A forward-backward stochastic analysis of diffusion flows

Pierre del Moral, S.S. Singh

A forward-backward stochastic analysis of diffusion flows

P. Del Moral[1] and S. S. Singh[2]

[1]INRIA, Bordeaux Research Center, Talence, France
[2]University of Cambridge, Department of Engineering, Cambridge, England.

Abstract

We present forward-backward stochastic semigroup formulae to analyse the difference of diffusion flows driven by different drift and diffusion functions. These formulae are expressed in terms of tangent and Hessian processes and can be interpreted as an extension of the Aleeksev-Gröbner lemma to diffusion flows. We present some natural spectral conditions that allows to derive in a direct way a series of uniform estimates with respect to the time horizon. We illustrate the impact of these results in the context of diffusion perturbation theory, interacting diffusions and discrete time approximations.

Keywords : Stochastic flows, variational equations, tangent and Hessian processes, perturbation semigroups, Aleeksev-Gröbner lemma, Skorohod stochastic integral, Malliavin differential, Bismut-Elworthy-Li formulae.

Mathematics Subject Classification : 47D07, 93E15, 60H07.

1 Introduction

Let $b_t(x)$ and $\sigma_t(x)$ be twice differentiable functions from $\mathbb{R}^d$ into $\mathbb{R}^d$ and $\mathbb{R}^{d\times r}$, for some parameters $d,r\geqslant 1$; We set $a_t(x):=\sigma_t(x)\ \sigma_t(x)'$ and we let W_t be a r-dimensional Brownian motion.

For any time horizon $s\geqslant 0$ we denote by $X_{s,t}(x)$ be the stochastic flow defined for any $t\in[s,\infty[$ and any starting point $X_{s,s}(x)=x\in\mathbb{R}^d$ by the stochastic differential equation

$$dX_{s,t}(x)=b_t\left(X_{s,t}(x)\right)\ dt+\sigma_t\left(X_{s,t}(x)\right)\ dW_t \tag{1.1}$$

We assume that $b_t(x)$ and $\sigma_t(x)$ have uniformly bounded first order derivatives. This condition is clearly met for linear Gaussian models as well as for the geometric Brownian motion. As it is well known, dynamical systems and hence stochastic models involving drift functions with quadratic growth require additional regularity conditions to ensure non explosion of the solution in finite time. It is also implicitly assumed that all functions (b_t,σ_t) are smooth w.r.t. the time parameter.

Let $\overline{X}_{s,t}(x)$ be the stochastic flow associated with a stochastic differential equation defined as (1.1) by replacing (b_t,σ_t) by some drift and diffusion functions $(\overline{b}_t,\overline{\sigma}_t)$ with the same regularity properties; we also set $\overline{a}_t:=\overline{\sigma}_t\,\overline{\sigma}_t'$. Constant diffusion functions $(\sigma_t,\overline{\sigma}_t)$ are defined by

$$\sigma_t(x)=\Sigma_t\quad\text{and}\quad\overline{\sigma}_t(x)=\overline{\Sigma}_t\quad\text{for some matrices }\Sigma_t\text{ and }\overline{\Sigma}_t. \tag{1.2}$$

In this context, we shall assume that Σ_t and $\overline{\Sigma}_t$ are uniformly bounded w.r.t. the time horizon.

The Markov transition semigroups associated with the flows $X_{s,t}(x)$ and $\overline{X}_{s,t}(x)$ are defined for any measurable function f on $\mathbb{R}^d$ by the formula

$$P_{s,t}(f)(x) := \mathbb{E}\left(f(X_{s,t}(x))\right) \quad \text{and} \quad \overline{P}_{s,t}(f)(x) := \mathbb{E}\left(f(\overline{X}_{s,t}(x))\right)$$

One natural question is to express $(X_{s,t} - \overline{X}_{s,t})$ and $(P_{s,t} - \overline{P}_{s,t})$ in terms of the difference of functions

$$\Delta a_t := a_t - \overline{a}_t \qquad \Delta b_t := b_t - \overline{b}_t \quad \text{and} \quad \Delta\sigma_t = \sigma_t - \overline{\sigma}_t \tag{1.3}$$

The functions $-\Delta b_t$ and $-\Delta\sigma_t$ can be interpreted as a local perturbation of the drift and the diffusion functions of the stochastic flow $X_{s,t}$.

Another challenging problem is to estimate the difference between the stochastic flows $X_{s,t}$ and $\overline{X}_{s,t}$ and their transition probabilities *uniformly w.r.t. the time horizon.*

These important questions arise in a variety of domains including in stochastic perturbation theory as well as in the stability and the qualitative theory of stochastic systems.

Classical analytic estimates on the difference between the stochastic flows driven by different drift and diffusion functions are often much too large for most diffusion processes of practical interest. In some instances none of the diffusion flows are stable. In this context, any local perturbation of the stochastic model propagates so that any global error estimate eventually tends to ∞ as the time horizon $t \to \infty$. Whenever one of the stochastic flows is stable, classical perturbation bounds combining Lipschitz type inequalities with Gronwall lemma [8, 18] yield exceedingly pessimistic global estimates that grows exponentially fast w.r.t. the. time horizon. Notice that an exponential type estimate of the form $e^{\lambda t}$ for some parameter $\lambda > 0$ and some time horizon t s.t. $\lambda\, t \geqslant 199$ would induce an error bound larger than the estimated number 10^{86} of elementary particles of matters in the visible universe. As mentioned in [21] in the context of Euler scheme type approximations of deterministic dynamical systems, one may encounter situations where $\lambda = 10^8$ and $t = 10^2$ and the resulting exponential bounds does not belong to numerical analysis.

The article presents a forward-backward stochastic perturbation formula that expresses the difference between the stochastic flows $X_{s,t}$ and $\overline{X}_{s,t}$ and their transition probabilities in terms of tangent and Hessian processes. We also provide some natural spectral conditions that allows to derive in a direct way a series of more realistic uniform estimates with respect to the time horizon.

We illustrate the impact of these results in the context of diffusion perturbation theory, interacting diffusions and discrete time approximations.

We end this section with some basic notation necessary for the statement of our main results. When there are no confusions we slight abuse notation and we denote by $\|.\|$ any (equivalent) norm on some finite dimensional vector space over $\mathbb{R}$ as well as the uniform norm $\|f\| := \sup_{t,x}\|f_t(x)\|$ and $\|f(x)\| := \sup_t \|f_t(x)\|$ of some multivariate function $f_t(x)$ on $([0,\infty[\times\mathbb{R}^d)$. For any $n \geqslant 1$ we also set

$$|||f(x)|||_n := \sup_{s\leqslant t} \mathbb{E}\left(\|f_t(\overline{X}_{s,t}(x))\|^n\right)^{1/n} \tag{1.4}$$

In the further development of the article we represent the gradient of a real valued function as a column vector, or equivalently as the transpose of the differential-Jacobian operator which is as any cotangent represented by a row vector. The gradient and the Hessian of a column vector valued function as tensors of type $(1,1)$ and $(2,1)$, see for instance (1.17) and (1.18).

We also denote by κ, κ_n and $\kappa_{\delta,n}$ some constants that depends on some parameters n and (δ, n) but they don't depend on the time horizon, nor on the space variable.

1.1 Statement of some main results

To describe our result in an intuitive and informal way, we slice the interval $[0, t]$ into small pieces of step size $\delta s \simeq 0$ and we use the backward approximations

$$\begin{aligned} & X_{s,t}(x) - X_{s-\delta s,t}(x) = X_{s,t}(x) - X_{s,t}(X_{s-\delta s,s}(x)) \\ & \simeq X_{s,t}(x) - X_{s,t}\left(x + b_s(x)\ \delta s + \sigma_s(x)\ \left(W_s - W_{s-\delta s}\right)\right) \\ & \simeq -\left[\left(\nabla X_{s,t}(x)'\ b_s(x) + \frac{1}{2}\ \nabla^2 X_{s,t}(x)'\ a_s(x)\right)\ \delta s + \nabla X_{s,t}(x)'\sigma_s(x)\ \left(W_s - W_{s-\delta s}\right)\right] \end{aligned} \tag{1.5}$$

The above approximations are rigorously justified in section 3.1 and lead to the backward stochastic flow evolution equation

$$d_s X_{s,t}(x) = -\left[\left(\nabla X_{s,t}(x)'\ b_s(x) + \frac{1}{2}\ \nabla^2 X_{s,t}(x)'\ a_s(x)\right)\ ds + \nabla X_{s,t}(x)'\sigma_s(x)\ dW_s\right] \tag{1.6}$$

In the above display, $d_s X_{s,t}(x)$ stands for the increment of $X_{s,t}(x)$ w.r.t. the variable s. Combining the backward formula (1.6) with Ito's lemma for any $s \leqslant u \leqslant t$ we check the interpolation formula

$$\begin{aligned} & d_u\left(X_{u,t} \circ \overline{X}_{s,u}\right)(x) \\ & = \left(d_u X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right) + \left(\nabla X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)'\ d_u \overline{X}_{s,u}(x) + \frac{1}{2}\ \left(\nabla^2 X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)'\ \overline{a}_u(\overline{X}_{s,u}(x))\ du \end{aligned} \tag{1.7}$$

This yields basically the following interpolation theorem.

