

Hypocoercivity with Schur complements

E. Bernard,¹ M. Fathi,² A. Levitt^{3,1} and G. Stoltz^{1,3}

1: CERMICS, Ecole des Ponts, Marne-la-Vallée, France

2: LPSM and LJLL, Université de Paris, France

3: MATHERIALS project-team, Inria Paris, France

April 23, 2022

Abstract

We propose an approach to obtaining explicit estimates on the resolvent of hypocoercive operators by using Schur complements, rather than from an exponential decay of the evolution semigroup combined with a time integral. We present applications to Langevin-like dynamics and Fokker–Planck equations, as well as the linear Boltzmann equation (which is also the generator of randomized Hybrid Monte Carlo in molecular dynamics). In particular, we make precise the dependence of the resolvent bounds on the parameters of the dynamics and on the dimension. We also highlight the relationship of our method with other hypocoercive approaches.

1 Introduction

Degenerate dynamics appear in many contexts. Two prominent examples are Langevin-type dynamics in molecular dynamics (whose associated generators are Fokker–Planck operators), and Boltzmann-type equations in the kinetic theory of fluids. From an analytical viewpoint, these operators have a degenerate dissipative structure. For instance, Fokker–Planck operators are partial differential operators with degenerate second order derivatives, in contrast to generators associated with non-degenerate diffusive dynamics which have full second order derivatives. The key point in proving the longtime convergence of degenerate dynamics is to retrieve some dissipation in all degrees of freedom by a combination of the transport part of the evolution and the degenerate diffusion – as provided by the hypocoercive techniques reviewed below.

We first briefly describe the two models which motivate our study (Langevin-type dynamics and linear Boltzmann equation), and then provide a review of existing hypocoercive approaches. We then turn to our motivation, namely providing bounds on the resolvent of the partial differential operator under consideration directly, without going through kinetic estimates on the evolution semigroups; and describe our approach based on Schur complements.

Paradigmatic hypocoercive dynamics. Our first motivation for studying hypocoercive operators stems from molecular dynamics [41, 82, 57, 60], the computational armed wing of statistical physics [9]. One family of dynamics commonly used to compute average properties according to the Boltzmann–Gibbs measure are Langevin dynamics. These dynamics evolve the positions q and momenta p of given particles interacting via a potential energy function $V(q)$ according to the following stochastic differential equation:

$$\begin{cases} dq_t = p_t dt, \\ dp_t = -\nabla V(q_t) dt - \gamma p_t dt + \sigma dW_t, \end{cases}$$

where W_t is a standard Brownian motion and γ, σ are positive numbers. The first result on the convergence of Langevin-like dynamics to the stationary measure is [81], which however does not

provide explicit convergence rates. The generator associated with Langevin dynamics

$$\mathcal{L}_{\text{Lang}} = p \cdot \nabla_q - \nabla V(q) \cdot \nabla_p - \gamma p \cdot \nabla_p + \frac{\sigma^2}{2} \Delta_p$$

is degenerate since second derivatives in q are missing.

Another classical example of hypocoercive equation is the linear Boltzmann equation modelling the behaviour of material particles interacting with the environment and subjected to a potential V . The behavior of the particles is described in an average way by the density $f(t, y, v)$ which gives the probability to observe a particle at position y with velocity v at time t . This density evolves as

$$\frac{\partial f}{\partial t}(t, y, v) - \nabla V(y) \cdot \nabla_v f(t, y, v) + v \cdot \nabla_y f(t, y, v) = (Mf)(t, y, v). \quad (1)$$

Here, M is the collision operator modelling the interactions between the particles and the environment, which leads to some dissipative structure in the velocities. In its simplest form, it is constructed from an integral operator with kernel k :

$$(Mf)(t, y, v) = \int_{\mathbb{R}^d} k(y, v, w) f(t, y, w) dw - \sigma(y, v) f(t, y, v),$$

with the equilibrium condition

$$\int_{\mathbb{R}^d} k(y, v, w) dv = \sigma(y, w),$$

meaning that there is neither creation nor annihilation of particles. A special case is to choose for $k(y, v, w)$ a positive multiple of a Gaussian density with identity covariance in the v variable (independent of y, w), which corresponds in molecular dynamics to the generator of the so-called randomized Hybrid Monte Carlo method [17]. The operator

$$\mathcal{L}_{\text{Boltz}} = \nabla V \cdot \nabla_v - v \cdot \nabla_y + M$$

generates a strongly continuous semigroup in L^2 , and under various conditions on k and σ , it decays exponentially fast towards a global equilibrium in long time. This exponential decay arises from some coupling between the Hamiltonian transport part and the dissipation in the velocities. Besides the hypocoercive techniques reviewed below, let us mention other convergence results specific to this equation, based on spectral methods [83, 84], control theory [46] or the theory of positive semigroups (see [13, 66] and references therein). However, these methods are not constructive and therefore do not provide explicit quantitative estimates of the exponential rate of decay. This is in contrast with hypocoercive approaches, which allow to obtain quantitative bounds by a clever combination of the transport and dissipation part of the dynamics.

A literature review on convergence results for operators with degenerate dissipation.

Let us list various approaches to study the exponential convergence of semigroups associated with hypoelliptic or degenerate generators, in a somewhat chronological order (see also the recent review [50]):

- A first set of results is based on Lyapunov techniques [86, 63, 75], the typical Lyapunov function being the total energy of the system plus some term coupling positions and momenta. This approach was mostly used for Fokker–Planck operators such as the generators of Langevin dynamics, but it can also be used for specific Boltzmann-type operators [17, 24, 22]. The corresponding convergence rates are however usually not very explicit in term of the parameters of the dynamics, because it is difficult to quantify the constants appearing in the so-called minorization condition in terms of the parameters of the dynamics (such as the friction γ for Fokker–Planck operators associated with Langevin dynamics). There were however recent attempts at providing more quantitative minorization conditions, using tools from Malliavin calculus [39].

- Subelliptic estimates [37, 51, 47] allow to obtain detailed information on the spectrum of Fokker–Planck operators (discrete nature, localization in a cusp region). The spectrum is completely known for systems with zero potential on a torus [55], or for quadratic potential energy functions (see [65, 73]).
- H^1 hypocoercivity was pioneered in [80, 70] and later abstractified in [85]. The application of this theory to Langevin dynamics allows to quantify the convergence rates in terms of the parameters of the dynamics; see for instance [45] for the Hamiltonian limit and [58, 60] for the overdamped limit. The method can be extended to Generalized Langevin dynamics [72] and certain piecewise deterministic Markov processes (whose generators are a specific form of Boltzmann operators) [28]. It can also be used to study the discretization of Fokker–Planck equations [34, 42].

Exponential convergence in H^1 can be established using only L^2 bounds on the initial data by hypoelliptic regularization [49]. The latter technique can also be adapted to discretization schemes [74].

- Entropic estimates for Fokker–Planck operators were initiated in [29] and abstractified in [85], though under conditions stronger than those for H^1 . Recently, it was shown in [27] how to remove the assumption that the Hessian of the potential is bounded. Entropic hypocoercivity for the linear Boltzmann equation was studied in [40, 69], and for linearized BGK models (that generalize the linear Boltzmann equation) in spatial dimension one in [1].
- A more direct route to proving the convergence in L^2 was first proposed in [48], then extended in [32, 33], and revisited in [43] where domain issues of the operators at play are addressed. This method can be applied to Fokker–Planck operators and Boltzmann-type operators. It is based on a modification of the L^2 scalar product with some regularization operator. This more direct approach makes it even easier to quantify convergence rates; see [31, 43] for studies on the dependence of parameters such as friction in Langevin dynamics, as well as [2, 5] for sharp estimates for equilibrium Langevin dynamics and a harmonic potential energy function.

The approach can be perturbatively extended to nonequilibrium situations [20, 79, 54] and to discretizations of the generator, either via spectral methods [76] or with finite volume schemes [14]. It also allows to consider non-quadratic kinetic energies (for which the associated generator may fail to be hypoelliptic) [78]. It can be applied to various dynamics, such as Adaptive Langevin [59] or certain piecewise deterministic Markov processes [3]. Finally, a combination of this approach with a mode-by-mode analysis in Fourier space was introduced in [19] to establish polynomial decay estimates for dynamics without confinement, and extended in [18] to cover fractional diffusion operators. Alternatively, it is also possible to rely on weak Poincaré estimates to obtain convergence rates when there is no confinement [44].

- It was shown recently how to use techniques from Γ_2 calculus for degenerate operators corresponding to Langevin dynamics [10, 67]. This approach allows to study more degenerate dynamics corresponding to the evolution of chains of oscillators [64].
- Fully probabilistic techniques, based on clever coupling strategies, can also be used to obtain the exponential convergence of the law of Langevin processes to their stationary state [36]. One interest of this approach is that the drift needs not be gradient, in contrast to standard analytical approaches for which the analytical expression of the invariant measure should be known in order to separate the symmetric and antisymmetric parts of the generator under consideration. This coupling approach can be combined with ideas from Γ_2 calculus [26].
- Finally, it was recently shown how to directly obtain L^2 estimates without changing the scalar product, relying on a space-time Poincaré inequality to conclude to an exponential convergence in time of the evolution semigroup. The method was initially developed for the generators of Langevin dynamics [4, 23], but can be extended to certain Boltzmann-type operators [62] (see Appendix A below for a more detailed account).

Note that degenerate norms can be considered for approaches based on H^1 , L^2 or Wasserstein norms, as initially done in [80], and recently used in [10, 61, 54]. It is also possible to extend various convergence results to singular interaction potentials [53, 52, 11, 21], and systems with several conservation laws [25]. Let us finally mention that some approaches are related one to another, such as Lyapunov techniques and estimates based on Poincaré inequalities [7].

Motivations. Our aim in this work is to directly obtain bounds on the resolvent \mathcal{L}^{-1} , as for subelliptic estimates [37, 51, 47], without going through kinetic estimates on the semigroup $e^{t\mathcal{L}}$ which usually require a modified scalar product (often involving a small parameter which makes the final estimates on the resolvent less quantitative). The motivation for obtaining such resolvent bounds stems from the fact they are a key element for proving that a Central Limit Theorem holds for time averages along realizations of Langevin-like dynamics, and for obtaining estimates on the associated asymptotic variance; see for example [15] which shows that a sufficient condition for the Central Limit Theorem to hold in this context is that $-\mathcal{L}^{-1}$ is well defined on a subspace of $L^2(\mu)$, where μ is the invariant measure of the underlying stochastic dynamics. Note that it is also possible to use hypocoercive results to obtain non-asymptotic concentration inequalities for time averages beyond the Central Limit Theorem [16, 68].

From a more algebraic perspective, our motivation was to understand the origin of the expression of the regularization operator which appears in the modified L^2 scalar product in [48, 32, 33], as well as the algebraic manipulations of [4, 23], which bear some similarity with some computations in [48, 32, 33]. Our hope is to extract a structure as general as possible, which encompasses many hypocoercive dynamics such as underdamped Langevin with non-quadratic kinetic energies [78], the linear Boltzmann equation or randomized Hybrid Monte Carlo with jump processes on the momenta [17], Adaptive Langevin dynamics with a dynamical friction [52, 59], Langevin dynamics with extra non-reversible perturbation [35], etc.

Saddle-point problems and Schur complements. We now motivate the technical approach we use, which is based on Schur complements once the operator under consideration has been written in a form reminiscent of a saddle-point problem. We illustrate the idea on a very simple example of stationary equation displaying some hypocoercive structure, namely the stationary Stokes equation:

$$\begin{aligned}\nabla \cdot u &= 0, \\ -\alpha \Delta u + \nabla p &= f,\end{aligned}$$

where $\alpha > 0$ is a viscosity constant, $u : \Omega \rightarrow \mathbb{R}^d$ is the velocity field and $p : \Omega \rightarrow \mathbb{R}$ is the pressure field, with Ω a subset of \mathbb{R}^d ($d = 2$ or 3). For simplicity, we take for Ω the torus $(\mathbb{R}/(2\pi\mathbb{Z}))^d$ with periodic boundary conditions, and add the constraints $\int_{\Omega} p = 0$, $\int_{\Omega} u = 0$. In a vector representation, the equation can be rewritten as $\mathcal{L}_{\text{Stokes}}(p, u) = (0, f)$, with

$$\mathcal{L}_{\text{Stokes}} = \begin{pmatrix} 0 & \nabla^T \\ \nabla & -\alpha \Delta \end{pmatrix}.$$

Expanding p, u and f into Fourier series, the equation is decoupled for each mode $n \in \mathbb{Z}^d$, resulting in the $(1 + d)$ -dimensional algebraic equation $\mathcal{L}_{\text{Stokes}, n}(p_n, u_n) = (0, f_n)$, where

$$\mathcal{L}_{\text{Stokes}, n} = \begin{pmatrix} 0 & in^T \\ in & \alpha|n|^2 \text{Id}_d \end{pmatrix}.$$

This $(1 + d) \times (1 + d)$ matrix is coercive in the last d variables (associated with the velocity field u), but not in the first one (associated with the pressure field p). However, we can first use the second equation to solve u_n as a function of p_n :

$$u_n = \alpha^{-1}|n|^{-2}(f_n - in p_n).$$

Next, replacing in the first equation, we obtain

$$\mathfrak{S}_n p_n = in \cdot (\alpha^{-1} |n|^{-2} f_n),$$

where $\mathfrak{S}_n = in^T \alpha^{-1} |n|^{-2} in = -\alpha^{-1}$ is the *Schur complement*. This shows that $\mathcal{L}_{\text{Stokes},n}$ is invertible for all $n \in \mathbb{Z} \setminus \{0\}$; together with the zero-mean constraints on u and p , this implies that the original problem has a unique solution. This scheme of proof generalizes to other types of boundary conditions or equations (see for instance [38, Section 2.4]).

The Stokes equation is an example of a more general class of problems characterized by the block operator structure

$$L = \begin{pmatrix} 0 & B^T \\ -B & A \end{pmatrix}. \quad (2)$$

This type of problems arises for instance as the Hessian of the Lagrangian of a minimization problem with linear constraints. The expression of L corresponds to a splitting of the variables as (u, p) , where u are the primal variables and p the dual variables (Lagrange multipliers). In this case, solutions are saddle points of the Lagrangian, hence the general name of *saddle-point problems* (see [12] for a comprehensive review). As for the Stokes equation above, when A is invertible, one can solve the equation for the second variable as a function of the first; in the finite-dimensional case, this shows that the matrix L is invertible if and only if the Schur complement $\mathfrak{S} = B^T A^{-1} B$ is, which is equivalent to B having full column rank.

