

Conservative stochastic differential equations: Mathematical and numerical analysis

Erwan Faou¹ & Tony Lelièvre²

¹*INRIA, Projet-team Ipso, ENS Cachan Bretagne, Avenue Robert Schumann, 35170 Bruz, FRANCE*

²*CERMICS, université Paris-Est, ENPC, 6-8 avenue Blaise Pascal, Champs sur Marne, 77455 Marne La Vallée Cédex 2, FRANCE*

*CERMICS — ENPC
6 et 8 avenue Blaise Pascal
Cité Descartes - Champs sur Marne
77455 Marne la Vallée Cedex 2*

<http://cermics.enpc.fr>

Abstract

We consider stochastic differential equations on the whole space possessing an invariant along their solutions. The stochastic dynamics therefore evolves on a hypersurface of the ambient Euclidean space. Using orthogonal coordinate systems, we show the existence and uniqueness of smooth solutions of the Kolmogorov equation under some ellipticity conditions over the invariant hypersurfaces. If we assume moreover the existence of an invariant measure, we show the exponential convergence of the solution towards its average. In a second part, we consider numerical approximation of the stochastic differential equation, and show the convergence and numerical ergodicity of a class of projected schemes. In the context of molecular dynamics, this yields numerical schemes that are ergodic with respect to the microcanonical measure over isoenergy surfaces.
MSC numbers: 60H10, 60H30, 58J65, 65C20.

Contents

1	Introduction	1
2	Analysis of the Kolmogorov operators	5
2.1	Orthogonal coordinate system	5
2.2	Existence of solutions	10
2.3	Invariant measure and exponential convergence	12
3	Numerical analysis	18
3.1	Convergence result	20
3.2	The ergodic case.	24
4	Applications	25
4.1	Two examples of conservative SDEs	26
4.1.1	Stochastic shakers	26
4.1.2	Projected gradient dynamics	27
4.2	Numerical schemes	29
4.2.1	Methods for stochastic shakers.	29
4.2.2	Projected gradient dynamics schemes.	30
4.3	Numerical results	31

1 Introduction

We consider stochastic differential equations in \mathbb{R}^N in the Stratonovich sense, of the form

$$dX(t) = f(X(t))dt + \sigma(X(t)) \circ dW(t),$$

written in coordinates

$$dX^i(t) = f^i(X(t))dt + \sum_{\ell=1}^D \sigma_{[\ell]}^i(X(t)) \circ dW^{[\ell]}(t), \quad (1.1)$$

where $f(x) = (f^i(x))_{i=1}^N$ is an N -dimensional vector field, and $\sigma(x) = (\sigma_{[\ell]}^i(x))$, $i = 1, \dots, N$, $\ell = 1, \dots, D$ is a $N \times D$ matrix. The vector $W(t)$ with components $W^{[\ell]}$, $\ell = 1, \dots, D$, is a standard Brownian motion in \mathbb{R}^D . To this equation corresponds the Itô equation: For $i = 1, \dots, N$,

$$dX^i(t) = \left(f^i(X(t)) + \sum_{j=1}^N \sum_{\ell=1}^D \frac{1}{2} (\sigma_{[\ell]}^j \partial_j \sigma_{[\ell]}^i)(X(t)) \right) dt + \sum_{\ell=1}^D \sigma_{[\ell]}^i(X(t)) dW^{[\ell]}(t) \quad (1.2)$$

where $\partial_j \sigma_{[\ell]}^i$ denotes the derivative with respect to x^j of the function $\sigma_{[\ell]}^i(t, x)$, $x = (x^1, \dots, x^N)$.

We assume in this work that the SDE (1.1) is *conservative* in the following sense: There exists a function $H(x)$ defined on \mathbb{R}^N such that

$$\forall x \in \mathbb{R}^N, \quad \forall \ell = 1, \dots, D, \quad \langle \nabla H(x), f(x) \rangle = 0 \quad \text{and} \quad \langle \nabla H(x), \sigma_{[\ell]}(x) \rangle = 0 \quad (1.3)$$

where for all $\ell = 1, \dots, D$, $\sigma_{[\ell]}(x)$ is the N -dimensional vector with components $\sigma_{[\ell]}^i$, and where $\langle \cdot, \cdot \rangle$ denotes the Euclidean scalar product on \mathbb{R}^N . The vector field $\nabla H(x)$ is the gradient vector of H with respect to the coordinates $x = (x^1, \dots, x^N)$.

Using Itô calculus, it is straightforward to show that if $H(x)$ is smooth enough on \mathbb{R}^N , we have for all solution $X(t)$ of (1.1)

$$\forall t \geq 0, \quad H(X(t)) = H(X(0)). \quad (1.4)$$

This expresses that the function H is an invariant of (1.1).

Under ellipticity assumptions on the vector fields $\sigma_{[\ell]}$, we show below that the SDE (1.1) exhibits a unique invariant measure over the manifolds

$$\Sigma_z = \{x \in \mathbb{R}^N | H(x) = z\}. \quad (1.5)$$

In the context of molecular dynamics, if the equation (1.1) is *volume preserving* over the whole space \mathbb{R}^N , then the invariant measure turns to be the *microcanonical* probability measure (see for instance [1, 11]),

$$d\nu(z) = \left(\int_{\Sigma_z} \frac{d\Sigma_z}{\|\nabla H\|} \right)^{-1} \frac{d\Sigma_z}{\|\nabla H\|} \quad (1.6)$$

where $d\Sigma_z$ is the surface measure over the isosurface Σ_z induced by the Euclidean metric of \mathbb{R}^N .

For a general function $H(x)$, it is very difficult in practice to exhibit an atlas of the manifolds $\Sigma_z = \{x \in \mathbb{R}^N | H(x) = z\}$. The derivation of a numerical scheme to approximate the solution of (1.1) cannot thus rely on the choice of a “good” atlas as in [13], and one has to rely on projection step to remain on the manifold. The numerical analysis then has to be made directly in \mathbb{R}^N and the dynamics of (1.1) has to be understood in a vicinity of a given hypersurface Σ_z rather than on the manifold itself.

The Kolmogorov equation associated with the SDE (1.1) writes, using the summation convention of covariant and contravariant indices

$$\partial_t v = Lv \quad \text{where} \quad Lv = f^i \partial_i v + \frac{1}{2} \sum_{\ell=1}^D \sigma_{[\ell]}^i \partial_i (\sigma_{[\ell]}^j \partial_j v), \quad (1.7)$$

where $v(t, x)$ is a function depending on the time $t \geq 0$ and $x \in \mathbb{R}^N$. We recall that a solution of (1.7) is given by $v(t, x) = \mathbb{E}(v(0, X(t)) | X(0) = x)$.

After a possible translation of the function $H(x)$, we can always assume that the initial value of (1.1) lies on the manifold $\Sigma_0 = \{x \in \mathbb{R}^N | H(x) = 0\}$. In this work, we will show that we can define smooth solutions of (1.7) in a neighborhood of this manifold by considering an orthogonal parametrisation of \mathbb{R}^N by the hypersurface Σ_0 which is supposed to be a $N - 1$ dimensional submanifold of \mathbb{R}^N , and the parameter $z = H(x)$ for $x \in \mathbb{R}^N$. In such a way, we can parametrise a domain of \mathbb{R}^N by a product $\Sigma_0 \times (-\varepsilon, \varepsilon)$ for small ε , provided that Σ_0 is smooth enough and compact (note that this last hypothesis could be weakened in the present analysis, but it is in general fulfilled for the problem of computing averages in molecular dynamics).

In this new coordinate system, the Kolmogorov operator can be seen as an *intrinsic* operator on Σ_0 involving the covariant derivative and curvature terms on Σ_0 . The coordinate z is then only a smooth parameter in the operator. In particular, the derivative with respect to the global variable z does not come into play in the definition of the Kolmogorov operator in orthogonal coordinate. By standard arguments, we therefore can define a smooth solution in the neighborhood defined by $\Sigma_0 \times (-\varepsilon, \varepsilon)$, provided that the collection $\sigma_{[\ell]}(x)$ satisfy conditions so that the operator L can be viewed as elliptic operator on Σ_0 , uniformly in z .

We then consider the case where (1.1) is possess a family of invariant measures on Σ_z , depending smoothly on z . An important case where this situation occurs is when (1.1) is *volume preserving* which means that the constants are in the kernel of the adjoint of the operator (1.7) which writes

$$L^T v = -\partial_i (f^i v) + \frac{1}{2} \sum_{\ell=1}^D \partial_i (\sigma_{[\ell]}^i \partial_j (\sigma_{[\ell]}^j v)). \quad (1.8)$$

We will refer to this operator as the Fokker-Planck operator associated with (1.1). Note that (1.1) will be volume preserving in the case where

$$\partial_i f^i(x) = 0 \quad \text{and} \quad \forall \ell = 1, \dots, D, \quad \partial_i \sigma_{[\ell]}^i(x) = 0, \quad (1.9)$$

but this condition is not necessary (see the examples in Section 4).

The motivation for considering such volume and energy preserving systems arises from computational problems of NVE averages in molecular dynamics (namely averages with respect to the microcanonical measure (1.6), see for instance [1, 11]). In this situation, for d interacting particles, the phase space variable decomposes into $x = (p_1, \dots, p_d, q_1, \dots, q_d)$ where each $q_i \in \mathbb{R}^3$ denotes the position of a particle, and p_i its momentum. The energy hamiltonian $H(p, q)$ is then written

$$H(p, q) = \sum_{i=1}^d \frac{p_i^T p_i}{2m_i} + U(q_1, \dots, q_d) \quad (1.10)$$

where m_i is the mass of the i -th particle and U is the potential function.

The principle of molecular dynamics is to simulate numerically the solution of the corresponding hamiltonian equations

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}(p, q) \quad \text{and} \quad \frac{dp}{dt} = -\frac{\partial H}{\partial q}(p, q) \quad (1.11)$$

which define a volume and energy preserving (deterministic) system. The *ergodic hypothesis* states that, if $(p(t), q(t))$ denotes a solution of (1.11) evolving on a hypersurface Σ_z corresponding to the Hamiltonian (1.10), we have for all function g defined on Σ_z ,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T g(p(t), q(t)) dt = \int_{\Sigma_z} g(y) d\nu(z) \quad (1.12)$$

where the measure $d\nu(z)$ in the right-hand side is the microcanonical measure (1.6).

However, it is well-known that this hypothesis failed in general, in particular for integrable or close to integrable systems exhibiting hidden stable invariants preserved by symplectic numerical methods (see [14, 5] and the references therein). At variance the numerical schemes we present below are shown to be ergodic with respect to the microcanonical measure.

Another context where such families of measures appear is free energy computations by thermodynamic integration [7]. In this case, one needs to sample a Boltzmann-Gibbs probability measure conditioned to a fixed value of a reaction coordinate: $\exp(-\beta V) |\nabla \xi|^{-1} dS_z$ where $V : \mathbb{R}^N \rightarrow \mathbb{R}$ is the energy, $\xi : \mathbb{R}^N \rightarrow \mathbb{R}$ is the reaction coordinate and $S_z = \{x \in \mathbb{R}^N | \xi(x) = z\}$ (ξ plays here the role of the above function H).

In a first part of this paper, under an ellipticity assumption on the vector fields $\sigma_{[\ell]}$ and in the case where (1.1) admits an invariant measure, we show the exponential convergence of the solution of (1.7) towards its average over the microcanonical measure, uniformly in z for z in a small interval. We moreover show that all the tangent space derivatives of $v(t, x)$ decay exponentially in time. This implies in particular the ergodicity of the exact solution of (1.1) and this gives the equivalent of (1.12) for the stochastic flow of (1.1) in the volume preserving case.

In a second part of this work, we consider numerical schemes satisfying in essence two conditions: Consistency with (1.1) and preservation of the energy $H(x)$. Using the results obtained in Section 2, we show the weak convergence of such numerical schemes, which gives a new way of computing NVE averages in the context of molecular dynamics. We conclude by giving examples of schemes satisfying the above conditions.

The method and analysis we use are closely related to previous works by Talay and co-workers [22, 13, 20, 21]. The particularity here lies in a systematic use of differential geometry to understand the properties of the Kolmogorov operator on a hypersurface of the ambient space. Let us also mention the analysis made in [17, 16] where constrained symplectic SDEs and appropriate discretization schemes are introduced.

In the following, for a given $\varepsilon > 0$, we define the domain

$$\Omega^\varepsilon = \{x \in \mathbb{R}^N | H(x) < \varepsilon\}.$$

It defines a neighborhood of the hypersurface Σ_0 .

2 Analysis of the Kolmogorov operators

The goal of this section is to derive expressions of the operators L and L^T in terms of intrinsic objects defined on the hypersurfaces Σ_z .