Theorem 1.1. *We have the forward-backward stochastic interpolation formula*

$$X_{s,t}(x) - \overline{X}_{s,t}(x) = T_{s,t}(\Delta a, \Delta b)(x) + S_{s,t}(\Delta\sigma)(x) \tag{1.8}$$

with the stochastic process

$$\begin{aligned} & T_{s,t}(\Delta a, \Delta b)(x) \\ & := \int_s^t \left[\left(\nabla X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)'\ \Delta b_u(\overline{X}_{s,u}(x)) + \frac{1}{2}\ \left(\nabla^2 X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)'\ \Delta a_u(\overline{X}_{s,u}(x))\right]\ du \end{aligned}$$

and the Skorohod stochastic integral fluctuation term given by

$$S_{s,t}(\Delta\sigma)(x) := \int_s^t\ \left(\nabla X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)'\ \Delta\sigma_u(\overline{X}_{s,u}(x))\ dW_u \tag{1.9}$$

Under some regularity conditions that depend on some parameters $n \geqslant 2$ and $\delta \in]0,1[$ we have the uniform estimates

$$\begin{aligned} & \mathbb{E}\left[\|X_{s,t}(x) - \overline{X}_{s,t}(x)\|^n\right]^{1/n} \\ & \leqslant \kappa_{\delta,n}\ \left(|||\Delta a(x)|||_{2n/(1+\delta)} + |||\Delta b(x)|||_{2n/(1+\delta)} + |||\Delta\sigma(x)|||_{2n/\delta}\ (1 \vee \|x\|)\right) \end{aligned} \tag{1.10}$$

For constant diffusion functions (1.2), for any $n \geqslant 2$ we have

$$\mathbb{E}\left[\|X_{s,t}(x) - \overline{X}_{s,t}(x)\|^n\right]^{1/n} \leqslant \kappa_n\ \left(|||\Delta b(x)|||_n + \|\Sigma - \overline{\Sigma}\|\right) \tag{1.11}$$

A more detailed proof of (1.8) is provided in section 3.

The estimates (1.10) come from (1.29) and (5.11) and they rely on the regularity conditions $(M)_{2n/\delta}$ and $(T)_{2n/(1-\delta)}$ discussed in section 1.3. A more detailed proof of (1.10) is provided in the appendix, on page 37. The estimates (1.11) come from (1.32) and (5.13) and they rely on the regularity condition $(T)_2$.

Notice that (1.6) can also be deduced from (1.7) replacing $\overline{X}_{s,u}$ by the stochastic flow $X_{s,u}$. Up to some technicalities, the forward-backward stochastic interpolation formula (1.8) is a direct consequence of the second order approximation discussed in (1.5). Several extensions to more general stochastic perturbation processes can be developed. For instance, suppose we are given some stochastic processes $\overline{Y}_{s,t}(x) \in \mathbb{R}^d$ and $\overline{Z}_{s,t}(x) \in \mathbb{R}^{d\times r}$ adapted to the filtration of the Brownian motion W_t, and let $\overline{X}_{s,t}(x)$ be the stochastic flow defined by the stochastic differential equation

$$d\overline{X}_{s,t}(x) = \overline{Y}_{s,t}(x)\ dt + \overline{Z}_{s,t}(x)\ dW_t \tag{1.12}$$

In this situation, the interpolation formula (1.7) remains valid when $\overline{a}_u(\overline{X}_{s,u}(x))$ is replaced by the stochastic matrices $\overline{Z}_{s,t}(x)\overline{Z}_{s,t}(x)'$. This yields without further work the forward-backward stochastic interpolation formula (1.8) with the local perturbations

$$\begin{array}{rcl}
\Delta b_u(\overline{X}_{s,u}(x)) & := & b_u(\overline{X}_{s,u}(x)) - \overline{Y}_{s,u}(x) \\
\Delta \sigma_u(\overline{X}_{s,u}(x)) & := & \sigma_u(\overline{X}_{s,u}(x)) - \overline{Z}_{s,u}(x) \quad \text{and} \quad \Delta a_u(\overline{X}_{s,u}(x)) := a_u(\overline{X}_{s,u}(x)) - \overline{Z}_{s,u}(x)\overline{Z}_{s,u}(x)'
\end{array}$$

The forward-backward stochastic interpolation formula (1.8) can also be extended to more general classes of stochastic flows on abstract state spaces. For instance the recent article [22] provides a deterministic first order version of (1.8) on abstract Banach spaces. The articles [4, 5] also discuss similar interpolation formulae for mean field particle systems and deterministic nonlinear measure valued semigroups. The stability properties of these abstract models depends on the problem at hand. To focus on the main ideas without clouding the article with unnecessary technical details and sophisticated mathematical tools and abstract regularity conditions we have chosen to concentrate the article on diffusion flows on Euclidian spaces with simple and easily checked regularity conditions.

We mention that condition $(M)_n$ is a technical condition that ensures that the n-absolute moments of the flows $X_{s,t}$ and $\overline{X}_{s,t}$ are uniformly bounded w.r.t. the time horizon. More explicit sufficient polynomial growth conditions on the drift and diffusion functions are discussed in section 1.3. Condition $(T)_n$ is a spectral condition on the gradient of the drift and diffusion matrices of the stochastic flows. For instance, for constant diffusion functions the spectral condition $(T)_n$ is met for any $n \geqslant 2$ as soon as the following log-norm conditions are met

$$\nabla b_t + (\nabla b_t)' \leqslant -2\lambda\ I \quad \text{and} \quad \nabla\overline{b}_t + (\nabla\overline{b}_t)' \leqslant -2\overline{\lambda}\ I \quad \text{for some } \lambda \wedge \overline{\lambda} > 0. \tag{1.13}$$

For constant diffusion functions (1.2) s.t. $\Sigma_t = \overline{\Sigma}_t$ whenever the l.h.s. log-norm condition in (1.13) is satisfied, for any $n \geqslant 2$ we also have the uniform estimate

$$\mathbb{E}\left[\|X_{s,t}(x) - \overline{X}_{s,t}(x)\|^n\right]^{1/n} \leqslant \kappa\ \|\!|\Delta b(x)|\!\|_n$$

The above assertion is a direct consequence of the estimates (1.32).

When $\sigma = 0$ the flow $X_{s,t}(x)$ is deterministic so that the Skorohod fluctuation term (1.9) reduces to the traditional Ito stochastic integral

$$S_{s,t}(\Delta\sigma)(x) := -\int_s^t\ (\nabla X_{u,t})\,(\overline{X}_{s,u}(x))'\ \overline{\sigma}_u(\overline{X}_{s,u}(x))\ dW_u \tag{1.14}$$

In this situation, whenever the l.h.s. log-norm condition in (1.13) is satisfied combining the Burkholder-Davis-Gundy inequality with the generalized Minkowski inequality for any $n \geqslant 2$ we have

$$\mathbb{E}\left(\|S_{s,t}(\Delta\sigma)(x)\|^n\right)^{2/n} \leqslant (e/2)\ n^2 \int_s^t e^{-2\lambda(t-u)}\ \mathbb{E}\left(\|\overline{\sigma}_u(\overline{X}_{s,u}(x))\|_F^n\right)^{2/n}\ du \leqslant \kappa\ n^2\ \|\!|\overline{\sigma}(x)|\!\|_n^2$$

For linear drift functions of the form $b_t(x) = B_t\ x$, the l.h.s. condition in (1.13) indicates that the log-norm of B_t is uniformly bounded by $-\lambda < 0$. In addition, whenever $\sigma = 0$ the tangent process $\nabla X_{s,t}(x)$ satisfies a time-varying deterministic linear dynamical system

$$\partial_t \nabla X_{s,t}(x) = \nabla X_{s,t}(x)\ B_t'$$

The asymptotic behavior of these processes cannot be characterized by the statistical properties of the spectral abscissa of the matrices B_t. Indeed, unstable semigroups associated with time-varying (deterministic) matrices B_t with negative eigenvalues are exemplified in [10, 31]. Conversely, stable semigroups with B_t having positive eigenvalues are given by Wu in [31]. The uniform log-norm condition (1.13) provides a rather weak and readily verifiable condition.

The rest of the article is organized as follows.

Section 2 provides some basic tools associated with the first and second variational equations associated with a diffusion flow. We also present some quantitative estimates of the tangent and the Hessian processes.

Section 3 also presents an extended version of the above formula to multivariate functions of stochastic flows, see for instance theorem 3.1. Apart more complex and sophisticated tensor notation, the stochastic analysis of these multivariate formulae follows the same arguments as the ones used in the proof of theorem 1.1. Thus, we have chosen to concentrate this introduction on the perturbation analysis of stochastic flows.

Some extensions of the stochastic interpolation formula (1.8) are discussed in section 4, including non Markov perturbations and jump diffusion perturbation processes.

Section 5 is dedicated to the analysis of the Skorohod fluctuation process introduced in (1.9).

Section 6 presents some illustrations of the forward-backward interpolation formulae discussed in the present article in the context of diffusion perturbation theory, interacting diffusions and discrete time approximations.

We end this section with some comments and comparisons of of results with existing literature on the subject.

The forward-backward formula (1.8) can be seen as an extension of theorem 6.1 in [27] on two-sided stochastic integrals to diffusion flows. The terminology " two-sided " coined by the authors in [27] comes from the fact that the integrand of the Skorohod integral depend on the past as well as on the future of the history generated by the Brownian motion.

Formula (1.8) can also be interpreted as an extension of Aleeksev-Gröbner lemma [1, 17, 22] as well as an extended version of the variation-of-constant and related Gronwall type lemma [8, 18] to diffusion processes. In this connection we underline that the forward-backward formula (1.8) differs from the stochastic Gronwall lemma presented in [28] based on particular classes of stochastic linear inequalities that doesn't involve Skorohod type integrals.