The main result of this paper is an infinite-dimensional generalization of this simple approach that covers the case of Fokker–Planck and Boltzmann-type operators. In these applications, the block structure corresponds to the separation between the indirectly dissipated variables (position) and the directly dissipated ones (momentum). The algebraic structure is the same as the one above (compare (2) and (4)), but the analytical structure (bounds on operators) that we use is adapted to the specific case of kinetic operators.

Outline of the work. From a technical point of view, our approach based on Schur complements allows to obtain a formal expression of the resolvent from a decomposition of the operator under study according to the kernel and image of the degenerate symmetric part of the operator, which is at the origin of the degenerate dissipation. To our knowledge, Schur complements were not used to understand the behavior of hypocoercive operators, although they are well known in the mathematical physics literature where they are used for spectral analysis (see for instance the review [77]). As already hinted at above, the main benefit of our approach is that it avoids the appearance of somewhat uncontrolled prefactors in the bounds on the resolvent, in contrast to estimates obtained from the exponential decay of the evolution semigroup; see in particular Theorem 1 for general estimates, and for instance Corollary 1 and Proposition 3 for illustrative applications of this general result.

We start by presenting the method in an abstract setting in Section 2, and then apply it to various dynamics in Section 3. We discuss some relationship with the recent works [4, 23] in Appendix A.

Extensions and future work tracks. This work calls for various extensions and refinements, on which we are currently working, including in particular:

- (i) obtaining resolvent estimates for $(z - \mathcal{L})^{-1}$, for negative real parts of z sufficiently small compared to the imaginary values of z (as already obtained with subelliptic estimates [37, 51, 47]);
- (ii) extending the approach to more degenerate dynamics, starting with the Generalized Langevin dynamics [72], and if possible addressing oscillator chains.

2 Abstract resolvent estimates

2.1 Schur decomposition of the generator

We consider the Hilbert space $L^2(\mu)$ for some Boltzmann–Gibbs probability measure $\mu(dx)$ on a configuration space $\mathcal{X} = \mathbb{R}^D$ (the analysis could however be straightforwardly extended to situations such as $\mathcal{X} = \mathbb{T}^D$ with $\mathbb{T} = \mathbb{R}/\mathbb{Z}$). The space $L^2(\mu)$ is equipped with its canonical scalar product, denoted by $\langle \cdot, \cdot \rangle$. The induced norm is denoted by $\|\cdot\|$. We consider a stochastic dynamics with generator

$$\mathcal{L} = \mathcal{S} + \mathcal{A},$$

seen as an unbounded operator on $L^2(\mu)$, where

$$\mathcal{S} = \frac{\mathcal{L} + \mathcal{L}^*}{2}, \quad \mathcal{A} = \frac{\mathcal{L} - \mathcal{L}^*}{2}$$

are respectively the symmetric (dissipation) and antisymmetric (transport) parts of \mathcal{L} . Adjoint operators are considered with respect to the scalar product $\langle \cdot, \cdot \rangle$. We assume that the space \mathcal{C} of smooth and uniformly bounded functions with uniformly bounded derivatives is a core for both \mathcal{A} and \mathcal{S} . In the following, with some abuse of notation, we denote by \mathcal{A} and \mathcal{S} the closure of these operators.

We assume that μ is invariant by the dynamics, which translates into:

$$\forall \varphi \in \mathcal{C}, \quad \int_{\mathcal{X}} \mathcal{L}\varphi d\mu = 0.$$

We also assume that $\mathcal{L}\mathbf{1} = 0$ (which for instance is obviously the case for generators associated with diffusion processes; note that for hypocoercive dynamics, the assumption of the existence of an invariant probability measure implies that no blowup can occur). We define the subspace of functions of $L^2(\mu)$ with average 0 with respect to μ :

$$\mathcal{H} = \left\{ \varphi \in L^2(\mu) \left| \int_{\mathcal{X}} \varphi(x) \mu(dx) = \langle \varphi, \mathbf{1} \rangle = 0 \right. \right\}.$$

In the orthogonal decomposition $L^2(\mu) = \text{Ran}(\mathbf{1}) \oplus \mathcal{H}$, the operator \mathcal{L} can therefore be written as

$$\mathcal{L} = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{L}|_{\mathcal{H}} \end{pmatrix},$$

where $\mathcal{L}|_{\mathcal{H}}$ denotes the restriction of \mathcal{L} to \mathcal{H} . In particular, the inverse of \mathcal{L} cannot be defined on the whole space $L^2(\mu)$, but only on \mathcal{H} . In the sequel, we work on the Hilbert space \mathcal{H} only. We still denote by $\mathcal{L}, \mathcal{S}, \mathcal{A}$ the restrictions of $\mathcal{L}, \mathcal{S}, \mathcal{A}$ to \mathcal{H} .

Our first structural assumption is the following.

Assumption 1. *There exists an orthogonal projector Π_0 such that $\Pi_0\mathcal{C} \subset \mathcal{C}$ and*

$$\Pi_0\mathcal{A}\Pi_0 = 0, \quad \mathcal{S}\Pi_0 = \Pi_0\mathcal{S} = 0. \quad (3)$$

We denote by $\Pi_{\star} = 1 - \Pi_0$ the orthogonal projector complementary to Π_0 , and by

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_{\star}$$

the resulting decomposition of the Hilbert space. In the sequel, when T is an operator on \mathcal{H} and α, β labels, we use the notation $\mathcal{H}_{\alpha} = \text{Ran}(\Pi_{\alpha})$ for the subspaces of \mathcal{H} and

$$T_{\alpha\beta} = \Pi_{\alpha}T\Pi_{\beta} : \mathcal{H}_{\beta} \rightarrow \mathcal{H}_{\alpha},$$

for the restrictions (blocks) of T . When $T_{\alpha\beta}$ is a bounded operator, we denote its operator norm by

$$\|T_{\alpha\beta}\| = \sup_{f \in \mathcal{H}_\beta \setminus \{0\}} \frac{\|T_{\alpha\beta}f\|}{\|f\|}.$$

In view of Assumption 1, the operator \mathcal{L} can be written on $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_+$ as

$$\mathcal{L} = \begin{pmatrix} 0 & \mathcal{L}_{0+} \\ \mathcal{L}_{+0} & \mathcal{L}_{++} \end{pmatrix} = \begin{pmatrix} 0 & \mathcal{A}_{0+} \\ \mathcal{A}_{+0} & \mathcal{L}_{++} \end{pmatrix} \quad (4)$$

with $\mathcal{A}_{0+} = -\mathcal{A}_{+0}^*$. We next *formally* compute the action of \mathcal{L}^{-1} by solving the following linear system for a given couple $(\phi_0, \phi_+) \in \mathcal{H}_0 \times \mathcal{H}_+$:

$$\begin{pmatrix} 0 & \mathcal{A}_{0+} \\ \mathcal{A}_{+0} & \mathcal{L}_{++} \end{pmatrix} \begin{pmatrix} u_0 \\ u_+ \end{pmatrix} = \begin{pmatrix} \phi_0 \\ \phi_+ \end{pmatrix}.$$

The second line leads to $u_+ = \mathcal{L}_{++}^{-1}(\phi_+ - \mathcal{A}_{+0}u_0)$, and then the first line to $u_0 = (\mathcal{A}_{0+}\mathcal{L}_{++}^{-1}\mathcal{A}_{+0})^{-1}(-\phi_0 + \mathcal{A}_{0+}\mathcal{L}_{++}^{-1}\phi_+)$. Therefore,

$$\mathcal{L}^{-1} = \begin{pmatrix} \mathfrak{S}_0^{-1} & -\mathfrak{S}_0^{-1}\mathcal{A}_{0+}\mathcal{L}_{++}^{-1} \\ -\mathcal{L}_{++}^{-1}\mathcal{A}_{+0}\mathfrak{S}_0^{-1} & \mathcal{L}_{++}^{-1} + \mathcal{L}_{++}^{-1}\mathcal{A}_{+0}\mathfrak{S}_0^{-1}\mathcal{A}_{0+}\mathcal{L}_{++}^{-1} \end{pmatrix} \quad (5)$$

where

$$\mathfrak{S}_0 = -\mathcal{A}_{0+}\mathcal{L}_{++}^{-1}\mathcal{A}_{+0} = \mathcal{A}_{+0}^*\mathcal{L}_{++}^{-1}\mathcal{A}_{+0}$$

is the *Schur complement* associated with the decomposition (4). The subscript 0 in \mathfrak{S}_0 emphasizes that this operator acts on \mathcal{H}_0 .

To give a meaning to the computations leading to (5), we need to ensure that both \mathcal{L}_{++} and \mathfrak{S}_0 are invertible, and that the operators appearing in (5) are bounded. We start with conditions ensuring the invertibility of \mathcal{L}_{++} .

Assumption 2. *There exists $s > 0$ such that $-\mathcal{S} \geq s\Pi_+$ in the sense of symmetric operators.*

This shows in particular that $-\mathcal{L}_{++}$ is coercive on \mathcal{H}_+ , and therefore invertible, with

$$\|\mathcal{L}_{++}^{-1}\| \leq s^{-1}.$$

It is clear from (4) that if \mathcal{A}_{+0} has a non trivial kernel, then \mathcal{L} cannot be invertible. We therefore need a quantitative assumption about the injectivity of this operator (a property called “macroscopic coercivity” in [33]).

Assumption 3 (Macroscopic coercivity). *There exists $a > 0$ such that*

$$\forall \varphi \in \mathcal{H}_0 \cap \mathcal{C}, \quad \|\mathcal{A}_{+0}\varphi\| \geq a\|\varphi\|. \quad (6)$$

Let us next introduce the range of \mathcal{A}_{+0} :

$$\mathcal{H}_1 = \text{Ran}(\mathcal{A}_{+0}) = \overline{\{\mathcal{A}_{+0}\varphi, \varphi \in \mathcal{H}_0 \cap \mathcal{C}\}}^{\mathcal{H}_+}.$$

Assumption 3 ensures that $-\mathcal{A}_{0+}\mathcal{A}_{+0} = \mathcal{A}_{+0}^*\mathcal{A}_{+0} \geq a^2$, so that $-\mathcal{A}_{0+}\mathcal{A}_{+0}$ is invertible on \mathcal{H}_0 . It is then easily seen that the operator

$$\Pi_1 = \mathcal{A}_{+0}(\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1}\mathcal{A}_{+0}^* \quad (7)$$

is an orthogonal projector, and that its range is \mathcal{H}_1 , so that Π_1 is the orthogonal projector onto \mathcal{H}_1 .

We further decompose \mathcal{H}_* orthogonally as $\mathcal{H}_* = \mathcal{H}_1 \oplus \mathcal{H}_2$, which boils down to defining $\mathcal{H}_2 = (\mathcal{H}_0 \oplus \mathcal{H}_1)^\perp$. We denote by Π_2 the orthogonal projector onto \mathcal{H}_2 . We therefore have the following decomposition of \mathcal{H} :

$$\mathcal{H} = \mathcal{H}_0 \oplus \underbrace{\mathcal{H}_1 \oplus \mathcal{H}_2}_{=\mathcal{H}_*},$$

which induces the following decomposition of the generator:

$$\mathcal{L} = \begin{pmatrix} 0 & \mathcal{A}_{01} & 0 \\ \mathcal{A}_{10} & \mathcal{L}_{11} & \mathcal{L}_{12} \\ 0 & \mathcal{L}_{21} & \mathcal{L}_{22} \end{pmatrix}, \quad \mathcal{A}_{01} = -\mathcal{A}_{10}^*.$$

The operator $\mathcal{A}_{10} : \mathcal{H}_0 \rightarrow \mathcal{H}_1$ is injective by Assumption 3, and surjective by definition of \mathcal{H}_1 . It is therefore invertible, and the norm of its inverse \mathcal{A}_{10}^{-1} is bounded by $1/a$ in view of (6). It can be easily checked that the explicit expression of the inverse is

$$\mathcal{A}_{10}^{-1} = (\mathcal{A}_{10}^* \mathcal{A}_{10})^{-1} \mathcal{A}_{10}^* = (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \mathcal{A}_{10}^*.$$

Remark 1. Note that \mathcal{A}_{10}^{-1} is quite similar to the regularization operator $R = (1 + \mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \mathcal{A}_{+0}^*$ introduced in [48, 32, 33] to modify the scalar product on $L^2(\mu)$ for proving hypocoercive properties. More precisely, the main steps of the proof in [32, 33] are to introduce a modified $L^2(\mu)$ scalar product, induced by the modified norm $\langle M_\varepsilon f, f \rangle$ with

$$M_\varepsilon = \begin{pmatrix} 1 & \varepsilon R \\ \varepsilon R^* & 1 \end{pmatrix}$$

and equivalent to the canonical one; and to prove that $-\mathcal{L}$ is coercive with respect to this modified scalar product, i.e. $\langle -M_\varepsilon \mathcal{L} f, f \rangle \geq \kappa \langle M_\varepsilon f, f \rangle$ for some $\kappa > 0$. The proof can be extended to any regularization operator of the form $R = (\eta + \mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \mathcal{A}_{+0}^*$ with $\eta > 0$. The inverse \mathcal{A}_{10}^{-1} is recovered in the limit $\eta \rightarrow 0$. Our algebraic derivation therefore provides a complementary viewpoint on the origin of the regularization operator in [48, 32, 33].

We next remark that

$$\mathfrak{S}_0 = \mathcal{A}_{+0}^* \mathcal{L}_{++}^{-1} \mathcal{A}_{+0} = \mathcal{A}_{10}^* [\mathcal{L}_{++}^{-1}]_{11} \mathcal{A}_{10}. \quad (8)$$

Both operators $-\mathcal{L}_{++}$ and $-\mathcal{L}_{22}$ are coercive and hence invertible. We can then consider a second Schur complement to compute the action of $[\mathcal{L}_{++}^{-1}]_{11}$:

$$[\mathcal{L}_{++}^{-1}]_{11} = \mathfrak{S}_1^{-1}, \quad \mathfrak{S}_1 = \mathcal{L}_{11} - \mathcal{L}_{12} \mathcal{L}_{22}^{-1} \mathcal{L}_{21}.$$

Therefore, $\mathfrak{S}_0 = \mathcal{A}_{10}^* \mathfrak{S}_1^{-1} \mathcal{A}_{10}$, and so

$$\begin{aligned} \mathfrak{S}_0^{-1} &= \mathcal{A}_{10}^{-1} (\mathcal{L}_{11} - \mathcal{L}_{12} \mathcal{L}_{22}^{-1} \mathcal{L}_{21}) \mathcal{A}_{10}^{-*} \\ &= (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \mathcal{A}_{10}^* (\mathcal{L}_{11} - \mathcal{L}_{12} \mathcal{L}_{22}^{-1} \mathcal{L}_{21}) \mathcal{A}_{10} (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1}. \end{aligned} \quad (9)$$

We introduce a final structural assumption, which simplifies the action of some of the operators.