2.1 Orthogonal coordinate system

We can write the Kolmogorov and Fokker-Planck operators as

$$\begin{aligned} Lv &= f^i \nabla_i v + \sum_{\ell=1}^D \frac{1}{2} \sigma_{[\ell]}^i \nabla_i (\sigma_{[\ell]}^j \nabla_j v). \\ \text{and } L^T v &= -\nabla_i (f^i v) + \sum_{\ell=1}^D \frac{1}{2} \nabla_i (\sigma_{[\ell]}^i \nabla_j (\sigma_{[\ell]}^j v)) \end{aligned} \tag{2.1}$$

respectively, where ∇ is the Riemannian connection of \mathbb{R}^N associated with the Euclidean product $\langle \cdot, \cdot \rangle$. The advantage of these expressions is that they are

expressed only in term of globally defined objects: The covariant derivative and the metric tensor. In the Euclidean coordinate system $\{x^1, \dots, x^N\}$, the metric tensor is the identity tensor, and this expression reduces to (1.7). For a general local coordinate system $\{y^1, \dots, y^N\}$, the metric is represented by the matrix $g_{ij}(y)$, $i, j = 1, \dots, N$ which are the component of the metric tensor in the coordinate system $\{y^i\}$. With this metric tensor are associated the Christoffel symbols (see for instance [4])

$$\Gamma_{ij}^k = \frac{1}{2}g^{km}(\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}), \quad (2.2)$$

where ∂_i denotes the derivative with respect to y^i , where g^{km} are the components of the inverse of the metric tensor defined by $g^{km}g_{ml} = \delta_l^k$ the Kronecker tensor.

The coordinate system $\{y^i\}$ induces a local basis $\frac{\partial}{\partial y^i}$ in the corresponding tangent bundle. For a vector field $T = T^i \frac{\partial}{\partial y^i}$, the covariant derivative is the (1, 1) tensor fields with components

$$\nabla_i T^j = \partial_i T^j + \Gamma_{im}^j T^m.$$

Now an expression like $\nabla_i T^i$ is the contraction of the previous tensor, and is thus a function equal to

$$\nabla_i T^i = \partial_i T^i + \Gamma_{im}^i T^m.$$

It is worth noticing that the right-hand side in this expression *a priori* depends on the local coordinate system $\{y^i\}$ while the left-hand side denotes a function defined globally on \mathbb{R}^N . For a function v , the covariant derivative reduces to $\nabla_i v = \partial_i v$.

In a local coordinate system, we thus see that (2.1) can be written

$$Lv = f^i \partial_i v + \sum_{\ell=1}^D \frac{1}{2} \sigma_{[\ell]}^i \partial_i (\sigma_{[\ell]}^j \partial_j v).$$

and

$$\begin{aligned} L^T v &= -\partial_i (f^i v) - \Gamma_{im}^i f^m v \\ &+ \frac{1}{2} \sum_{\ell=1}^D \partial_i (\sigma_{[\ell]}^i \partial_j (\sigma_{[\ell]}^j v)) + \frac{1}{2} \sum_{\ell=1}^D \partial_i (\sigma_{[\ell]}^i \Gamma_{jk}^j \sigma_{[\ell]}^k v) \\ &+ \frac{1}{2} \sum_{\ell=1}^D \Gamma_{im}^i \sigma_{[\ell]}^m \partial_j (\sigma_{[\ell]}^j v) + \frac{1}{2} \sum_{\ell=1}^D \Gamma_{im}^i \sigma_{[\ell]}^m \Gamma_{jk}^j \sigma_{[\ell]}^k v. \end{aligned} \quad (2.3)$$

(Recall that for Euclidean coordinates, all the Christoffel symbols vanish).

We introduce now an orthogonal coordinate system that will reflect the foliation of Ω^ε by the isosurfaces Σ_z defined by (1.5).

In the following, we make the following assumption on H :

Hypothesis 1 *The function H is smooth, and there exists $\varepsilon > 0$ such that for all $z \in (-\varepsilon, \varepsilon)$, Σ_z is compact, and for all $x \in \Sigma_z$, $\nabla H(x) \neq 0$. This implies that for $z \in (-\varepsilon, \varepsilon)$, Σ_z is a smooth submanifold of \mathbb{R}^N of dimension $N - 1$.*

Note that $\Omega^\varepsilon := \bigcup_{z \in (-\varepsilon, \varepsilon)} \Sigma_z$. We then denote by $F_s(x)$ the flow of the ordinary differential equation

$$\frac{d}{ds} F_s(x) = \frac{\nabla H(F_s(x))}{\|\nabla H(F_s(x))\|^2}, \quad F_0(x) = x \in \mathbb{R}^N, \quad (2.4)$$

where $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^N . Up to a possible reduction of ε , Hypothesis 1 implies that for all $x \in \Omega^\varepsilon$, $F_s(x)$ is well defined for $|s| < \varepsilon$.

Let (U, φ) be a local chart on Σ_0 with $U \subset \Sigma_0$. The application $\varphi : \Sigma_0 \supset U \rightarrow \varphi(U) \subset \mathbb{R}^{N-1}$ induces local coordinates $(y^\alpha)_{\alpha=1}^{N-1} \in \varphi(U) \subset \mathbb{R}^{N-1}$. In the following, Greek indices will refer to $N - 1$ dimensional indices running from 1 to $N - 1$. We define the application

$$\varphi(U) \times (-\varepsilon, \varepsilon) \ni (y^\alpha, z) \mapsto \Phi(y^\alpha, z) := F_z \circ \varphi^{-1}(y^\alpha). \quad (2.5)$$

This is a local coordinate system over the domain Ω^ε . By construction, we have

$$\begin{aligned} \partial_z [H(\Phi(y^\alpha, z))] &= \langle \nabla H(\Phi(y^\alpha, z)), \partial_z \Phi(y^\alpha, z) \rangle \\ &= \langle \nabla H(\Phi(y^\alpha, z)), \nabla H(\Phi(y^\alpha, z)) \rangle / \|\nabla H(\Phi(y^\alpha, z))\|^2 \\ &= 1 \end{aligned}$$

whence for all $z \in (-\varepsilon, \varepsilon)$ and all $y^\alpha \in U$,

$$H(\Phi(y^\alpha, z)) = z. \quad (2.6)$$

Hence, the inverse application of (2.5) is the map

$$\Omega^\varepsilon \ni x \mapsto (\varphi(F_{-H(x)}(x)), H(x)) \in \varphi(U) \times (-\varepsilon, \varepsilon) \quad (2.7)$$

and it is clear that to an atlas on Σ_0 corresponds an atlas on Ω^ε made of the charts (2.7).

For given $(y^\alpha, z) \in \varphi(U) \times (-\varepsilon, \varepsilon)$, we denote for simplicity by (Y_α, Y_z) the local basis of \mathbb{R}^N corresponding to the local basis $(\frac{\partial}{\partial y^\alpha}, \frac{\partial}{\partial z})$ and we identify the vector fields Y_i , $i = \alpha, z$ with the vectors of \mathbb{R}^N , $Y_\alpha = \partial_\alpha \Phi$ and $Y_z = \partial_z \Phi$.

The metric tensor $g_{ij}(y^\alpha, z)$ splits into the “surfacic” components $g_{\alpha\beta}$ and the “normal” components $g_{\alpha z}$ and g_{zz} . By definition we have

$$\begin{aligned} g_{zz}(y^\alpha, z) &= \langle Y_z(y^\alpha, z), Y_z(y^\alpha, z) \rangle \\ &= \langle \partial_z \Phi(y^\alpha, z), \partial_z \Phi(y^\alpha, z) \rangle \\ &= 1 / \|\nabla H(\Phi(y^\alpha, z))\|^2. \end{aligned} \tag{2.8}$$

From (2.6), we deduce that for all β ,

$$\langle \nabla H(\Phi(y^\alpha, z)), \partial_\beta \Phi(y^\alpha, z) \rangle = 0.$$

Dividing this expression by $\|\nabla H(\Phi(y^\alpha, z))\|^2$, we obtain

$$\langle \partial_z \Phi(y^\alpha, z), \partial_\beta \Phi(y^\alpha, z) \rangle = 0.$$

which means that

$$g_{\beta z} = \langle Y_\beta, Y_z \rangle = 0.$$

This expression justifies the name “orthogonal coordinate system” on Ω^ε .

Finally, we see that

$$g_{\alpha\beta}(y^\alpha, z) = \langle Y_\alpha, Y_\beta \rangle = \langle \partial_\alpha \Phi, \partial_\beta \Phi \rangle$$

expands as

$$g_{\alpha\beta}(y^\alpha, z) = a_{\alpha\beta} + \sum_{k \geq 1} \frac{z^k}{k!} \partial_z^k g_{\alpha\beta}(y^\alpha, 0) \tag{2.9}$$

where $a_{\alpha\beta} = \langle \partial_\alpha \varphi^{-1}, \partial_\beta \varphi^{-1} \rangle$ is the metric tensor on Σ_0 (the expansion in z comes from the analyticity in z of the solution $F_z(x)$ of (2.4)).

Using this orthogonal coordinate system, it is easy to show that each tensor field on Ω^ε can be decomposed into several tensor fields on Σ_0 depending smoothly on $z \in (-\varepsilon, \varepsilon)$. For instance, if u^i is a vector field on Ω^ε , it can be represented by a couple $(u^\alpha(\mathbf{y}, z), u^z(\mathbf{y}, z)) \in \mathcal{C}^\infty((-\varepsilon, \varepsilon), \Gamma(T^1\Sigma_0) \times \mathcal{C}^\infty(\Sigma_0))$ where $\Gamma(T^1\Sigma_0)$ denotes the space of vector fields on U and where $\mathbf{y} = (y^\alpha)$ denotes a point in Σ_0 . The proof of this result relies on the special choice of the atlas defined by (2.7) and this situation is very similar to the case of shells (see [9, 8]).

In the orthogonal coordinate system defined above, the local charts (2.7) define local coordinates on open subset of Ω^ε . We denote by

$$dy^i(\mathbf{y}, z) := (dy^\alpha(\mathbf{y}, z), dz(\mathbf{y}, z))$$

the dual basis of the vector field basis $Y_i(\mathbf{y}, z) := (Y_\alpha(\mathbf{y}, z), Y_z(\mathbf{y}, z))$ defined above (we identify here y^N with z). Hence, local basis for tensor fields on Σ_0 can be expressed as tensor products of $dy^\alpha(\mathbf{y}, 0)$ and $Y_\alpha(\mathbf{y}, 0)$.

As mentioned above, the components of Christoffel symbols do not define a tensor field on a given manifold. However, using the fact that the metric tensor satisfies $g_{\alpha z} = 0$ in orthogonal coordinates, we can easily show the following result:

Lemma 2.1 *Let $(\mathbf{y}, z) = (y^\alpha, z)$ be a local orthogonal coordinate system induces by a local chart (U, φ) on Σ_0 and let Γ_{ij}^k be the Christoffel symbols (2.2). Then the expressions*

$$\begin{aligned} & \Gamma_{z\alpha}^\beta(\mathbf{y}, z)Y_\beta(0) \otimes dy^\alpha(\mathbf{y}, 0) \\ & \Gamma_{\alpha\beta}^z(\mathbf{y}, z)dy^\beta(\mathbf{y}, 0) \otimes dy^\alpha(\mathbf{y}, 0) \\ & \Gamma_{\alpha z}^z(\mathbf{y}, z)dy^\alpha(\mathbf{y}, 0) \\ & \Gamma_{zz}^\alpha(\mathbf{y}, z)Y_\alpha(\mathbf{y}, 0) \end{aligned}$$

all define tensor fields on Σ_0 depending smoothly on z .

We do not give a proof here, as it is very similar to the situation of shells (see [9, 8] for similar statements). Roughly speaking, it states that freezing one or two coefficients of the Christoffel symbol to z in orthogonal coordinates yields tensor fields.

Let u^i be a smooth vector field on Ω^ε . The covariant derivative of u^i in a local coordinate system is given by

$$\nabla_i u^j = \partial_i u^j + \Gamma_{ik}^j u^k$$

and defines a tensor field on Ω^ε .

In a local orthogonal basis, this tensor field can be decomposed into four parts:

$$\nabla_z u^z = \partial_z u^z + \Gamma_{zz}^z(\mathbf{y}, z)u^z \quad (2.10)$$

which defines a function in $\mathcal{C}^\infty((-\varepsilon, \varepsilon), \mathcal{C}^\infty(\Sigma_0))$, and similarly

$$\nabla_z u^\alpha = \partial_z u^\alpha + \Gamma_{z\beta}^\alpha(\mathbf{y}, z)u^\beta + \Gamma_{zz}^\alpha(\mathbf{y}, z)u^z. \quad (2.11)$$

which defines an element of $\mathcal{C}^\infty((-\varepsilon, \varepsilon), \Gamma(T^1\Sigma_0))$ after using the previous Lemma.