This backward perturbation formula has been used in [3, 5] in the context of nonlinear diffusions and their mean field type interacting particle interpretations. Similar backward-error second order expansions have been used [4] in the context of Feynman-Kac interacting jumps. The discrete time version of this backward perturbation semigroup methodology can also be found in chapter 7 in [12], a well as in the articles [13, 14, 15].

In a more recent independent work [20], the authors also present a weak formulation of the above formula using a rather sophisticated discrete time approximation analysis to reformulate Skorohod stochastic integration. The article [20] doesn't provide any uniform estimates w.r.t. the time horizon, nor any almost sure estimates.

One crucial advantage of the backward perturbation analysis discussed above is that it enters the stability properties of the tangent and the Hessian flow to the estimation of the differences of the stochastic flows. This important property allows to derive several uniform estimates with respect to the time horizon.

1.2 Some basic notation

We introduce some matrix notation needed from the onset. We denote by $\|A\|_2 := \lambda_{max}(AA')^{1/2}$, resp. $\|A\|_F = \mathrm{Tr}(AA')^{1/2}$ and $\rho(A) = \lambda_{max}((A+A')/2)$ the spectral norm, the Frobenius norm, and the logarithmic norm of some matrix A, where A' stands for the transpose of A, and $\lambda_{max}(.)$ the maximal eigenvalue. With a slight abuse of notation, we denote by I the identity $(d \times d)$-matrix, for any $d \geqslant 1$. We also denote by $\|.\|$ any (equivalent) norm on some finite dimensional vector space over $\mathbb{R}$.

For any symmetric positive definite matrices Q_1, Q_2 we recall the Ando-Hemmen inequality

$$\|Q_1^{1/2} - Q_2^{1/2}\| \leqslant \left[\lambda_{min}^{1/2}(Q_1) + \lambda_{min}^{1/2}(Q_2)\right]^{-1} \|Q_1 - Q_2\| \tag{1.15}$$

for any unitary invariant matrix norm $\|.\|$. In the above display, $\lambda_{min}(.)$ the minimal eigenvalue. We also have the square root inequality

$$Q_1 \geqslant Q_2 \Longrightarrow Q_1^{1/2} \geqslant Q_2^{1/2} \tag{1.16}$$

See for instance Theorem 6.2 on page 135 in [19], as well as Proposition 3.2 in [2]. A proof of (1.16) can be found in [7]. We let $[n]$ stands for the set of n multiple indexes $i = (i_1, \ldots, i_n) \in \mathcal{I}^n$ over some finite set $\mathcal{I}$. We denote by $(A_{i,j})_{(i,j)\in[p]\times[q]}$ the entries of a (p,q)-tensor A index set $\mathcal{I}$.

For a given (p_1, q)-tensor A and a given (p_2, q) tensor B over a common index set $\mathcal{I}$, AB' is a (p_1, p_2)-tensor with entries given by

$$\forall (i,j) \in [p_1] \times [p_2] \qquad (AB')(i,j) = \sum_{k\in[q]} A(i,k)B(j,k)$$

We consider the Frobenius inner product

$$\langle A, B\rangle_F = \mathrm{Tr}(AB') \quad \text{and the norm} \quad \|A\|_F = \sqrt{\mathrm{Tr}(AA')}$$

For any (p,q)-tensors A and B we also check the Cauchy-Schwartz inequality

$$\langle A, B\rangle_F^2 \leqslant \|A\|_F \ \|B\|_F \quad \text{and} \quad \|A\|_2 \leqslant \|A\|_F \leqslant \mathrm{Card}(\mathcal{I})^p \ \|A\|_2 \quad \text{with} \quad \|A\|_2 := \lambda_{max}(AA')^{1/2}$$

For any tensors A, B with appropriate dimensions we have the inequality

$$\|AB\|_F \leqslant \|A\|_F \ \|B\|_F$$

Given some tensor valued function $T : (t,x) \mapsto T_t(x)$ we also set

$$\|T\|_F := \sup_{t,x} \|T_t(x)\|_F \qquad \|T\|_2 := \sup_{t,x} \|T_t(x)\|_2 \qquad \text{and} \qquad \|T\| := \sup_{t,x} \|T_t(x)\|$$

Given some smooth function $h(x)$ from $\mathbb{R}^p$ into $\mathbb{R}^q$ we denote by

$$\nabla h = \left[\nabla h^1, \ldots, \nabla h^q\right] \quad \text{with} \quad \nabla h^i = \left[\begin{array}{c} \partial_{x_1} h^i \\ \vdots \\ \partial_{x_p} h^i \end{array}\right] \tag{1.17}$$

the gradient (p, q)-matrix associated with the column vector-valued function $h = (h^i)_{1 \leqslant i \leqslant q}$. In the same vein, we denote by

$$\nabla^2 h = \left[\nabla^2 h^1, \ldots, \nabla^2 h^q\right] \quad \text{with} \quad \nabla^2 h^i = \left[\begin{array}{ccc} \partial_{x_1, x_1} h^i & \ldots & \partial_{x_1, x_p} h^i \\ \vdots & \ldots & \vdots \\ \partial_{x_p, x_1} h^i & \ldots & \partial_{x_p, x_p} h^i \end{array}\right] \tag{1.18}$$

the Hessian $(2,1)$-tensor associated with the function $h = (h^i)_{1 \leqslant i \leqslant q}$. For any $n \geqslant 1$ we let $\mathcal{P}_n(\mathbb{R}^d)$ be the convex set of probability measures μ_1, μ_2 on $\mathbb{R}^d$ with absolute n-th moment and equipped with the Wasserstein distance of order n denoted by

$$\mathbb{W}_n(\mu_1, \mu_2) := \inf \mathbb{E}(\|X_1 - X_2\|^n)^{1/n}$$

In the above display the infimum is taken over all pair or random variables (X_1, X_2) with marginal distributions (μ_1, μ_2). The stochastic transition semigroups associated with the flows $X_{s,t}(x)$ and $\overline{X}_{s,t}(x)$ are defined for any measurable function f on $\mathbb{R}^d$ by the formulae

$$\mathbb{P}_{s,t}(f)(x) := f(X_{s,t}(x)) \quad \text{and} \quad \overline{\mathbb{P}}_{s,t}(f)(x) := f(\overline{X}_{s,t}(x))$$

Given some smooth function column vector-valued function $f = (f^i)_{1 \leqslant i \leqslant p}$ we write $\mathbb{P}_{s,t}(f)$ and $P_{s,t}(f)$ the column vector-valued function with entries $\mathbb{P}_{s,t}(f^i)$ and $P_{s,t}(f^i)$.

In the same vein, we write $\mathbb{P}_{s,t}(\nabla f)$ and and $\mathbb{P}_{s,t}(\nabla^2 f)$ the $(1,1)$ and $(2,1)$-tensor valued functions with entries given by

$$\mathbb{P}_{s,t}(\nabla f)(x)_{i,k} := \mathbb{P}_{s,t}(\partial_i f^k)(x) \quad \text{and} \quad \mathbb{P}_{s,t}(\nabla^2 f)(x)_{(i,j),k} := \mathbb{P}_{s,t}(\partial_{i,j} f^k)(x)$$

We also consider the random $(2,1)$ and $(2,2)$-tensors given by

$$\begin{array}{rcl} \nabla^2 X_{s,t}(x)_{(i,j),k} & = & \partial_{i,j} X^k_{s,t}(x) = \left[\nabla^2 X_{s,t}(x)\right]'_{k,(i,j)} \\ \left[\nabla X_{s,t}(x) \otimes \nabla X_{s,t}(x)\right]_{(i,j),(k,l)} & = & \nabla X_{s,t}(x)_{i,k} \nabla X_{s,t}(x)_{j,l} = \left[\nabla X_{s,t}(x) \otimes \nabla X_{s,t}(x)\right]'_{(k,l),(i,j)} \end{array}$$

Throughout the rest of the article, unless otherwise is stated we write $\kappa, \kappa_\epsilon, \kappa_n, \kappa_{n,\epsilon}$ some constants whose values may vary from line to line but they only depend on the parameters $n \geqslant 0$ and $\epsilon > 0$ as well as on the parameters of the model. We also use the letters $c, c_\epsilon, c_n, c_{n,\epsilon}$ to denote universal constants that doesn't depend on the parameters of the model. Importantly these contants don't depend on the time horizon. We also consider the uniform log-norm parameters

$$\rho(\nabla\sigma)^2 := \sum_{1 \leqslant k \leqslant d} \sup_{t,x} \rho_\star(\nabla\sigma_k)^2 \qquad \rho_\star(\nabla\sigma_k) := \sup_t \rho_\star(\nabla\sigma_{t,k}(x)) \qquad \rho_\star(\nabla\sigma) := \sup_k \rho_\star(\nabla\sigma_k(x)) \tag{1.19}$$

and the parameters $\chi(b, \sigma)$ defined by

$$\chi(b, \sigma) := c + \|\nabla^2 b\| + \|\nabla^2 \sigma\|^2 + \rho_\star(\nabla\sigma)^2 \tag{1.20}$$

1.3 Regularity conditions and preliminary results

We consider two different type of regularity conditions $(\mathcal{M})_n$ and $(\mathcal{T})_n$ indexed by $n \geqslant 2$.

