains defined in assumption 1(b) with boundaries  $\partial \mathcal{O}_k = \{q \in \mathcal{O} : U(q) = k\}$ , as well as the sequence of boundary-value problems

$$\begin{cases} (1 - L_{\text{OD}})\psi_k = \phi & \text{on } \mathcal{O}_k, \\ \psi_k = 0 & \text{on } \partial \mathcal{O}_k. \end{cases} \quad (49)$$

By assumption 1(d), there exists  $k_* \in \mathbb{N}$  such that  $|\nabla U| > 0$  on the boundary  $\partial \mathcal{O}_k$  for all  $k \geq k_*$  and recall that  $U \in C^\infty(\mathcal{O})$ . The boundary  $\partial \mathcal{O}_k$  is therefore  $C^\infty$  for  $k \geq k_*$ . Since the operator  $1 - L_{\text{OD}}$  is uniformly elliptic and the domain  $\mathcal{O}_k$  is bounded, there exists a unique solution  $\psi_k \in C^\infty(\mathcal{O}_k) \cap C_b(\overline{\mathcal{O}_k})$  to the boundary value problem (49) (see for instance [20, chapter 6]). Using Itô's formula applied to  $e^{-t} \psi_k(q_t)$  and a nearly identical line of reasoning as in the proof of [43, theorem 9.1.1], it follows that this unique solution  $\psi_k$  is given by the Feynman–Kac formula

$$\psi_k(q) = \mathbf{E}_q \left[ \int_0^{\sigma_k} e^{-s} \phi(q_s) ds \right],$$

where  $\sigma_k = \inf\{t \geq 0 : q_t \notin \mathcal{O}_k\}$ .

We next claim that if  $\psi$  is defined by the formula (48), then  $\psi_k \rightarrow \psi$  pointwise in  $\mathcal{O}$  as  $k \rightarrow \infty$ . Let  $q \in \mathcal{O}$  and consider  $k_* \in \mathbb{N}$  such that  $q \in \mathcal{O}_k$  for all  $k \geq k_*$  (such an integer  $k_*$  exists by assumption 1). It then follows that, for all  $k \geq k_*$ ,

$$|\psi(q) - \psi_k(q)| = \left| \mathbf{E}_q \left[ \int_{\sigma_k}^\infty e^{-s} \phi(q_s) ds \right] \right| \leq \|\phi\|_\infty \mathbf{E}_q(e^{-\sigma_k}). \quad (50)$$

Since  $q_t$  is a non-explosive process,  $\sigma_k \uparrow \infty$ ,  $\mathbf{P}_q$ -almost surely as  $k \rightarrow \infty$ . Using the dominated convergence theorem on the right-hand side of (50), it follows that  $\psi_k(q) \rightarrow \psi(q)$  as  $k \rightarrow \infty$ , thus establishing the claimed pointwise convergence.

The next step is to show that  $\psi$  solves (47) in the sense of distributions on  $\mathcal{O}$ . To see this, let  $k_* \in \mathbb{N}$  be arbitrary and suppose that  $\varphi \in C_c^\infty(\mathcal{O}_{k_*})$ . Note that  $\psi_k$  and  $\psi$  are both globally bounded by  $\|\phi\|_\infty$ . Letting  $(1 - L_{\text{OD}})^\dagger = 1 - \Delta_q - \text{div}_q(\nabla U \cdot)$  denote the formal adjoint of  $(1 - L_{\text{OD}})$  with respect to the Lebesgue measure on  $\mathcal{O}$  and recalling that  $\psi_k \rightarrow \psi$  pointwise on  $\mathcal{O}$ , we thus obtain by the dominated convergence theorem again:

$$\begin{aligned} \int_{\mathcal{O}} \psi (1 - L_{\text{OD}})^\dagger \varphi &= \lim_{k \rightarrow \infty} \int_{\mathcal{O}_{k_*}} \psi_k (1 - L_{\text{OD}})^\dagger \varphi \\ &= \lim_{k \rightarrow \infty} \int_{\mathcal{O}_{k_*}} [(1 - L_{\text{OD}}) \psi_k] \varphi = \int_{\mathcal{O}_{k_*}} \phi \varphi. \end{aligned}$$

This shows indeed that  $\psi$  solves (47) in the sense of distributions. By results of elliptic regularity, this functions also solves (47) in the classical sense on  $\mathcal{O}$ .

Finally, to establish the uniqueness of solutions of (47) belonging to  $C_b^\infty(\mathcal{O})$ , let  $\psi_1, \psi_2 \in C_b^\infty(\mathcal{O})$  satisfy (47) and fix  $q \in \mathcal{O}$ . Applying Itô's formula to  $e^{-t}(\psi_1(q_t) - \psi_2(q_t))$  and evaluating this quantity at time  $t \wedge \sigma_k$ , we find that, for all  $t \geq 0$ ,  $k \in \mathbb{N}$ , and  $i = 1, 2$ ,

$$\begin{aligned} \mathbf{E}_q [e^{-t \wedge \sigma_k} \psi_i(q_{t \wedge \sigma_k})] &= \psi_i(q) + \mathbf{E}_q \left[ \int_0^{t \wedge \sigma_k} (L_{\text{OD}} - 1) \psi_i(q_s) ds \right] \\ &= \psi_i(q) - \mathbf{E}_q \left[ \int_0^{t \wedge \sigma_k} \phi(q_s) ds \right]. \end{aligned}$$

Therefore,

$$\forall t \geq 0, \quad \forall k \in \mathbb{N}, \quad \psi_1(q) - \psi_2(q) = \mathbf{E}_q [e^{-t \wedge \sigma_k} (\psi_1(q_{t \wedge \sigma_k}) - \psi_2(q_{t \wedge \sigma_k}))].$$

Hence, we can bound the difference above as

$$|\psi_1(q) - \psi_2(q)| \leq (\|\psi_1\|_\infty + \|\psi_2\|_\infty) \mathbf{E}_q [e^{-t \wedge \sigma_k}].$$



Taking  $t \rightarrow \infty$  and then  $k \rightarrow \infty$ , and using the nonexplosivity of the dynamics, we find that  $\psi_1(q) = \psi_2(q)$ . Since  $q \in \mathcal{O}$  was arbitrary, we conclude that  $\psi_1 = \psi_2$  on all of  $\mathcal{O}$ .  $\square$

**Remark 12.** Note that the proof of the previous result gives a natural way to define the action of the operator

$$A = (1 - L_{\text{OD}})^{-1} \Pi L_{\text{H}} \quad (51)$$

on the core of functions  $\phi \in C^\infty(\mathcal{X})$  such that  $\Pi L_{\text{H}} \phi \in C_b^\infty(\mathcal{O})$ , via the stochastic representation

$$(A\phi)(q) := \mathbf{E}_q \left[ \int_0^\infty e^{-s} (\Pi L_{\text{H}} \phi)(q_s) ds \right], \quad (52)$$

where  $q_t$  satisfies (11). In particular,  $A\phi \in C_b^\infty(\mathcal{O})$ . Observe that the vector space of such functions  $\phi$  is dense in  $L^2(d\mu)$  since it contains  $C_c^\infty(\mathcal{X})$  as well as the constant functions, hence also  $C_{c,0}^\infty(\mathcal{X})$ . The operator  $A$  above can then be extended to an operator on  $L^2(d\mu)$  once we show it is bounded on its dense domain; see lemma 3 below for the latter result.

We next establish some moments estimates for  $|\nabla U|$  with respect to the measure  $\mu_{\text{OD}}$ , which will prove crucial to control derivatives of  $\chi_n$ . Such estimates also allow us to obtain some control on the derivatives of the solution  $\psi \in C_b^\infty(\mathcal{O})$  to (47). Part (b) of the lemma below is used crucially in the final line of the proof of proposition 5.