In the following, we denote by D_α the covariant derivative on Σ_0 , associated with the metric $a_{\alpha\beta}$. As $\partial_\alpha u^z = D_\alpha u^z$ because u^z is a scalar function, we obtain that the expression

$$\nabla_\alpha u^z = D_\alpha u^z + \Gamma_{\alpha\beta}^z(\mathbf{y}, z)u^\beta + \Gamma_{\alpha z}^z(\mathbf{y}, z)u^z \quad (2.12)$$

defines an element of $\mathcal{C}^\infty((-\varepsilon, \varepsilon), \Gamma(T_1\Sigma_0))$, where $\Gamma(T_1\Sigma_0)$ denotes the space of 1-form fields on Σ_0 .

Finally, we write

$$\nabla_\alpha u^\beta = \partial_\alpha u^\beta + \Gamma_{\alpha\sigma}^\beta(\mathbf{y}, z)u^\sigma + \Gamma_{\alpha z}^\beta(\mathbf{y}, z)u^z. \quad (2.13)$$

But using the expansion (2.9) it can be shown (see [9]) that the expression

$$\Lambda_{\alpha\beta}^{\sigma}(\mathbf{y}, z) = \Gamma_{\alpha\beta}^{\sigma}(\mathbf{y}, z) - \Gamma_{\alpha\beta}^{\sigma}(\mathbf{y}, 0)$$

defines a tensor field on Σ_0 . Note that $\Gamma_{\alpha\beta}^{\sigma}(\mathbf{y}, 0)$ are the Christoffel symbols on Σ_0 associated with the metric $a_{\alpha\beta}$. Consequently, we can write (2.13) as

$$\nabla_{\alpha} u^{\beta} = D_{\alpha} u^{\beta} + \Lambda_{\alpha\sigma}^{\beta}(\mathbf{y}, z) u^{\sigma} + \Gamma_{\alpha z}^{\beta}(\mathbf{y}, z) u^z, \quad (2.14)$$

which now makes sense in $\mathcal{C}^{\infty}((-\varepsilon, \varepsilon), \Gamma(T_1^1 \Sigma_0))$ where $\Gamma(T_1^1 \Sigma_0)$ denotes the space of tensor fields of type (1, 1) on Σ_0 . Note that $\Lambda_{\alpha\beta}^{\sigma}(\mathbf{y}, 0) = 0$ and that $\Lambda_{\alpha\beta}^{\sigma}(\mathbf{y}, z)$ is analytic in z .

2.2 Existence of solutions

We express now the operators L and L^T in orthogonal coordinate system. Let f^i be a vector field on Ω^{ε} such that $\langle \nabla H, f \rangle = 0$. Written in an orthogonal coordinate system, this implies that $f^z = 0$. In the following, we write $f^{\beta}(\mathbf{y}, z)$ the tangential part of the vector field f viewed as a vector field on the manifold Σ_0 depending smoothly on z . We hence have

$$f^i \nabla_i v = f^{\alpha}(\mathbf{y}, z) D_{\alpha} v(\mathbf{y}, z).$$

Using the equations (2.10) and (2.14), we find

$$\begin{aligned} \nabla_j (f^j v) &= \nabla_{\alpha} (f^{\alpha} v) + \nabla_z (f^z v) \\ &= D_{\alpha} (f^{\alpha} v) + \Lambda_{\alpha\sigma}^{\alpha} (f^{\sigma} v) \end{aligned} \quad (2.15)$$

where all the tensors are evaluated at the point $(\mathbf{y}, z) \in \Sigma_0 \times (-\varepsilon, \varepsilon)$. Similarly, we can write

$$\nabla_i (f^i \nabla_j (f^j v)) = D_{\alpha} (f^{\alpha} \nabla_j (f^j v)) + \Lambda_{\alpha\sigma}^{\alpha} (f^{\sigma} \nabla_j (f^j v))$$

whence using (2.15)

$$\begin{aligned} \nabla_i (f^i \nabla_j (f^j v)) &= D_{\alpha} (f^{\alpha} D_{\beta} (f^{\beta} v)) + D_{\alpha} (f^{\alpha} \Lambda_{\beta\sigma}^{\beta} (f^{\sigma} v)) \\ &\quad + \Lambda_{\alpha\sigma}^{\alpha} (f^{\sigma} D_{\beta} (f^{\beta} v)) + \Lambda_{\alpha\sigma}^{\alpha} (f^{\sigma} \Lambda_{\beta\sigma}^{\beta} (f^{\sigma} v)). \end{aligned}$$

In the following, we denote by \mathbf{D} the intrinsic operator with components D_{α} in a local coordinate system. Using the previous expression, we can easily show the following result:

Theorem 2.2 *In orthogonal coordinates, the Kolmogorov operator L viewed as an operator on the manifold $\Sigma_0 \times (-\varepsilon, \varepsilon)$ writes*

$$L(\mathbf{y}, z; \mathbf{D})v = f^\alpha(\mathbf{y}, z)D_\alpha v + \frac{1}{2} \sum_{\ell=1}^D \sigma_{[\ell]}^\alpha(\mathbf{y}, z)D_\alpha(\sigma_{[\ell]}^\beta(\mathbf{y}, z)D_\beta v)$$

where $\sigma_{[\ell]}^\alpha(\mathbf{y}, z)$ and $f^\alpha(\mathbf{y}, z)$ are the components of the vector fields $\sigma_{[\ell]}$ and f in orthogonal coordinates. Similarly, the Fokker-Planck operator L^T , writes in orthogonal coordinates

$$L^T(\mathbf{y}, z; \mathbf{D})v = \frac{1}{2} \sum_{\ell=1}^D D_\alpha(\sigma_{[\ell]}^\alpha(\mathbf{y}, z)D_\beta(\sigma_{[\ell]}^\beta(\mathbf{y}, z)v)) - D_\alpha(f^\alpha(\mathbf{y}, z)v) + P(\mathbf{y}, z; \mathbf{D})v$$

$P(\mathbf{y}, z; \mathbf{D})$ is an intrinsic operator of order 1 on Σ_0 depending analytically of z and such that

$$P(\mathbf{y}, 0; \mathbf{D}) = 0.$$

In the rest of this work, we make the following hypothesis:

Hypothesis 2 *For all one-form field $\xi_\alpha \in \Gamma(T_1^*\Sigma_0)$ and all $(\mathbf{y}, z) \in \Sigma_0 \times (-\varepsilon, \varepsilon)$,*

$$\sum_{\ell=1}^D \sigma_{[\ell]}^\alpha(\mathbf{y}, z)\sigma_{[\ell]}^\beta(\mathbf{y}, z)\xi_\alpha\xi_\beta \geq \Lambda \xi^\alpha\xi_\alpha \quad (2.16)$$

for some Λ independent on z .

This means that $L(\mathbf{y}, z; \mathbf{D})$ is elliptic on Σ_0 , uniformly in $z \in (-\varepsilon, \varepsilon)$.

For a function φ on Σ_0 , and $\ell \geq 0$, we denote by

$$|\mathbf{D}^\ell \varphi|^2 = D^{\alpha_1}D^{\alpha_2} \dots D^{\alpha_\ell} \varphi D_{\alpha_1}D_{\alpha_2} \dots D_{\alpha_\ell} \varphi$$

the semi norm of order ℓ of φ . Note that the contravariant indices denote a multiplication by the inverse of the metric tensor $a^{\alpha\beta}$ on Σ_0 : $D^\alpha = a^{\alpha\beta}D_\beta$.

We recall that for $k \geq 0$ the Sobolev space $H^k(\Sigma_0)$ on Σ_0 is defined as the space of functions $\varphi : \Sigma_0 \rightarrow \mathbb{R}$ such that for all ℓ , $0 \leq \ell \leq k$,

$$\|\mathbf{D}^\ell \varphi\|_{L^2} = \int_{\Sigma_0} |\mathbf{D}^\ell \varphi| d\Sigma_0 < \infty.$$

The corresponding norm is written

$$\|\varphi\|_{H^k} = \sum_{\ell=0}^k \|\mathbf{D}^\ell \varphi\|_{L^2}.$$

On the manifold Σ_0 , the covariant derivative operator commutes with the contraction by the metric tensor: $D^\alpha \varphi = a^{\alpha\beta} D_\beta \varphi = D_\beta a^{\alpha\beta} \varphi$. However, the covariant derivative does not commute with itself. For a given one-form fields u_α on Σ_0 , we have for instance

$$D_\alpha D_\beta u_\nu - D_\beta D_\alpha u_\nu = R_{\alpha\beta\delta\nu} u^\delta \quad (2.17)$$

where $u^\delta = a^{\delta\sigma} u_\sigma$ is the contravariant vector-field corresponding to u_α , and where $R_{\alpha\beta\delta\nu}$ is the *curvature* tensor on Σ_0 depending only on the metric tensor $a_{\alpha\beta}$ (see for instance [4]).

Theorem 2.3 *Let $v_0 \in \mathcal{C}^\infty(\Sigma_0 \times (-\varepsilon, \varepsilon))$. Under the hypothesis 2, the equation*

$$\partial_t v(t, \mathbf{y}, z) = L(\mathbf{y}, z; \mathbf{D})v(t, \mathbf{y}, z), \quad v(0) = v_0 \quad (2.18)$$

possesses a unique solution $v(t, \mathbf{y}, z) \in \mathcal{C}^\infty((0, +\infty) \times \Sigma_0 \times (-\varepsilon, \varepsilon))$.

Proof. The result is a consequence of a priori bounds for the solution of (2.18). Let $(\cdot, \cdot)_{L^2}$ denote the L^2 product on Σ_0 . If v is a solution, we can write

$$\begin{aligned} \frac{1}{2} \partial_t \|v\|_{L^2}^2 &= (L(\mathbf{y}, z; \mathbf{D})v, v)_{L^2} \\ &= -\frac{1}{2} \sum_{\ell=1}^D \sigma_{[\ell]}^\alpha(\mathbf{y}, z) \sigma_{[\ell]}^\beta(\mathbf{y}, z) D_\alpha v D_\beta v + (Q(\mathbf{y}, z; \mathbf{D})v, v)_{L^2} \end{aligned}$$

where $Q(\mathbf{y}, z; \mathbf{D})$ is an operator of order 1 on Σ_0 depending smoothly on z . This shows that for constants C , c and b depending on ε , we have

$$\frac{1}{2} \partial_t \|v\|_{L^2}^2 \leq -C \|\mathbf{D}v\|_{L^2}^2 + c \|\mathbf{D}v\|_{L^2} \|v\|_{L^2} + b \|v\|_{L^2}^2,$$

whence

$$\partial_t \|v\|_{L^2}^2 \leq M \|v\|_{L^2}^2,$$

for a constant M depending on ε . This shows the a priori estimate

$$\|v\|_{L^2} \leq e^{tM} \|v_0\|_{L^2}.$$

Using (2.17), we obtain similar estimates for the derivatives $\partial_z^k \mathbf{D}^\ell v$ which satisfy equations of the form

$$\partial_t \partial_z^k \mathbf{D}^\ell v = L(\mathbf{y}, z; \mathbf{D}) \partial_z^k \mathbf{D}^\ell v + R(\mathbf{y}, z; \mathbf{D}, \partial_z)v$$

where R is an operator of order $k + \ell - 1$ in ∂_z and D_α . We hence easily obtain a priori estimates for these derivatives by induction. The uniqueness is clear, as the equation (2.18) is linear. \blacksquare