$(\mathcal{M})_n$: *There exists some parameters* $\kappa_n \geqslant 0$ *such that for any* $x \in \mathbb{R}^d$ *we have*

$$m_n(x) := \sup_{s \leqslant t} \mathbb{E}\left(\|X_{s,t}(x)\|^n\right)^{1/n} \leqslant \kappa_n \ (1 \vee \|x\|)$$

$(\mathcal{T})_n$: *There exists some parameter* $\lambda_A > 0$ *such that*

$$A_t := \nabla b_t + (\nabla b_t)' + \sum_{1 \leqslant k \leqslant r} \nabla \sigma_{k,t} (\nabla \sigma_{k,t})' \leqslant -2\lambda_A \ I \tag{1.21}$$

In addition, the following condition is satisfied

$$\lambda_A(n) := \lambda_A - \frac{d(n-2)}{2} \ \rho_\star(\nabla\sigma)^2 > 0 \tag{1.22}$$

Let $\overline{A}_t$ be the symmetric matrix defined as A_t by replacing in (1.21) the drift and the diffusion matrix (b_t, σ_t) by $(\overline{b}_t, \overline{\sigma}_t)$. We denote by $(\overline{\mathcal{M}})_n$ and $(\overline{\mathcal{T}})_n$ the regularity conditions defined as $(\mathcal{M})_n$ and $(\mathcal{T})_n$ by replacing the functions (b_t, σ_t) by $(\overline{b}_t, \overline{\sigma}_t)$. We also let $\lambda_{\overline{A}}(n)$ the collection of parameters defined as $\lambda_A(n)$ by replacing the functions (b_t, σ_t) by $(\overline{b}_t, \overline{\sigma}_t)$.

We write $(M)_n$ when both conditions $(\mathcal{M})_n$ and $(\overline{\mathcal{M}})_n$ are satisfied as well as $(T)_n$ when conditions $(\mathcal{T})_n$ and $(\overline{\mathcal{T}})_n$ are met. When $(T)_n$ is satisfied we also set

$$\lambda_{A,\overline{A}}(n) := \lambda_A(n) \wedge \lambda_{\overline{A}}(n)$$

• In practice, the uniform moment condition $(\mathcal{M})_n$ is often checked using Lyapunov techniques. We can also use the following polynomial growth condition:

$(\mathcal{P})_n$: *There exists some parameters* $\alpha_i, \beta_i \geqslant 0$ *with* $i = 0, 1, 2$ *such that for any* $t \geqslant 0$ *and any* $x \in \mathbb{R}^d$ *we have*

$$\|\sigma_t(x)\|_F^2 \leqslant \alpha_0 + \alpha_1\|x\| + \alpha_2\|x\|^2 \quad \textit{and} \quad \langle x, b_t(x)\rangle \leqslant \beta_0 + \beta_1\|x\| - \beta_2\|x\|^2 \tag{1.23}$$

for some norm $\|\sigma_t(x)\|$ *of the matrix-valued diffusion function. In addition, we have*

$$\beta_2(n) := \beta_2 - \frac{(n-1)}{2} \ \alpha_2 > 0$$

For any $n \geqslant 2$ we have

$$(\mathcal{P})_n \quad \Longrightarrow \quad (\mathcal{M})_n \quad \text{with} \quad \kappa_n = 1 + \frac{(\gamma_1 + (n-2)\alpha_1) + (\gamma_0 + (n-2)\alpha_0)^{1/2}}{2\beta_2(n)^{1/2}} \tag{1.24}$$

The proof of the above assertion follows standard stochastic calculations, thus it is housed in the appendix, on page 31.

• For one-dimensional geometric Brownian motions the condition $(\mathcal{P})_n$ is a sufficient and necessary condition for the existence of uniformly bounded absolute n-moments. In this case $(\mathcal{T})_n$ coincides with $(\mathcal{P})_n$ by setting

$$\lambda_A = \beta_2 - \alpha_2/2 \quad \text{and} \quad \alpha_2 = \rho_\star(\nabla\sigma)^2$$

• Whenever condition $(M)_n$ is met for some $n \geqslant 2$, we also check the uniform estimates

$$\sup_{s\leqslant u\leqslant t} \mathbb{E}\left(\|[X_{u,t}\circ \overline{X}_{s,u}](x)\|^n\right)^{1/n} \leqslant \kappa_n\ (1+\|x\|) \tag{1.25}$$

In this situation, with the $\mathbb{L}_n$-norms $|||.|||_n$ introduced in (1.4) we also have that

$$|||\Delta b(x)|||_n \leqslant \kappa_{1,n}(1\vee\|x\|) \quad \text{and} \quad |||\Delta a(x)|||_{n/2} \leqslant \kappa_{2,n}\ (1\vee\|x\|)^2 \tag{1.26}$$

• Condition $(\mathcal{T})_n$ ensures the exponential decays of the absolute and uniform n-moments of the tangent and the Hessian processes; that is, when $(\mathcal{T})_n$ is met for some $n\geqslant 2$ we have that

$$\mathbb{E}\left(\|\nabla X_{s,t}(x)\|^n\right)^{1/n} \vee \mathbb{E}\left(\|\nabla^2 X_{s,t}(x)\|^n\right)^{1/n} \leqslant \kappa_n\ e^{-\lambda(n)(t-s)} \quad \text{for some} \quad \lambda(n)>0 \tag{1.27}$$

A more precise statement is provided in proposition 2.2 and proposition 2.9. These uniform estimates clearly implies that for any $n\geqslant 2$ and $s\leqslant u\leqslant t$ we have

$$\mathbb{E}\left(\|(\nabla X_{u,t})(\overline{X}_{s,u}(x))\|^n\right)^{1/n} \vee \mathbb{E}\left(\|(\nabla^2 X_{u,t})(\overline{X}_{s,u}(x))\|^n\right)^{1/n} \leqslant \kappa_n\ e^{-\lambda(n)(t-u)} \tag{1.28}$$

with the same parameters $(\kappa_n,\lambda(n))$ as in (1.27). Thus, applying the generalized Minkowski inequality to (1.8) whenever $(\mathcal{T})_{n/\delta}$ is met for some $\delta\in]0,1[$ and $n\geqslant 2$ we have

$$\begin{aligned}&\mathbb{E}\left[\|\left[X_{s,t}(x)-\overline{X}_{s,t}(x)\right]-S_{s,t}(\Delta\sigma)(x)\|^n\right]^{1/n}\\ &\leqslant \frac{\kappa_{n/\delta}}{\lambda(n/\delta)}\ \left(|||\Delta b(x)|||_{n/(1-\delta)}+|||\Delta a(x)|||_{n/(1-\delta)}\right) \quad \text{with } (\kappa_n,\lambda(n)) \text{ given in (1.27).}\end{aligned} \tag{1.29}$$

• As expected, the case $\nabla\sigma=0$ plays a particular role. For instance whenever $(\mathcal{T})_2$ is met we have the almost sure and uniform gradient estimates

$$\|\nabla X_{s,t}\|_2 := \sup_x \|\nabla X_{s,t}(x)\|_2 \leqslant e^{-\lambda_A(t-s)} \tag{1.30}$$

In addition, we have the almost sure and uniform Hessian estimates

$$\|\nabla^2 X_{s,t}\|_F := \sup_x \|\nabla^2 X_{s,t}(x)\|_F \leqslant \frac{d}{\lambda_A}\ \|\nabla^2 b\|_F\ e^{-\lambda_A(t-s)} \tag{1.31}$$

A proof of the above estimates is provided in the beginning of section 2.1 and section 2.2. In this situation, whenever $(\mathcal{T})_2$ is met we have

$$\mathbb{E}\left[\|\left[X_{s,t}(x)-\overline{X}_{s,t}(x)\right]-S_{s,t}(\Delta\sigma)(x)\|^n\right]^{1/n} \leqslant \kappa\ \left(|||\Delta b(x)|||_n+|||\Delta a(x)|||_n\right) \tag{1.32}$$

For instance, for Langevin diffusions associated with some convex potential function U we have

$$\begin{aligned}&b=-\nabla U \qquad \nabla\sigma=0 \quad \text{and} \quad \nabla^2 U\geqslant \lambda\ I\\ &\Longrightarrow \quad \|\nabla X_{s,t}\|_2\leqslant e^{-\lambda(t-s)} \quad \text{and} \quad \|\nabla^2 X_{s,t}\|_F \leqslant \frac{d}{\lambda}\ \|\nabla^3 U\|_F\ e^{-\lambda(t-s)}\end{aligned} \tag{1.33}$$

• In practice, it is often easier to work with $a_t(x)$ than $\sigma_t(x)$. Next we discuss some ways of estimating $\Delta\sigma_t$ in terms of Δa_t and inversely. To this end, we further assume that the following ellipticity condition is satisfied

$$a_t(x)\geqslant \upsilon\ I \quad \text{and} \quad \overline{a}_t(x)\geqslant \upsilon\ I \quad \text{for some parameter } \upsilon>0. \tag{1.34}$$

In this situation, using (1.15) and (1.16) we check that

$$\|\Delta\sigma_t(x)\| \leqslant \frac{1}{\sqrt{\upsilon}}\ \|\Delta a_t(x)\| \quad \text{and} \quad \|\sigma_t(x)\| \leqslant \|\sigma_t(0)\| + \frac{1}{\sqrt{\upsilon}}\ [\|a_t(x)\| + \|a_t(0)\|] \tag{1.35}$$

This provides a way to estimate the growth of $\sigma_t(x)$ in terms of the one of $a_t(x)$. For instance the estimate (1.10) combined with (1.35) implies that

$$\mathbb{E}\left[\|X_{s,t}(x) - \overline{X}_{s,t}(x)\|^n\right]^{1/n} \leqslant \kappa_{\delta,n}\ \left(\|\!|\Delta b(x)|\!\|_{2n/(1+\delta)} + \|\!|\Delta a(x)|\!\|_{2n/\delta}(1 \vee \|x\|)\right)$$