**Lemma 2.** Suppose that  $U$  satisfies assumptions 1 and 2. Then,

- (a)  $|\nabla U| \in L^r(d\mu_{\text{OD}})$  for any  $r \geq 1$ ;
- (b) For any  $\phi \in C_b^\infty(\mathcal{O})$  and  $r \geq 1$ , the unique solution  $\psi \in C_b^\infty(\mathcal{O})$  to (47) belongs to  $H^1(d\mu_{\text{OD}}) \cap H^1(|\nabla U|^r d\mu_{\text{OD}})$ .
- (c) For any  $\delta > 0$ ,  $e^{-\delta U} \in L^1(dq)$ .

**Proof.** To prove (a), fix  $r \geq 3$  and consider the sequence of functions

$$\varphi_n(q, p) = p \cdot \nabla U |\nabla U|^{r-2} \chi_n(U),$$

where  $\chi_n$  is defined in (45). A simple computation shows that

$$L\varphi_n(q, p) = \sum_{i,j=1}^d p_i p_j \partial_{q_j} (\partial_{q_i} U |\nabla U|^{r-2} \chi_n(U)) - |\nabla U|^r \chi_n(U) - \gamma \varphi_n(q, p).$$

Note that the function  $L\varphi_n$  belongs to  $L^1(\mu)$  since  $\chi_n(U)$  has compact support. Now, since the integral of  $p_i p_j$  against  $\mu_{\text{OD}}$  is  $\delta_{ij}$  because  $\mu_{\text{OD}}$  is a standard centered Gaussian, while  $\varphi_n$  has average 0 with respect to  $\mu$ ,

$$0 = \int_{\mathcal{X}} L\varphi_n d\mu = \int_{\mathcal{O}} \sum_{j=1}^d \partial_{q_j} [\partial_{q_j} U |\nabla U|^{r-2} \chi_n(U)] - |\nabla U|^r \chi_n(U) d\mu_{\text{OD}}.$$

Since  $\chi_n' \leq 0$  and in view of assumption 2, there exists  $C > 0$  such that

$$\sum_{j=1}^d \partial_{q_j} [\partial_{q_j} U |\nabla U|^{r-2} \chi_n(U)] \leq \left( \frac{1}{2} |\nabla U|^r + C |\nabla U|^{r-2} \right) \chi_n(U).$$

Combining the previous two estimates we find that

$$\forall n \geq 1, \quad \frac{1}{2} \int_{\mathcal{O}} |\nabla U|^r \chi_n(U) \, d\mu_{\text{OD}} \leq C \int_{\mathcal{O}} |\nabla U|^{r-2} \chi_n(U) \, d\mu_{\text{OD}}.$$

In view of the following lower bound, obtained by a Hölder inequality,

$$\int_{\mathcal{O}} |\nabla U|^r \chi_n(U) \, d\mu_{\text{OD}} \geq \left( \int_{\mathcal{O}} |\nabla U|^{r-2} \chi_n(U) \, d\mu_{\text{OD}} \right)^{r/(r-2)} \left( \int_{\mathcal{O}} \chi_n(U) \, d\mu_{\text{OD}} \right)^{-2/(r-2)},$$

we find that

$$\left( \int_{\mathcal{O}} |\nabla U|^{r-2} \chi_n(U) \, d\mu_{\text{OD}} \right)^{2/(r-2)} \leq 2C \left( \int_{\mathcal{O}} \chi_n(U) \, d\mu_{\text{OD}} \right)^{2/(r-2)} \leq 2C,$$

from which the result (a) follows.

For part (b), let us fix  $\phi \in C_b^\infty(\mathcal{O})$  and consider the unique solution  $\psi \in C_b^\infty(\mathcal{O})$  to (47). In the remainder of this proof, because all functions are functions of  $q$  only, we use  $\|\cdot\|$  to denote the canonical norm on  $L^2(d\mu_{\text{OD}})$ . We first show that  $\psi \in H^1(d\mu_{\text{OD}})$ . A Cauchy–Schwarz inequality and integration by parts lead to

$$\begin{aligned} \|\psi\| \|\phi\| &\geq \int_{\mathcal{O}} \phi \psi \chi_n(U) \, d\mu_{\text{OD}} = \int_{\mathcal{O}} (1 - L_{\text{OD}}) \psi \cdot \psi \chi_n(U) \, d\mu_{\text{OD}} \\ &= \int_{\mathcal{O}} \psi^2 \chi_n(U) \, d\mu_{\text{OD}} + \int_{\mathcal{O}} \nabla_q \psi \cdot \nabla_q (\psi \chi_n(U)) \, d\mu_{\text{OD}} \\ &= \int_{\mathcal{O}} \psi^2 \chi_n(U) \, d\mu_{\text{OD}} + \int_{\mathcal{O}} |\nabla_q \psi|^2 \chi_n(U) \, d\mu_{\text{OD}} - \frac{1}{2} \int_{\mathcal{O}} \psi^2 L_{\text{OD}}(\chi_n(U)) \, d\mu_{\text{OD}}. \end{aligned}$$

The last integral on the right-hand side converges to 0 as  $n \rightarrow +\infty$  by a dominated convergence argument, since  $\psi$  is bounded, while  $L_{\text{OD}}(\chi_n(U)) = \chi_n''(U)|\nabla U|^2 + \chi_n'(U)(\Delta U - |\nabla U|^2)$  converges pointwise to 0 and  $|\nabla U|^2, \Delta U \in L^1(d\mu_{\text{OD}})$  by assumption 2 and part (a). We therefore obtain the bound  $\|\psi\|^2 + \|\nabla_q \psi\|^2 \leq \|\psi\| \|\phi\|$  as  $n \rightarrow +\infty$ , which proves that  $\psi \in H^1(d\mu_{\text{OD}})$ .

To see that  $\psi \in H^1(|\nabla U|^r d\mu_{\text{OD}})$  for any  $r \geq 1$ , we may assume without loss of generality that  $r \geq 3$ . By manipulations similar to the ones used to prove the  $H^1(\mu_{\text{OD}})$  bound, we obtain

$$\begin{aligned} \|\psi |\nabla U|^r\| \|\phi\| &\geq \int_{\mathcal{O}} (1 - L_{\text{OD}}) \psi \cdot \psi |\nabla U|^r \chi_n(U) \, d\mu_{\text{OD}} \\ &= \int_{\mathcal{O}} \psi^2 \chi_n(U) |\nabla U|^r \, d\mu_{\text{OD}} + \int_{\mathcal{O}} |\nabla_q \psi|^2 |\nabla U|^r \chi_n(U) \, d\mu_{\text{OD}} \\ &\quad + \int_{\mathcal{O}} \psi \nabla_q \psi \cdot \nabla [|\nabla U|^r \chi_n(U)] \, d\mu_{\text{OD}}. \end{aligned}$$

The last integral can be bounded as

$$\begin{aligned} & \left| \int_{\mathcal{O}} \psi \nabla_q \psi \cdot \nabla [|\nabla U|^r \chi_n(U)] \, d\mu_{\text{OD}} \right| \\ & \leq \left| \int_{\mathcal{O}} \psi \nabla_q \psi \cdot \nabla [|\nabla U|^r] \chi_n(U) \, d\mu_{\text{OD}} \right| + \left| \int_{\mathcal{O}} \psi \nabla_q \psi \cdot \nabla U |\nabla U|^{r-1} \chi'_n(U) \, d\mu_{\text{OD}} \right| \\ & \leq \|\nabla_q \psi\| (\|\psi \nabla [|\nabla U|^r]\| + \|\psi \nabla U |\nabla U|^{r-1} \chi'_n(U)\|). \end{aligned}$$

By part (a) and assumption 2, the last term in the above inequality converges to 0 as  $n \rightarrow +\infty$ , while the other term is bounded. We therefore obtain, in the limit  $n \rightarrow +\infty$ ,

$$\begin{aligned} & \|\psi\|_{L^2(|\nabla U|^r \, d\mu_{\text{OD}})}^2 + \|\nabla_q \psi\|_{L^2(|\nabla U|^r \, d\mu_{\text{OD}})}^2 \\ & \leq \|\psi |\nabla U|^r\| \|\phi\| + \|\nabla_q \psi\| \|\psi \nabla (|\nabla U|^r)\| \\ & \leq \|\psi\|_{\infty} (\|\nabla U|^r\| \|\phi\| + \|\nabla_q \psi\| \|\nabla (|\nabla U|^r)\|) < +\infty, \end{aligned}$$

which concludes the proof of part (b).