2.3 Invariant measure and exponential convergence

Throughout the rest of this section, we assume moreover that there exists a smooth function $\rho(x) > 0$ defined on Ω^ε such that for all $x \in \Omega^\varepsilon$, we have

$$L^T \rho = 0. \quad (2.19)$$

Note that in the case where $\rho \equiv \mathbb{1}$ the constant function equals to 1 on Ω^ε , then (2.19) expresses that the equation (1.1) preserves the volume. In general, this hypothesis states that ρ is an invariant measure for (1.1). Indeed, for $x \in \Omega^\varepsilon$, let $X(t, x)$ denote the solution of (1.1) starting at x for $t = 0$. Using (1.4), we have $X(t, x) \in \Omega^\varepsilon$ for all time $t \geq 0$. Then for all function $\varphi(x)$ with compact support in Ω^ε , we have using the Kolmogorov equation

$$\begin{aligned} \frac{d}{dt} \int_{\Omega^\varepsilon} \mathbb{E}(\varphi(X(t, x))\rho(x) dx &= \langle L(x^i; \partial_i)\mathbb{E}(\varphi(X(t, x)), \rho(x)) \rangle_{L^2(\Omega^\varepsilon)} \\ &= \langle \mathbb{E}(\varphi(X(t, x)), L^T(x^i; \partial_i)\rho(x)) \rangle_{L^2(\Omega^\varepsilon)} \\ &= 0 \end{aligned}$$

where $\langle \cdot, \cdot \rangle_{L^2(\Omega^\varepsilon)}$ denotes the L^2 product in Cartesian coordinates on Ω^ε , and where $L^T(x^i; \partial_i)$ denotes the adjoint of the operator $L(x^i; \partial_i)$ given by (1.7) in Cartesian coordinates. In a normal orthogonal coordinate system (y^α, z) , the volume form dx is transformed into the measure $\sqrt{|g_{ij}(\mathbf{y}, z)|} dy^\alpha \wedge dz$ where $|g_{ij}(z)|$ denotes the determinant of the metric tensor $g_{ij}(\mathbf{y}, z)$. Using the equations (2.8) and (2.9), this measure can be written

$$\frac{1}{\|\nabla H\|} d\Sigma_z \wedge dz \quad (2.20)$$

where $d\Sigma_z$ is the volume form $\sqrt{|g_{\alpha\beta}(\mathbf{y}, z)|} dy^\alpha$ induced by the Euclidean measure in \mathbb{R}^N on the surface Σ_z . Hence, we can write for all $x = (\mathbf{y}, z) \in \Sigma_0 \times (-\varepsilon, \varepsilon)$,

$$\rho(x) dx = \mu(\mathbf{y}, z) d\Sigma_0 \wedge dz$$

where

$$\mu(\mathbf{y}, z) = \frac{\rho(\mathbf{y}, z)}{\|\nabla H(\mathbf{y}, z)\|} \sqrt{\frac{|g_{\alpha\beta}(\mathbf{y}, z)|}{|g_{\alpha\beta}(\mathbf{y}, 0)|}} \quad (2.21)$$

defines a density function on Σ_0 depending smoothly on z . In the sequel, we will often ignore the dependency in $\mathbf{y} \in \Sigma_0$, and write $\mu(z)$ for $\mu(\mathbf{y}, z)$. Shortly, we can write

$$\rho(x) dx = d\mu(z) \wedge dz$$

where $d\mu(z) = \mu(z)d\Sigma_0$. Note that in the case where $\rho \equiv \mathbb{1}$, then up to a multiplication by a constant, the measure $d\mu(0)$ is the microcanonical measure on Σ_0 (see (1.12)).

Proposition 2.4 *For all $(\mathbf{y}, z) \in \Sigma_0 \times (-\varepsilon, \varepsilon)$, let $L(\mathbf{y}, z; \mathbf{D})^*$ denote the adjoint of $L(\mathbf{y}, z; \mathbf{D})$ with respect to the L^2 product on the hypersurface Σ_0 . Let $\mu(\mathbf{y}, z)$ the function defined by (2.21). Then for all $(\mathbf{y}, z) \in \Sigma_0 \times (-\varepsilon, \varepsilon)$, we have*

$$L(\mathbf{y}, z; \mathbf{D})^* \mu(\mathbf{y}, z) = 0. \quad (2.22)$$

Remark 2.5 Before giving the proof, let us notice that the operators L^* (the adjoint in $L^2(\Sigma_0)$) and L^T (the adjoint in $L^2(\Omega^\varepsilon)$) do differ in general. ■

Proof. Let φ be a function defined on $\Omega^\varepsilon \simeq \Sigma_0 \times (-\varepsilon, \varepsilon)$. Let ρ be the function satisfying (2.19). We have in Cartesian coordinates

$$\langle \rho, L(x^i; \partial_i)\varphi \rangle_{L^2(\Omega^\varepsilon)} = \langle L^T(x^i; \partial_i)\rho, \varphi \rangle_{L^2(\Omega^\varepsilon)} = 0.$$

Written in orthogonal coordinates $(\mathbf{y}, z) = (x^i)$, this equation reads

$$\int_{\Sigma_0 \times (-\varepsilon, \varepsilon)} (L(\mathbf{y}, z; \mathbf{D})\varphi(\mathbf{y}, z))\mu(\mathbf{y}, z)d\Sigma_0 dz = 0.$$

We take now $\psi(\mathbf{y}, z) = v(\mathbf{y})\chi(z)$ where v is a function on Σ_0 and $\chi(z)$ a function on $(-\varepsilon, \varepsilon)$. We then have using Theorem 2.2

$$L(\mathbf{y}, z; \mathbf{D})\varphi(\mathbf{y}, z) = \chi(z)L(\mathbf{y}, z; \mathbf{D})v(\mathbf{y}).$$

Ignoring the argument \mathbf{y} in the equation, we get

$$\int_{-\varepsilon}^{\varepsilon} \chi(z)dz \int_{\Sigma_0} \mu(z)(L(z; \mathbf{D})v)d\Sigma_0 = 0.$$

As $\chi(z)$ is an arbitrary function of $z \in (-\varepsilon, \varepsilon)$, this implies that for all $z \in (-\varepsilon, \varepsilon)$ and all $v \in C^\infty(\Sigma_0)$,

$$\int_{\Sigma_0} \mu(z)(L(z; \mathbf{D})v)d\Sigma_0 = \int_{\Sigma_0} v(L(z; \mathbf{D})^*\mu(z))d\Sigma_0 = 0.$$

This yields the result. ■

Lemma 2.6 Let $v(t, \mathbf{y}, z)$ be a solution of the equation $\partial_t v = L(\mathbf{y}, z; \mathbf{D})v$. For all $z \in (-\varepsilon, \varepsilon)$, and all $t \geq 0$, we have that

$$\partial_t \int_{\Sigma_0} v(t, \mathbf{y}, z)^2 d\mu(\mathbf{y}, z) = - \sum_{\ell=1}^D \int_{\Sigma_0} \left(\sigma_{[\ell]}^\alpha(\mathbf{y}, z) D_\alpha v(t, \mathbf{y}, z) \right)^2 d\mu(\mathbf{y}, z). \quad (2.23)$$

Proof. For all smooth function v defined on $\Sigma_0 \times (-\varepsilon, \varepsilon)$, we have on the manifold $\Sigma_0 \times (-\varepsilon, \varepsilon)$ (ignoring again the argument $\mathbf{y} \in \Sigma_0$),

$$L(z; \mathbf{D})v^2 = 2v f^\alpha(z) D_\alpha v + \sum_{\ell=1}^D \sigma_{[\ell]}^\alpha(z) D_\alpha (v \sigma_{[\ell]}^\beta(z) D_\beta v)$$

which implies that

$$L(z; \mathbf{D})v^2 = 2vL(z; \mathbf{D})v + \sum_{\ell=1}^D \left(\sigma_{[\ell]}^\alpha(z) D_\alpha v \right)^2.$$

Now if v is a solution of $\partial_t v = Lv$, we have

$$\begin{aligned}\partial_t \int_{\Sigma_0} v^2 d\mu(z) &= 2\langle vL(z; \mathbf{D})v, \mu(z) \rangle \\ &= \langle L(z; \mathbf{D})v^2, \mu(z) \rangle - \sum_{\ell=1}^D \int_{\Sigma_0} \left(\sigma_{[\ell]}^\alpha(z) D_\alpha v \right)^2 d\mu(z),\end{aligned}$$

and the result follows from $L(z; \mathbf{D})^* \mu(z) = 0$. \blacksquare

In the following, we denote by $\|v\|_{\mu(z)}$ the weighted norm

$$\|v\|_{\mu(z)}^2 = \int_{\Sigma_0} v^2 d\mu(z). \quad (2.24)$$

Note that the norms $\|\cdot\|_{L^2(\Sigma_0)}$ and $\|\cdot\|_{\mu(z)}$ are not the same. However, as ρ is a positive function, Σ_0 is compact and $z \in (-\varepsilon, \varepsilon)$, there exist two positive constants c and C such that for all $z \in (-\varepsilon, \varepsilon)$,

$$c \leq \mu(z) \leq C \quad (2.25)$$

on Σ_0 and hence the norms $\|\cdot\|_{L^2(\Sigma_0)}$ and $\|\cdot\|_{\mu(z)}$ are in fact equivalent uniformly in z .

Lemma 2.7 *Under the hypothesis 2, there exists a constant λ_0 depending on Σ_0 and ε such that for all function $v(z) \in \mathcal{C}^0((-\varepsilon, \varepsilon), H^1(\Sigma_0))$ satisfying*

$$\forall z \in (-\varepsilon, \varepsilon), \quad \int_{\Sigma_0} v(z) d\mu(z) = 0, \quad (2.26)$$

then we have

$$\forall z \in (-\varepsilon, \varepsilon), \quad \lambda_0 \|v(z)\|_{\mu(z)}^2 \leq \sum_{\ell=1}^D \|\sigma_{[\ell]}^\alpha(z) D_\alpha v(z)\|_{\mu(z)}^2. \quad (2.27)$$

Proof. As Σ_0 is a compact manifold, there exists a constant $a > 0$ such that for all function $\varphi \in H^1(\Sigma_0)$ satisfying

$$\langle \varphi \rangle_0 := \int_{\Sigma_0} \varphi d\Sigma_0 = 0, \quad (2.28)$$

we have the Poincaré-Wirtinger inequality (see [4])

$$a \|\varphi\|_{L^2(\Sigma_0)}^2 \leq \|\mathbf{D}\varphi\|_{L^2(\Sigma_0)}^2.$$

Using (2.16) and (2.25), we have for all such function φ

$$\sum_{\ell=1}^D \|\sigma_{[\ell]}^\alpha(z) D_\alpha \varphi\|_{\mu(z)}^2 \geq \Lambda \|\mathbf{D}\varphi\|_{\mu(z)}^2 \geq c\Lambda \|\mathbf{D}\varphi\|_{L^2(\Sigma_0)}^2 \quad (2.29)$$

and hence using again (2.25), we see that there exists a constant $b > 0$ such that for all $z \in (-\varepsilon, \varepsilon)$, and all φ satisfying (2.28),

$$b\|\varphi\|_{\mu(z)}^2 \leq \sum_{\ell=1}^D \|\sigma_{[\ell]}^\alpha(z) D_\alpha \varphi\|_{\mu(z)}^2.$$

Now if $v(z) \in C^0((-\varepsilon, \varepsilon), H^1(\Sigma_0))$, satisfies (2.26), we can write for all $z \in (-\varepsilon, \varepsilon)$,

$$b\|v(z) - \langle v(z) \rangle_0\|_{\mu(z)}^2 \leq \sum_{\ell=1}^D \|\sigma_{[\ell]}^\alpha(z) D_\alpha v(z)\|_{\mu(z)}^2. \quad (2.30)$$

where $\langle v(z) \rangle_0$ denotes the average (2.28) of $v(z)$. Let $z \in (-\varepsilon, \varepsilon)$, and assume that ψ satisfies $\int_{\Sigma_0} \psi d\mu(z) = 0$. We have that

$$\begin{aligned} \|\psi\|_{\mu(z)}^2 &= \|\psi - \langle \psi \rangle_0 + \langle \psi \rangle_0\|_{\mu(z)}^2 \\ &= \|\psi - \langle \psi \rangle_0\|_{\mu(z)}^2 + 2 \int_{\Sigma_0} (\psi - \langle \psi \rangle_0) \langle \psi \rangle_0 d\mu(z) + \langle \psi \rangle_0^2 \int_{\Sigma_0} d\mu(z) \\ &= \|\psi - \langle \psi \rangle_0\|_{\mu(z)}^2 - \langle \psi \rangle_0^2 \int_{\Sigma_0} d\mu(z) \leq \|\psi - \langle \psi \rangle_0\|_{\mu(z)}^2. \end{aligned}$$

Applying this formula to $\psi = v(z)$ and combining with (2.30) then yield the result with $\lambda_0 = b$. \blacksquare

Theorem 2.8 *Let $v(t, \mathbf{y}, z)$ be a solution of the equation $\partial_t v = Lv$ with initial value $v_0(z) \in C^\infty((-\varepsilon, \varepsilon) \times \Sigma_0)$. Let $\mu(z)$ the function defined in (2.21) and $\|\cdot\|_{\mu(z)}$ the corresponding weighed L^2 norm. Under the assumption (2.16), for all $z \in (-\varepsilon, \varepsilon)$, and all $t \geq 0$, we have that*

$$\|v(t, z) - \int v(0, z) d\mu(z)\|_{\mu(z)} \leq \|v(0, z) - \int v(0, z) d\mu(z)\|_{\mu(z)} \exp(-\gamma_0 t) \quad (2.31)$$

where $\gamma_0 = \lambda_0/2$, λ_0 being the constant appearing in (2.27). Moreover, for all $J \in \mathbb{N}$, there exist constants γ_J and C_J such that for all $z \in (-\varepsilon, \varepsilon)$,

$$\|\mathbf{D}^J v(t, z)\|_{\mu(z)} \leq C_J \exp(-\gamma_J t) \quad (2.32)$$

Proof. We have using (2.22)

$$\partial_t \int_{\Sigma_0} v(t, z) d\mu(z) = (L(z; \mathbf{D})v(t, z), \mu(z))_{L^2} = (v(t, z), L(z; \mathbf{D})^* \mu(z))_{L^2} = 0.$$

This shows that the average of $v(t, z)$ with respect to the measure $d\mu(z)$ is constant with respect to t . Now we easily see that $L(z; \mathbf{D})\mathbf{1} = 0$. Hence the function $w(t, z) = v(t, z) - \int v(t, z) d\mu(z)$ satisfies the equation $\partial_t w = L(z; \mathbf{D})w$. The equation (2.31) is then an easy consequence of the two previous Lemmas.