We can estimate Δa_t in terms of $\Delta\sigma_t$ using the inequality

$$\|\Delta a_t(x)\| \leqslant \|\Delta\sigma_t(x)\|\ [\|\sigma_t(x)\| + \|\overline{\sigma}_t(x)\|]$$

- Assume that $(\overline{\mathcal{M}})_n$ is satisfied for some $n \geqslant 1$. Also let $f_t(x)$ be some multivariate function such that

$$\|f(0)\| := \sup_t \|f_t(0)\| < \infty \quad \text{and} \quad \|\nabla f\| := \sup_{t,x} \|\nabla f_t(x)\| < \infty$$

we have the estimates

$$\|\!|f(x)|\!\|_n \leqslant \|f(0)\| + \|\nabla f\|\ \overline{m}_n(x) \quad \text{and therefore} \quad \|\!|f(x)|\!\|_n \leqslant \kappa_n\ (\|f(0)\| + \|\nabla f\|)$$

2 Variational equations

2.1 The tangent process

The gradient $\nabla X_{s,t}(x)$ of the diffusion flow $X_{s,t}(x)$ is given by the gradient $(d \times d)$-matrix

$$d\,\nabla X_{s,t}(x) = \nabla X_{s,t}(x)\ \left[\nabla b_t\left(X_{s,t}(x)\right)\ dt + \sum_{1\leqslant k\leqslant r} \nabla\sigma_{t,k}\left(X_{s,t}(x)\right)\ dW_t^k\right]$$

After some calculations we check that

$$d\left[\nabla X_{s,t}(x)\,\nabla X_{s,t}(x)'\right] = \nabla X_{s,t}(x)\ A_t\left(X_{s,t}(x)\right)\ \nabla X_{s,t}(x)'\ dt + dM_{s,t}(x) \tag{2.1}$$

with the symmetric matrix valued martingale $M_{s,t}(x)$ given by

$$dM_{s,t}(x) := \sum_{1\leqslant k\leqslant r} \nabla X_{s,t}(x)\left[\nabla\sigma_{t,k}\left(X_{s,t}(x)\right) + \nabla\sigma_{t,k}\left(X_{s,t}(x)\right)'\right]\nabla X_{s,t}(x)'\ dW_t^k$$

Whenever $(\mathcal{T})_2$ is met, we have the following uniform estimate

$$(\mathcal{T})_2 \Longrightarrow \mathbb{E}\left(\|\nabla X_{s,t}(x)\|_2^2\right)^{1/2} \leqslant \mathbb{E}\left(\|\nabla X_{s,t}(x)\|_F^2\right)^{1/2} \leqslant \sqrt{d}\ e^{-\lambda_A(t-s)} \tag{2.2}$$

In addition, when $\nabla\sigma = 0$ the martingale $M_{s,t}(x) = 0$ is null and (1.30) is a consequence of (2.1).

Using the Taylor expansion

$$X_{s,t}(x) - X_{s,t}(y) = \int_0^1\ \nabla X_{s,t}(\epsilon x + (1-\epsilon)y)'(x-y)\ d\epsilon$$

$$\Longrightarrow \|X_{s,t}(x) - X_{s,t}(y)\|^2 \leqslant \left[\int_0^1\ \|\nabla X_{s,t}(\epsilon x + (1-\epsilon)y)\|_2^2\ d\epsilon\right]\ \|x-y\|^2$$

we readily check the following proposition.

Proposition 2.1. *Assume $(\mathcal{T})_2$ is satisfied. In this situation, we have*

$$\mathbb{E}\left(\|X_t(x)-X_t(y)\|^2\right)^{1/2} \leqslant \sqrt{d}\ e^{-\lambda_A(t-s)}\quad \|x-y\| \tag{2.3}$$

In addition, we have the almost sure estimate

$$\nabla\sigma=0 \Longrightarrow \|X_{s,t}(x)-X_{s,t}(y)\| \leqslant e^{-\lambda_A(t-s)}\quad \|x-y\| \tag{2.4}$$

These contraction inequalities quantify the stability of the stochastic flow $X_{s,t}(x)$ w.r.t. the initial state x. For instance, the estimate (2.3) ensures that the Markov transition semigroup is exponentially stable; that is, we have that

$$\mathbb{W}_2\left(\mu_0 P_{s,t},\mu_1 P_{s,t}\right) \leqslant c\ \exp\left[-\lambda_A(t-s)\right]\ \mathbb{W}_2\left(\mu_0,\mu_1\right) \tag{2.5}$$

For the Langevin diffusions discussed in (1.33) the stochastic flow is time homogeneous; that is we have that $X_{s,t}=X_{t-s}:=X_{0,(t-s)}$ and $P_{s,t}=P_{t-s}:=P_{0,(t-s)}$. In addition when $\sigma(x)=\sigma\ I$, the diffusion flow $X_t(x)$ has a single invariant measure on $\mathbb{R}^d$ given by the Boltzmann-Gibbs measure

$$\pi(dx)=\frac{1}{Z}\ \exp\left(-\frac{2}{\sigma^2}\,U(x)\right)\,dx \quad\text{with}\quad Z:=\int\ e^{-\frac{2}{\sigma^2}U(x)}\ dx \tag{2.6}$$

In this situation, using (1.33) for any $n\geqslant 1$ we check that

$$\nabla^2 U \geqslant \lambda\ I \quad\Longrightarrow\quad \mathbb{W}_n\left(\mu P_{s,t},\pi\right) \leqslant \exp\left[-\lambda(t-s)\right]\ \mathbb{W}_n\left(\mu,\pi\right)$$

Taking the trace in (2.1) we also find that

$$d\,\|\nabla X_{s,t}(x)\|_F^2=\operatorname{Tr}\left[\nabla X_{s,t}(x)\ A_t\left(X_{s,t}(x)\right)\ \nabla X_{s,t}(x)'\right]\ dt+dN_{s,t}(x)$$

with the martingale

$$dN_{s,t}(x)=\sum_k\ \operatorname{Tr}\left(\nabla X_{s,t}(x)\left[\nabla\sigma_{t,k}\left(X_{s,t}(x)\right)+\nabla\sigma_{t,k}\left(X_{s,t}(x)\right)'\right]\nabla X_{s,t}(x)'\right)\ dW_t^k$$

Observe that

$$\partial_t\langle N_{s,.}(x)\rangle_t=\sum_k\ \operatorname{Tr}\left(\nabla X_{s,t}(x)\left[\nabla\sigma_{t,k}\left(X_{s,t}(x)\right)+\nabla\sigma_{t,k}\left(X_{s,t}(x)\right)'\right]\nabla X_{s,t}(x)'\right)^2$$

This implies that

$$\begin{aligned}\partial_t\mathbb{E}\left(\|\nabla X_{s,t}(x)\|_F^4\right)&=2\ \mathbb{E}\left(\|\nabla X_{s,t}(x)\|_F^2\ \operatorname{Tr}\left[\nabla X_{s,t}(x)\ A_t\left(X_{s,t}(x)\right)\ \nabla X_{s,t}(x)'\right]\right)\\&\quad+\sum_k\mathbb{E}\left(\operatorname{Tr}\left(\nabla X_{s,t}(x)\left[\nabla\sigma_{t,k}\left(X_{s,t}(x)\right)+\nabla\sigma_{t,k}\left(X_{s,t}(x)\right)'\right]\nabla X_{s,t}(x)'\right)^2\right)\end{aligned}$$

Whenever $(\mathcal{T})_2$ is met, we have the estimate

$$\partial_t\mathbb{E}\left(\|\nabla X_{s,t}(x)\|_F^4\right) \leqslant -4\left[\lambda_A-\rho(\nabla\sigma)^2\right]\ \mathbb{E}\left(\|\nabla X_{s,t}(x)\|_F^4\right)$$

with the uniform log-norm parameter $\rho(\nabla\sigma$ defined in (1.19). This yields the estimate

$$\partial_t\mathbb{E}\left(\|\nabla X_{s,t}(x)\|_F^4\right)^{1/4} \leqslant -\left[\lambda_A-\rho(\nabla\sigma)^2\right]\ \mathbb{E}\left(\|\nabla X_{s,t}(x)\|_F^4\right)^{1/4}$$

More generally, we readily check the following result.