In order to deduce part (c) of the result, it is enough to show that  $e^{\eta U} \in L^1(d\mu_{\text{OD}})$  for any  $\eta \in (0, 1)$ . We do this by applying [23, proposition 5.1] and showing that, for any  $\eta \in (0, 1)$ , it holds  $L_{\text{OD}}(e^{\eta U}) \rightarrow -\infty$  as  $U \rightarrow \infty$ . Fixing  $\eta \in (0, 1)$ , a direct calculation shows that

$$\frac{L_{\text{OD}}(e^{\eta U})}{\eta e^{\eta U}} = -(1 - \eta)|\nabla U|^2 + \Delta U,$$

the function on the right-hand side going to  $-\infty$  as  $U(q) \rightarrow +\infty$  by assumptions 1(d) and 2.  $\square$

## 5.2. $L^2$ -bounds and the elliptic regularity estimate

We deduce here various  $L^2$  estimates involving the operator  $A$ , as well as an elliptic regularity estimate similar to the one in [17], from the estimates proved in section 5.1.

**Lemma 3.** *Suppose that  $U$  satisfies assumptions 1 and 2. Then, for any  $\phi \in C_c^\infty(\mathcal{X})$ ,*

$$\|A\phi\| \leq \frac{1}{2} \|(1 - \Pi)\phi\|, \quad (53)$$

$$|\langle L^* A \phi, \phi \rangle| \leq \|(1 - \Pi)\phi\|^2, \quad (54)$$

$$|\langle A L_{\text{OU}} \phi, \phi \rangle| \leq \frac{1}{2} \|(1 - \Pi)\phi\| \|\Pi\phi\|. \quad (55)$$

The extension of the same inequalities to functions in  $C_{c,0}^\infty(\mathcal{X})$  (corollary 1 below) is immediate since functions in this space are a constant shift of functions in  $C_c^\infty(\mathcal{X})$ , and  $Ac = L_{\text{OU}}c = Lc = 0$  for any constant function  $c$ .

**Corollary 1.** *Suppose that  $U$  satisfies assumptions 1 and 2. Then for any  $\phi \in C_{c,0}^\infty(\mathcal{X})$  the estimates (53)–(55) are also satisfied.*

**Remark 13.** Note that, by the definition of  $A$  given in remark 12, part (a) of proposition 3 readily follows from the above result by extending  $A$  to be a bounded operator on  $L^2(d\mu)$ .

**Proof of lemma 3.** From an algebraic viewpoint, the proof follows the proof of [17, lemma 1]. Let  $\phi \in C_c^\infty(\mathcal{X})$ . We start by proving (53). By proposition 4, we know that  $\psi = A\phi \in C_b^\infty(\mathcal{O})$ . For simplicity of notation in the arguments that follow, we introduce  $\langle f, g \rangle_n := \langle f, g\chi_n(H) \rangle$  and  $\|\psi\|_n^2 = \langle \psi, \psi \rangle_n$ . Noting that  $\psi + (L_H\Pi)^*(L_H\Pi)\psi = -(L_H\Pi)^*\phi$ , we obtain

$$\|\psi\|_n^2 + \langle (L_H\Pi)^*L_H\Pi\psi, \psi \rangle_n = -\langle (L_H\Pi)^*\phi, \psi \rangle_n. \quad (56)$$

First observe that (using that  $\psi$  is a function of  $q$  only, so that  $\Pi(\psi\chi_n(H)) = \psi\Pi\chi_n(H)$ )

$$\langle (L_H\Pi)^*L_H\Pi\psi, \psi \rangle_n = \|L_H\Pi\psi\|_{L^2(\Pi\chi_n(H)d\mu)}^2 + \langle L_H\Pi\psi, \psi L_H\Pi\chi_n(H) \rangle =: T_1(n) + T_2(n).$$

Recalling  $\Pi\psi = \psi$ , and using  $2\psi L_H\psi = L_H(\psi^2)$  as well as an integration by parts, we obtain

$$2T_2(n) = -\langle \psi^2, L_H^2\Pi\chi_n(H) \rangle.$$

Now,

$$\begin{aligned} L_H^2\Pi\chi_n(H) &= L_H(p^T\Pi[\chi_n'(H)\nabla U]) \\ &= -|\nabla U|^2\Pi[\chi_n'(H)] + p^T(\Pi[\chi_n'(H)]\nabla^2 U + \Pi[\chi_n''(H)]\nabla U \otimes \nabla U)p. \end{aligned}$$

Applying lemma 2 and recalling that  $\psi \in C_b^\infty(\mathcal{O})$ , we see that  $T_2(n) \rightarrow 0$  as  $n \rightarrow \infty$ . On the other hand,

$$-\langle (L_H\Pi)^*\phi, \psi \rangle_n = -\langle (1 - \Pi)\phi, (L_H\Pi\psi)\Pi\chi_n(H) \rangle - \langle \phi, \psi L_H\Pi\chi_n(H) \rangle =: -T_3(n) - T_4(n),$$

where we used that  $\Pi L_H\Pi = 0$  in the first equality. By the same reasoning used above to show that  $T_2(n) \rightarrow 0$  as  $n \rightarrow \infty$  via lemma 2, it also holds that  $T_4(n) \rightarrow 0$  as  $n \rightarrow \infty$ . Putting these estimates together and using Young's inequality for products, we find that

$$\begin{aligned} \|\psi\|_n^2 &= -\|L_H\Pi\psi\|_{L^2(\Pi\chi_n(H)d\mu)}^2 \\ &\quad - \langle (1 - \Pi)\phi, (L_H\Pi\psi)\Pi\chi_n(H) \rangle - T_2(n) - T_4(n) \\ &\leq -\|L_H\Pi\psi\|_{L^2(\Pi\chi_n(H)d\mu)}^2 + \frac{1}{4}\|(1 - \Pi)\phi\|_{L^2(\Pi\chi_n(H)d\mu)}^2 \\ &\quad + \|L_H\Pi\psi\|_{L^2(\Pi\chi_n(H)d\mu)}^2 - T_2(n) - T_4(n) \\ &\leq \frac{1}{4}\|(1 - \Pi)\phi\|^2 + |T_2(n)| + |T_4(n)|, \end{aligned}$$

where on the last line we used the fact that  $\chi_n(H) \in [0, 1]$ . The claimed estimate (53) follows by taking the limit  $n \rightarrow \infty$ .