Assumption 4. There exists an involution \mathcal{R} on \mathcal{H} (i.e. a bounded operator for which $\mathcal{R}^2 = 1$) such that

$$\mathcal{R} \Pi_0 = \Pi_0 \mathcal{R} = \Pi_0, \quad \mathcal{R} \mathcal{S} \mathcal{R} = \mathcal{S}, \quad \mathcal{R} \mathcal{A} \mathcal{R} = -\mathcal{A}.$$

This assumption is motivated by the physical notion of microscopic reversibility for kinetic dynamics, where \mathcal{R} can be chosen as the momentum-reversal operator (see Section 3). Assumption 4 implies in particular that

$$\mathcal{R} \mathcal{L} \mathcal{R} = \mathcal{L}^*. \quad (10)$$

Since $\mathcal{R}\Pi_+ = \mathcal{R}(1 - \Pi_0) = \Pi_+\mathcal{R}$, it holds

$$\mathcal{R}\mathcal{A}_{+0} = -\mathcal{A}_{+0}, \quad (11)$$

and therefore $\mathcal{R}\Pi_1 = -\Pi_1$. Similarly, $\Pi_1\mathcal{R} = -\Pi_1$. The operator \mathcal{R} can therefore be decomposed as

$$\mathcal{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \mathcal{R}_{22} \end{pmatrix}.$$

It then follows from (10) that \mathcal{L}_{11} is symmetric, *i.e.* $\mathcal{L}_{11} = \mathcal{S}_{11}$. In view of the latter equality and of (9), we finally assume the following to ensure the invertibility of the Schur complement \mathfrak{S}_0 .

Assumption 5. *The operators \mathcal{S}_{11} and $\mathcal{L}_{21}\mathcal{A}_{10}(\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1}$ are bounded.*

As we shall see in Section 3, Assumptions 1 to 5 are satisfied for generators of Langevin-like dynamics, under suitable conditions on the potential and kinetic energies. The final result can then be summarized as follows.

Theorem 1. *Suppose that Assumptions 1 to 5 hold true. Then, \mathcal{L} is invertible on \mathcal{H} and*

$$\|\mathcal{L}^{-1}\| \leq 2 \left(\frac{\|\mathcal{S}_{11}\|}{a^2} + \frac{\|\mathcal{R}_{22}\| \|\mathcal{L}_{21}\mathcal{A}_{10}(\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1}\|^2}{s} \right) + \frac{3}{s}.$$

Proof. From (9) and since $\mathcal{L}_{11} = \mathcal{S}_{11}$, we obtain the following bound:

$$\|\mathfrak{S}_0^{-1}\| \leq \|\mathcal{A}_{10}^{-1}\mathcal{S}_{11}\mathcal{A}_{10}^{-*}\| + \left\| (\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1} \mathcal{A}_{10}^* \mathcal{L}_{12} \right\| \|\mathcal{L}_{22}^{-1}\| \left\| \mathcal{L}_{21}\mathcal{A}_{10}(\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1} \right\|.$$

Since $(\mathcal{L}_{12})^* = (\mathcal{L}^*)_{21} = (\mathcal{R}\mathcal{L}\mathcal{R})_{21} = -\mathcal{R}_{22}\mathcal{L}_{21}$,

$$\left\| (\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1} \mathcal{A}_{10}^* \mathcal{L}_{12} \right\| = \left\| (\mathcal{L}_{12})^* \mathcal{A}_{10}(\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1} \right\| \leq \|\mathcal{R}_{22}\| \left\| \mathcal{L}_{21}\mathcal{A}_{10}(\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1} \right\|.$$

From Assumption 2 we have $\|\mathcal{L}_{22}^{-1}\| \leq s^{-1}$. Therefore,

$$\|\mathfrak{S}_0^{-1}\| \leq \frac{\|\mathcal{S}_{11}\|}{a^2} + \frac{\|\mathcal{R}_{22}\| \|\mathcal{L}_{21}\mathcal{A}_{10}(\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1}\|^2}{s}. \quad (12)$$

We now turn to the other terms in (5). Note first that $\mathcal{R}_{++}\mathcal{L}_{++}\mathcal{R}_{++} = \mathcal{L}_{++}^*$, so that $\mathcal{R}_{++}\mathcal{L}_{++}^{-1}\mathcal{R}_{++} = \mathcal{L}_{++}^{-*}$. In view of (8) and (11), the Schur complement \mathfrak{S}_0 is in fact symmetric and can be written as $\mathfrak{S}_0 = \mathcal{A}_{+0}^* [\mathcal{L}_{++}^{-1}]_s \mathcal{A}_{+0}$, where $T_s = (T + T^*)/2$ denotes the symmetric part of an operator T . Since $[T^{-1}]_s = T^{-*}T_sT^{-1}$, the operator $[-\mathcal{L}_{++}^{-1}]_s = -\mathcal{L}_{++}^{-*}\mathcal{S}_{++}\mathcal{L}_{++}^{-1}$ is a positive bounded self-adjoint operator, with a well-defined square root. The Schur complement is therefore symmetric negative and can be factored as

$$\mathfrak{S}_0 = -Q^*Q, \quad Q = [-\mathcal{L}_{++}^{-1}]_s^{1/2} \mathcal{A}_{+0}.$$

Consider now the second term in the lower right corner of (5):

$$\mathcal{L}_{++}^{-1} \mathcal{A}_{+0} \mathfrak{S}_0^{-1} \mathcal{A}_{+0}^* \mathcal{L}_{++}^{-1} = - \underbrace{\mathcal{L}_{++}^{-1} [-\mathcal{L}_{++}^{-1}]_s^{-1/2}}_{T_1} \underbrace{Q(Q^*Q)^{-1}Q^*}_{T_2} \underbrace{[-\mathcal{L}_{++}^{-1}]_s^{-1/2} \mathcal{L}_{++}^{-1}}_{T_3}.$$

The operator T_2 is an orthogonal projector, hence has a norm bounded by 1. We bound T_3 by noting that

$$T_3^*T_3 = \mathcal{L}_{++}^{-*} [-\mathcal{L}_{++}^{-1}]_s^{-1} \mathcal{L}_{++}^{-1} = -2 (\mathcal{L}_{++} [\mathcal{L}_{++}^{-1} + \mathcal{L}_{++}^*] \mathcal{L}_{++}^*)^{-1} = -2(\mathcal{L}_{++} + \mathcal{L}_{++}^*)^{-1} = -\mathcal{S}_{++}^{-1},$$

and therefore $\|T_3\| \leq 1/\sqrt{s}$ by Assumption 2. Similarly, $\|T_1\| \leq 1/\sqrt{s}$, so that $\|\mathcal{L}_{++}^{-1} \mathcal{A}_{+0} \mathfrak{S}_0^{-1} \mathcal{A}_{+0}^* \mathcal{L}_{++}^{-1}\| \leq 1/s$. To bound the operator in the upper right corner of (5), we remark that $\mathfrak{S}_0^{-1} \mathcal{A}_{+0}^* \mathcal{L}_{++}^{-1} = -(Q^*Q)^{-1} Q^* T_3$, so that (using $\|A\|^2 = \|AA^*\|$ with $A = (Q^*Q)^{-1} Q^*$)

$$\|\mathfrak{S}_0^{-1} \mathcal{A}_{+0}^* \mathcal{L}_{++}^{-1}\| \leq \frac{1}{\sqrt{s}} \|(Q^*Q)^{-1} Q^*\| = \frac{1}{\sqrt{s}} \sqrt{\|(Q^*Q)^{-1}\|} = \sqrt{\frac{\|\mathfrak{S}_0^{-1}\|}{s}}.$$

The same bound holds for the operator in the lower left corner of (5).

By gathering all the estimates, we finally obtain

$$\begin{aligned} \|\mathcal{L}^{-1}\| &\leq \|(\mathcal{L}^{-1})_{00}\| + \|(\mathcal{L}^{-1})_{0+}\| + \|(\mathcal{L}^{-1})_{+0}\| + \|(\mathcal{L}^{-1})_{++}\| \\ &\leq \|\mathfrak{S}_0^{-1}\| + 2\sqrt{\frac{\|\mathfrak{S}_0^{-1}\|}{s}} + \frac{2}{s} \leq 2\|\mathfrak{S}_0^{-1}\| + \frac{3}{s}, \end{aligned}$$

which, together with (12), gives the desired bound. \square

3 Applications and extensions

3.1 Langevin dynamics

For Langevin dynamics (also known as underdamped Langevin dynamics, or kinetic Langevin dynamics in some communities), the reference measure is the phase-space Boltzmann–Gibbs measure defined on $\mathcal{X} = \mathcal{D} \times \mathbb{R}^d$ with $\mathcal{D} = \mathbb{T}^d$ or \mathbb{R}^d , which reads

$$\mu(dq dp) = \nu(dq) \kappa(dp), \quad \nu(dq) = Z_\nu^{-1} e^{-\beta V(q)} dq, \quad \kappa(dp) = Z_\kappa^{-1} e^{-\beta U(p)} dp, \quad (13)$$

where $Z_\nu, Z_\kappa \in (0, +\infty)$ are normalization constants ensuring that ν, κ are probability measures. The dynamics reads, for some positive friction $\gamma > 0$,

$$\begin{cases} dq_t = \nabla U(p_t) dt, \\ dp_t = -\nabla V(q_t) dt - \gamma \nabla U(p_t) dt + \sqrt{\frac{2\gamma}{\beta}} dW_t. \end{cases}$$

Dynamics in an extended space can also be considered, as we shall see in Section 3.3. Note also that while for the classical Langevin dynamics the kinetic energy is simply $U(p) = |p|^2/2$, we consider a general kinetic energy U in order to emphasize the structure of the dynamics. In particular, ∇U can vanish on an open set, in which case the generator associated with the corresponding Langevin dynamics is not hypoelliptic [78]. The antisymmetric part of the generator is the generator of the Hamiltonian dynamics

$$\mathcal{A} = \mathcal{L}_{\text{ham}} = \nabla U(p)^T \nabla_q - \nabla V(q)^T \nabla_p, \quad (14)$$

while the symmetric part is (the subscript 'FD' stands for 'fluctuation/dissipation')

$$\mathcal{S} = \gamma \mathcal{L}_{\text{FD}}, \quad \mathcal{L}_{\text{FD}} = -\nabla U(p)^T \nabla_p + \frac{1}{\beta} \Delta_p = -\frac{1}{\beta} \nabla_p^* \nabla_p. \quad (15)$$

When the kinetic energy is quadratic, the choice (15) corresponds to considering an Ornstein–Uhlenbeck process on the momenta. The projector Π_0 is

$$(\Pi_0 \varphi)(q) = \int_{\mathbb{R}^d} \varphi(q, p) \kappa(dp). \quad (16)$$

Simple computations show that Assumptions 1 and 4 hold true when U is even, upon choosing the momentum reversal operator $\mathcal{R}f(q, p) = f(q, -p)$ in Assumption 4. Note that \mathcal{R} is unitary when U is even, in which case $\|\mathcal{R}_{22}\| = 1$.

3.1.1 General estimates

We assume that ν and κ satisfy Poincaré inequalities. As we will see below, we also need growth conditions on the potential in order to apply some results from [33]; as well as moment estimates for derivatives of U . To state them, we denote by $\partial_p^\alpha = \partial_{p_1}^{\alpha_1} \dots \partial_{p_d}^{\alpha_d}$ and $|\alpha| = \alpha_1 + \dots + \alpha_d$ for any $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$.

Assumption 6. *The function V is smooth and such that $e^{-\beta V} \in L^1(\mathcal{D})$. There exists $K_\nu > 0$ such that, for any $\phi \in H^1(\nu)$,*

$$\left\| \phi - \int_{\mathcal{D}} \phi d\nu \right\|_{L^2(\nu)} \leq \frac{1}{K_\nu} \|\nabla_q \phi\|_{L^2(\nu)}.$$

Moreover, there exist $c_1 > 0$, $c_2 \in [0, 1]$ and $c_3 > 0$ such that

$$\Delta V \leq c_1 d + \frac{c_2 \beta}{2} |\nabla V|^2, \quad |\nabla^2 V|^2 := \sum_{i,j=1}^d |\partial_{x_i} \partial_{x_j} V|^2 \leq c_3^2 (d + |\nabla V|^2), \quad (17)$$

where $|\cdot|$ denotes the standard Euclidean norm on \mathbb{R}^d .

Assumption 7. *The function U is smooth and even, and such that $e^{-\beta U} \in L^1(\mathbb{R}^d)$. There exists $K_\kappa > 0$ such that, for any $\varphi \in H^1(\kappa)$,*

$$\left\| \varphi - \int_{\mathbb{R}^d} \varphi d\kappa \right\|_{L^2(\kappa)} \leq \frac{1}{K_\kappa} \|\nabla_p \varphi\|_{L^2(\kappa)}.$$

Moreover, the kinetic energy U is such that $\partial_p^\alpha U$ belongs to $L^2(\kappa)$ for any $|\alpha| \leq 3$, and $(\partial_p^\alpha U)(\partial_p^{\alpha'} U)$ is in $L^2(\kappa)$ for any $|\alpha| \leq 2$ and $|\alpha'| = 1$.