Proposition 2.2. *When condition $(\mathcal{T})_2$ is met, for any $n\geqslant 2$ we have*

$$\lambda_A>(n-2)\rho(\nabla\sigma)^2/2 \Longrightarrow \mathbb{E}\left(\|\nabla X_{s,t}(x)\|_F^n\right)^{1/n} \leqslant \sqrt{d}\ e^{-\left[\lambda_A-(n-2)\rho(\nabla\sigma)^2/2\right](t-s)} \tag{2.7}$$

2.2 The Hessian process

We have the matrix diffusion equation

$$\begin{aligned}
&d\,\nabla^2 X_{s,t}(x)\\
&= \left[\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]\nabla^2 b_t(X_{s,t}(x))+\nabla^2 X_{s,t}(x)\nabla b_t(X_{s,t}(x))\right]dt+d\mathcal{M}_{s,t}(x)
\end{aligned}$$

with the null matrix initial condition $\nabla^2 X_{s,s}(x)=0$ and the matrix-valued martingale

$$d\mathcal{M}_{s,t}(x)=\sum_{1\leqslant k\leqslant d}\left(\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]\nabla^2\sigma_{t,k}(X_{s,t}(x))+\nabla^2 X_{s,t}(x)\nabla\sigma_{t,k}(X_{s,t}(x))\right)\,dW_t^k$$

Consider the tensor functions

$$\upsilon_t:=\sum_{1\leqslant k\leqslant d}(\nabla^2\sigma_{t,k})\,(\nabla^2\sigma_{t,k})'\quad\text{and}\quad\tau_t:=\nabla^2 b_t+\sum_{1\leqslant k\leqslant d}(\nabla^2\sigma_{t,k})\,(\nabla\sigma_{t,k})'\tag{2.8}$$

After some computations, we check that

$$\begin{aligned}
&d\left[\nabla^2 X_{s,t}(x)\nabla^2 X_{s,t}(x)'\right]\\
&=\Big\{\left[\nabla^2 X_{s,t}(x)\;A_t(X_{s,t}(x))\;\nabla^2 X_{s,t}(x)'\right]+2\left[\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]\;\tau_t(X_{s,t}(x))\;\nabla^2 X_{s,t}(x)'\right]_{sym}\\
&\qquad+\left[\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]\upsilon_t(X_{s,t}(x))\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]'\right]\Big\}\,dt+d\mathcal{N}_{s,t}(x)
\end{aligned}$$

with the tensor-valued martingale

$$\begin{aligned}
d\mathcal{N}_{s,t}(x)=2\sum_{1\leqslant k\leqslant d}&\big\{\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]\;\nabla^2\sigma_{t,k}(X_{s,t}(x))\;\nabla^2 X_{s,t}(x)'\\
&+\nabla^2 X_{s,t}(x)\;\nabla\sigma_{t,k}(X_{s,t}(x))\;\nabla^2 X_{s,t}(x)'\big\}_{sym}\;dW_t^k
\end{aligned}$$

When $\nabla\sigma=0$ the above equation reduces to

$$\begin{aligned}
&\partial_t\left[\nabla^2 X_{s,t}(x)\nabla^2 X_{s,t}(x)'\right]\\
&=\left[\nabla^2 X_{s,t}(x)\;A_t(X_{s,t}(x))\;\nabla^2 X_{s,t}(x)'\right]+2\left[\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]\;\nabla^2 b_t(X_{s,t}(x))\;\nabla^2 X_{s,t}(x)'\right]_{sym}
\end{aligned}$$

Whenever $(\mathcal{T})_2$ is met, taking the trace in the above display we check that

$$\partial_t\,\|\nabla^2 X_{s,t}(x)\|_F^2\leqslant-2\lambda_A\;\|\nabla^2 X_{s,t}(x)\|_F^2+2\|\nabla^2 b\|_F\;\|\nabla X_{s,t}(x)\|_F^2\;\|\nabla^2 X_{s,t}(x)\|_F$$

This yields the estimate

$$\partial_t\,\|\nabla^2 X_{s,t}(x)\|_F\leqslant-\lambda_A\;\|\nabla^2 X_{s,t}(x)\|_F^2+\|\nabla^2 b\|_F\;\|\nabla X_{s,t}(x)\|_F^2$$

Using (1.30) this implies that

$$\|\nabla^2 X_{s,t}(x)\|_F\leqslant\|\nabla^2 b\|_F^2\;e^{-\lambda_A(t-s)}\int_s^t e^{\lambda_A(u-s)}\;\|\nabla X_{s,u}(x)\|_F^2\;du\leqslant\frac{d}{\lambda_A}\;\|\nabla^2 b\|_F\;e^{-\lambda_A(t-s)}$$

This ends the proof of the almost sure estimate (1.31).

For more general models, we have that

$$
\begin{aligned}
&d\,\|\nabla^2 X_{s,t}(x)\|_F^2\\
&=\left\{\mathrm{Tr}\left[\nabla^2 X_{s,t}(x)\ A_t(X_{s,t}(x))\ \nabla^2 X_{s,t}(x)'\right]+2\mathrm{Tr}\left[\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]\ \tau_t(X_{s,t}(x))\ \nabla^2 X_{s,t}(x)'\right]\right.\\
&\qquad\left.+\mathrm{Tr}\left[\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]\upsilon_t(X_{s,t}(x))\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]'\right]\right\}dt+dM_{s,t}(x)
\end{aligned}
$$

with a continuous martingale $M_{s,t}(x)$ with angle bracket

$$
\begin{aligned}
&\partial_t\langle M_{s,\cdot}(x)\rangle_t\\
&=4\sum_{1\leqslant k\leqslant d}\mathrm{Tr}\left\{\left[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)\right]\ \nabla^2\sigma_{t,k}(X_{s,t}(x))\ \nabla^2 X_{s,t}(x)'\right.\\
&\qquad\qquad\left.+\nabla^2 X_{s,t}(x)\ \nabla\sigma_{t,k}(X_{s,t}(x))\ \nabla^2 X_{s,t}(x)'\right\}^2
\end{aligned}
$$

Proposition 2.3. *Assume $(\mathcal{T})_n$ is met. In this situation, for any $\epsilon>0$ s.t. $\lambda_A(n)>\epsilon$ we have*

$$
\mathbb{E}\left(\|\nabla^2 X_{s,t}(x)\|_F^n\right)^{1/n}\leqslant n\ \epsilon^{-1}\ \chi(b,\sigma)\ \exp\left(-\left[\lambda_A(n)-\epsilon\right](t-s)\right)\tag{2.9}
$$

with the parameters $\chi(b,\sigma)$ and $\lambda_A(n)$ defined in (1.20) and (1.22).

In the above display, $\rho_\star(\nabla\sigma)$ stands for the defined in (1.19). The proof of the above estimate is rather technical, thus it is housed in the appendix on page 34

2.3 Bismut-Elworthy-Li formulae

The gradient semigroup formulae discussed in (3.5) are valid for smooth test functions. We further assume that ellipticity condition (1.34) is met. In this situation, we can extend gradient semigroup formulae to measurable functions using the Bismut-Elworthy-Li formula

$$
\nabla P_{s,t}(f)(x)=\mathbb{E}\left(f(X_{s,t}(x))\ \tau^{\omega}_{s,t}(x)\right)\tag{2.10}
$$

with the stochastic process

$$
\tau^{\omega}_{s,t}(x):=\int_s^t\ \partial_u\omega_{s,t}(u)\ \nabla X_{s,u}(x)\ a_u(X_{s,u}(x))^{-1/2}\ dW_u
$$

The above formula is valid for any function $\omega_{s,t}:u\in[s,t]\mapsto\omega_{s,t}(u)\in\mathbb{R}$ of the following form

$$
\omega_{s,t}(u)=\varphi\left((u-s)/(t-s)\right)\ \Longrightarrow\partial_u\omega_{s,t}(u)=\frac{1}{t-s}\ \partial\varphi\left((u-s)/(t-s)\right)\tag{2.11}
$$

for some non decreasing differentiable function φ on $[0,1]$ with bounded continuous derivatives and such that

$$
(\varphi(0),\varphi(1))=(0,1)\Longrightarrow\omega_{s,t}(t)-\omega_{s,t}(s)=1
$$

Whenever $(\mathcal{T})_2$ is met, combining (2.2) with (2.10), for any f s.t. $\|f\|\leqslant 1$ we check that

$$
\begin{aligned}
\|\nabla P_{s,t}(f)\|^2&\leqslant\ \mathbb{E}\left(\|\tau^{\omega}_{s,t}(x)\|^2\right)\\
&\leqslant\ \kappa_1\int_s^t e^{-2\lambda_A(u-s)}\ \|\partial_u\omega^{s,t}(u)\|^2\ du=\ \frac{\kappa_1}{t-s}\int_0^1 e^{-2\lambda_A(t-s)v}\ \left(\partial\varphi(v)\right)^2\ dv
\end{aligned}
$$

Let φ_ϵ with $\epsilon\in]0,1[$ be some differentiable function on $[0,1]$ null on $[0,1-\epsilon]$ and such that $|\partial\varphi_\epsilon(u)|\leqslant c/\epsilon$ and $(\varphi_\epsilon(1-\epsilon),\varphi(1))=(0,1)$. In this situation, we check that

$$\|\nabla P_{s,t}(f)\|^2\leqslant\ \frac{\kappa_2}{\epsilon^2}\ \frac{1}{t-s}\ \int_{1-\epsilon}^1\ e^{-2\lambda_A(t-s)v}\ dv$$

from which we find the rather crude uniform estimate

$$\|\nabla P_{s,t}(f)\|\leqslant\frac{\kappa}{\epsilon}\ \frac{1}{\sqrt{t-s}}\ e^{-\lambda_A(1-\epsilon)(t-s)} \tag{2.12}$$

In the same vein, for any $s\leqslant u\leqslant t$ we have the formulae

$$\begin{aligned}\nabla^2P_{s,t}(f)(x)&=\mathbb{E}\left(f(X_{s,t}(x))\ \tau_{s,t}^{[2],\omega}(x)+\nabla X_{s,t}(x)\,\nabla f(X_{s,t}(x))\ \tau_{s,t}^{\omega}(x)'\right) &(2.13)\\&=\mathbb{E}\left(f(X_{s,t}(x))\ \left[\tau_{s,u}^{[2],\omega}(x)+\nabla X_{s,u}(x)\ \tau_{u,t}^{\omega}(X_{s,u}(x))\,\tau_{s,u}^{\omega}(x)'\right]\right) &(2.14)\end{aligned}$$

with the process

$$\begin{aligned}&\tau_{s,t}^{[2],\omega}(x)\\&:=\int_s^t\ \partial_u\omega_{s,t}(u)\ \left[\nabla^2X_{s,u}(x)\ a_u(X_{s,u}(x))^{-1/2}+(\nabla X_{s,u}(x)\otimes\nabla X_{s,u}(x))\ \ (\overline{\nabla}a_u^{-1/2})(X_{s,u}(x))\right]\ dW_u\end{aligned}$$