We next claim that

$$\forall \phi \in C_c^\infty(\mathcal{X}), \quad \|L_H\Pi A\phi\| = \|L_H A\phi\| \leq \|(1 - \Pi)\phi\|. \quad (57)$$

To prove this inequality, we use Young's inequality to estimate  $T_3$  as

$$\begin{aligned} |T_3(n)| &\leq \frac{1}{2}\|(1 - \Pi)\phi\|_{L^2(\Pi\chi_n(H)d\mu)}^2 + \frac{1}{2}\|L_H\Pi\psi\|_{L^2(\Pi\chi_n(H)d\mu)}^2 \\ &\leq \frac{1}{2}\|(1 - \Pi)\phi\|^2 + \frac{1}{2}\|L_H\Pi\psi\|_{L^2(\Pi\chi_n(H)d\mu)}^2, \end{aligned}$$

where on the last line we used the fact that  $\chi_n(H) \in [0, 1]$ . Then,

$$\begin{aligned} \|L_H \Pi \psi\|_{L^2(\Pi \chi_n(H) d\mu)}^2 + \|\psi\|_n^2 &= -T_2(n) - T_3(n) - T_4(n) \\ &\leq \frac{1}{2} \|(1 - \Pi)\phi\|^2 + \frac{1}{2} \|L_H \Pi \psi\|_{L^2(\Pi \chi_n(H) d\mu)}^2 \\ &\quad + |T_2(n)| + |T_4(n)|. \end{aligned}$$

The bound (57) follows after taking  $n \rightarrow \infty$ . Since  $\Pi A = A$ , this finally gives

$$\begin{aligned} \forall \phi \in C_c^\infty(\mathcal{X}), \quad |\langle L_H A \phi, \phi \rangle| &= |\langle L_H A \phi, (1 - \Pi)\phi \rangle| = |\langle L^* A \phi, (1 - \Pi)\phi \rangle| \\ &\leq \|(1 - \Pi)\phi\|^2. \end{aligned} \quad (58)$$

Finally, a short calculation shows that  $AL_{OU} = -A$ , so that, for any  $\phi \in C_c^\infty(\mathcal{X})$ ,

$$|\langle AL_{OU} \phi, \phi \rangle| = |\langle AL_{OU} \phi, \Pi \phi \rangle| = |\langle A \phi, \Pi \phi \rangle| \leq \frac{1}{2} \|(1 - \Pi)\phi\| \|\Pi \phi\|.$$

This concludes the proof of the lemma.  $\square$

We next turn to the elliptic regularity estimate. Note that the proof follows arguments similar to those in [9], but with a slightly different application of the more general bound in assumption 2. We also refer the reader to [5, lemma 4] for a related result.

**Proposition 5.** *Suppose that the potential  $U$  satisfies assumptions 1 and 2. Let  $c_1 \in (0, 1)$ ,  $C_2 > 0$  be such that the bound in (27) is satisfied and let  $\varepsilon > 0$  satisfy (28). Let  $C_\varepsilon$  be the corresponding constant for this choice of  $\varepsilon$  in assumption 2, and set  $\xi_\varepsilon = \eta_\varepsilon^2/2$  with  $\eta_\varepsilon$  defined in (29). Then, for any  $\phi \in C_b^\infty(\mathcal{O})$ ,*

$$\|\nabla_q^2 \psi\|_{L^2(d\mu_{OD})}^2 \leq \xi_\varepsilon \|\phi\|_{L^2(d\mu_{OD})}^2,$$

where  $\psi \in C_b^\infty(\mathcal{O})$  is the unique classical bounded solution of  $(1 - L_{OD})\psi = \phi$  as given by proposition 4.

**Proof of proposition 5.** Since all the functions which appear in this proof are functions of  $q$  only, and since the norm  $\|\cdot\|$  coincides with the norm  $\|\cdot\|_{L^2(d\mu_{OD})}$  in this case, we use in this proof the notation  $\|\cdot\|$  for  $\|\cdot\|_{L^2(d\mu_{OD})}$ . We first prove that

$$\forall \varphi \in C_c^\infty(\mathcal{O}), \quad \|\nabla_q^2 \varphi\|^2 \leq \xi_\varepsilon \|(1 - L_{OD})\varphi\|^2. \quad (59)$$

We then conclude the desired estimate using a density argument afforded by lemma 2.

Fix  $\varphi \in C_c^\infty(\mathcal{O})$ . Bochner's formula gives

$$\|\nabla_q^2 \varphi\|^2 = \sum_{i,j=1}^d \left\| \partial_{q_i q_j}^2 \varphi \right\|^2 = \|\nabla_q^* \cdot \nabla_q \varphi\|^2 - \langle \nabla_q \varphi, \nabla^2 U(q) \nabla_q \varphi \rangle.$$

Using assumption 2 with the choice of  $\varepsilon, C_\varepsilon$  as in the statement of the result leads to

$$\|\nabla_q^2 \varphi\|^2 \leq \|\nabla_q^* \cdot \nabla_q \varphi\|^2 + \varepsilon \|\nabla U(q)\| \|\nabla_q \varphi\|^2 + C_\varepsilon \|\nabla_q \varphi\|^2. \quad (60)$$

To control  $\|\nabla U(q)\| \|\nabla_q \varphi\|$ , we use the structure of the Gibbs measure  $\mu$  to write, for any  $f \in C_c^\infty(\mathcal{O})$ ,

$$\|f \nabla U(q)\|^2 = \langle \nabla_q(f^2), \nabla U(q) \rangle + \langle f^2, \Delta U(q) \rangle = 2 \langle \nabla_q f, f \nabla U(q) \rangle + \langle f^2, \Delta U(q) \rangle.$$

With the previous inequality, the estimate (27) and Young's inequality, we obtain, for any  $\eta \in (0, 1 - c_1)$ ,

$$\|f \nabla U(q)\|^2 \leq (c_1 + \eta) \|f \nabla U(q)\|^2 + \eta^{-1} \|\nabla_q f\|^2 + C_2 \|f\|^2,$$

so that, by rearranging the previous expression,

$$\|f \nabla U(q)\|^2 \leq \frac{\eta^{-1}}{1 - c_1 - \eta} \|\nabla_q f\|^2 + \frac{C_2}{1 - c_1 - \eta} \|f\|^2. \quad (61)$$

The above estimate can be extended by density to any  $f \in H^1(\mu_{\text{OD}})$ . Now, if  $\varphi \in C_c^\infty(\mathcal{O})$ , then for all  $\zeta > 0$  we define  $g_\zeta := \sqrt{|\nabla_q \varphi|^2 + \zeta} \in H^1(\mu_{\text{OD}})$ . Therefore, by applying (61) to  $g_\zeta$ ,

$$\|\nabla U\| \|\nabla_q \varphi\|^2 \leq \|\nabla U(q)\| g_\zeta^2 \leq \frac{\eta^{-1}}{1 - c_1 - \eta} \|\nabla_q^2 \varphi\|^2 + \frac{C_2}{1 - c_1 - \eta} \|g_\zeta\|^2.$$

Taking  $\zeta \rightarrow 0$ , we arrive at the following inequality

$$\|\nabla U\| \|\nabla_q \varphi\|^2 \leq \frac{\eta^{-1}}{1 - c_1 - \eta} \|\nabla_q^2 \varphi\|^2 + \frac{C_2}{1 - c_1 - \eta} \|\nabla_q \varphi\|^2.$$