To prove (2.32), let us consider first the case where $J = 1$.
From Lemma 2.6 and (2.16), we have

$$\partial_t \|w(t, z)\|_{\mu(z)}^2 \leq -\Lambda \|\mathbf{D}w(t, z)\|_{\mu(z)}^2.$$

Multiplying this equation by $e^{\delta t}$ with $\delta > 0$ and integrating from 0 to T , we obtain

$$\int_0^T e^{\delta t} \partial_t \|w(t, z)\|_{\mu(z)}^2 dt + \Lambda \int_0^T e^{\delta t} \|\mathbf{D}w(t, z)\|_{\mu(z)}^2 dt \leq 0.$$

After integration by part, this shows that

$$\begin{aligned} e^{\delta T} \|w(T, z)\|_{\mu(z)}^2 + \Lambda \int_0^T e^{\delta t} \|\mathbf{D}w(t, z)\|_{\mu(z)}^2 dt \\ \leq \|w(0, z)\|_{\mu(z)}^2 + \delta \int_0^T e^{\delta t} \|w(t, z)\|_{\mu(z)}^2 dt. \end{aligned}$$

Using (2.31) and taking $\delta < \gamma_0$ shows that there exists a constant M such that for all $z \in (-\varepsilon, \varepsilon)$,

$$\int_0^\infty e^{\delta t} \|\mathbf{D}w(t, z)\|_{\mu(z)}^2 dt \leq M. \quad (2.33)$$

It is clear that the operator $L(z; \mathbf{D})$ acts on tensor fields, and we have

$$\mathbf{D}_\alpha L(z; \mathbf{D})w = L(z; \mathbf{D})(\mathbf{D}_\alpha w) + [\mathbf{D}_\alpha, L(z; \mathbf{D})]w$$

where $[\mathbf{D}_\alpha, L(z; \mathbf{D})]$ is an operator of order 1 on Σ_0 depending smoothly on z , and involving the curvature tensor on Σ_0 (see (2.17)). Hence $\mathbf{D}_\alpha w(t, v)$ satisfies the equation

$$\partial_t \mathbf{D}_\alpha w = L(z; \mathbf{D})(\mathbf{D}_\alpha w) + [\mathbf{D}_\alpha, L(z; \mathbf{D})]w.$$

Computations similar to those made in Lemma 2.6 show that there exist positive constants c_1, c_2 and c_3 independent on z such that

$$\partial_t \|\mathbf{D}w(t, z)\|_{\mu(z)}^2 \leq -c_1 \|\mathbf{D}^2 w(t, z)\|_{\mu(z)}^2 + c_2 \|\mathbf{D}w(t, z)\|_{\mu(z)}^2 + c_3 \|w(t, z)\|_{\mu(z)}^2.$$

Multiplying by $e^{\gamma_1 t}$ with $0 < \gamma_1 < \delta$ and integrating from 0 to T yields

$$\begin{aligned} \int_0^T e^{\gamma_1 t} \partial_t \|\mathbf{D}w(t, z)\|_{\mu(z)}^2 dt + c_1 \int_0^T e^{\gamma_1 t} \|\mathbf{D}^2 w(t, z)\|_{\mu(z)}^2 dt \\ \leq c_2 \int_0^T e^{\gamma_1 t} \|\mathbf{D}w(t, z)\|_{\mu(z)}^2 dt + c_3 \int_0^T e^{\gamma_1 t} \|w(t, z)\|_{\mu(z)}^2 dt \leq C_0 \end{aligned} \quad (2.34)$$

for some constant C_0 independent of T and z , where we used (2.33), (2.31) and the fact that $\gamma_1 < \delta < \gamma_0$. We thus have

$$e^{\gamma_1 T} \|\mathbf{D}w(T, z)\|_{\mu(z)}^2 \leq C_0 + \|\mathbf{D}w(0, z)\|_{\mu(z)}^2 + \gamma_1 \int_0^T e^{\gamma_1 t} \|\mathbf{D}w(t, z)\|_{\mu(z)}^2 dt$$

and using again (2.33) we conclude that

$$\|\mathbf{D}w(T, z)\|_{\mu(z)}^2 \leq C_1 \exp(-\gamma_1 T)$$

for a positive constant C_1 independent of T and z . This shows (2.32) for $J = 1$, as $\mathbf{D}v = \mathbf{D}w$. Moreover, we derive from (2.34) that

$$\int_0^\infty e^{\gamma_1 t} \|\mathbf{D}^2 w(t, z)\|_{\mu(z)}^2 dt \leq C_0.$$

The result is then easily shown by induction using computations similar to the previous ones, with the fact that for all $J \in \mathbb{N}$,

$$\mathbf{D}^J L(z; \mathbf{D}) = L(z; \mathbf{D})\mathbf{D}^J + [\mathbf{D}^J, L(z; \mathbf{D})]$$

where $[\mathbf{D}^J, L(z; \mathbf{D})]$ is an operator of order $J - 1$ on Σ_0 and depending smoothly on z . ■

As Σ_0 is a compact manifold, using (2.25) and Sobolev embedding Theorems (see [4]) we immediately get the following result:

Corollary 2.9 *Let $v(t, \mathbf{y}, z)$ be a solution of the equation $\partial_t v = Lv$ with initial value $v_0(\mathbf{y}, z) \in C^\infty(\Sigma_0 \times (-\varepsilon, \varepsilon))$. Under the assumption (2.16), for all $J \in \mathbb{N}$, there exists constants ν_J and M_J depending only on Σ_0 and ε such that for all $(\mathbf{y}, z) \in \Sigma_0 \times (-\varepsilon, \varepsilon)$, and all $t \geq 0$,*

$$|\mathbf{D}^J v(t, \mathbf{y}, z)| \leq M_J \exp(-\nu_J t) \quad (2.35)$$

Remark 2.10 The previous result give uniform bounds with respect to $z \in (-\varepsilon, \varepsilon)$. This fact is not necessary for the analysis of the properties of the *projected* numerical schemes defined below, which requires only the previous estimates on the manifold Σ_0 . However, we believe that these uniform bounds should be necessary to understand the good behaviour of non projected schemes, see for instance [16]. ■

3 Numerical analysis

We now consider the discretisation of the stochastic differential equation (1.1). Following [20], we define the sequence

$$\{U_p^{[\ell]}, \quad \ell = 1, \dots, D, \quad p \in \mathbb{N}\}$$

of i.i.d. random variables satisfying the following conditions

$$\forall p \in \mathbb{N}, \quad \forall \ell = 1, \dots, D, \quad \mathbb{E}(U_p^{[\ell]})^m = \begin{cases} 0 & \text{if } m = 1, 3 \text{ or } 5 \\ 1 & \text{if } m = 2 \end{cases} \quad (3.1)$$

and the condition

$$\mathbb{E}(U_p^{[\ell]})^4 < \infty \quad \text{and} \quad \mathbb{E}(U_p^{[\ell]})^6 < \infty.$$

Notice that these last two conditions are in particular satisfied if $|U_p^{[\ell]}| \leq C$ a.s. for a constant C independent of p and ℓ . Examples of such family of random variables can be found in [20].

Let $X_0 \in \Sigma_0$ and a time step $h \leq h_0$. We assume that we can define a family $\{X_p\}$ of random variables satisfying the following conditions:

(C1) $\forall p \in \mathbb{N}, \quad X_p \in \Sigma_0$

(C2) Let $X_p(\theta) = X_p + \theta(X_{p+1} - X_p), \quad \theta \in (0, 1)$. Then

$$\forall p \in \mathbb{N}, \quad \forall \theta \in (0, 1), \quad |H(X_p(\theta))| < \varepsilon.$$

(C3) In Cartesian coordinates we have

$$\begin{aligned} X_{p+1}^i &= X_p^i + \sqrt{h} \sum_{\ell=1}^D \sigma_{[\ell]}^i(X_p) U_p^{[\ell]} \\ &\quad + h \left(f^i(X_p) + \frac{1}{2} \sum_{\ell=1}^D \sigma_{[\ell]}^j(X_p) \partial_j \sigma_{[\ell]}^i(X_p) \right) \\ &\quad + h w_p^i + h^{3/2} r_p^i + h^2 s_p^i \end{aligned}$$

where for all p , if \mathcal{F}_p denotes the σ -algebra generated by the family $\{X_0, \dots, X_p\}$, we assume that w_p^i, r_p^i and s_p^i are \mathcal{F}_p measurable, and satisfy

- w_p^i is a *even* polynomial in $(U_p^{[1]}, \dots, U_p^{[D]})$ of degree less than 2 with coefficients depending on X_p , and such that $\mathbb{E}(w_p^i | X_p) = 0$.
- r_p^i is a *odd* polynomial in $(U_p^{[1]}, \dots, U_p^{[D]})$ of degree less than 3 with coefficients depending on X_p , and such that $\mathbb{E}(r_p^i | X_p) = 0$.
- $\mathbb{E}(|s_p^i| | X_p) < R$,

where R does not depend on $p \in \mathbb{N}$ and on $h \leq h_0$.

The condition **(C1)** will be fulfilled for projection schemes in general, while the condition **(C2)** is equivalent to the fact that the corresponding piecewise trajectory lies in the neighbourhood Ω^ε of Σ_0 : It will be fulfilled for h sufficiently small in the case where the random variables $U_p^{[\ell]}$ are uniformly a.s. bounded.

The condition **(C3)** expresses the fact that the scheme is consistent with (1.1). Note that this condition implies in particular that for a given $\ell \in \{1, \dots, D\}$, we have using (3.1)

$$\mathbb{E}(w_p^i U_p^{[\ell]} | X_p) = 0. \tag{3.2}$$

Note moreover, that the first three terms in the expression of $X_{p+1} - X_p$ define a vector in Ω^ε attached to the point $X_p \in \Sigma_0$. Using (2.10), (2.12) and (2.14), the

fact that $\sigma_{[\ell]}^z = f^z = 0$ and the fact that $\Lambda_{\alpha\sigma}^\beta(\mathbf{y}, 0) = 0$, we see that in orthogonal coordinate system, the tangential part of this vector is

$$(X_{p+1} - X_p)^\alpha = \sqrt{h} \sum_{\ell=1}^D \sigma_{[\ell]}^\alpha U_p^{[\ell]} + h \left(f^\alpha + \frac{1}{2} \sum_{\ell=1}^D \sigma_{[\ell]}^\beta D_\beta \sigma_{[\ell]}^\alpha \right) \in T_{X_p} \Sigma_0 \quad (3.3)$$

while its normal part is

$$(X_{p+1} - X_p)^z = \frac{h}{2} \sum_{\ell=1}^D \Gamma_{\alpha\beta}^z \sigma_{[\ell]}^\beta \sigma_{[\ell]}^\alpha \in T_{X_p} \Sigma_0^\perp \quad (3.4)$$

where all the functions are evaluated in X_p (for which we have $z = 0$).

3.1 Convergence result

Lemma 3.1 *Let v be a function defined on Σ_0 . Then we have*

$$\forall p \in \mathbb{N}, \quad \mathbb{E}(v(X_{p+1}) - v(X_p) | X_p) = hL(X_p; \mathbf{D})v(X_p) + h^2 \xi_p \quad (3.5)$$

with

$$\mathbb{E}|\xi_p| \leq C \sup_{J=1,\dots,4} \|\mathbf{D}^J v\|_{L^\infty(\Sigma_0)}$$

where C does not depend on p and of $h \leq h_0$.

Proof. With the notation of **(C3)**, the curve

$$(0, 1) \ni \theta \mapsto \alpha_p(\theta) = F_{-H(X_p(\theta))}(X_p(\theta))$$

is a well-defined curve on Σ_0 such that $\alpha_p(0) = X_p$ and $\alpha_p(1) = X_{p+1}$. We can write

$$v(X_{p+1}) - v(X_p) = \int_0^1 \frac{d}{d\theta} v(\alpha_p(\theta)) d\theta$$

whence

$$\begin{aligned} v(X_{p+1}) - v(X_p) &= \frac{d}{d\theta} v(\alpha_p(\theta))|_{\theta=0} + \frac{1}{2} \frac{d^2}{d\theta^2} v(\alpha_p(\theta))|_{\theta=0} + \frac{1}{6} \frac{d^3}{d\theta^3} v(\alpha_p(\theta))|_{\theta=0} \\ &\quad + \frac{1}{24} \int_0^1 (1-\theta)^3 \frac{d^4}{d\theta^4} v(\alpha_p(\theta)) d\theta. \end{aligned} \quad (3.6)$$

We can decompose the application $\theta \mapsto v(\alpha_p(\theta))$ as

$$v(\alpha_p(\theta)) = v \circ \Pi \circ X_p(\theta)$$

where $\Pi : \Omega^\varepsilon \rightarrow \Sigma_0$ is defined as $\Pi(x) = F_{-H(x)}(x)$. Let us denote by $(d\Pi)_i^\alpha(x)$ the jacobian matrix of Π acting from $T_x \Omega^\varepsilon$ to $T_{\Pi(x)} \Sigma_0$, and similarly, $(d^2\Pi)_{ij}^\alpha$

and $(d^3\Pi)_{ijk}^\alpha$ the second and third differential of Π . Here, the greek indices are $(N-1)$ -dimensional, while the roman indices are N -dimensional.