We refer to [6] for simple conditions on U and V that ensure that Poincaré inequalities hold. The scaling with respect to the dimension of the constants in the bounds (17) is motivated by the case of separable potentials for which $V(q) = v(q_1) + \dots + v(q_d)$ for some smooth one dimensional function v , which corresponds to tensorized probability measures. The bounds (17) then follow from the inequalities

$$v'' \leq c_1 + \frac{c_2 \beta}{2} (v')^2, \quad |v''|^2 \leq c_3^2 (1 + |v'|^2).$$

These bounds generally hold if v has polynomial growth for example. The scaling of the constants should be similar for particles on a lattice (such as one dimensional atom chains) with finite interaction ranges, or systems for which correlations between degrees of freedom are bounded with respect to the dimension, in the sense that each column/line of the matrix $\nabla^2 V$ has a finite number of nonzero entries. Note finally that a careful inspection of the proof of Lemma 2 below shows that it would be possible to choose $c_2 \in [0, 2)$, but the final estimates would be more cumbersome to write, which is why we stick to the condition $c_2 \in [0, 1]$.

Theorem 2. *Suppose that Assumptions 6 and 7 hold true. Then, the resolvent of the generator of the Langevin dynamics satisfies the following bound:*

$$\begin{aligned} \|\mathcal{L}^{-1}\| &\leq \frac{2\beta\gamma}{\lambda_{\min}(\mathcal{M})K_\nu^2} \|\Pi_1 \mathcal{L}_{\text{FD}} \Pi_1\| \\ &+ \frac{4\beta}{\gamma K_\kappa^2} \left(\frac{3}{4} + \left\| \Pi_+ \mathcal{L}_{\text{ham}} \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \right\|^2 + \gamma^2 \left\| \Pi_2 \mathcal{L}_{\text{FD}} \Pi_1 \mathcal{L}_{\text{ham}} \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \right\|^2 \right), \end{aligned} \quad (18)$$

where $\lambda_{\min}(\mathcal{M}) > 0$ denotes the smallest eigenvalue of the symmetric positive definite matrix

$$\mathcal{M} = \int_{\mathbb{R}^d} \nabla^2 U d\kappa = \beta \int_{\mathbb{R}^d} \nabla U \otimes \nabla U d\kappa.$$

Note that $\|\mathcal{L}^{-1}\|$ indeed scales as $\min(\gamma, \gamma^{-1})$, as noted in various previous works [31, 43, 54] (upon possibly integrating in time the bounds on the evolution semigroup).

Proof. We apply Theorem 1, and check to this end that Assumptions 2, 3 and 5 hold true. The Poincaré inequality in the momentum variable implies that Assumption 2 holds with

$$s = \gamma \frac{K_\kappa^2}{\beta}.$$

To check Assumptions 3 and 5, we use $\partial_{p_i}^* = -\partial_{p_i} + \beta \partial_{p_i} U$ and $\partial_{q_i}^* = -\partial_{q_i} + \beta \partial_{q_i} V$ to rewrite the generators as

$$\mathcal{S} = \gamma \mathcal{L}_{\text{FD}} = -\frac{\gamma}{\beta} \sum_{i=1}^d \partial_{p_i}^* \partial_{p_i}, \quad \mathcal{A} = \mathcal{L}_{\text{ham}} = \frac{1}{\beta} \sum_{i=1}^d \partial_{p_i}^* \partial_{q_i} - \partial_{q_i}^* \partial_{p_i}.$$

We also use the following rules

$$[\partial_{p_i}, \partial_{q_j}] = 0, \quad \partial_{p_i} \Pi_0 = 0, \quad \Pi_0 \partial_{p_i}^* = 0, \quad [\partial_{p_i}, \partial_{p_j}^*] = \beta \partial_{p_i, p_j}^2 U,$$

to obtain

$$\mathcal{A}_{+0}^* \mathcal{A}_{+0} = \frac{1}{\beta^2} \sum_{i=1}^d \sum_{j=1}^d \Pi_0 \partial_{q_i}^* \partial_{p_i} \partial_{p_j}^* \partial_{q_j} \Pi_0 = \frac{1}{\beta} \Pi_0 \nabla_q^* \mathcal{M} \nabla_q \Pi_0 = \frac{1}{\beta} \sum_{i,j=1}^d \mathcal{M}_{ij} \Pi_0 \partial_{q_i}^* \partial_{q_j} \Pi_0. \quad (19)$$

To prove that \mathcal{M} is a symmetric positive definite matrix, we write, for $\xi \in \mathbb{R}^d$,

$$\xi^T \mathcal{M} \xi = \beta \int_{\mathbb{R}^d} |\xi^T \nabla U|^2 d\kappa.$$

If the latter quantity was 0 for some element $\xi \in \mathbb{R}^d \setminus \{0\}$, then ∇U would be 0 in the direction of ξ , and hence U would be constant in this direction, which would contradict the assumption $e^{-\beta U} \in L^1(\mathbb{R}^d)$. We then obtain, for a smooth function φ with compact support and average 0 with respect to μ ,

$$\|\mathcal{A}_{+0} \varphi\|_{L^2(\mu)}^2 = \frac{1}{\beta} \int_{\mathcal{D}} (\nabla_q \Pi_0 \varphi)^T \mathcal{M} \nabla_q \Pi_0 \varphi d\nu \geq \frac{\lambda_{\min}(\mathcal{M})}{\beta} \|\nabla_q \Pi_0 \varphi\|_{L^2(\nu)}^2 \geq \frac{\lambda_{\min}(\mathcal{M}) K_\nu^2}{\beta} \|\Pi_0 \varphi\|_{L^2(\mu)}^2,$$

where we used the Poincaré inequality for ν in the last step. This shows that Assumption 3 holds true with

$$a = K_\nu \sqrt{\frac{\lambda_{\min}(\mathcal{M})}{\beta}}.$$

We finally turn to Assumption 5. Note first that

$$-\mathcal{A}_{+0}^* \mathcal{L}_{\text{FD}} \mathcal{A}_{+0} = \frac{1}{\beta} \nabla_q^* \mathcal{N} \nabla_q \Pi_0, \quad \mathcal{N}_{ij} = \int_{\mathbb{R}^d} [(\nabla^2 U)^2]_{ij} d\kappa = \sum_{k=1}^d \int_{\mathbb{R}^d} (\partial_{p_i, p_k}^2 U) (\partial_{p_k, p_j}^2 U) d\kappa,$$

so that, in view of (7) and (19),

$$-\Pi_1 \mathcal{L}_{\text{FD}} \Pi_1 = \frac{1}{\beta} \sum_{i,j=1}^d T_{ij} \tilde{T}_{ij},$$

with

$$T_{ij} = \partial_{p_i}^* \Pi_0 \partial_{p_j}, \quad \tilde{T}_{ij} = \partial_{q_i} (\nabla_q^* \mathcal{M} \nabla_q)^{-1} \nabla_q^* \mathcal{N} \nabla_q (\nabla_q^* \mathcal{M} \nabla_q)^{-1} \partial_{q_j}^*,$$

where \tilde{T}_{ij} is seen as an operator on $L^2(\nu)$. Note that T_{ij} is bounded since it is the composition of operators of the form $\Pi_0 \partial_{p_k}$ and their adjoints, such operators being bounded because, by an integration by parts,

$$\Pi_0 \partial_{p_k} \varphi = \beta \int_{\mathbb{R}^d} \varphi \partial_{p_k} U d\kappa,$$

so that

$$\|\Pi_0 \partial_{p_k} \varphi\|_{L^2(\mu)} \leq \beta \|\partial_{p_k} U\|_{L^2(\mu)} \|\varphi\|_{L^2(\mu)}.$$

The operators \tilde{T}_{ij} are also easily seen to be bounded, as a linear combination of composition of operators of the form $(\nabla_q^* \mathcal{M} \nabla_q)^{-1/2} \partial_{q_j}^*$ (which are bounded from $L^2(\nu)$ to $L_0^2(\nu)$, the subspace of functions of functions of $L^2(\nu)$ with average 0 with respect to ν) and their adjoints. This allows to conclude that $\mathcal{S}_{11} = \gamma \Pi_1 \mathcal{L}_{\text{FD}} \Pi_1$ is bounded.

Let us next consider

$$\mathcal{L}_{21} \mathcal{A}_{10} (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} = \Pi_2 \mathcal{S} \mathcal{A} \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} + \Pi_2 \mathcal{A}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1}. \quad (20)$$

We start by the second operator on the right hand side of the latter equality. Since $\Pi_2 = \Pi_2 \Pi_+$, it suffices in fact to prove that $\Pi_+ \mathcal{A}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1}$ is bounded. First, for a smooth function φ with compact support, we compute $\mathcal{A}^2 \Pi_0 \varphi = \nabla U^T (\nabla_q^2 \Pi_0 \varphi) \nabla U - \nabla V^T (\nabla^2 U) \nabla_q \Pi_0 \varphi$, and therefore

$$\Pi_+ \mathcal{A}^2 \Pi_0 \varphi = \sum_{i,j=1}^d \mathcal{U}_{ij} \partial_{q_i, q_j}^2 \Pi_0 \varphi - \mathcal{V}_{ij} (\partial_{q_i} V) \partial_{q_j} \Pi_0 \varphi, \quad (21)$$

with

$$\mathcal{U}_{ij} = \left[(\partial_{p_i} U) (\partial_{p_j} U) - \frac{1}{\beta} \mathcal{M}_{ij} \right], \quad \mathcal{V}_{ij} = \partial_{p_i, p_j}^2 U - \mathcal{M}_{ij}. \quad (22)$$

The boundedness of the operator $\Pi_+ \mathcal{A}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1}$ then follows from the fact that the operators $\Pi_+ \partial_{q_i, q_j}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1}$ and $\Pi_+ (\partial_{q_i} V) \partial_{q_j} \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1}$ are bounded (in view of (17), see respectively [33, Proposition 5] and [78, Lemma A.4]; see also the proof of Proposition 1 below).

Let us finally bound the first operator on the right hand side of (20), using the following equality as operators from \mathcal{H}_0 to \mathcal{H}_2 :

$$\Pi_2 \mathcal{S} \mathcal{A} \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} = -\frac{\gamma}{\beta^2} \sum_{i=1}^d \Pi_2 \mathcal{T}_i \partial_{q_i} (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1}, \quad \mathcal{T}_i = \sum_{j=1}^d \partial_{p_j}^* \partial_{p_j} \partial_{p_i}^* \Pi_0.$$

The result then follows from the fact that $\partial_{q_i} (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1}$ is bounded from \mathcal{H}_0 to \mathcal{H} , and

$$\mathcal{T}_i = \beta \sum_{j=1}^d \left[\beta (\partial_{p_j} U) \left(\partial_{p_i, p_j}^2 U \right) - \partial_{p_i, p_j, p_j}^3 U \right] \Pi_0$$

is a bounded operator on $L^2(\mu)$, because of the $L^2(\kappa)$ -bounds on the derivatives of U (see Assumption 7). The estimate (18) finally follows by Theorem 1. \square

3.1.2 Scaling with the dimension

The aim of this section is to make precise the dependence on the dimension of the various terms in (18). We consider to this end the simple situation where the kinetic energy is quadratic and where all degrees of freedom are associated with the same mass $m > 0$, *i.e.*

$$U(p) = \sum_{i=1}^d u(p_i), \quad u(p) = \frac{p^2}{2m}. \quad (23)$$

In this case, $\kappa(dp)$ is a tensor product of Gaussian measures with variances m/β , so that the Poincaré constant is $K_\kappa^2 = \beta/m$.

Corollary 1. *Suppose that Assumption 6 holds true. Then, the resolvent of the generator of the Langevin dynamics with quadratic kinetic energy (23) satisfies the following bound:*

$$\|\mathcal{L}^{-1}\| \leq \frac{2\beta\gamma}{K_\nu^2} + \frac{4m}{\gamma} \left(\frac{3}{4} + \left\| \Pi_+ \mathcal{L}_{\text{ham}}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \right\|^2 \right). \quad (24)$$

The operator norm on the right hand side can be bounded using Proposition 1 below, or Proposition 2 in Section 3.1.4 under some extra assumptions on the potential V . Let us emphasize that a particularly nice feature of the estimate (24) is that there is no uncontrolled prefactor in the estimate, as usually obtained by writing the resolvent as the time integral of the semigroup.

Proof. First, $\mathcal{M} = m^{-1}\text{Id}$ so that $\lambda_{\min}(\mathcal{M}) = m^{-1}$. Consider next $\Pi_1 \mathcal{L}_{\text{FD}} \Pi_1$. Since $\mathcal{L}_{\text{FD}} \mathcal{L}_{\text{ham}} \Pi_0 = -\mathcal{L}_{\text{ham}} \Pi_0 / m$ and therefore $\mathcal{A}_{+0}^* \mathcal{L}_{\text{FD}} \mathcal{A}_{+0} = -\mathcal{A}_{+0}^* \mathcal{A}_{+0} / m$, it holds $\mathcal{L}_{\text{FD}} \Pi_1 = \Pi_1 \mathcal{L}_{\text{FD}} \Pi_1 = -\Pi_1 / m$. Since Π_1 is an orthogonal projector, we finally obtain $\|\Pi_1 \mathcal{L}_{\text{FD}} \Pi_1\| \leq m^{-1}$. Moreover, $\Pi_2 \mathcal{L}_{\text{FD}} \Pi_1 = 0$ since $\mathcal{L}_{\text{FD}} \Pi_1 = -\Pi_1 / m$ by the above computations. The estimates finally follow from (18). \square

In addition to Assumption 6, and similarly to what is done in [23], we consider two possible conditions on the potential energy: (i) V is convex; (ii) the Hessian of V is bounded below as $\nabla_q^2 V \geq -K\text{Id}$ for some $K \in \mathbb{R}_+$. Notice that the second condition implies the first one when $K = 0$. The following result, whose proof is postponed to Section 3.1.3, is key in understanding the dependence on the dimension of the bound provided by (24).

Proposition 1. *Suppose that Assumption 6 holds. Then,*

$$\left\| \Pi_+ \mathcal{L}_{\text{ham}}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \right\|^2 \leq 2 \left(C + \frac{C'}{K_\nu^2} \right),$$

where C and C' can be chosen as:

- (i) If V is convex, then $C = 1$ and $C' = 0$;
- (ii) If $\nabla_q^2 V \geq -K\text{Id}$ for some $K \geq 0$, then $C = 1$ and $C' = K$;
- (iii) In the general case, $C = 2$ and $C' = 2c_3 \left[\sqrt{d} + 2 \max \left(\frac{8c_3}{\beta^2}, \sqrt{\frac{c_1 d}{\beta}} \right) \right]$.