Combining this estimate with (60) gives

$$\|\nabla_q^2 \varphi\|^2 \leq \|\nabla_q^* \cdot \nabla_q \varphi\|^2 + \left( C_\varepsilon + \frac{\varepsilon C_2}{1 - c_1 - \eta} \right) \|\nabla_q \varphi\|^2 + \frac{\varepsilon \eta^{-1}}{1 - c_1 - \eta} \|\nabla_q^2 \varphi\|^2,$$

and, after rearranging the various terms and provided  $\eta(1 - c_1 - \eta) > \varepsilon$ ,

$$\|\nabla_q^2 \varphi\|^2 \leq \left( 1 - \frac{\varepsilon \eta^{-1}}{1 - c_1 - \eta} \right)^{-1} \left[ \|\nabla_q^* \cdot \nabla_q \varphi\|^2 + \left( C_\varepsilon + \frac{\varepsilon C_2}{1 - c_1 - \eta} \right) \|\nabla_q \varphi\|^2 \right].$$

Choosing  $\eta = (1 - c_1)/2$ , and noting that

$$\|(1 + \nabla_q^* \cdot \nabla_q) \varphi\|^2 = \|\varphi\|^2 + 2 \|\nabla_q \varphi\|^2 + \|\nabla_q^* \cdot \nabla_q \varphi\|^2,$$

the inequality reduces to

$$\begin{aligned}\|\nabla_q^2 \varphi\|^2 &\leq \max \left\{ 1, \frac{1}{2} \left( C_\varepsilon + \frac{2C_2\varepsilon}{1-c_1} \right) \right\} \left( 1 - \frac{4\varepsilon}{(1-c_1)^2} \right)^{-1} \|(1 + \nabla_q^* \cdot \nabla_q) \varphi\|^2 \\ &= \xi_\varepsilon \|(1 + \nabla_q^* \cdot \nabla_q) \varphi\|^2.\end{aligned}\quad (62)$$

Since  $-L_{\text{OD}} = \nabla_q^* \cdot \nabla_q$ , this indeed provides the claimed bound (59).

Now consider  $\phi \in C_b^\infty(\mathcal{O})$  and  $\psi \in C_b^\infty(\mathcal{O})$  the unique classical bounded solution of  $(1 - L_{\text{OD}})\psi = \phi$ . Plugging  $\varphi = \psi\chi_n(U) \in C_c^\infty(\mathcal{O})$  into (62) gives

$$\forall n \geq 1, \quad \|\nabla_q^2(\psi\chi_n(U))\|^2 \leq \xi_\varepsilon \|(1 - L_{\text{OD}})(\psi\chi_n(U))\|^2.$$

In view of lemma 2(b) and assumption 2, we therefore obtain, by dominated convergence,

$$\begin{aligned}\lim_{n \rightarrow \infty} \|\nabla_q^2(\psi\chi_n(U))\|^2 &= \|\nabla_q^2 \psi\|^2 \leq \xi_\varepsilon \|\phi\|^2 \\ &= \xi_\varepsilon \lim_{n \rightarrow \infty} \|(1 - L_{\text{OD}})(\psi\chi_n(U))\|^2,\end{aligned}$$

which concludes the proof.  $\square$

With the previous estimate at hand, we conclude this subsection with one final estimate. This is the final ingredient needed in the proof of proposition 3 parts (b) and (c).

**Lemma 4.** *Let  $U$  satisfy assumptions 1 and 2, and let  $\eta_\varepsilon$  be as in (29). Then,*

$$\forall \phi \in C_{c,0}^\infty(\mathcal{X}), \quad |\langle AL_H(1 - \Pi)\phi, \phi \rangle| \leq \eta_\varepsilon \|(1 - \Pi)\phi\| \|\Pi\phi\|. \quad (63)$$

**Proof.** Let  $\phi \in C_{c,0}^\infty(\mathcal{X})$ . Recall that  $L_{\text{OD}} = -\nabla_q^* \cdot \nabla_q$ . A direct calculation yields

$$[(AL_H)(1 - \Pi)]^* = ((p \cdot \nabla_q)^2 - \Delta_q) (1 + \nabla_q^* \cdot \nabla_q)^{-1} \Pi.$$

Hence

$$\begin{aligned}|\langle AL_H(1 - \Pi)\phi, \phi \rangle| &= |\langle (1 - \Pi)\phi, [AL_H(1 - \Pi)]^* \phi \rangle| \\ &= \left| \left\langle (1 - \Pi)\phi, ((p \cdot \nabla_q)^2 - \Delta_q) (1 + \nabla_q^* \cdot \nabla_q)^{-1} \Pi\phi \right\rangle \right| \\ &\leq \|(1 - \Pi)\phi\| \left\| ((p \cdot \nabla_q)^2 - \Delta_q) (1 + \nabla_q^* \cdot \nabla_q)^{-1} \Pi\phi \right\|.\end{aligned}$$

Since  $\psi = (1 + \nabla_q^* \cdot \nabla_q)^{-1} \Pi \phi \in C_b^\infty(\mathcal{O})$  is a function of the variable  $q$  only, we have, from the structure of  $\mu$ ,

$$\begin{aligned}
 \|((p \cdot \nabla_q)^2 - \Delta_q) \psi\|^2 &= \left\| \sum_{i,j=1}^d p_i p_j \partial_{q_i q_j}^2 \psi - \Delta_q \psi \right\|^2 \\
 &= \|\Delta_q \psi\|^2 + \sum_{i,j,k,\ell=1}^d \int_{\mathcal{X}} p_i p_j p_k p_\ell \left( \partial_{q_i q_j}^2 \psi \right) \left( \partial_{q_k q_\ell}^2 \psi \right) d\mu \\
 &\quad - 2 \sum_{i,j=1}^d \left\langle p_i p_j \partial_{q_i q_j}^2 \psi, \Delta_q \psi \right\rangle \\
 &= \|\Delta_q \psi\|^2 + \left( \sum_{i=j,k=\ell} + \sum_{i=k,j=\ell} + \sum_{i=\ell,j=k} - 2 \sum_{i=j=k=\ell} \right) \\
 &\quad \times \int_{\mathcal{X}} p_i p_j p_k p_\ell \left( \partial_{q_i q_j}^2 \psi \right) \left( \partial_{q_k q_\ell}^2 \psi \right) d\mu - 2 \sum_{i=1}^d \left\langle p_i^2 \partial_{q_i}^2 \psi, \Delta_q \psi \right\rangle \\
 &= -\|\Delta_q \psi\|^2 + \sum_{i,j=1}^d \int_{\mathcal{X}} p_i^2 p_j^2 \left[ 2 \left( \partial_{q_i q_j}^2 \psi \right)^2 \right. \\
 &\quad \left. + \left( \partial_{q_i}^2 \psi \right) \left( \partial_{q_j}^2 \psi \right) \right] d\mu - 2 \sum_{i=1}^d \int_{\mathcal{X}} p_i^4 \left( \partial_{q_i}^2 \psi \right)^2 d\mu \\
 &= 2 \sum_{i,j=1}^d \left\| \partial_{q_i q_j}^2 \psi \right\|^2 + \sum_{i,j=1}^d \left\langle \partial_{q_i}^2 \psi, \partial_{q_j}^2 \psi \right\rangle - \|\Delta_q \psi\|^2 \\
 &= 2 \|\nabla_q^2 \psi\|^2 \leq 2 \xi_\varepsilon \|\Pi \phi\|^2,
 \end{aligned}$$

where the last inequality follows by proposition 5. Also, in the above, we can calculate the terms involving products of the  $p_i$ 's since  $\psi$  is a function of  $q$  only, and  $\mu$  is a product measure with a standard Gaussian in  $p$  as its  $p$ -marginal. In particular, along the  $i = j$  terms of the final summation we recover the fourth-moment of the Gaussian measure in  $p$  (i.e. a coefficient of 3), while for the  $i \neq j$  terms of the final sum we recover the variance, which is 1. It thus now follows  $|\langle AL_H(1 - \Pi)\phi, \phi \rangle| \leq \sqrt{2\xi_\varepsilon} \|(1 - \Pi)\phi\| \|\Pi\phi\| = \eta_\varepsilon \|(1 - \Pi)\phi\| \|\Pi\phi\|$ , as claimed.  $\square$

### 5.3. Proof of proposition 3

We conclude this section by giving a proof of proposition 3. This is the last theoretical step in the proof of our main general results, namely theorem 1 part (a) and theorem 2 part (a).