In the above display $\overline{\nabla}a_u^{-1/2}$ stands for the tensor function

$$(\overline{\nabla}a_u^{-1/2}(x))_{(i,j),k}:=\partial_{x_i}a_u^{-1/2}(x)_{j,k}=-\left(a_u^{-1/2}(x)\left[\partial_{x_i}a_u^{1/2}(x)\right]a_u^{-1/2}(x)\right)_{j,k}$$

Observe that

$$(1.34)\Longrightarrow\sup_i\|\partial_{x_i}a_u^{-1/2}(x)\|\leqslant c\ \|\nabla\sigma\|/\upsilon$$

Whenever $(\mathcal{T})_2$ is met, using the estimate (2.3) for any $\epsilon\in]0,1[$

$$\|\nabla^2P_{s,t}(f)\|\leqslant\frac{\kappa}{\epsilon}\ \frac{1}{\sqrt{t-s}}\ e^{-\lambda_A(t-s)(1-\epsilon)}\ \ (\|f\|+\|\nabla f\|) \tag{2.15}$$

In the same vein, using (2.14) for any $u\in]s,t[$ and any bounded measurable function f s.t. $\|f\|\leqslant1$ we also check the rather crude uniform estimate

$$\|\nabla^2P_{s,t}(f)\|\leqslant\frac{\kappa_1}{\epsilon}\ \frac{1}{\sqrt{u-s}}\ e^{-\lambda_A(u-s)(1-\epsilon)}+\frac{\kappa_2}{\epsilon^2}\ \frac{1}{\sqrt{(t-u)(u-s)}}\ e^{-\lambda_A(u-s)}\ e^{-\lambda_A(t-s)(1-\epsilon)}$$

Choosing $u=s+(1-\epsilon)(t-s)$ in the above display we check that for any $\epsilon\in]0,1[$ we obtain the uniform estimate

$$\|\nabla^2P_{s,t}(f)\|\leqslant\frac{\kappa}{\epsilon}\ \frac{1}{\sqrt{t-s}}\ e^{-\lambda_A(t-s)(1-\epsilon)}\ \left(1+\frac{1}{\epsilon}\ \frac{1}{\sqrt{t-s}}\ e^{-\lambda_A(t-s)(1-\epsilon)}\right)\ \|f\| \tag{2.16}$$

The extended versions of the above formulae in the context of diffusions on differentiable manifolds can be found in the series of articles [6, 9, 16, 23, 29].

3 Backward semigroup analysis

3.1 A multivariate Skorohod-Aleeksev-Gröbner formulae

For any given time horizon $s \leqslant t$ we have the rather well known backward stochastic flow equation

$$X_{s,t}(x) = x + \int_s^t \left[\nabla X_{u,t}(x)'\ b_u(x) + \frac{1}{2}\ \ \nabla^2 X_{u,t}(x)'\ a_u(x)\ \right]\ du + \int_s^t \nabla X_{u,t}(x)' \sigma_u(x)\ \widehat{d} W_u$$

In the above display, $\widehat{d} W_u$ stands for the backward integration notation, so that the r.h.s. term in the above formula is a square integrable backward martingale. A detailed proof of the above formula in the context of nonlinear diffusion can be found in the appendix of [3], see also the original proof given in [11]. In a more synthetic form, the above backward formula reduces to (1.6).

Formally, the idea is to consider the discrete time interval $[s,t]_h := \{u_0, \ldots, u_{n-1}\}$ associated with some refining time mesh $u_{i+1} = u_i + h$ from $u_0 = s$ to $u_n = t$, for some time step $h > 0$. In this notation, for any $[s,t]_h$ we have

$$\begin{aligned} & X_{u+h,t} \circ X_{s,u+h} - X_{u,t} \circ X_{s,u} = 0 \\ \Longleftrightarrow\ & (X_{u+h,t} - X_{u,t}) \circ X_{s,u} = -X_{u+h,t} \circ (X_{s,u+h} - X_{s,u}) \end{aligned}$$

We obtain formally (1.6) by applying the Ito's formula to the r.h.s. in the above display, and then letting $s = u$.

The second term in (1.7) is understood as the increment of a Skorohod stochastic integral or equivalently as a two-sided stochastic integral [27] defined by the $\mathbb{L}_2$-convergence formula

$$\int_s^t (\nabla X_{u,t})\, (\overline{X}_{s,u}(x))'\ d\overline{X}_{s,u}(x) := \lim_{h \to 0} \sum_{u \in [s,t]_h} (\nabla X_{u+h,t})\, (\overline{X}_{s,u}(x))'\ (\overline{X}_{s,u+h}(x) - \overline{X}_{s,u}(x)) \tag{3.1}$$

Consider the interpolating flow

$$Z^{s,t}\ :\ u \in [s,t]\ \mapsto\ Z_u^{s,t} := X_{u,t} \circ \overline{X}_{s,u} \quad \Longrightarrow \quad Z_s^{s,t} - Z_t^{s,t} = \Delta X_{s,t}$$

In this notation, formulae (1.7) reduces to

$$\begin{aligned} d_u Z_u^{s,t}(x) = & -(\nabla X_{u,t})\, (\overline{X}_{s,u}(x))'\ \left[\Delta b_u(\overline{X}_{s,u}(x))\ du + \Delta\sigma_u(\overline{X}_{s,u}(x))\ dW_u\right] \\ & -\frac{1}{2}\ \ \left(\nabla^2 X_{u,t}\right) (\overline{X}_{s,u}(x))' \Delta a_u(\overline{X}_{s,u}(x))\ du \end{aligned}$$

from which we readily check the first assertion (1.8) in theorem 1.1.

In the same vein, for any twice differentiable function f from $\mathbb{R}^d$ into $\mathbb{R}$ we check that

$$\begin{aligned} & d_u \left[(f \circ X_{u,t})(\overline{X}_{s,u}(x))\right] \\ & = -\langle (\nabla(f \circ X_{u,t}))(\overline{X}_{s,u}(x)), \Delta b_u(\overline{X}_{s,u}(x))\rangle - \frac{1}{2}\ \mathrm{Tr}\left[(\nabla^2(f \circ X_{u,t}))(\overline{X}_{s,u}(x))\ \Delta a_u(\overline{X}_{s,u}(x))\right]\ du \\ & \qquad -(\nabla(f \circ X_{u,t}))(\overline{X}_{s,u}(x))\ \Delta\sigma_u(\overline{X}_{s,u}(x)) dW_u \end{aligned}$$

More generally, we readily check the following theorem.

Theorem 3.1. *For any $p \geqslant 1$ and any twice differentiable function f from $\mathbb{R}^d$ into $\mathbb{R}^p$ we have the forward-backward multivariate interpolation formula*

$$\mathbb{P}_{s,t}(f)(x) - \overline{\mathbb{P}}_{s,t}(f)(x) = \mathbb{T}_{s,t}(f, \Delta a, \Delta b)(x) + \mathbb{S}_{s,t}(f, \Delta\sigma)(x) \tag{3.2}$$

with the stochastic integro-differential operator

$$\begin{aligned}&\mathbb{T}_{s,t}(f, \Delta a, \Delta b)(x)\\ &:= \int_s^t \left[\nabla \mathbb{P}_{u,t}(f)(\overline{X}_{s,u}(x))' \ \Delta b_u(\overline{X}_{s,u}(x)) + \frac{1}{2}\ \nabla^2 \mathbb{P}_{u,t}(f)(\overline{X}_{s,u}(x))' \ \Delta a_u(\overline{X}_{s,u}(x))\right] du\end{aligned} \tag{3.3}$$

and the Skorohod stochastic integral term given by

$$\mathbb{S}_{s,t}(f, \Delta\sigma)(x) := \int_s^t \nabla \mathbb{P}_{u,t}(f)(\overline{X}_{s,u}(x))' \ \Delta\sigma_u(\overline{X}_{s,u}(x)) \ dW_u \tag{3.4}$$

Using elementary differential calculus, for twice differentiable (column vector-valued) function f from $\mathbb{R}^d$ into $\mathbb{R}^p$ we readily check the gradient and the Hessian formulae

$$\begin{aligned}\nabla \mathbb{P}_{s,t}(f)(x) &= \nabla X_{s,t}(x) \ \mathbb{P}_{s,t}(\nabla f)(x)\\ \nabla^2 \mathbb{P}_{s,t}(f)(x) &= [\nabla X_{s,t}(x) \otimes \nabla X_{s,t}(x)] \ \mathbb{P}_{s,t}(\nabla^2 f)(x) + \nabla^2 X_{s,t}(x) \ \mathbb{P}_{s,t}(\nabla f)(x)\end{aligned} \tag{3.5}$$

Observe that (3.2) coincides with (1.8) for the identity function; that is, we have that

$$f(x) = x \Longrightarrow \mathbb{T}_{s,t}(f, \Delta a, \Delta b) = T_{s,t}(\Delta a, \Delta b) \quad \text{and} \quad \mathbb{S}_{s,t}(f, \Delta\sigma) = S_{s,t}(\Delta\sigma)$$

The above discussion shows that the analysis of the differences of the stochastic semigroups $(\mathbb{P}_{s,t} - \overline{\mathbb{P}}_{s,t})$ in terms of the tangent and the Hessian processes is essentially the same as the one of the difference of the stochastic flows $(X_{s,t} - \overline{X}_{s,t})$. For instance, the estimates stated in theorem 1.1 can be easily extended at the level of the stochastic semigroups using the discussion provided section 5.3.