The scaling with respect to the dimension of the bound for item (iii) above is better than the one in [23], because we directly integrate into our assumption (17) how the prefactors in the estimates should depend on the dimension, and then carefully track this dependency in the proofs of Lemmas 1 and 2 below (the latter lemma being directly obtained from [85, Lemma A.24]).

To finally discuss more precisely the dependence on the dimension of the upper bound (24), we would need to understand how the Poincaré constant K_ν and constants c_1, c_2, c_3 in Assumption 6 depend on the dimension:

- (i) As discussed after Assumption 6, c_1, c_2, c_3 can be chosen independently of the dimension for simple systems (in particular for separable potentials, in which case the Langevin dynamics can be seen as the juxtaposition of d one dimensional Langevin dynamics).
- (ii) When $\nabla^2 V \geq R\text{Id}$ with $R > 0$, the Poincaré constant K_ν^2 is bounded from below by the positive constant βR by the celebrated Bakry–Emery criterion [8], independently of the dimension. For separable potentials, the Poincaré constant for the full measure is dimension-free as long as the marginal distribution of any single position itself satisfies a Poincaré inequality (with some uniform lower bound on the Poincaré constant), by the tensorization property. Sufficient conditions for dimension-free constants for functional inequalities when the potentials have weak enough interactions have been well-studied, see for example [87, 71].

These considerations allow to further make precise the estimates on the resolvent.

3.1.3 Proof of Proposition 1

We start by noting that, for the quadratic kinetic energy (23), the expressions (21)-(22) simplify as

$$\forall 1 \leq i, j \leq d, \quad \mathcal{U}_{ij} = \frac{p_i p_j}{m^2} - \frac{1}{m\beta} \delta_{i,j}, \quad \mathcal{V}_{ij} = 0,$$

and

$$\Pi_+ \mathcal{L}_{\text{ham}}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \varphi = \sum_{i,j=1}^d \mathcal{U}_{ij} \partial_{q_i, q_j}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \varphi.$$

Besides, a simple computation gives $\|\mathcal{U}_{ij}\|_{L^2(\kappa)}^2 = 1/(m^2 \beta^2)$ for $i \neq j$, and $\|\mathcal{U}_{ii}\|_{L^2(\kappa)}^2 = 2/(m^2 \beta^2)$, as well as $\langle \mathcal{U}_{ij}, \mathcal{U}_{kl} \rangle_{L^2(\kappa)} = 0$ when $\{i, j\} \neq \{k, l\}$. Therefore, for $\varphi \in L^2(\mu)$,

$$\begin{aligned} \left\| \Pi_+ \mathcal{L}_{\text{ham}}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \varphi \right\|^2 &= \sum_{i,j=1}^d \|\mathcal{U}_{ij}\|_{L^2(\kappa)}^2 \left\| \partial_{q_i, q_j}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \varphi \right\|_{L^2(\nu)}^2 \\ &\leq \frac{2}{m^2 \beta^2} \sum_{i,j=1}^d \left\| \partial_{q_i, q_j}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \varphi \right\|_{L^2(\nu)}^2 = \frac{2}{m^2 \beta^2} \left\| \nabla_q^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \varphi \right\|_{L^2(\nu)}^2, \end{aligned}$$

so that, using $\mathcal{A}_{+0}^* \mathcal{A}_{+0} = (m\beta)^{-1} \nabla_q^* \nabla_q \Pi_0$,

$$\left\| \Pi_+ \mathcal{L}_{\text{ham}}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \varphi \right\|^2 \leq 2 \left\| \nabla_q^2 (\nabla_q^* \nabla_q)^{-1} \Pi_0 \varphi \right\|_{L^2(\nu)}^2.$$

The proof is then directly obtained from the following technical result (where we denote by $C_c^\infty(\mathcal{D})$ the set of C^∞ functions with compact support).

Lemma 1. *It holds, with the same constants $C, C' \in \mathbb{R}_+$ as in Proposition 1,*

$$\forall u \in C_c^\infty(\mathcal{D}), \quad \left\| \nabla^2 u \right\|_{L^2(\nu)}^2 = \sum_{i,j=1}^d \left\| \partial_{q_i, q_j}^2 u \right\|_{L^2(\nu)}^2 \leq C \left\| \nabla_q^* \nabla_q u \right\|_{L^2(\nu)}^2 + C' \left\| \nabla_q u \right\|_{L^2(\nu)}^2.$$

To prove this result, we need the following lemma.

Lemma 2. *Suppose that Assumption 6 holds. Then,*

$$\forall \phi \in H^1(\nu), \quad \left\| \phi \nabla V \right\|_{L^2(\nu)}^2 \leq \frac{16}{\beta^2} \left\| \nabla \phi \right\|_{L^2(\nu)}^2 + \frac{4dc_1}{\beta} \left\| \phi \right\|_{L^2(\nu)}^2. \quad (25)$$

Proof of Lemma 2. We closely follow the proof of [85, Lemma A.24], precisely keeping track of all constants (which leads to estimates with a better scaling than in [23, Lemma 2.2]). First, for a given function $\varphi \in C_c^\infty(\mathcal{D})$,

$$\left\| \phi \nabla V \right\|_{L^2(\nu)}^2 = \int_{\mathcal{D}} \phi^2 |\nabla V|^2 d\nu = \frac{1}{\beta} \int_{\mathcal{D}} \operatorname{div} (\phi^2 \nabla V) d\nu = \frac{2}{\beta} \int_{\mathcal{D}} \phi \nabla \phi \cdot \nabla V d\nu + \frac{1}{\beta} \int_{\mathcal{D}} \phi^2 \Delta V d\nu,$$

so that, in view of Assumption 6,

$$\left\| \phi \nabla V \right\|_{L^2(\nu)}^2 \leq \frac{2}{\beta} \left\| \nabla \phi \right\|_{L^2(\nu)} \left\| \phi \nabla V \right\|_{L^2(\nu)} + \frac{c_1 d}{\beta} \left\| \phi \right\|_{L^2(\nu)}^2 + \frac{c_2}{2} \left\| \phi \nabla V \right\|_{L^2(\nu)}^2.$$

By a Young inequality with parameter $\eta > 0$, we then obtain

$$\left(1 - \frac{c_2}{2} - \frac{\eta}{\beta} \right) \left\| \phi \nabla V \right\|_{L^2(\nu)}^2 \leq \frac{1}{\beta \eta} \left\| \nabla \phi \right\|_{L^2(\nu)}^2 + \frac{c_1 d}{\beta} \left\| \phi \right\|_{L^2(\nu)}^2.$$

Choosing $\eta = \beta/4$ and recalling that $c_2 \in [0, 1]$ provides the claimed estimate by the density of $C_c^\infty(\mathcal{D})$ in $H^1(\nu)$. \square

We can finally turn to the proof of Lemma 1.

Proof of Lemma 1. We fix $u \in C_c^\infty(\mathcal{D})$ and follow the proof of [23, Lemma 2.3], which relies on Bochner's formula

$$\sum_{i,j=1}^d \left| \partial_{q_i, q_j}^2 u \right|^2 = \nabla_q u \cdot \nabla_q (\nabla_q^* \nabla_q u) - (\nabla_q u)^T \nabla_q^2 V \nabla_q u - \nabla_q^* \nabla_q \left(\frac{|\nabla_q u|^2}{2} \right),$$

to write

$$\sum_{i,j=1}^d \left\| \partial_{q_i, q_j}^2 u \right\|_{L^2(\nu)}^2 = \left\| \nabla_q^* \nabla_q u \right\|_{L^2(\nu)}^2 - \int_{\mathcal{D}} (\nabla_q u)^T \nabla_q^2 V \nabla_q u \, d\nu. \quad (26)$$

This already gives the desired result when either V is convex (*i.e.* $\nabla^2 V \geq 0$) or $\nabla^2 V \geq -K \text{Id}$.

Let us next consider the general case given by Assumption 6. Note that $|\nabla^2 V| \leq c_3(\sqrt{d} + |\nabla V|)$. We first obtain from (26) that

$$\left\| \nabla_q^2 u \right\|_{L^2(\nu)}^2 \leq \left\| \nabla_q^* \nabla_q u \right\|_{L^2(\nu)}^2 + c_3 \sqrt{d} \left\| \nabla_q u \right\|_{L^2(\nu)}^2 + c_3 \left\| \nabla_q u \right\|_{L^2(\nu)} \left\| |\nabla_q u| |\nabla_q V| \right\|_{L^2(\nu)}. \quad (27)$$

Note that, from (25),

$$\begin{aligned} \left\| |\nabla_q u| |\nabla_q V| \right\|_{L^2(\nu)}^2 &= \sum_{i=1}^d \left\| \partial_{q_i} u \nabla_q V \right\|_{L^2(\nu)}^2 \leq \sum_{i=1}^d \frac{16}{\beta^2} \left\| \nabla(\partial_{q_i} u) \right\|_{L^2(\nu)}^2 + \frac{4dc_1}{\beta} \left\| \partial_{q_i} u \right\|_{L^2(\nu)}^2 \\ &\leq \frac{16}{\beta^2} \left\| \nabla_q^2 u \right\|_{L^2(\nu)}^2 + \frac{4dc_1}{\beta} \left\| \nabla_q u \right\|_{L^2(\nu)}^2. \end{aligned}$$

Using a Young inequality to bound the last term on the right hand side of (27) with the inequality above then leads to

$$\left\| \nabla_q^2 u \right\|_{L^2(\nu)}^2 \leq \left\| \nabla_q^* \nabla_q u \right\|_{L^2(\nu)}^2 + c_3 \left(\sqrt{d} + \frac{1}{2\eta} \right) \left\| \nabla_q u \right\|_{L^2(\nu)}^2 + \frac{c_3 \eta}{2} \left(\frac{16}{\beta^2} \left\| \nabla_q^2 u \right\|_{L^2(\nu)}^2 + \frac{4dc_1}{\beta} \left\| \nabla_q u \right\|_{L^2(\nu)}^2 \right),$$

and finally

$$\left(1 - \frac{8c_3 \eta}{\beta^2} \right) \left\| \nabla_q^2 u \right\|_{L^2(\nu)}^2 \leq \left\| \nabla_q^* \nabla_q u \right\|_{L^2(\nu)}^2 + c_3 \left(\sqrt{d} + \frac{1}{2\eta} + \frac{2\eta dc_1}{\beta} \right) \left\| \nabla_q u \right\|_{L^2(\nu)}^2.$$

We next choose $\eta = \beta^2 \varepsilon / (8c_3)$ for some $\varepsilon \in (0, 1)$ to be determined, so that

$$\left\| \nabla_q^2 u \right\|_{L^2(\nu)}^2 \leq \frac{1}{1 - \varepsilon} \left[\left\| \nabla_q^* \nabla_q u \right\|_{L^2(\nu)}^2 + L_{d,\varepsilon} \left\| \nabla_q u \right\|_{L^2(\nu)}^2 \right],$$

with

$$L_{d,\varepsilon} = c_3 \left(\sqrt{d} + \frac{\mathbf{a}}{\varepsilon} + \mathbf{b} \varepsilon \right), \quad \mathbf{a} = \frac{4c_3}{\beta^2}, \quad \mathbf{b} = \frac{\beta c_1 d}{4c_3}.$$

When d is large, one should choose ε small enough in order for $\mathbf{b} \varepsilon$ to have a limited growth. Such a choice implies that the denominator $1 - \varepsilon$ in the expression of $L_{d,\varepsilon}$ will be close to 1. In order to balance the diverging term \mathbf{a}/ε as $\varepsilon \rightarrow 0$, one should set $\varepsilon = \sqrt{\mathbf{a}/\mathbf{b}}$. We however want to keep $\varepsilon \leq 1/2$ in all cases, which suggests choosing

$$\varepsilon = \min \left(\frac{1}{2}, \sqrt{\frac{\mathbf{a}}{\mathbf{b}}} \right),$$

in which case (using that $\mathbf{b} \leq 4\mathbf{a}$ when $\varepsilon = 1/2$)

$$L_{d,\varepsilon} = c_3 \left[\sqrt{d} + 2 \max \left(2\mathbf{a}, \sqrt{\mathbf{a}\mathbf{b}} \right) \right] = c_3 \left[\sqrt{d} + 2 \max \left(\frac{8c_3}{\beta^2}, \sqrt{\frac{c_1 d}{\beta}} \right) \right].$$

The conclusion then follows from the fact that $1 - \varepsilon \geq 1/2$. \square

3.1.4 Improving the dependence on the dimension under a log-Sobolev inequality

It is possible to improve the explicit dependence on the dimension from \sqrt{d} to $\log d$ in item (iii) of Proposition 1, but under some extra assumptions on the potential, and if we strengthen the Poincaré inequality assumption into a logarithmic Sobolev inequality.

Definition 1. A probability measure ν on \mathcal{D} is said to satisfy a logarithmic Sobolev inequality with constant C_{LSI} if

$$\forall f \in C_c^\infty(\mathcal{D}), \quad \text{Ent}_\nu(f^2) := \int_{\mathcal{D}} f^2 \log f^2 d\nu - \|f\|_{L^2(\nu)}^2 \log \|f\|_{L^2(\nu)}^2 \leq C_{\text{LSI}} \|\nabla f\|_{L^2(\nu)}^2.$$

A Poincaré inequality can be derived by linearizing this inequality, so it is strictly stronger than the Poincaré inequality we used above. However, it is still valid for uniformly log-concave distributions, and tensorizes with dimension-free constants. We refer to [56] for background information, and more general sufficient conditions for LSI to hold. To state the result giving bounds on the resolvent with respect to the dimension, we define, for a matrix $M \in \mathbb{R}^{d \times d}$, the operator norm

$$\|M\|_{\mathcal{B}(\ell^2)} = \sup_{|\xi| \leq 1} |M\xi|.$$

We also denote by $|\xi|_\infty = \max(|\xi_1|, \dots, |\xi_d|)$ the ℓ^∞ norm on \mathbb{R}^d .