**Proof of proposition 3.** Part (a) of the result is a direct consequence of lemma 3 and a density argument. Let us next turn to part (b) of the result. For any  $\phi \in C_b^\infty(\mathcal{X})$  and  $n \in \mathbb{N}$ ,



we define (recalling the definition (45) for  $\chi_n(H)$ )

$$P_t^n \phi = \chi_n(H) P_t \phi, \quad \mathcal{P} \phi = \int_{\mathcal{X}} \phi \, d\mu, \quad \mathcal{P}^\perp \phi = \phi - \mathcal{P} \phi. \quad (64)$$

Consider now  $\phi \in C_{c,0}^\infty(\mathcal{X})$ . Since  $A \mathcal{P}^\perp \varphi = A \varphi$  for any  $\varphi \in C_c^\infty(\mathcal{X})$ , and recalling that  $\frac{d}{dt} P_t \phi = L P_t \phi = P_t L \phi \in C^\infty([0, +\infty) \times \mathcal{X})$  by hypoellipticity [32],

$$\frac{d}{dt} \langle A \mathcal{P}^\perp P_t^n \phi, \mathcal{P}^\perp P_t^n \phi \rangle = T_1 + T_2 + T_3,$$

with

$$\begin{aligned} T_1 &= \langle A[(L P_t \phi) \chi_n(H)], \mathcal{P}^\perp P_t^n \phi \rangle, \\ T_2 &= \langle A \mathcal{P}^\perp P_t^n \phi, (L P_t \phi) \chi_n(H) \rangle, \\ T_3 &= -\mathcal{P} [A P_t^n \phi] \mathcal{P} [(L P_t \phi) \chi_n(H)]. \end{aligned}$$

Using the identity  $L(fg) = fLg + gLf + 2\gamma \nabla_p f \cdot \nabla_p g$ , we can write

$$\begin{aligned} T_1 &= \langle AL \mathcal{P}^\perp P_t^n \phi, \mathcal{P}^\perp P_t^n \phi \rangle - \langle A[P_t \phi L \chi_n(H)], \mathcal{P}^\perp P_t^n \phi \rangle \\ &\quad - 2\gamma \langle A[\nabla_p(P_t \phi) \cdot \nabla_p \chi_n(H)], \mathcal{P}^\perp P_t^n \phi \rangle \\ &=: T_1' + R_1(n) + R_2(n). \end{aligned}$$

In order to help estimate  $R_2(n)$ , note that, for  $f \in C^\infty(\mathcal{X})$  and  $g \in C_c^\infty(\mathcal{X})$ ,

$$A(\nabla_p f \cdot \nabla_p g) = -\sum_{i=1}^d (1 - L_{\text{OD}})^{-1} \partial_{q_i}^* \Pi [\partial_{p_i} (\nabla_p f \cdot \nabla_p g)].$$

Next, using nearly identical arguments to those in lemma 3, observe that the operators  $\mathcal{T}_i = (1 - L_{\text{OD}})^{-1} \partial_{q_i}^*$  are bounded on  $L^2(d\mu_{\text{OD}})$  with norms smaller than 1 since

$$\begin{aligned} \sum_{i=1}^d \mathcal{T}_i \mathcal{T}_i^* &= \sum_{i=1}^d (1 - L_{\text{OD}})^{-1} \partial_{q_i}^* [(1 - L_{\text{OD}})^{-1} \partial_{q_i}^*]^* \\ &= (1 - L_{\text{OD}})^{-1} \left( \sum_{i=1}^d \partial_{q_i}^* \partial_{q_i} \right) (1 - L_{\text{OD}})^{-1} \\ &= -L_{\text{OD}} (1 - L_{\text{OD}})^{-2} \leq 1. \end{aligned}$$

On the other hand, integrating by parts in the  $p$  variable produces

$$\Pi [\partial_{p_i} (\nabla_p f \cdot \nabla_p g)] = \Pi [f \nabla_p^* \cdot (p_i \nabla_p g)] = \Pi [f(p_i p \cdot \nabla_p g - p_i \Delta_p g - \partial_{p_i} g)].$$

Putting these estimates together with lemma 3 and  $\|P_t \phi\|_\infty \leq \|\phi\|_\infty$ , we can estimate  $R_1(n)$  and  $R_2(n)$  as

$$|R_1(n)| + |R_2(n)| \leq c_n^1 \|\phi\|_\infty \|\mathcal{P}^\perp P_t^n \phi\| \leq c_n^1 \|\phi\|_\infty^2$$

where we used the fact that  $\chi_n \in [0, 1]$  and introduced

$$c_n^1 = \frac{1}{2} \|L\chi_n(H)\| + 2\gamma \sum_{i=1}^d \|p_i (p^2 - d - 1) \chi_n'(H) + p_i p^2 \chi_n''(H)\|.$$

By construction of  $\chi_n(H)$  and the moment estimates of lemma 2, a simple argument based on dominated convergence shows that  $c_n^1 \rightarrow 0$  as  $n \rightarrow \infty$ .

Turning to the term  $T_2$ , a similar argument yields

$$\begin{aligned} T_2 &= \langle L^* A \mathcal{P}^\perp P_t^n \phi, \mathcal{P}^\perp P_t^n \phi \rangle - \langle A \mathcal{P}^\perp P_t^n \phi, P_t \phi L\chi_n(H) \rangle \\ &\quad - 2\gamma \langle A \mathcal{P}^\perp P_t^n \phi, \nabla_p(P_t \phi) \cdot \nabla_p \chi_n(H) \rangle \\ &= \langle L^* A \mathcal{P}^\perp P_t^n \phi, \mathcal{P}^\perp P_t^n \phi \rangle - \langle A \mathcal{P}^\perp P_t^n \phi, P_t \phi L\chi_n(H) \rangle \\ &\quad - 2\gamma \langle A \mathcal{P}^\perp P_t^n \phi, (P_t \phi)(p \cdot \nabla_p - \Delta_p) \chi_n(H) \rangle \\ &=: T_2' + R_3(n) + R_4(n), \end{aligned}$$

with  $|R_3(n)| + |R_4(n)| \leq c_n^2 \|\phi\|_\infty^2$  where  $c_n^2 \rightarrow 0$  as  $n \rightarrow \infty$ . Finally, since  $\mathcal{P}[(LP_t \phi)\chi_n(H)] = \mathcal{P}[P_t \phi L^* \chi_n(H)]$ , the term  $T_3$  can be estimated as

$$|T_3| \leq \frac{1}{2} \|P_t^n \phi\| \|\phi\|_\infty \|L^* \chi_n(H)\| \leq c_3^n \|\phi\|_\infty^2,$$

with  $c_3^n \rightarrow 0$  as  $n \rightarrow +\infty$  in view of lemmas 2 and 3.