We have

$$\frac{d}{d\theta}v(\alpha_p(\theta)) = D_\alpha v(\Pi \circ X_p(\theta))(d\Pi)_i^\alpha(X_p(\theta))(X_{p+1} - X_p)^i.$$

For $\theta = 0$, this yields

$$\frac{d}{d\theta}v(\alpha_p(\theta))|_{\theta=0} = D_\alpha v(X_p)(d\Pi)_i^\alpha(X_p)(X_{p+1} - X_p)^i.$$

In a normal coordinate system, the application Π simply reads $(\mathbf{y}, z) \mapsto \mathbf{y}$, and thus its restriction to Σ_0 is the identity. Hence, if T is a tangent vector to Σ_0 viewed as vector in \mathbb{R}^N , we have as $X_p \in \Sigma_0$,

$$(d\Pi)_i^\alpha(X_p)T^i = T^\alpha$$

where T^α denote the components of T in a local basis of $T_{X_p}\Sigma_0$. Using 3.3 we have

$$\begin{aligned} (d\Pi)_i^\alpha(X_p)(X_{p+1} - X_p)^i &= \sqrt{h} \sum_{\ell=1}^D \sigma_{[\ell]}^\alpha(X_p) U_p^{[\ell]} \\ &+ h \left(f^\alpha(X_p) + \frac{1}{2} \sum_{\ell=1}^D \sigma_{[k]}^\beta(X_p) D_\beta \sigma_{[\ell]}^\alpha(X_p) \right) \\ &+ h(d\Pi)_i^\alpha(X_p)w_p^i + h^{3/2}(d\Pi)_i^\alpha(X_p)r_p^i + h^2(d\Pi)_i^\alpha(X_p)s_p^i. \end{aligned}$$

Denoting the conditional expectation $\mathbb{E}(\cdot | X_p)$ by \mathbb{E}^{X_p} , this implies that

$$\mathbb{E}^{X_p} \frac{d}{d\theta}v(\alpha_p(\theta))|_{\theta=0} = h \left(f^\alpha + \frac{1}{2} \sum_{\ell=1}^D \sigma_{[\ell]}^\beta D_\beta \sigma_{[\ell]}^\alpha \right) D_\alpha v + h^2 \xi_p^{(1)}$$

where all the functions are evaluated in X_p , and where

$$\mathbb{E}|\xi_p^{(1)}| \leq R \|Dv\|_{L^\infty}.$$

For the second derivative, we have

$$\begin{aligned} \frac{d^2}{d\theta^2}v(\alpha_p(\theta)) &= D_\alpha v(\Pi \circ X_p(\theta))(d^2\Pi)_{ij}^\alpha(X_p(\theta))(X_{p+1} - X_p)^i (X_{p+1} - X_p)^j \\ &+ D_\beta D_\alpha v(\Pi \circ X_p(\theta))[(d\Pi)_i^\alpha(X_p(\theta))(X_{p+1} - X_p)^i][(d\Pi)_j^\beta(X_p(\theta))(X_{p+1} - X_p)^j]. \end{aligned} \tag{3.7}$$

Let T be a vector in \mathbb{R}^N . In an orthogonal coordinate system around X_p , let T^α be its components in $T_{X_p}\Sigma_0$, and T^z its components along the normal to $T_{X_p}\Sigma_0$. We have

$$(d^2\Pi)_{ij}^\alpha(X_p)T^i T^j = (d^2\Pi)_{\sigma\beta}^\alpha(X_p)T^\sigma T^\beta + 2(d^2\Pi)_{\beta z}^\alpha(X_p)T^\beta T^z + (d^2\Pi)_{zz}^\alpha(X_p)T^z T^z.$$

But as the restriction of Π to Σ_0 is the identity, we have $(d^2\Pi)_{\sigma\beta}^\alpha(X_p)T^\sigma T^\beta = 0$. Using **(C3)**, **(3.3)**, **(3.4)** and **(3.2)** we easily see that

$$|\mathbb{E}^{X_p}(X_{p+1} - X_p)^\beta (X_{p+1} - X_p)^z| \leq Ch^2$$

for a constant C independent of p .

Hence, using the fact that $\mathbb{E}^{X_p}|(X_{p+1} - X_p)^z| = \mathcal{O}(h)$ we get

$$\mathbb{E}^{X_p}|(d^2\Pi)_{ij}^\alpha(X_p)(X_{p+1} - X_p)^i (X_{p+1} - X_p)^j| \leq Ch^2$$

for a constant C independent of p .

Similarly we have for the second part of **(3.7)**, for $\theta = 0$,

$$\begin{aligned} \mathbb{E}^{X_p} D_\beta D_\alpha v(X_p) (X_{p+1} - X_p)^\alpha (X_{p+1} - X_p)^\beta \\ = \sum_{\ell=1}^D \sigma_{[\ell]}^\beta(X_p) \sigma_{[\ell]}^\alpha(X_p) D_\beta D_\alpha v(X_p) + \mathcal{O}(h^2). \end{aligned}$$

This implies that

$$\mathbb{E}^{X_p} \frac{d^2}{d\theta^2} v(\alpha_p(\theta))|_{\theta=0} = \sum_{\ell=1}^D \sigma_{[\ell]}^\beta(X_p) \sigma_{[\ell]}^\alpha(X_p) D_\beta D_\alpha v + h^2 \xi_p^{(2)}$$

where the functions are evaluated in X_p , and with

$$\mathbb{E}|\xi_p^{(2)}| \leq C \sup_{J=1,2} \|D^J v\|_{L^\infty}.$$

Differentiating again the function $v(\alpha_p(\theta))$, we can see by similar computations that

$$\mathbb{E}^{X_p} \left| \frac{d^3}{d\theta^3} v(\alpha_p(\theta))|_{\theta=0} \right| \leq Ch^2 \sup_{J=1,2,3} \|D^J v\|_{L^\infty}$$

and that for all $\theta \in (0, 1)$,

$$\mathbb{E}^{X_p} \left| \frac{d^4}{d\theta^4} v(\alpha_p(\theta)) \right| \leq Ch^2 \sup_{J=1,\dots,4} \|D^J v\|_{L^\infty}.$$

Collecting together the previous formulas, we get

$$\begin{aligned} \mathbb{E}(v(X_{p+1} - X_p | X_p)) = \\ h \left(f^\alpha D_\alpha v + \frac{1}{2} \sum_{\ell=1}^D (\sigma_{[\ell]}^\beta D_\beta \sigma_{[\ell]}^\alpha) (D_\alpha v) + \frac{1}{2} \sum_{\ell=1}^D \sigma_{[\ell]}^\beta \sigma_{[\ell]}^\alpha D_\beta D_\alpha v \right) + h^2 \xi_p \end{aligned}$$

where ξ_p satisfies the estimate of the Lemma. This yields the result. \blacksquare

Theorem 3.2 *Under the hypothesis 2, let $X(t)$ be the solution of (1.1) starting at $x_0 \in \Sigma_0$, and let $(X_p)_{p \in \mathbb{N}}$ be a family of random variables satisfying **(C1)** – **(C3)** for $h \leq h_0$ and such that $X_0 = x_0$. Let g be a smooth function defined on*

Σ_0 and let $t_p = ph$ for $p \in \mathbb{N}$. Then we have for h_0 sufficiently small and $T > 0$,

$$\forall t_p \leq T, \quad \left| \mathbb{E}g(X_p) - \mathbb{E}g(X(t_p)) \right| \leq C(T)h \quad (3.8)$$

where $C(T)$ depends on T and h_0 .

Proof. Let $u(t, x)$ be the solution of the equation $\partial_t u = Lu$ on Ω^ε with the initial condition $u_0(\mathbf{y}, z) = g(\mathbf{y})$ written in an orthogonal coordinate system (see Theorem 2.3). Following [22], we write

$$\begin{aligned} g(X_p) &= u(0, X_p) \\ &= u(t_p, x_0) + \sum_{i=0}^{p-1} u(t_i, X_{p-i}) - u(t_i, X_{p-i-1}) \\ &\quad + \sum_{i=0}^{p-1} u(t_i, X_{p-i-1}) - u(t_{i+1}, X_{p-i-1}) \end{aligned}$$

Hence we have

$$\begin{aligned} \mathbb{E}g(X_p) - \mathbb{E}g(X(t_p)) &= \sum_{i=0}^{p-1} \mathbb{E}(\mathbb{E}(u(t_i, X_{p-i}) - u(t_i, X_{p-i-1}) | X_{p-i-1})) \\ &\quad + \sum_{i=0}^{p-1} \mathbb{E}(\mathbb{E}(u(t_i, X_{p-i-1}) - u(t_{i+1}, X_{p-i-1}) | X_{p-i-1})) \end{aligned}$$

Using the previous lemma, we have

$$\mathbb{E}(u(t_i, X_{p-i}) - u(t_i, X_{p-i-1}) | X_{p-i-1}) = hL(X_{p-i-1}; \mathbf{D})u(t_i, X_{p-i-1}) + h^2 \xi_{p,i}$$

with

$$\mathbb{E}|\xi_{p,i}| \leq C \sup_{J=1, \dots, 4} \|D^J u(t_i, \cdot)\|_{L^\infty(\Sigma_0)}. \quad (3.9)$$

Now using Taylor expansion, we easily get

$$\begin{aligned} u(t_i, X_{p-i-1}) - u(t_{i+1}, X_{p-i-1}) &= -hL(X_{p-i-1}; \mathbf{D})u(t_i, X_{p-i-1}) \\ &\quad - \frac{1}{2} \int_0^h (h-s)L(X_{p-i-1}; \mathbf{D})^2 u(t_i + s, X_{p-i-1}) ds. \end{aligned}$$

But as L is an operator of order 2 in \mathbf{D} , we have

$$\mathbb{E}(L(X_{p-i-1}; \mathbf{D})^2 u(t_i + s, X_{p-i-1}) | X_{p-i-1}) \leq C \sup_{J=1, \dots, 4} \|D^J u(t_i + s, \cdot)\|_{L^\infty(\Sigma_0)} \quad (3.10)$$

for some constant C independent of $p \in \mathbb{N}$ and $i \in \mathbb{N}$. Collecting together the previous results, we see that

$$\mathbb{E}g(X_p) - \mathbb{E}g(X(t_p)) = h^2 \sum_{i=0}^{p-1} \zeta_{p,i}$$

where $|\zeta_{p,i}| \leq C \sup_{J=1,\dots,4} \|D^J u(t_i, \cdot)\|_{L^\infty(\Sigma_0)}$ and this yields the result. \blacksquare

3.2 The ergodic case.

In this subsection, we assume that the hypothesis (2.19) is satisfied. We assume moreover that the family of random variables $\{X_p\}$ satisfies the additional hypothesis

(C4) For all $x_0 \in \Sigma_0$, and for all open set $\mathcal{W} \subset \Sigma_0$ then
if $X_0 = x_0$ $\mathbb{P}(\exists p \in \mathbb{N} \mid X_p \in \mathcal{W}) > 0$.

As Σ_0 is compact, this hypothesis ensures that the process $\{X_p\}$ possesses a unique invariant probability measure μ^h for which it is ergodic: For all bounded function g on Σ_0 ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{p=1}^N g(X_p) = \int_{\Sigma_0} g d\mu^h \quad \text{a.s.} \quad (3.11)$$

The following result refines the preceding convergence theorem. It shows that under the hypothesis (2.19) the constant in the convergence estimate (3.8) is uniform in T :

Lemma 3.3 *Under the hypothesis 2 and (2.19), let $X(t)$ be the solution of (1.1) starting at $x_0 \in \Sigma_0$, and let $(X_p)_{p \in \mathbb{N}}$ be a family of random variables satisfying (C1) – (C3) for $h \leq h_0$ and such that $X_0 = x_0$. Let g be a smooth function defined on Σ_0 and let $t_p = ph$ for $p \in \mathbb{N}$. Then we have for h_0 sufficiently small,*

$$\forall p \in \mathbb{N}, \quad \left| \mathbb{E} g(X_p) - \mathbb{E} g(X(t_p)) \right| \leq Ch$$

where C does not depend on p and $h \leq h_0$.