Proposition 2. Consider the quadratic kinetic energy (23). Assume that $e^{-\beta V} \in L^1(\mathcal{D})$, that ν satisfies a logarithmic Sobolev inequality, and that

$$\forall x \in \mathbb{R}^d, \quad \|\nabla^2 V(x)\|_{\mathcal{B}(\ell^2)} \leq c_3 (1 + |\nabla V(x)|_\infty). \quad (28)$$

Then the conclusion of Proposition 1 holds true with

$$C = 2, \quad C' = 2 \left(c_3 + \frac{1}{2c_3 C_{\text{LSI}}} \left[\log d + \log \left(\max_{i=1, \dots, d} \int e^{2c_3 C_{\text{LSI}} |\partial_i V|} d\nu \right) \right] \right).$$

Note that for this proposition to be useful, we implicitly assume the exponential integrability of the partial derivatives of V with respect to ν . If we look at nice enough separable potentials, the dependence of C' on the dimension is of order $\log d$, since the other quantities involved in the bound are dimension-free. Note that a logarithmic scaling in d is typically correct for maximas of d independent quantities [30, Lemma 2.2] (which here would correspond to the case of a separable potential). Note also that the assumption (28), while natural for separable potentials, is not invariant under a change of orthonormal basis.

Proof. The proof follows the same general approach as that of Proposition 1, but with the extra assumption on V , (27) can be replaced by

$$\|\nabla_q^2 u\|_{L^2(\nu)}^2 \leq \|\nabla_q^* \nabla_q u\|_{L^2(\nu)}^2 + c_3 \|\nabla_q u\|_{L^2(\nu)}^2 + c_3 \int |\nabla_q u|^2 |\nabla V|_\infty d\nu. \quad (29)$$

We cannot use Lemma 2 to control the last term in the above equation, but use instead the following classical entropy inequality: for g bounded measurable,

$$\int_{\mathcal{D}} f^2 g d\nu \leq \text{Ent}_\nu(f^2) + \|f\|_{L^2(\nu)}^2 \log \left(\int_{\mathcal{D}} e^g d\nu \right).$$

This inequality is an immediate consequence of the fact that the entropy is the Legendre transform of the log-Laplace functional, with respect to the duality between functions and measures, *i.e.*

$$\text{Ent}_\nu(f^2) = \sup \left\{ \int_{\mathcal{D}} f^2 g d\nu - \|f\|_{L^2(\nu)}^2 \log \left(\int_{\mathcal{D}} e^g d\nu \right), \quad g \text{ bounded measurable} \right\}.$$

As a consequence, we obtain a bound of the form

$$\int_{\mathcal{D}} |\nabla_q u|^2 |\nabla V|_{\infty} d\nu \leq \frac{1}{c} \text{Ent}_{\nu}(|\nabla_q u|^2) + \frac{1}{c} \|\nabla_q u\|_{L^2(\nu)}^2 \log \left(\int_{\mathcal{D}} e^{c|\nabla V|_{\infty}} d\nu \right),$$

with $c > 0$ a constant to be adjusted later. When ν satisfies a logarithmic Sobolev inequality, the first factor on the right-hand side can be controlled by $C_{\text{LSI}} \|\nabla_q^2 u\|_{L^2(\nu)}^2$, using the inequality $|\nabla|\nabla f||^2 \leq |\nabla^2 f|^2$. The second factor can be bounded as

$$\log \left(\int_{\mathcal{D}} e^{c|\nabla V|_{\infty}} d\nu \right) \leq \log \left(\sum_{i=1}^d \int_{\mathcal{D}} e^{c|\partial_{q_i} V|_{\infty}} d\nu \right) \leq \log d + \log \left(\max_{i=1, \dots, d} \int_{\mathcal{D}} e^{c|\partial_{q_i} V|} d\nu \right).$$

We therefore obtain the inequality

$$\begin{aligned} \|\nabla^2 u\|_{L^2(\nu)}^2 &\leq \|\nabla_q^* \nabla_q u\|_{L^2(\nu)}^2 + c_3 \|\nabla_q u\|_{L^2(\nu)}^2 \\ &\quad + \frac{c_3 C_{\text{LSI}}}{c} \|\nabla^2 u\|_{L^2(\nu)}^2 + \frac{1}{c} \|\nabla_q u\|_{L^2(\nu)}^2 \left[\log d + \log \left(\max_{i=1, \dots, d} \int_{\mathcal{D}} e^{c|\partial_{q_i} V|} d\nu \right) \right]. \end{aligned} \quad (30)$$

Taking $c = \varepsilon^{-1} c_3 C_{\text{LSI}}$ leads to

$$\|\nabla^2 u\|_{L^2(\nu)}^2 \leq \frac{1}{1 - \varepsilon} \|\nabla_q^* \nabla_q u\|_{L^2(\nu)}^2 + L'_{d, \varepsilon} \|\nabla_q u\|_{L^2(\nu)}^2,$$

with

$$L'_{d, \varepsilon} := \frac{1}{1 - \varepsilon} \left(c_3 + \frac{\varepsilon}{c_3 C_{\text{LSI}}} \left[\log d + \log \left(\max_{i=1, \dots, d} \int_{\mathcal{D}} e^{\varepsilon^{-1} c_3 C_{\text{LSI}} |\partial_{q_i} V|} d\nu \right) \right] \right).$$

We next set $\varepsilon = 1/2$ and follow the same reasoning as in the end of the proof of Proposition 1 to conclude. \square

3.2 Linear Boltzmann equation

We turn in this section to another paradigmatic hypocoercive dynamics, known as the linear Boltzmann equation in the community of researchers working on kinetic equations, while it corresponds to the randomized Hybrid Monte Carlo method in computational statistics and molecular simulation.

We consider the linear Boltzmann equation (1) with a more general kinetic energy $U(v)$ than $|v|^2/2$, for the choice

$$k(y, v, w) = \frac{\gamma}{Z_{\kappa}} e^{-\beta U(v)}.$$

This corresponds to the following dynamics

$$\frac{\partial f}{\partial t}(t, y, v) - \nabla V(y) \cdot \nabla_v f(t, y, v) + \nabla U(v) \cdot \nabla_y f(t, y, v) = \gamma \left[\frac{e^{-\beta U(v)}}{Z_{\kappa}} \int_{\mathbb{R}^d} f(t, y, w) dw - f(t, y, v) \right].$$

We next change the unknown function from $f(t, y, v)$ to $\varphi(t, y, v) = f(t, y, v)/f_{\infty}(y, v)$, with f_{∞} the steady-state of the above equation. In fact, f_{∞} is the density of (13), namely $f_{\infty}(y, v) = Z^{-1} e^{-\beta(V(y) + U(v))}$. We then obtain the following dynamics on φ :

$$\partial_t \varphi + \mathcal{L}_{\text{ham}} \varphi = \gamma(\Pi_0 - 1)\varphi,$$

with Π_0 defined in (16). This shows that the operator to consider for the longtime convergence of the linear Boltzmann equation is $\mathcal{L} = \mathcal{A} + \mathcal{S}$ with $\mathcal{A} = -\mathcal{L}_{\text{ham}}$ and

$$\mathcal{S} = \gamma(\Pi_0 - 1). \quad (31)$$

Remark 2. We can alternatively interpret the generator as the Fokker–Planck operator associated with a piecewise deterministic Markov process where Hamiltonian trajectories are interrupted at random exponential times by a resampling of the momenta according to κ . This corresponds to the so-called Randomized Hybrid Monte Carlo method [17].

The estimates on the resolvent of the generator obtained by the approach described in Section 2 are the following.

Proposition 3. Suppose that Assumption 6 holds true and that U is smooth and even, with $e^{-\beta U} \in L^1(\mathbb{R}^d)$. Then, the resolvent of the generator $\mathcal{L} = \mathcal{A} + \mathcal{S}$ with \mathcal{A} given by the opposite of the generator in (14) and \mathcal{S} given by (31) satisfies the following bound:

$$\|\mathcal{L}^{-1}\| \leq \frac{2\beta\gamma}{\lambda_{\min}(\mathcal{M})K_\nu^2} + \frac{2}{\gamma} \left(\frac{3}{2} + \left\| \Pi_+ \mathcal{L}_{\text{ham}}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \right\|^2 \right).$$

Note that, here as well, the upper bound on the resolvent of the generator scales as $\min(\gamma, \gamma^{-1})$. The scaling with respect to the dimension d can be made precise as in Sections 3.1.2 and 3.1.4.

Proof. All the computations of Section 3.1.1 are valid upon taking $s = \gamma$. The only changes that need to be made are in the verification of Assumption 5. Note first that $\Pi_1 \mathcal{S} \Pi_1 = \gamma \Pi_1 \Pi_+ \Pi_1 = \gamma \Pi_1$, which immediately implies that $\|\Pi_1 \mathcal{S} \Pi_1\| = \gamma$. Next, $\mathcal{S}_{21} = 0$, so that $\mathcal{S}_{21} \mathcal{A}_{10} (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} = 0$. The final estimate then directly follows from Theorem 1. \square

3.3 Adaptive Langevin dynamics

We show here how to apply the framework of Section 2 to Adaptive Langevin dynamics, which is a Langevin dynamics in which the friction is a dynamical variable $\xi \in \mathbb{R}$ following some Nosé–Hoover feedback dynamics. Therefore, $x = (q, p, \xi)$ and $\mathcal{X} = \mathcal{D} \times \mathbb{R}^d \times \mathbb{R}$. We consider for simplicity the case when

$$U(p) = \frac{|p|^2}{2} = \frac{1}{2} \sum_{i=1}^d p_i^2. \quad (32)$$

After some suitable normalization (see [59]), the invariant measure of the dynamics reads

$$\mu(dq dp d\xi) = Z^{-1} \exp \left(-\beta \left[\frac{|p|^2}{2} + V(q) + \frac{\xi^2}{2} \right] \right) dq dp d\xi,$$

and the generator is, for some $\varepsilon > 0$,

$$\mathcal{L} = \mathcal{L}_{\text{ham}} + \gamma \mathcal{L}_{\text{FD}} + \frac{1}{\varepsilon} \mathcal{L}_{\text{NH}},$$

with \mathcal{L}_{ham} and \mathcal{L}_{FD} defined respectively in (14) and (15), and

$$\mathcal{L}_{\text{NH}} = \left(|p|^2 - \frac{d}{\beta} \right) \partial_\xi - \xi p^T \nabla_p = \frac{1}{\beta^2} \left((\partial_\xi - \partial_\xi^*) \nabla_p^* \nabla_p + \Delta_p^* \partial_\xi - \Delta_p \partial_\xi^* \right).$$

Therefore, $\mathcal{S} = \gamma \mathcal{L}_{\text{FD}}$ as for standard Langevin dynamics, while $\mathcal{A} = \mathcal{L}_{\text{ham}} + \varepsilon^{-1} \mathcal{L}_{\text{NH}}$. The framework introduced in Section 2 allows to retrieve the scaling on the resolvent of the generator found in [59] using the techniques from [48, 33].

Proposition 4. Suppose that Assumption 6 holds true. Then there exists $C \in \mathbb{R}_+$ such that

$$\forall \gamma, \varepsilon > 0, \quad \|\mathcal{L}^{-1}\| \leq C \max \left(\gamma \varepsilon^2, \gamma, \frac{1}{\gamma}, \frac{1}{\gamma \varepsilon^2} \right).$$

Proof. Assumptions 1, 2 and 4 hold true as for Langevin dynamics. A simple computation next shows that

$$\mathcal{A}_{+0}^* \mathcal{A}_{+0} = \left(\frac{2d}{\beta^2 \varepsilon^2} \partial_\xi^* \partial_\xi + \frac{1}{\beta} \nabla_q^* \nabla_q \right) \Pi_0,$$

which, by tensorization of Poincaré inequalities, implies that Assumption 3 holds with

$$a^2 = \frac{1}{\beta} \min \left(\frac{2d}{\varepsilon^2}, K_\nu^2 \right).$$

Let us next turn to Assumption 5. Computations similar to the ones performed for Langevin dynamics (see the proof of Corollary 1) show that $\Pi_1 \mathcal{L}_{\text{FD}} \Pi_1$ is bounded and $\mathcal{S}_{21} \mathcal{A}_{12} \Pi_0 (B^* B)^{-1} = 0$ since $\mathcal{S}_{21} = \gamma \Pi_2 \mathcal{L}_{\text{FD}} \Pi_1 = 0$ for the choice (32). From the computations in the proof of [59, Lemma 2.7], there exists $R \in \mathbb{R}_+$ such that

$$\left\| \Pi_+ \mathcal{L}_{\text{ham}}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \right\| \leq R \max \left(1, \frac{1}{\varepsilon} \right).$$

An application of Theorem 1 then allows to conclude. \square

A Generalized Poincaré-type inequalities

We show here how to obtain Poincaré-type inequalities, as recently derived in [4] and revisited in [23], relying on the algebraic framework presented in Section 2. We assume in all this section that $\mathcal{A}\mathbf{1} = 0$ and that \mathcal{S} is a negative operator. Let us emphasize that some arguments are formal. The main interest of our presentation in our opinion is to make explicit the algebraic framework behind the estimates obtained in [4, 23] for Langevin dynamics, in order to extend the approach to other hypocoercive dynamics (as done in [62] for various models of piecewise deterministic Markov processes, as the one considered in Section 3.2). In particular, the framework used here allows to simplify some algebraic manipulations, as for instance in the proof of [23, Theorem 2].

We first show how to obtain Poincaré-type inequalities using the antisymmetric part of the generator, in a static setting where only spatial degrees of freedom are considered. We then give a somewhat abstract account of Poincaré-type inequalities in a space-time setting and recall how they are used to prove the exponential decay of the evolution semigroup. We do not provide new results in the space-time setting, but merely present the results of [4, 23] in the language of this work.

Poincaré-type inequality using the antisymmetric part of the generator. We start by proving the following Poincaré-type estimate. Recall that $\Pi_+ = 1 - \Pi_0$. We consider, in the result below, a slight modification of Assumption 5.

Proposition 5. *Suppose that Assumptions 1 and 3 hold true, and that the operators $(1 - \mathcal{S})^{1/2} \Pi_1$ and $\Pi_+ \mathcal{A}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1}$ are bounded. Then,*

$$\forall f \in C_c^\infty(\mathcal{X}), \quad \|f - \langle f, \mathbf{1} \rangle\| \leq C_1 \|(1 - \Pi_0)f\| + C_2 \|(1 - \mathcal{S})^{-1/2} \mathcal{A}f\|, \quad (33)$$

with

$$C_1 = 1 + \left\| \Pi_+ \mathcal{A}^2 \Pi_0 (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \right\| < +\infty, \quad C_2 = \left\| (1 - \mathcal{S}_{++})^{1/2} \mathcal{A}_{+0} (\mathcal{A}_{+0}^* \mathcal{A}_{+0})^{-1} \right\| < +\infty. \quad (34)$$

Note that, for Langevin dynamics, $\|(1 - \mathcal{S})^{-1/2} \cdot\|$ is a norm equivalent to the canonical norm on $L^2(\nu, H^{-1}(\kappa))$ (since $\mathcal{S} = \mathcal{L}_{\text{FD}}$ with \mathcal{L}_{FD} given by (15)), so that (33) corresponds to the result in [4, Theorem 1.2]. The assumption on $(1 - \mathcal{S})^{1/2} \Pi_1$ is automatically satisfied for operators associated with linear Boltzmann dynamics, or Langevin dynamics with quadratic kinetic energies (since $\mathcal{S} \Pi_1$ is proportional to Π_1 in these cases and hence bounded, see the proof of Corollary 1).