It remains at this stage to treat the terms  $T_1'$  and  $T_2'$ . Using lemmas 3 and 4 and the coercivity of  $AL_H\Pi$  (granted by (25) and extended to functions in  $L^2(d\mu)$  by a limiting argument),

$$\begin{aligned} T_1' + T_2' &= \langle AL_H\Pi \mathcal{P}^\perp P_t^n \phi, \mathcal{P}^\perp P_t^n \phi \rangle + \langle AL_H(1 - \Pi) \mathcal{P}^\perp P_t^n \phi, \mathcal{P}^\perp P_t^n \phi \rangle \\ &\quad + \gamma \langle AL_{OU} \mathcal{P}^\perp P_t^n \phi, \mathcal{P}^\perp P_t^n \phi \rangle + \langle L^* A \mathcal{P}^\perp P_t^n \phi, \mathcal{P}^\perp P_t^n \phi \rangle \\ &\leq -\frac{\rho}{\rho + 1} \|\Pi \mathcal{P}^\perp P_t^n \phi\|^2 + \left(\eta_\varepsilon + \frac{\gamma}{2}\right) \|(1 - \Pi) \mathcal{P}^\perp P_t^n \phi\| \|\Pi \mathcal{P}^\perp P_t^n \phi\| \\ &\quad + \|(1 - \Pi) \mathcal{P}^\perp P_t^n \phi\|^2. \end{aligned}$$

Combining these estimates, integrating from  $s$  to  $t$ , and then taking  $n \rightarrow \infty$  produces the following estimate

$$\begin{aligned} \langle AP_t \phi, P_t \phi \rangle &\leq \langle AP_s \phi, P_s \phi \rangle + \int_s^t \left[ -\frac{\rho}{\rho + 1} \|\Pi P_u \phi\|^2 + \left(\eta_\varepsilon + \frac{\gamma}{2}\right) \|(1 - \Pi) P_u \phi\| \|\Pi P_u \phi\| \right] du \\ &\quad + \int_s^t \|(1 - \Pi) P_u \phi\|^2 du, \end{aligned}$$

where we used that  $P_t^n \phi$  converges to  $P_t \phi$  in  $L^2(d\mu)$  by the dominated convergence theorem since  $P_t^n \phi$  converges to  $P_t \phi$  pointwise and both quantities are bounded by  $\|\phi\|_\infty$ . This concludes the proof of part (b).

For part (c), we introduce  $\psi = P_t \phi \in C_{0,b}^\infty(\mathcal{X})$ . For any  $n \in \mathbb{N}$ ,

$$\begin{aligned} \frac{d}{dt} \|P_t \phi\|_{L^2(W^* \chi_n(H) d\mu)}^2 &= 2 \int_{\mathcal{X}} \psi (L\psi) W^* \chi_n(H) d\mu \\ &= -2\gamma \int_{\mathcal{X}} |\nabla_p \psi|^2 W^* \chi_n(H) d\mu + \int_{\mathcal{X}} L(\psi^2) W^* \chi_n(H) d\mu. \end{aligned} \quad (65)$$

The second integral on the right-hand side of the above equality can be estimated as follows since  $W^* \in \mathcal{W}_{\alpha,\beta}(L^*)$ :

$$\begin{aligned} \int_{\mathcal{X}} L(\psi^2) W^* \chi_n(H) d\mu &= \int_{\mathcal{X}} \psi^2 L^* [W^* \chi_n(H)] d\mu \\ &= \int_{\mathcal{X}} \psi^2 (L^* W^*) \chi_n(H) d\mu + \int_{\mathcal{X}} \psi^2 W^* [L^* \chi_n(H)] d\mu + 2\gamma \int_{\mathcal{X}} \psi^2 \nabla_p W^* \cdot \nabla_p [\chi_n(H)] d\mu \\ &\leq \int_{\mathcal{X}} \psi^2 (-\alpha W^* + \beta) \chi_n(H) d\mu + \int_{\mathcal{X}} \psi^2 W^* [L^* \chi_n(H)] d\mu + 2\gamma \int_{\mathcal{X}} \psi^2 \nabla_p W^* \cdot \nabla_p [\chi_n(H)] d\mu. \end{aligned}$$

Since  $W^*$  is strongly integrable, lemma 2 implies that the last two integrals on the last line above converge to 0 as  $n \rightarrow \infty$ . To see why, let us for instance show how to establish this fact for the second integral on the right-hand side of the last inequality. Using  $L_H^* \chi_n(H) = 0$ ,

$$L^* \chi_n(H) = \gamma [(-|p|^2 + d) \chi_n'(H) + |p|^2 \chi_n''(H)].$$

Since  $W^*$  is strongly integrable, there exist  $C \in \mathbb{R}_+$  and  $\delta \in (0, 1)$  such that

$$\left| \int_{\mathcal{X}} \psi^2 W^* [L^* \chi_n(H)] d\mu \right| \leq C \|\psi\|_\infty^2 \int_{\mathcal{X}} (|\chi_n'(H)| + |\chi_n''(H)|) (1 + |p|^2) e^{-\delta H} dq dp.$$

Given the specific form of the Hamiltonian  $H$ , there exists  $\tilde{C} \in \mathbb{R}_+$  such that

$$\left| \int_{\mathcal{X}} \psi^2 W^* [L^* \chi_n(H)] d\mu \right| \leq \tilde{C} \|\psi\|_\infty^2 \int_{\mathcal{X}} (|\chi_n'(H)| + |\chi_n''(H)|) e^{-\delta H/2} dq dp.$$

Now, lemma 2(c) ensures that  $e^{-\delta H} \in L^1(dq dp)$ , so that the right-hand side of the above inequality converges to 0 as  $n \rightarrow \infty$  by the dominated convergence theorem. The other term is dealt with in a similar fashion.

The claimed result in (c) then follows by integrating (65) in time, and passing to the limit  $n \rightarrow +\infty$  (using the monotone and dominated convergence theorems).  $\square$

## Acknowledgments

DPH and EC graciously acknowledge support from Grants DMS-1612898 (DPH) and DMS-1855504 (DPH) from the National Science Foundation. The work of GS is funded in part by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 810367), and by the Agence Nationale de la Recherche, under Grant ANR-19-CE40-0010-01 (QuAMProcs). M Gordina was funded from Grant DMS-1954264 from the National Science Foundation.