Proof. The proof is similar to the proof of Theorem 3.2. The changes are the following: Eqn. (3.9) together with (2.35) imply that there exists positive constants γ and M independent of $h \leq h_0$, such that for all $p \in \mathbb{N}$ and $i \in \mathbb{N}$,

$$\mathbb{E} |\xi_{p,i}| \leq M \exp(-\gamma t_i).$$

Using the same estimates in (3.10), we see that

$$\mathbb{E} g(X_p) - \mathbb{E} g(X(t_p)) = h^2 \sum_{i=0}^{p-1} \zeta_{p,i}$$

where

$$|\zeta_{p,i}| \leq M \exp(-\gamma t_i)$$

after possible change of the constant M which does not depend on p , i and $h \leq h_0$. Hence we have

$$\begin{aligned} \mathbb{E} g(X_p) - \mathbb{E} g(X(t_p)) &\leq Mh^2 \sum_{i=0}^{\infty} \exp(-\gamma t_i) \\ &= Mh^2 \sum_{i=0}^{\infty} \exp(-\gamma h)^i \\ &= Mh^2 \frac{1}{1 - \exp(-\gamma h)} \\ &\leq Ch \end{aligned}$$

for some constant C independent of p , and $h \leq h_0$ sufficiently small. This shows the result. \blacksquare

Theorem 3.4 *Under the hypothesis 2 and (2.19), (C1) – (C4), let μ^h be the invariant measure on Σ_0 of the process $\{X_p\}$ for $h \leq h_0$. Let g be a smooth function defined on Σ_0 and let $t_p = ph$ for $p \in \mathbb{N}$. Then there exists a constant C such that*

$$\left| \int_{\Sigma_0} g d\mu^h - \int_{\Sigma_0} g d\nu(0) \right| \leq Ch$$

where $d\nu(0)$ is the probability measure

$$\left(\int_{\Sigma_0} \mu(\mathbf{y}, 0) d\Sigma_0 \right)^{-1} \mu(\mathbf{y}, 0) d\Sigma_0$$

on Σ_0 with $\mu(\mathbf{y}, 0)$ defined by (2.21).

Proof. Using Lemma 3.3, we have

$$\left| \mathbb{E} \frac{1}{N} \sum_{p=0}^N g(X_p) - \mathbb{E} \frac{1}{N} \sum_{p=0}^N g(X(t_p)) \right| \leq Ch. \quad (3.12)$$

Now, using the hypothesis 2 and (2.19), we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{p=0}^N g(X(t_p)) = \int_{\Sigma_0} g d\nu(0) \quad a.s.$$

Hence, taking the limit $N \rightarrow \infty$ in (3.12) yields the result. \blacksquare

4 Applications

In this Section, we give two classes of SDE of the form (1.1) satisfying the hypothesis (1.3), (2.19) with $\rho = \mathbf{1}$, and (2.16). We then give example of numerical schemes in both these cases, and show numerical results.

4.1 Two examples of conservative SDEs

4.1.1 Stochastic shakers

In the spirit of [10], we introduce the following SDE, constructed as Hamiltonian system with stochastic time dependent symplectic matrices.

Let $(G_{[\ell]})_{1 \leq \ell \leq N(N-1)/2}$ be the set of $N \times N$ skew symmetric matrices with coefficients satisfying $G_{[\ell]}^{ij} = 0$ for all $(i, j) \in \{1, \dots, N\}^2 \setminus \{(i(\ell), j(\ell)), (j(\ell), i(\ell))\}$, and $G_{[\ell]}^{i(\ell), j(\ell)} = -G_{[\ell]}^{j(\ell), i(\ell)} = 1$, where the indices $(i(\ell), j(\ell))$ are such that $\bigcup_{\ell=1}^{N(N-1)/2} (i(\ell), j(\ell)) = \bigcup_{1 \leq i < j \leq N} (i, j)$. The set of matrices $(G_{[\ell]})_{1 \leq \ell \leq N(N-1)/2}$ is a basis of the space of (real) skew symmetric matrices. In the case $N = 3$, we can choose, for instance, $(i(1), j(1)) = (1, 2)$, $(i(2), j(2)) = (1, 3)$ and $(i(3), j(3)) = (2, 3)$, which corresponds to

$$G^1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad G^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad G^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

We consider stochastic differential equations of the form

$$dX^i(t) = J^{ik} \partial_k H(X(t)) dt + \sum_{\ell=1}^D G_{[\ell]}^{ik} \partial_k H(X(t)) \circ dW_{[\ell]}(t) \quad (4.1)$$

where J is a $N \times N$ skew symmetric matrix and $D = N(N-1)/2$. This SDE is of the form (1.1), with $f^i(x)$ a Hamiltonian deterministic vector field, and $\sigma_{[\ell]}^i(x) = G_{[\ell]}^{ik} \partial_k H(x)$. The skew-symmetry of the matrices J and $G_{[\ell]}$ implies that the condition (1.3) is satisfied.

Moreover, it is easy to verify that this system satisfies the condition (1.9), and hence (2.19) with $\rho = \mathbb{1}$ and thus is volume preserving.

The following result shows that the condition (2.16) is fulfilled for the equation (4.1):

Lemma 4.1 *Assume that the hypothesis 1 is satisfied. Then there exists a constant $c > 0$ and $\varepsilon > 0$, such that, for all $z \in (-\varepsilon, \varepsilon)$, $x \in \Sigma_z$, and $\xi = (\xi_i)_{i=1}^N \in T_x \Sigma_z$, we have*

$$\sum_{\ell=1}^D \left(\sum_{i,k=1}^N \xi_i G_{[\ell]}^{ki} \partial_k H(x) \right)^2 \geq c \sum_{i=1}^N |\xi_i|^2. \quad (4.2)$$

Proof. By the definition of the set of matrices $(G_{[\ell]})_{1 \leq \ell \leq N(N-1)/2}$, the sum in Equation (4.2) can be written

$$\sum_{1 \leq i < k \leq N} \left(\xi_i \partial_k H - \xi_k \partial_i H \right)^2.$$

Let us denote by n_k the k -th component of $\nabla H/\|\nabla H\|$, *i.e.* $n_k = \partial_k H/\|\nabla H\|$. We have

$$\begin{aligned} \sum_{1 \leq i < k \leq N} (\xi_i n_k - \xi_k n_i)^2 &= \sum_{1 \leq i < k \leq N} \xi_i^2 n_k^2 + \xi_k^2 n_i^2 - 2\xi_i n_k \xi_k n_i \\ &= \sum_{k \neq i} \xi_i^2 n_k^2 - 2 \sum_{1 \leq i < k \leq N} \xi_i n_k \xi_k n_i. \end{aligned}$$

Writing

$$\sum_{k \neq i} \xi_i^2 n_k^2 = \sum_{k=1}^N \xi_k^2 \sum_{i \neq k} n_i^2$$

and using the fact that $\sum_k n_k^2 = 1$, we find that

$$\begin{aligned} \sum_{1 \leq i < k \leq N} (\xi_i n_k - \xi_k n_i)^2 &= \sum_{k=1}^N \xi_k^2 (1 - n_k^2) - 2 \sum_{1 \leq i < k \leq N} \xi_i n_k \xi_k n_i \\ &= \sum_{k=1}^N \xi_k^2 - \sum_{k=1}^N \xi_k^2 n_k^2 - 2 \sum_{1 \leq i < k \leq N} \xi_i n_k \xi_k n_i \\ &= \|\xi\|^2 - \left(\sum_{k=1}^N \xi_k n_k \right)^2. \end{aligned}$$

Now if $\xi \in T_x \Sigma_z$ we have by definition

$$\sum_k \xi_k \partial_k H = 0.$$

Gathering these results together, we find that

$$\sum_{1 \leq i < k \leq N} (\xi_i \partial_k H - \xi_k \partial_i H)^2 = \|\nabla H\|^2 \|\xi\|^2$$

and this yields the result with

$$c = \min_{z \in (-\varepsilon, \varepsilon)} \min_{x \in \Sigma_z} \|\nabla H(x)\|,$$

which is positive under the hypothesis **1**, for sufficiently small ε . ■

4.1.2 Projected gradient dynamics

We would like now to introduce another dynamics ergodic with respect to the microcanonical measure (see [7]).

For any point $x \in \mathbb{R}^N$, let us define the orthogonal projector on $T_x \Sigma_{H(x)}$:

$$P(x) = \text{Id} - \frac{\nabla H \otimes \nabla H}{\|\nabla H\|^2}(x). \quad (4.3)$$

In Cartesian coordinates, its components are written

$$P_j^i = \delta_j^i + \frac{(\nabla^i H)(\nabla_j H)}{\|\nabla H\|}.$$

Let us now consider the following SDE:

$$dX^i(t) = -P_j^i(X(t))\nabla^j \left(g \circ H + \ln \|\nabla H\| \right) (X(t)) dt + \sqrt{2}P_j^i(X(t)) \circ dW^j(t) \quad (4.4)$$

where $X^i(t) \in \mathbb{R}^N$, $W^i(t)$ is a N -dimensional Brownian motion, and $g : \mathbb{R} \rightarrow \mathbb{R}$ denotes any smooth function. This SDE is of the form (1.1) with $\sigma(x) = \sqrt{2}P(x)$, and $f(x) = -P(x)\nabla(g(H(x)) + \ln \|\nabla H(x)\|)$. Using the definition of $P(x)$, it is clear that (1.3) is satisfied.

Let us introduce the function V defined by:

$$V(x) = \left(g \circ H + \ln \|\nabla H\| \right) (x). \quad (4.5)$$

The fact that (2.19) is satisfied for $\rho = 1$ is equivalent to the equation¹

$$-\nabla_i P_j^i \nabla^j V + \nabla_i P_\ell^i \nabla^j P_j^\ell = 0. \quad (4.6)$$

But we have (see Lemma A.1 in [7] for the first equality)

$$\begin{aligned} P_\ell^i \nabla^j P_j^\ell &= -P_\ell^i \nabla^\ell \ln \|\nabla H\| \\ &= -P_\ell^i \nabla^\ell V, \end{aligned}$$

and hence the equation (4.6) is satisfied.

Finally, the equation (4.4) satisfies (2.16): We have for all $x \in \Sigma_z$, for all $\xi \in T_x \Sigma_z$

$$\sum_{j=1}^N (\xi_i P_j^i)^2 = \sum_{j=1}^N \xi_j^2,$$

which shows (2.16) with $\Lambda = 1/2$.

Remark 4.2 More generally (see [7]), using such projected SDE, it is possible to sample any measure of the form $\exp(-\beta V)|\nabla \xi|^{-1} dS_z$ where $V : \mathbb{R}^N \rightarrow \mathbb{R}$ is an energy, $\xi : \mathbb{R}^N \rightarrow \mathbb{R}$ is a reaction coordinate and $S_z = \{x \in \mathbb{R}^N | \xi(x) = z\}$. One needs to consider the solution $X(t)$ of the SDE

$$dX^i(t) = -P_j^i(X(t))\nabla^j V(X(t)) dt + \sqrt{2\beta^{-1}}P_j^i(X(t)) \circ dW^j(t)$$

¹Recall that in Cartesian coordinates we have $\nabla^i H = \nabla_i H = \partial_i H$, and similarly for the components of the projection operator $P_j^i = P^{ij} = P_i^j$

where $P(x)$ is the projector

$$P(x) = \text{Id} - \frac{\nabla\xi \otimes \nabla\xi}{\|\nabla\xi\|^2}(x).$$

■

4.2 Numerical schemes

We now give examples of schemes corresponding to both the previous systems of SDE.

4.2.1 Methods for stochastic shakers.

A general method for constructing schemes adapted to (4.1) is the following: For each discrete time step $t_p = ph$, draw a symplectic matrix $(hJ + \sqrt{h}G_{[\ell]}U_p^{[\ell]})$, where $U_p^{[\ell]}$ satisfies the hypothesis of Section 3, and then apply any energy preserving method to the deterministic system associated with this matrix. In order to be consistent with the SDE, the scheme has to be of order 2 which will be automatically the case if the method is symmetric. For instance, the following two schemes can be considered:

1. Symmetric projection. This is the scheme defined as follows: Assume that $X_p \in \Sigma_0$, and define X_{p+1} by

$$\begin{aligned} \tilde{X}_p &= X_p + \nabla H(X_p)^T \mu, \\ \tilde{X}_{p+1} &= \tilde{X}_p + (hJ + G_{[\ell]}U_p^{[\ell]})\nabla H\left(\frac{\tilde{X}_p + \tilde{X}_{p+1}}{2}\right), \\ X_{p+1} &= \tilde{X}_{p+1} + \nabla H(X_{p+1})^T \mu, \end{aligned} \quad (4.7)$$

where μ is chosen such that $X_{p+1} \in \Sigma_0$ (see [14]).