Proof. To prove (33), we first note that, since $\|f\| \leq \|(1 - \Pi_0)f\| + \|\Pi_0 f\|$, it is sufficient to establish that there exists $\tilde{C}_1 \in \mathbb{R}_+$ such that, for any $f \in C_c^\infty(\mathcal{X})$,

$$\|\Pi_0 f - \langle f, \mathbf{1} \rangle\| \leq \tilde{C}_1 \|(1 - \Pi_0)f\| + C_2 \|(1 - \mathcal{S})^{-1/2} \mathcal{A}f\|. \quad (35)$$

The inequality (33) then holds with $C_1 = 1 + \tilde{C}_1$. To obtain (35), we assume without loss of generality that $\langle f, \mathbf{1} \rangle_{L^2(\mu)} = 0$ and compute

$$\begin{aligned} \|\Pi_0 f\|^2 &= \langle \mathcal{A}_+ f, \mathcal{A}_+ (\mathcal{A}_+^* \mathcal{A}_+)^{-1} \Pi_0 f \rangle \\ &= \langle \mathcal{A} \Pi_0 f, \mathcal{A}_+ (\mathcal{A}_+^* \mathcal{A}_+)^{-1} \Pi_0 f \rangle \\ &= \langle \mathcal{A} f, \mathcal{A}_+ (\mathcal{A}_+^* \mathcal{A}_+)^{-1} \Pi_0 f \rangle - \langle \mathcal{A}(1 - \Pi_0)f, \mathcal{A}_+ (\mathcal{A}_+^* \mathcal{A}_+)^{-1} \Pi_0 f \rangle \\ &\leq \|(1 - \mathcal{S})^{-1/2} \mathcal{A}f\| \left\| (1 - \mathcal{S})^{1/2} \mathcal{A}_+ (\mathcal{A}_+^* \mathcal{A}_+)^{-1} \Pi_0 f \right\| \\ &\quad + \|(1 - \Pi_0)f\| \left\| \Pi_+ \mathcal{A}^2 \Pi_0 (\mathcal{A}_+^* \mathcal{A}_+)^{-1} \Pi_0 f \right\|. \end{aligned}$$

Note that the operator $(1 - \mathcal{S})^{1/2} \mathcal{A}_+ (\mathcal{A}_+^* \mathcal{A}_+)^{-1}$ can be written as the composition of $(1 - \mathcal{S})^{1/2} \Pi_1$ (bounded by assumption) and the bounded operator $\mathcal{A}_+ (\mathcal{A}_+^* \mathcal{A}_+)^{-1}$. This leads finally to (33). \square

Space-time Poincaré inequality (formal). We now turn to space-time domains. We fix a time $T > 0$ and consider the reference Hilbert space

$$\mathcal{H}_T = \left\{ \varphi \in L^2(\tilde{\mu}_T) \mid \langle \varphi, \mathbf{1} \rangle_{L^2(\tilde{\mu}_T)} = 0 \right\},$$

where

$$\tilde{\mu}_T(dt dx) = \lambda_T(dt) \mu(dx), \quad \lambda_T(dt) = \frac{1}{T} \mathbf{1}_{[0, T]}(t) dt.$$

The space-time Poincaré inequality proved in [4, 23] is the following: For any $T > 0$, there exist $C_{1,T}, C_{2,T} \in \mathbb{R}_+$ such that, for any $f \in C_c^\infty([0, T] \times \mathcal{X})$,

$$\left\| f - \langle f, \mathbf{1} \rangle_{L^2(\tilde{\mu}_T)} \right\|_{L^2(\tilde{\mu}_T)} \leq C_{1,T} \|(1 - \Pi_0)f\|_{L^2(\tilde{\mu}_T)} + C_{2,T} \|(1 - \mathcal{L}_{\text{FD}})^{-1/2} (-\partial_t + \mathcal{L}_{\text{ham}}) f\|_{L^2(\tilde{\mu}_T)},$$

with \mathcal{L}_{ham} and \mathcal{L}_{FD} defined in (14) and (15) respectively.

This suggests considering the following space-time Poincaré inequality for generators $\mathcal{L} = \mathcal{A} + \mathcal{S}$:

$$\left\| f - \langle f, \mathbf{1} \rangle_{L^2(\tilde{\mu}_T)} \right\|_{L^2(\tilde{\mu}_T)} \leq C_{1,T} \|(1 - \Pi_0)f\|_{L^2(\tilde{\mu}_T)} + C_{2,T} \|(1 - \mathcal{S})^{-1/2} (-\partial_t + \mathcal{A}) f\|_{L^2(\tilde{\mu}_T)}. \quad (36)$$

Formally, this corresponds to replacing \mathcal{L} in the statement of Proposition 5 by $-\partial_t + \mathcal{L}$, considered on the Hilbert space \mathcal{H}_T . The total antisymmetric part of the operator is then $-\partial_t + \mathcal{A}$. In order to derive an inequality such as (36), it formally suffices to replace the operator \mathcal{A}_+ in the derivation of (35) by $-\partial_t \Pi_0 + \mathcal{A}_+$. Let us emphasize that this operator does not correspond to the restriction of $-\partial_t + \mathcal{A} + \mathcal{S}$ from \mathcal{H}_0 to \mathcal{H}_+ because of the term $-\partial_t \Pi_0$. There are however some caveats in such a derivation, related to the boundary conditions in the time domain, which is why a careful treatment is required; see [23, Lemma 2.6].

Exponential convergence. Let us finally show how to formally deduce an exponential convergence from (36) (the complete rigorous argument would involve regularizations and truncations in order to give a meaning to all time derivatives) when Assumption 2 holds true. Recall that scalar products and associated norms are by default the ones on $L^2(\mu)$, unless explicitly mentioned otherwise.

In order to highlight the fact that it is possible to obtain lower bounds on the convergence rate with respect to some parameter in the dynamics, we consider the case when the generator is of the form

$$\mathcal{L} = \mathcal{A} + \gamma \mathcal{S},$$

for some parameter $\gamma > 0$. Consider $\varphi(t) = e^{t\mathcal{L}}\varphi_0$ for a given function $\varphi_0 \in \mathcal{H}$. Note first that $\langle \varphi(t), \mathbf{1} \rangle = 0$ for all $t \geq 0$, so that $\langle \varphi, \mathbf{1} \rangle_{L^2(\tilde{\mu}_T)} = 0$. Moreover, $t \mapsto \|\varphi(t)\|$ is non-increasing since

$$\frac{d}{dt} \left(\frac{1}{2} \|\varphi(t)\|^2 \right) = \gamma \langle \varphi(t), \mathcal{S}\varphi(t) \rangle \leq 0.$$

Finally, from the evolution equation satisfied by $\varphi(t)$,

$$-\gamma \mathcal{S}\varphi(t) = (-\partial_t + \mathcal{A})\varphi(t) = \Pi_*(-\partial_t + \mathcal{A})\varphi(t),$$

which implies, upon applying $(-\mathcal{S})^{-1/2}$ to both sides of the above equality (which is indeed possible in view of Assumption 2), that

$$\begin{aligned} -\gamma \langle \varphi(t), \mathcal{S}\varphi(t) \rangle &= \gamma \left\| (-\mathcal{S})^{1/2} \varphi(t) \right\|^2 = \frac{1}{\gamma} \left\| (-\mathcal{S})^{-1/2} (-\partial_t + \mathcal{A})\varphi(t) \right\|^2 \\ &\geq \frac{1}{\gamma} \left\| (1 - \mathcal{S})^{-1/2} (-\partial_t + \mathcal{A})\varphi(t) \right\|^2. \end{aligned}$$

Therefore, for any $\eta \in (0, 1)$,

$$\begin{aligned} \|\varphi(T)\|^2 - \|\varphi(0)\|^2 &= 2\gamma T \langle \varphi, \mathcal{S}\varphi \rangle_{L^2(\tilde{\mu}_T)} \\ &\leq -2\gamma T(1 - \eta)s \|\Pi_*\varphi\|_{L^2(\tilde{\mu}_T)}^2 - \frac{2T\eta}{\gamma} \left\| (1 - \mathcal{S})^{-1/2} (-\partial_t + \mathcal{A})\varphi \right\|_{L^2(\tilde{\mu}_T)}^2. \end{aligned}$$

The choice $\eta = \gamma^2 s C_{2,T}^2 / (\gamma^2 s C_{2,T}^2 + C_{1,T}^2)$ leads to

$$\begin{aligned} \|\varphi(T)\|_{L^2(\mu)}^2 - \|\varphi(0)\|_{L^2(\mu)}^2 &\leq -\frac{2\gamma s T}{\gamma^2 s C_{2,T}^2 + C_{1,T}^2} \left(C_{1,T}^2 \|(1 - \Pi_0)\varphi(t)\|_{L^2(\tilde{\mu}_T)}^2 + C_{2,T}^2 \left\| (1 - \mathcal{S})^{-1/2} (-\partial_t + \mathcal{A})\varphi(t) \right\|_{L^2(\tilde{\mu}_T)}^2 \right) \\ &\leq -\frac{\gamma s T}{\gamma^2 s C_{2,T}^2 + C_{1,T}^2} \left(C_{1,T} \|(1 - \Pi_0)\varphi(t)\|_{L^2(\tilde{\mu}_T)} + C_{2,T} \left\| (1 - \mathcal{S})^{-1/2} (-\partial_t + \mathcal{A})\varphi(t) \right\|_{L^2(\tilde{\mu}_T)} \right)^2 \\ &\leq -\frac{\gamma s T}{\gamma^2 s C_{2,T}^2 + C_{1,T}^2} \|\varphi\|_{L^2(\tilde{\mu}_T)}^2 \leq -\frac{\gamma s T}{\gamma^2 s C_{2,T}^2 + C_{1,T}^2} \|\varphi(T)\|_{L^2(\mu)}^2, \end{aligned}$$

where we used the Poincaré inequality (36), and the fact that $t \mapsto \|\varphi(t)\|_{L^2(\mu)}$ is non-increasing in the last step. By rewriting this final inequality as

$$\|\varphi(T)\|_{L^2(\mu)}^2 \leq \alpha_T \|\varphi(0)\|_{L^2(\mu)}^2, \quad 0 \leq \alpha_T = \left(1 + \frac{\gamma s T}{\gamma^2 s C_{2,T}^2 + C_{1,T}^2} \right)^{-1} < 1,$$

we can then conclude to an exponential convergence as in [4, 23]. It is also possible to make explicit the scaling with respect to the dimension and the friction coefficient γ in this setting, as done for Langevin dynamics in Section 3.1. More precisely, for $T > 0$ fixed, it holds

$$\alpha_T \underset{\gamma \rightarrow 0}{\sim} e^{-\gamma s T / C_{1,T}^2}, \quad \alpha_T \underset{\gamma \rightarrow +\infty}{\sim} e^{-T / (\gamma C_{2,T}^2)},$$

which shows that the convergence rate of the semigroup $e^{t\mathcal{L}}$ is indeed of order $\min(\gamma, \gamma^{-1})$.

Remark 3. If the constants $C_{1,T}, C_{2,T}$ in (36) were uniformly bounded by values $C_1, C_2 > 0$ with respect to $T > 0$, it would be possible to take the limit $T \rightarrow 0$ in the above estimates. This would lead to

$$\frac{d}{dt} \left(\|\varphi(t)\|_{L^2(\mu)}^2 \right) \leq -\frac{\gamma s}{\gamma^2 s C_2^2 + C_1^2} \|\varphi(t)\|_{L^2(\mu)}^2,$$

and hence to an exponential convergence without prefactor, which is clearly false since this would imply the coercivity of the generator for the canonical scalar product on $L^2(\mu)$. In [23], the constant $C_{1,T}$ in (36) involves a term scaling as $1/T$, which prevents taking the limit $T \rightarrow 0$. It can be traced back to the estimates on ψ'_2 in [23, Lemma 2.6].

Acknowledgements. We thank David Herzog for fruitful discussions on [4, 23], Francis Nier for pointing out [77], and the Heilbronn Institute for Mathematical Research in Bristol, UK, where this work was first presented in a workshop on hypocoercivity (3-4 March 2020). The work of A.L and G.S. was funded in part from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 810367). The work of G.S. was also funded by the Agence Nationale de la Recherche, under grant ANR-19-CE40-0010-01 (QuAMProcs). The work of M.F. was funded by the ANR, under grants ANR-17-CE40-0030 (EFI) and ANR-18-CE40-006 (MESA), as well as ANR-11-LABX-0040-CIMI within the program ANR-11-IDEX-0002-02. This work was done while M.F. was a guest of the teams Mokaplan and Materials at the INRIA Paris, whose support is gratefully acknowledged.

References

- [1] F. Achleitner, A. Arnold, and E. A. Carlen. On linear hypocoercive BGK models. In P. Gonçalves and A. Soares, editors, *From Particle Systems to Partial Differential Equations. III*, volume 162 of *Springer Proc. Math. Stat.*, pages 1–37. Springer, 2016.
- [2] F. Achleitner, A. Arnold, and D. Stürzer. Large-time behavior in non-symmetric Fokker-Planck equations. *Riv. Math. Univ. Parma*, 6(1):1–68, 2015.
- [3] C. Andrieu, A. Durmus, N. Nüsken, and J. Roussel. Hypocoercivity of piecewise deterministic Markov process-Monte Carlo. *arXiv preprint*, 1808.08592, 2018. To appear in *Annals of Applied Probability*.
- [4] S. Armstrong and J.-C. Mourrat. Variational methods for the kinetic Fokker-Planck equation. *arXiv preprint*, 1902.04037, 2019.
- [5] A. Arnold, C. Schmeiser, and B. Signorello. Propagator norm and sharp decay estimates for Fokker-Planck equations with linear drift. *arXiv preprint*, 2003.01405, 2020.
- [6] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality in a large class of probability measures including log-concave cases. *Elec. Comm. Prob.*, 13:60–66, 2008.
- [7] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3):727–759, 2008.
- [8] D. Bakry and M. Emery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX*, volume 1123 of *Lecture Notes in Mathematics*, pages 177–206. Springer Verlag, 1985.
- [9] R. Balian. *From Microphysics to Macrophysics. Methods and Applications of Statistical Physics*, volume I - II. Springer, 2007.
- [10] F. Baudoin. Bakry-Émery meet Villani. *J. Funct. Anal.*, 273(7):2275–2291, 2017.