## ORCID iDs

Evan Camrud  <https://orcid.org/0000-0003-2561-3164>

David P Herzog  <https://orcid.org/0000-0003-1692-7399>

Gabriel Stoltz  <https://orcid.org/0000-0002-2797-5938>

Maria Gordina  <https://orcid.org/0000-0002-9003-7362>

## References

- [1] Albritton D, Armstrong S, Mourrat J-C and Novack M 2019 Variational methods for the kinetic Fokker–Planck equation (arXiv:1902.04037)
- [2] Andrieu C, Durmus A, Nüsken N and Roussel J 2021 Hypocoercivity of piecewise deterministic Markov process-Monte Carlo *Ann. Appl. Probab.* **31** 2478–517
- [3] Bakry D, Barthe F, Cattiaux P and Guillin A 2008 A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case *Electron. Commun. Probab.* **13** 60–6
- [4] Baudoin F, Gordina M and Herzog D P 2021 Gamma calculus beyond Villani and explicit convergence estimates for Langevin dynamics with singular potentials *Arch. Ration. Mech. Anal.* **241** 765–804
- [5] Bernard E, Fathi M, Levitt A and Stoltz G 2021 Hypocoercivity with Schur complements *Ann. Henri Lebesgue* (to appear)
- [6] Bhattacharya R N 1982 On the functional central limit theorem and the law of the iterated logarithm for Markov processes *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* **60** 185–201
- [7] Bou-Rabee N and Sanz-Serna J M 2017 Randomized Hamiltonian Monte Carlo *Ann. Appl. Probab.* **27** 2159–94
- [8] Cancès E, Legoll F and Stoltz G 2007 Theoretical and numerical comparison of some sampling methods for molecular dynamics *ESAIM Math. Model. Numer. Anal.* **41** 351–89
- [9] Cao Y, Lu J and Wang L 2019 On explicit  $L^2$ -convergence rate estimate for underdamped Langevin dynamics (arXiv:1908.04746)
- [10] Cattiaux P, Guillin A, Monmarché P and Zhang C 2019 Entropic multipliers method for Langevin diffusion and weighted log Sobolev inequalities *J. Funct. Anal.* **277** 108288
- [11] Chandra R, Azizi L and Cripps S 2017 Bayesian neural learning via Langevin dynamics for chaotic time series prediction *Neural Information Processing (Lecture Notes in Computer Science)* (Cham: Springer) pp 564–73
- [12] Conrad F and Grothaus M 2010 Construction, ergodicity and rate of convergence of  $N$ -particle Langevin dynamics with singular potentials *J. Evol. Equ.* **10** 623–62
- [13] Cooke B, Herzog D P, Mattingly J C, McKinley S A and Schmidler S C 2017 Geometric ergodicity of two-dimensional Hamiltonian systems with a Lennard-Jones-like repulsive potential *Commun. Math. Sci.* **15** 1987–2025
- [14] Cuneo N, Eckmann J-P, Hairer M and Rey-Bellet L 2018 Non-equilibrium steady states for networks of oscillators *Electron. J. Probab.* **23** 55
- [15] Dolbeault J, Klar A, Mouhot C and Schmeiser C 2013 Exponential rate of convergence to equilibrium for a model describing fiber lay-down processes *Appl. Math. Res. eXpress* **2013** 165–75
- [16] Dolbeault J, Mouhot C and Schmeiser C 2009 Hypocoercivity for kinetic equations with linear relaxation terms *C. R. Math. Acad. Sci. Paris* **347** 511–6
- [17] Dolbeault J, Mouhot C and Schmeiser C 2015 Hypocoercivity for linear kinetic equations conserving mass *Trans. Am. Math. Soc.* **367** 3807–28
- [18] Durrant J D and McCammon J A 2011 Molecular dynamics simulations and drug discovery *BMC Biol.* **9** 1–9
- [19] Eckmann J-P and Hairer M 2003 Spectral properties of hypoelliptic operators *Commun. Math. Phys.* **235** 233–53
- [20] Evans L C 2010 *Partial Differential Equations (Graduate Studies in Mathematics)* (vol 19) 2nd edn (Providence, RI: American Mathematical Society)
- [21] Grothaus M and Stilgenbauer P 2015 A hypocoercivity related ergodicity method for singularly distorted non-symmetric diffusions *Integr. Equ. Oper. Theory* **83** 331–79

- [22] Grothaus M and Stilgenbauer P 2016 Hilbert space hypocoercivity for the Langevin dynamics revisited *Methods Funct. Anal. Topol.* **22** 152–68
- [23] Hairer M and Mattingly J C C 2009 Slow energy dissipation in anharmonic oscillator chains *Commun. Pure Appl. Math.* **62** 999–1032
- [24] Hairer M and Mattingly J C C 2011 Yet another look at Harris’ ergodic theorem for Markov chains *Random Fields and Applications VI (Progress in Probability)* vol 63 (Basel: Springer) pp 109–17
- [25] Hairer M and Pavliotis G A 2008 From ballistic to diffusive behavior in periodic potentials *J. Stat. Phys.* **131** 175–202
- [26] Helffer B and Nier F 2005 *Hypoelliptic Estimates and Spectral Theory for Fokker–Planck Operators and Witten Laplacians (Lecture Notes in Mathematics)* vol 1862 (Berlin: Springer)
- [27] Hérau F 2006 Hypocoercivity and exponential time decay for the linear inhomogeneous relaxation Boltzmann equation *Asymptot. Anal.* **46** 349–59
- [28] Hérau F 2007 Short and long time behavior of the Fokker–Planck equation in a confining potential and applications *J. Funct. Anal.* **244** 95–118
- [29] Hérau F and Nier F 2004 Isotropic hypoellipticity and trend to equilibrium for the Fokker–Planck equation with a high-degree potential *Arch. Ration. Mech. Anal.* **171** 151–218
- [30] Herzog D P and Mattingly J C 2019 Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials *Commun. Pure Appl. Math.* **72** 2231–55
- [31] Hoogerbrugge P J and Koelman J M V A 1992 Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics *Europhys. Lett.* **19** 155
- [32] Hörmander L 1967 Hypoelliptic second order differential equations *Acta Math.* **119** 147–71
- [33] Iacobucci A, Olla S and Stoltz G 2019 Convergence rates for nonequilibrium Langevin dynamics *Ann. Math. Québec* **43** 73–98
- [34] Khasminskii R 2011 *Stochastic Stability of Differential Equations* vol 66 (Heidelberg: Springer)
- [35] Kliemann W 1987 Recurrence and invariant measures for degenerate diffusions *Ann. Probab.* **15** 690–707
- [36] Leimkuhler B and Matthews C 2015 *Molecular Dynamics (Interdisciplinary Applied Mathematics)* vol 39 (Cham: Springer)
- [37] Li C, Chen C, Carlson D and Carin L 2016 Preconditioned stochastic gradient Langevin dynamics for deep neural networks *Proc. AAAI Conf. Artificial Intelligence* vol 30
- [38] Lu J and Wang L 2021 On explicit  $L^2$ -convergence rate estimate for piecewise deterministic Markov processes in MCMC algorithms *Ann. Appl. Probab.* (to appear)
- [39] Lu Y and Mattingly J C 2019 Geometric ergodicity of Langevin dynamics with Coulomb interactions *Nonlinearity* **33** 675
- [40] Mattingly J C, Stuart A M and Higham D J 2002 Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise *Stoch. Process. Appl.* **101** 185–232
- [41] Meyn S P and Tweedie R L 1993 Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes *Adv. Appl. Probab.* **25** 518–48
- [42] Mouhot C and Neumann L 2006 Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus *Nonlinearity* **19** 969–98
- [43] Øksendal B 2003 *Stochastic Differential Equations* (Heidelberg: Springer)
- [44] Bellet L R 2006 Ergodic properties of Markov processes *Open Quantum Systems II (Lecture Notes in Mathematics)* vol 1881 ed S Attal, A Joye and C-A Pillet (Berlin: Springer) pp 1–39
- [45] Roussel J and Stoltz G 2018 Spectral methods for Langevin dynamics and associated error estimates *ESAIM Math. Model. Numer. Anal.* **52** 1051–83
- [46] Roussel J and Stoltz G 2019 A perturbative approach to control variates in molecular dynamics *Multiscale Model. Simul.* **17** 552–91
- [47] Talay D 2002 Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme *Markov Processes Relat. Fields* **8** 163–98
- [48] Teh Y W, Thiery A H and Vollmer S J 2016 Consistency and fluctuations for stochastic gradient Langevin dynamics *J. Mach. Learn. Res.* **17** 1–33
- [49] Tropper M M 1977 Ergodic and quasideterministic properties of finite-dimensional stochastic systems *J. Stat. Phys.* **17** 491–509
- [50] Tuckerman M 2010 *Statistical Mechanics: Theory and Molecular Simulation* (New York: Oxford University Press)
- [51] Villani C 2009 Hypocoercivity *Memoirs Am. Math. Soc.* **202** 950

- [52] Welling M and Teh Y W 2011 Bayesian learning via stochastic gradient Langevin dynamics *Proc. 28th Int. Conf. Machine Learning (ICML-11)* pp 681–8
- [53] Wu L 2001 Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems *Stoch. Process. Appl.* **91** 205–38