2. Symmetric discrete gradient. This is the scheme defined by

$$X_{p+1} = X_p + (hJ + G_{[\ell]}U_p^{[\ell]})\bar{\nabla}H(X_p, X_{p+1}), \quad (4.8)$$

where $\bar{\nabla}H(\hat{y}, y)$ is a function satisfying:

$$\begin{aligned} \bar{\nabla}H(\hat{y}, y)^T(\hat{y} - y) &= H(\hat{y}) - H(y), \\ \bar{\nabla}H(y, y) &= \nabla H(y). \end{aligned}$$

With these properties, we can easily prove that $H(X_{p+1}) = H(X_p)$. We can take for instance

$$\bar{\nabla}H(\hat{y}, y) = \nabla H(\bar{y}) + \frac{H(\hat{y}) - H(y) - \nabla H(\bar{y})^T \Delta y}{\|\Delta y\|^2} \Delta y$$

with $\bar{y} = \frac{1}{2}(\hat{y} + y)$ and $\Delta y = \hat{y} - y$. This scheme, named *midpoint discrete gradient* was introduced by Gonzalez [12].

It can be checked that these schemes satisfy the hypothesis (C1) to (C4) for h sufficiently small provided the random variables $U_p^{[\ell]}$ are a.s. bounded.

4.2.2 Projected gradient dynamics schemes.

For the discretization of the SDE (4.4), we note that it can be rewritten in the following form:

$$dX(t) = -\nabla V(X(t)) dt + \sqrt{2}dW(t) + \nabla H(X(t))d\Lambda(t), \quad (4.9)$$

where $\Lambda(t)$ is a real valued process, which can be interpreted as the Lagrange multiplier associated with the constraint $H(X(t)) = H(X_0)$. This process can be decomposed into two parts:

$$d\Lambda(t) = d\Lambda^m(t) + d\Lambda^f(t). \quad (4.10)$$

The martingale part $\Lambda^m(t)$ is

$$d\Lambda^m(t) = -\sqrt{2} \frac{\nabla H}{\|\nabla H\|^2}(X(t)) \cdot dW(t), \quad (4.11)$$

where \cdot implicitly denotes the Itô product. The bounded variation part $\Lambda^f(t)$ is

$$d\Lambda^f(t) = \frac{\nabla H}{\|\nabla H\|^2}(X(t)) \cdot \nabla V(X(t)) dt + \frac{\nabla H}{\|\nabla H\|^2}(X(t)) \cdot \Gamma(X(t)) dt, \quad (4.12)$$

where $\Gamma(x)$ denotes the mean curvature vector to $\Sigma_{H(x)}$ at point x :

$$\Gamma = -\nabla \cdot \left(\frac{\nabla H}{\|\nabla H\|} \right) \frac{\nabla H}{\|\nabla H\|}. \quad (4.13)$$

We can now consider two discretizations of (4.4) which can be shown to be consistent (see [7, 15]). The first one is:

$$\begin{cases} X_{n+1} = X_n - \nabla V(X_n) h + \sqrt{2h} U_n + \Delta\Lambda_{n+1} \nabla H(X_{n+1}), \\ \text{where } \Delta\Lambda_{n+1} \text{ is such that } H(X_{n+1}) = H_0, \end{cases} \quad (4.14)$$

where h is the time step and U_n is a $3N$ -dimensional standard Gaussian random vector satisfying the hypothesis in the beginning of Section 3. Notice that (4.14) admits a natural variational interpretation, since X_{n+1} can be seen as the closest point on the submanifold Σ_z to the predicted position $X_n - \nabla V(X_n) \Delta t + \sqrt{2h} U_n$.

The real $\Delta\Lambda_{n+1}$ is then the Lagrange multiplier associated with the constraint $H(X_{n+1}) = H_0$.

Another possible discretization of (4.4) is

$$\begin{cases} X_{n+1} = X_n - \nabla V(X_n) h + \sqrt{2h} U_n + \Delta\Lambda_{n+1} \nabla H(X_n), \\ \text{where } \Delta\Lambda_{n+1} \text{ is such that } H(X_{n+1}) = H_0. \end{cases} \quad (4.15)$$

Although this scheme is not naturally associated with a variational principle, it may be more practical since its formulation is more explicit.

It can be checked that these schemes satisfy the hypothesis (C1) to (C4) for h sufficiently small provided the random variables U_n are a.s. bounded.

These numerical schemes for constrained Brownian dynamics are in the spirit of the so-called RATTLE [3] and SHAKE [19] algorithms classically used for constrained Hamiltonian dynamics, and also related with the algorithms proposed in [23, 2, 18].

4.3 Numerical results

We consider the following hamiltonian in dimension $N = 4$:

$$H(x_1, x_2, x_3, x_4) = \frac{1}{2}(x_1^2 + x_2^2) + V(x_3, x_4) \quad (4.16)$$

where

$$V(x_3, x_4) = (x_3^2 - 1)^2 + (x_4 + x_3^2 - 1)^2.$$

We are interesting in computing the averages of the following eight observables $(x_3, x_3^2, x_3^4, x_4, x_4^2, x_1^2, x_2^2, V(x_3, x_4))$ with respect to the NVE measure (1.6). To this aim, we simulate approximations X_p , $p \geq 1$ of the previous systems, and compute the time average (3.11) for sufficiently large N .

The hamiltonian (4.16) represents a particle in dimension 2, with momentum $(p_x, p_y) = (x_1, x_2)$ and position $(q_x, q_y) = (x_3, x_4)$ in the double-well potential V . It is possible to calculate analytically the exact averages of the observables mentioned above. We refer to Section 3.4 in [6].

In Table 1 and 2, we show the result for the method (4.8). The initial conditions are such that $x_1 = 0.5$, $x_2 = -0.5$, $x_3 = -1$ and $x_4 = 0.5$, so that the initial energy level is $H = 0.5$. The 95% confidence interval is obtained with 30 independent runs. The time step is h , and the number of points in the trajectory is represented by the number N .

Observables	Exact	Approx.	$h = 1.10^{-3}, N = 5.10^5$
x_3	-0.94459	-0.94580	[-0.94651, -0.94509]
x_3^2	0.92843	0.93064	[0.92930, 0.93196]
x_3^4	0.98964	0.9937	[0.99110, 0.99625]
x_4	0.07156	0.07075	[0.06825, 0.07326]
x_4^2	0.25517	0.25470	[0.25392, 0.25548]
x_1^2	0.24482	0.24496	[0.24432, 0.24560]
x_2^2	0.24482	0.24528	[0.24465, 0.24591]
$V(x_3, x_4)$	0.25517	0.25488	[0.25453, 0.25522]

Table 1: midpoint discrete gradient algorithm

Observables	Exact	Approx.	$h = 5.10^{-4}, N = 10^6$
x_3	-0.94459	-0.94458	[-0.94511, -0.94404]
x_3^2	0.92843	0.92850	[0.92750, 0.92950]
x_3^4	0.98964	0.99009	[0.98814, 0.99203]
x_4	0.07156	0.07201	[0.07011, 0.07391]
x_4^2	0.25517	0.25550	[0.25476, 0.25625]
x_1^2	0.24482	0.24444	[0.24389, 0.24500]
x_2^2	0.24482	0.24448	[0.24384, 0.24512]
$V(x_3, x_4)$	0.25517	0.25554	[0.25517, 0.25591]

Table 2: midpoint discrete gradient algorithm

In Tables 3, we give the results obtained with a slight modification of the dynamics (4.14), namely:

$$\left\{ \begin{array}{l} X_{n+1} = X_n - \nabla V(X_n) h + \sqrt{2h} U_n \\ \quad + \frac{\nabla V \cdot \nabla H}{\|\nabla H\|^2} \nabla H(X_n) - \sqrt{2h} \frac{U_n \cdot \nabla H}{\|\nabla H\|^2} \nabla H(X_n) + \Delta \Lambda_{n+1} \nabla H(X_{n+1}), \\ \text{where } \Delta \Lambda_{n+1} \text{ is such that } H(X_{n+1}) = H_0, \end{array} \right. \quad (4.17)$$

The two additional terms are easy to compute, do not modify the consistency of the scheme, and improve the convergence in the projection step. In practice, we use an Uzawa algorithm to perform the projection step. The initial values are the same as in the previous section

Observables	Exact	Approx	$h = 5.10^{-4}, N = 10^6$
x_3	-0.94459	-0.94094	[-0.94288, -0.93900]
x_3^2	0.92843	0.92051	[0.91688, 0.92414]
x_3^4	0.98964	0.97029	[0.96342, 0.97715]
x_4	0.07156	0.07372	[0.06666, 0.00808]
x_4^2	0.25517	0.26230	[0.26071, 0.26390]
x_1^2	0.24482	0.25074	[0.24910, 0.25238]
x_2^2	0.24482	0.25223	[0.25074, 0.25372]
$V(x_3, x_4)$	0.25517	0.24851	[0.24766, 0.24936]

Table 3: projected gradient algorithm

Acknowledgment

This work is supported by the ANR INGEMOL of the French Ministry of Research. The authors are glad to thank Arnaud Debussche and Denis Talay for helpful comments during the preparation of this work.

References

- [1] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, Oxford, 1987.
- [2] S.A. Allison and J.A. McCammon. Transport properties of rigid and flexible macromolecules by brownian dynamics simulation. *Biopolymers*, 23(1):167–187, 1984.
- [3] H.C. Andersen. Rattle: a “velocity” version of the Shake algorithm for molecular dynamics calculations. *J. Comput. Phys.*, 52:24–34, 1983.
- [4] T. Aubin. *Nonlinear analysis on manifolds. Monge-Ampère equations*. Springer, 1982.
- [5] E. Cancès, F. Castella, P. Chartier, E. Faou, C. Le Bris, F. Legoll, and G. Turinici. Long-time averaging using symplectic solvers with application to molecular dynamics. *Numer. Math.*, 100, 2005.
- [6] E. Cancès, F. Castella, P. Chartier, E. Faou, C. Le Bris, F. Legoll, and G. Turinici. High-order averaging schemes with error bounds for thermodynamical properties calculations by molecular dynamics simulations. *J. Chem. Phys.*, 121(21):10346–10355, 2004.
- [7] G. Ciccotti, T. Lelièvre, and E. Vanden-Eijnden. Projection of diffusions on submanifolds: Application to mean force computation. *Communications on Pure and Applied Mathematics*, 61:371–408, 2008.
- [8] E. Faou. Développements asymptotiques dans les coques linéairement élastiques. Thèse, Université de Rennes 1, 2000.
- [9] E. Faou. Elasticity on a thin shell: Formal series solution. *Asymptotic Analysis*, 31:317–361, 2002.
- [10] E. Faou. Nosé-Hoover dynamics in a shaker. *J. Chem. Phys.*, 124:184104, 2006.
- [11] D. Frenkel and B. Smit. *Understanding molecular dynamics: from algorithms to applications*. Academic Press, London, 2002.

- [12] O. Gonzalez. Time integration and discrete Hamiltonian systems. *J. Non-linear Sci.*, 6:449–467, 1996.
- [13] A. Gorod and D. Talay. Approximation of Lyapunov exponents of nonlinear stochastic differential equations. *SIAM J. Appl. Math.*, 56(2):627–650, 1996.
- [14] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics 31. Springer, Berlin, 2002.
- [15] C. Le Bris, T. Lelièvre, and E. Vanden-Eijnden. Analysis of some discretization schemes for constrained stochastic differential equations. *Preprint*, 2008.
- [16] G.N. Milstein, Y.M. Repin, and M.V. Tretyakov. Numerical methods for stochastic systems preserving symplectic structure. *SIAM J. Numer. Anal.*, 40(4):1583–1604, 2002.
- [17] G.N. Milstein, Y.M. Repin, and M.V. Tretyakov. Symplectic integration of Hamiltonian systems with additive noise. *SIAM J. Numer. Anal.*, 39(6):2066–2088, 2002.
- [18] H.C. Öttinger. Brownian dynamics of rigid polymer chains with hydrodynamic interactions. *Phys. Rev. E*, 50(4):2696–2701, 1994.
- [19] J.P. Ryckaert, G. Ciccotti, and H.J.C. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n -alkanes. *J. Comput. Phys.*, 23:327–342, 1977.
- [20] D. Talay. Second order discretization schemes of stochastic differential system for the computation of the invariant law. *Stochastics and Stochastic Reports*, 29 (1):13–36, 1990.
- [21] D. Talay. Approximation of upper lyapunov exponents of bilinear stochastic differential systems. *SIAM J. Numer. Anal.*, 28(4):1141–1164, 1991.
- [22] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.
- [23] W.F. Van Gunsteren and H.J.C. Berendsen. Algorithms for brownian dynamics. *Mol. Phys.*, 45(3):637–647, 1982.