- [11] F. Baudoin, M. Gordina, and D. P. Herzog. Gamma calculus beyond Villani and explicit convergence estimates for Langevin dynamics with singular potentials. *Arch. Ration. Mech. Anal.*, 241(2):765–804, 2021.
- [12] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta numerica*, 14:1, 2005.
- [13] É. Bernard and F. Salvarani. On the exponential decay to equilibrium of the degenerate linear Boltzmann equation. *J. Funct. Anal.*, 265(9):1934–1954, 2013.
- [14] M. Bessemoulin-Chatard, M. Herda, and T. Rey. Hypocoercivity and diffusion limit of a finite volume scheme for linear kinetic equations. *Math. Comp.*, 89:1093–1133, 2020.
- [15] R.N. Bhattacharya. On the functional Central Limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrscheinlichkeit*, 60(2):185–201, 1982.
- [16] J. Birrell and L. Rey-Bellet. Concentration inequalities and performance guarantees for hypocoercive MCMC samplers. *arXiv preprint*, 1907.11973, 2019.
- [17] N. Bou-Rabee and J. M. Sanz-Serna. Randomized Hamiltonian Monte Carlo. *Ann. Appl. Probab.*, 27(4):2159–2194, 2017.
- [18] E. Bouin, J. Dolbeault, L. Lafèche, and C. Schmeiser. Fractional hypocoercivity. *arXiv preprint*, 1911.11020, 2019.
- [19] E. Bouin, J. Dolbeault, S. Mischler, C. Mouhot, and C. Schmeiser. Hypocoercivity without confinement. *Pure Appl. Anal.*, 2(2):203–232, 2020.
- [20] E. Bouin, F. Hoffmann, and C. Mouhot. Exponential decay to equilibrium for a fiber lay-down process on a moving conveyor belt. *SIAM J. Math. Anal.*, 49(4):3233–3251, 2017.
- [21] E. Camrud, D. P. Herzog, G. Stoltz, and M. Gordina. Weighted L^2 -contractivity of Langevin dynamics with singular potentials. *arXiv preprint*, 2104.10574, 2021.
- [22] J. Canizo, C. Cao, J. Evans, and H. Yoldas. Hypocoercivity of linear kinetic equations via Harris’s theorem. *Kinetic and Related Models*, 13(1):97–128, 2020.
- [23] Y. Cao, J. Lu, and L. Wang. On explicit L^2 -convergence rate estimate for underdamped Langevin dynamics. *arXiv preprint*, 1908.04746, 2019.
- [24] E. A. Carlen, R. Esposito, J. L. Lebowitz, R. Marra, and C. Mouhot. Approach to the steady state in kinetic models with thermal reservoirs at different temperatures. *J. Stat. Phys.*, 172(2):522–543, 2018.
- [25] K. Carrapatoso, J. Dolbeault, F. Hérau, S. Mischler, C. Mouhot, and C. Schmeiser. Special modes and hypocoercivity for linear kinetic equations with several conservation laws and a confining potential. *arXiv preprint*, 2105.04855, 2021.
- [26] P. Cattiaux and A. Guillin. Semi log-concave Markov diffusions. In *Séminaire de Probabilités XLVI*, volume 2123 of *Lecture Notes in Mathematics*, pages 231–292. Springer, Cham, 2014.
- [27] P. Cattiaux, A. Guillin, P. Monmarché, and C. Zhang. Entropic multipliers method for Langevin diffusion and weighted log Sobolev inequalities. *J. Funct. Anal.*, 277(11), 2019.
- [28] G. Deligiannidis, D. Paulin, and A. Doucet. Randomized Hamiltonian Monte Carlo as scaling limit of the Bouncy Particle Sampler and dimension-free convergence rates. *arXiv preprint*, 1808.04299, 2018. To appear in *Annals of Applied Probability*.
- [29] L. Desvillettes and C. Villani. On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: The linear Fokker–Planck equation. *Comm. Pure Appl. Math.*, 54(1):1–42, 2001.

- [30] L. Devroye and G. Lugosi. *Combinatorial Methods in Density Estimation*. Springer Series in Statistics. Springer-Verlag, New York, 2001.
- [31] J. Dolbeault, A. Klar, C. Mouhot, and C. Schmeiser. Exponential rate of convergence to equilibrium for a model describing fiber lay-down processes. *Appl. Math. Res. eXpress*, 2013(2):165–175, 2013.
- [32] J. Dolbeault, C. Mouhot, and C. Schmeiser. Hypocoercivity for kinetic equations with linear relaxation terms. *C. R. Math. Acad. Sci. Paris*, 347(9-10):511–516, 2009.
- [33] J. Dolbeault, C. Mouhot, and C. Schmeiser. Hypocoercivity for linear kinetic equations conserving mass. *Trans. AMS*, 367(6):3807–3828, 2015.
- [34] G. Dujardin, F. Hérau, and P. Lafitte. Coercivity, hypocoercivity, exponential time decay and simulations for discrete Fokker–Planck equations. *Numer. Math.*, 144:615–697, 2020.
- [35] A. B. Duncan, N. Nüsken, and G. A. Pavliotis. Using perturbed underdamped Langevin dynamics to efficiently sample from probability distributions. *J. Stat. Phys.*, 169(6):1098–1131, 2017.
- [36] A. Eberle, A. Guillin, and R. Zimmer. Coupling and quantitative contraction rates for Langevin dynamics. *Ann. Probab.*, 47(4):1982–2010, 2019.
- [37] J.-P. Eckmann and M. Hairer. Spectral properties of hypoelliptic operators. *Commun. Math. Phys.*, 235:233–253, 2003.
- [38] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.
- [39] J. Evans. Hypocoercivity in Wasserstein-1 for the kinetic Fokker-Planck equation via Malliavin calculus. *arXiv preprint*, 1810.01324, 2018.
- [40] J. Evans. Hypocoercivity in Phi-entropy for the linear relaxation Boltzmann equation on the torus. *SIAM J. Math. Anal.*, 53(2):1357–1378, 2021.
- [41] D. Frenkel and B. Smit. *Understanding Molecular Simulation, From Algorithms to Applications (2nd ed.)*. Academic Press, 2002.
- [42] E. H. Georgoulis. Hypocoercivity-compatible finite element methods for the long-time computation of Kolmogorov’s equation. *SIAM J. Numer. Anal.*, 59(1):173–194, 2021.
- [43] M. Grothaus and P. Stilgenbauer. Hilbert space hypocoercivity for the Langevin dynamics revisited. *Methods Funct. Anal. Topology*, 22(2):152–168, 2016.
- [44] M. Grothaus and F.-Y. Wang. Weak Poincaré inequalities for convergence rate of degenerate diffusion processes. *Ann. Probab.*, 47(5):2930–2952, 2019.
- [45] M. Hairer and G. Pavliotis. From ballistic to diffusive behavior in periodic potentials. *J. Stat. Phys.*, 131:175–202, 2008.
- [46] D. Han-Kwan and M. Léautaud. Geometric analysis of the linear Boltzmann equation I. Trend to equilibrium. *Ann. PDE*, 1(1):3, 2015.
- [47] B. Helffer and F. Nier. *Hypoelliptic Estimates and Spectral Theory for Fokker–Planck Operators and Witten Laplacians*, volume 1862 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2005.
- [48] F. Hérau. Hypocoercivity and exponential time decay for the linear inhomogeneous relaxation Boltzmann equation. *Asymptot. Anal.*, 46(3-4):349–359, 2006.

- [49] F. Hérau. Short and long time behavior of the Fokker–Planck equation in a confining potential and applications. *J. Funct. Anal.*, 244(1):95–118, 2007.
- [50] F. Hérau. Introduction to hypocoercive methods and applications for simple linear inhomogeneous kinetic models. In *Lectures on the Analysis of Nonlinear Partial Differential Equations. Part 5*, volume 5 of *Morningside Lect. Math.*, pages 119–147. Int. Press, Somerville, MA, 2018.
- [51] F. Hérau and F. Nier. Isotropic hypoellipticity and trend to equilibrium for the Fokker–Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.*, 171:151–218, 2004.
- [52] D. P. Herzog. Exponential relaxation of the Nosé–Hoover thermostat under Brownian heating. *Commun. Math. Sci.*, 16(8):2231–2260, 2018.
- [53] D. P. Herzog and J. C. Mattingly. Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials. *Comm. Pure Appl. Math.*, 72(10):2231–2255, 2019.
- [54] A. Iacobucci, S. Olla, and G. Stoltz. Convergence rates for nonequilibrium Langevin dynamics. *Ann. Math. Québec*, 43(1):73–98, 2019.
- [55] S. M. Kozlov. Effective diffusion for the Fokker–Planck equation. *Math. Notes*, 45(5-6):360–368, 1989.
- [56] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Mathematics*, pages 120–216. Springer, Berlin, 1999.
- [57] B. Leimkuhler and C. Matthews. *Molecular Dynamics*, volume 39 of *Interdisciplinary Applied Mathematics*. Springer, 2015.
- [58] B. Leimkuhler, C. Matthews, and G. Stoltz. The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.*, 36(1):13–79, 2016.
- [59] B. Leimkuhler, M. Sachs, and G. Stoltz. Hypocoercivity properties of adaptive Langevin dynamics. *SIAM J. Appl. Math.*, 80(3):1197–1222, 2020.
- [60] T. Lelièvre and G. Stoltz. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numerica*, 25:681–880, 2016.
- [61] V. Letizia and S. Olla. Non-equilibrium isothermal transformations in a temperature gradient from a microscopic dynamics. *Ann. Probab.*, 45(6A):3987–4018, 2017.
- [62] J. Lu and L. Wang. On explicit L^2 -convergence rate estimate for piecewise deterministic Markov processes. *arXiv preprint*, 2007.14927, 2020. to appear in *Annals of Applied Probability*.
- [63] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.
- [64] A. Menegaki. Quantitative rates of convergence to non-equilibrium steady state for a weakly anharmonic chain of oscillators. *J. Stat. Phys.*, 181(1):53–94, 2020.
- [65] G. Metafune, D. Pallara, and E. Priola. Spectrum of Ornstein–Uhlenbeck operators in L^p spaces with respect to invariant measures. *J. Funct. Anal.*, 196(1):40–60, 2002.
- [66] M. Mokhtar-Kharroubi. On L^1 exponential trend to equilibrium for conservative linear kinetic equations on the torus. *J. Funct. Anal.*, 266(11):6418–6455, 2014.

- [67] P. Monmarché. Generalized Γ calculus and application to interacting particles on a graph. *Potential Analysis*, 50(3):439–466, 2019.
- [68] P. Monmarché. L^2 hypocoercivity, deviation bounds, hitting times and Lyapunov functions. *arXiv preprint*, 1911.01748, 2019. To appear in *Annales Mathématiques Blaise Pascal*.
- [69] P. Monmarché. A note on Fisher information hypocoercive decay for the linear Boltzmann equation. *Anal. Math. Phys.*, 11:1, 2021.
- [70] C. Mouhot and L. Neumann. Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus. *Nonlinearity*, 19(4):969–998, 2006.
- [71] F. Otto and M. G. Reznikoff. A new criterion for the logarithmic Sobolev inequality and two applications. *J. Funct. Anal.*, 243(1):121–157, 2007.
- [72] M. Ottobre and G. A. Pavliotis. Asymptotic analysis for the generalized Langevin equation. *Nonlinearity*, 24(5):1629–1653, 2011.
- [73] G. A. Pavliotis. *Stochastic Processes and Applications*, volume 60 of *Texts in Applied Mathematics*. Springer, New York, 2014.
- [74] A. Porretta and E. Zuazua. Numerical hypocoercivity for the Kolmogorov equation. *Math. Comp.*, 86(303):97–119, 2017.
- [75] L. Rey-Bellet. Ergodic properties of Markov processes. In S. Attal, A. Joye, and C.-A. Pillet, editors, *Open Quantum Systems II*, volume 1881 of *Lecture Notes in Mathematics*, pages 1–39. Springer, 2006.
- [76] J. Roussel and G. Stoltz. Spectral methods for Langevin dynamics and associated error estimates. *ESAIM Math. Model. Numer. Anal.*, 52(3):1051–1083, 2018.
- [77] J. Sjöstrand and M. Zworski. Elementary linear algebra for advanced spectral problems. *Ann. Inst. Fourier (Grenoble)*, 57(7):2095–2141, 2007.
- [78] G. Stoltz and Z. Trstanova. Langevin dynamics with general kinetic energies. *Multiscale Model. Sim.*, 16(2):777–806, 2018.
- [79] G. Stoltz and E. Vanden-Eijnden. Longtime convergence of the Temperature-Accelerated Molecular Dynamics method. *Nonlinearity*, 31(8):3748–3769, 2018.
- [80] D. Talay. Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Proc. Rel. Fields*, 8:163–198, 2002.
- [81] M. M. Tropper. Ergodic and quasideterministic properties of finite-dimensional stochastic systems. *J. Statist. Phys.*, 17(6):491–509, 1977.
- [82] M. Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford University Press, 2010.
- [83] S. Ukai. On the existence of global solutions of mixed problem for non-linear Boltzmann equation. *Proc. Japan Acad.*, 30:179–184, 1970.
- [84] I. Vidav. Spectra of perturbed semigroups with applications to transport theory. *J. Math. Anal. Appl.*, 30:264–279, 1970.
- [85] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950), 2009.
- [86] L. Wu. Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stoch. Proc. Appl.*, 91(2):205–238, 2001.

- [87] N. Yoshida. The equivalence of the log-Sobolev inequality and a mixing condition for unbounded spin systems on the lattice. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(2):223–243, 2001.