The $\mathbb{L}_2$-norm of the Skorohod stochastic integrals in (1.8) and (3.2) are uniformly estimated as soon as the pair of drift and diffusion functions (b_t, σ_t) and $(\overline{b}, \overline{\sigma}_t)$ satisfy condition $(\mathcal{T})_2$. For a more thorough discussion we refer to section 5.1, see for instance the $\mathbb{L}_n$-norm estimates presented in theorem 5.2.

3.2 Some semigroup perturbation formulae

Besides the fact that the Skorohod integral in the r.h.s. of (3.2) is not a martingale (w.r.t. the Brownian motion filtration) it is centered. Thus, taking the expectation in the univariate version of (3.2) we obtain the following interpolation semigroup decomposition.

Corollary 3.2. *For any twice differentiable function f from $\mathbb{R}^d$ into $\mathbb{R}$ with bounded derivatives we have the forward-backward semigroup interpolation formula*

$$\begin{aligned}P_{s,t}(f)(x) - \overline{P}_{s,t}(f)(x) &= \int_s^t \mathbb{E}\left(\langle \nabla P_{u,t}(f)(\overline{X}_{s,u}(x)), \Delta b_u(\overline{X}_{s,u}(x))\rangle\right) du\\ &\quad + \frac{1}{2} \int_s^t \mathbb{E}\left(\mathrm{Tr}\left[\nabla^2 P_{u,t}(f)(\overline{X}_{s,u}(x)) \ \Delta a_u(\overline{X}_{s,u}(x))\right]\right) du\end{aligned} \tag{3.6}$$

In addition, under some appropriate regularity conditions for any differentiable function f such that $\|f\| \leqslant 1$ and $\|\nabla f\| \leqslant 1$ we have the uniform estimate

$$|P_{s,t}(f)(x) - \overline{P}_{s,t}(f)(x)| \leqslant \kappa \ \left[\lvert\lvert\lvert \Delta a(x) \rvert\rvert\rvert_1 + \lvert\lvert\lvert \Delta b(x) \rvert\rvert\rvert_1\right] \tag{3.7}$$

Rewritten in terms of the infinitesimal generators $(L_t, \overline{L}_t)$ of the stochastic flows $(X_{s,t}, \overline{X}_{s,t})$ we recover the rather well known semigroup perturbation formula

$$P_{s,t} = \overline{P}_{s,t} + \int_s^t \overline{P}_{s,u}(L_u - \overline{L}_u)P_{u,t}\ du$$

Now we come to the proof of (3.7). Whenever $(\mathcal{T})_2$ is met, combining (2.12) with (2.15) for any differentiable function f s.t. $\|f\| \leqslant 1$ and $\|\nabla f\| \leqslant 1$ and for any $\epsilon \in]0,1[$ we check that

$$|P_{s,s+t}(f)(x) - \overline{P}_{s,s+t}(f)(x)| \leqslant \frac{\kappa}{\epsilon}\ \left[\vert\!\vert\!\vert \Delta a(x) \vert\!\vert\!\vert_1 + \vert\!\vert\!\vert \Delta b(x) \vert\!\vert\!\vert_1 \right] \int_0^t \frac{1}{\sqrt{u}}\ e^{-\lambda_A(1-\epsilon)u}\ du$$

This ends the proof of (3.7). ■

After some elementary manipulations the forward-backward interpolation formula (3.6) yields the following corollary.

Corollary 3.3. *Let X_t and $\overline{X}_t$ be some ergodic diffusions associated with some time homogeneous drift and diffusion functions (b, σ) and $(\overline{b}, \overline{\sigma})$. The invariant probability measures π and $\overline{\pi}$ of X_t and $\overline{X}_t$ are connected for any twice differentiable function f from $\mathbb{R}^d$ into $\mathbb{R}$ with bounded derivatives by the following interpolation formula*

$$(\pi - \overline{\pi})(f) = \int_0^\infty \mathbb{E}\left(\left\langle \nabla P_t(f)(\overline{Y}), \Delta b(\overline{Y}) \right\rangle + \frac{1}{2}\ \mathrm{Tr}\left[\nabla^2 P_t(f)(\overline{Y})\ \Delta a(\overline{Y}) \right] \right)\ dt \tag{3.8}$$

In the above display $\overline{Y}$ stands for a random variable with distribution $\overline{\pi}$ and P_t stands for the Markov transition semigroup of the process X_t.

The formula (3.8) can be used to estimate the invariant measure of a stochastic flow associated with some perturbations of the drift and the diffusion function.

For instance, for homogeneous Langevin diffusions X_t associated with some convex potential function U we have

$$b = -\nabla U \quad \text{and} \quad \sigma = I \quad \Longrightarrow \quad \pi(dx) \propto \exp\left(-2\,U(x)\right)\ dx$$

In the above display, dx stands for the Lebesgue measure on $\mathbb{R}^d$. In this situation, using (3.8), for any ergodic diffusion flow $\overline{X}_t$ with some drift $\overline{b}$ and an unit diffusion matrix we have

$$\overline{\pi}(f) = \pi(f) + \int_0^\infty \mathbb{E}\left(\left\langle (\overline{b} + \nabla U)(\overline{Y}), \nabla P_t(f)(\overline{Y}) \right\rangle \right)\ dt$$

Notice that the above formula is implicit as the r.h.s. term depends on $\overline{\pi}$. By symmetry arguments, we also have the following more explicit perturbation formula

$$\overline{\pi}(f) = \pi(f) + \int_0^\infty \mathbb{E}\left(\left\langle (\overline{b} + \nabla U)(Y), \nabla \overline{P}_t(f)(Y) \right\rangle \right)\ dt$$

In the above display Y stands for a random variable with distribution π and $\overline{P}_t$ stands for the Markov transition semigroup of the process $\overline{X}_t$.

4 Some extensions

4.1 Non Markovian perturbations

Assume that $\sigma = I$ and the regularity condition $(\mathcal{T})_2$ is met. Also suppose $\overline{X}_{s,t}(x)$ is given by a stochastic differential equation of the form (1.12) with $r = d$ and $\overline{Z}_{s,t}(x) = I$. Arguing as above, we have

$$X_{s,t}(x) - \overline{X}_{s,t}(x) = \int_s^t \left(\nabla X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)' \left(b_u(\overline{X}_{s,u}(x)) - \overline{Y}_{s,u}(x)\right) du \tag{4.1}$$

Combining (1.30) with the generalized Minkowski inequality, we check the following proposition.

Proposition 4.1. *Assume that $(\mathcal{T})_2$ is met for some $\lambda_A > 0$. In this situation, for any $1 \leqslant n \leqslant \infty$ we have the estimates*

$$\mathbb{E}\left[\|X_{s,t}(x) - \overline{X}_{s,t}(x)\|^n\right]^{1/n} \leqslant \int_s^t e^{-\lambda_A(t-u)} \ \mathbb{E}\left[\|b_u(\overline{X}_{s,u}(x)) - \overline{Y}_{s,u}(x)\|\right]^{1/n} du \tag{4.2}$$

In the same vein, we have

$$P_{s,t}(f)(x) - \overline{P}_{s,t}(f)(x) = \int_s^t \mathbb{E}\left(\langle \nabla P_{u,t}(f)(\overline{X}_{s,u}(x)), b_u(\overline{X}_{s,u}(x)) - \overline{Y}_{s,u}(x)\rangle\right) du \tag{4.3}$$

For instance, for the Langevin diffusion discussed in (1.33) and (2.6) the weak expansion (4.3) implies that

$$[\pi\overline{P}_{s,t} - \pi](f) = \int_s^t \int \pi(dx) \ \mathbb{E}\left(\langle \nabla P_{t-u}(f)(\overline{X}_{s,u}(x)), \nabla U(\overline{X}_{s,u}(x)) + \overline{Y}_{s,u}(x)\rangle\right) du \tag{4.4}$$

This yields the $\mathbb{W}_1$-Wasserstein estimate

$$\mathbb{W}_1(\pi\overline{P}_{s,t}, \pi)| \leqslant \int_s^t e^{-\lambda_A(t-u)} \int \pi(dx) \ \mathbb{E}\left(\|\nabla U(\overline{X}_{s,u}(x)) + \overline{Y}_{s,u}(x)\|\right) du$$

Combining (2.12) with (4.4), for any $\epsilon \in]0,1[$ we also have the total variation norm estimate

$$\|\pi\overline{P}_{s,t} - \pi\|_{tv} \leqslant \frac{c}{\epsilon} \int_s^t \frac{1}{\sqrt{t-u}} \ e^{-\lambda_A(1-\epsilon)(t-u)} \left[\int \pi(dx) \ \mathbb{E}\left(\|\nabla U(\overline{X}_{s,u}(x)) + \overline{Y}_{s,u}(x)\|\right)\right] du \tag{4.5}$$

4.2 Jump diffusion perturbations

Assume that $\sigma = 0$ and $b = \overline{b}$. In this situation, the Skorohod fluctuation term (1.9) reduces to the Ito stochastic integral defined in (1.14). In addition, (1.8) reduces to the interpolation formula

$$\overline{X}_{s,t}(x) - X_{s,t}(x) = \int_s^t \left[\frac{1}{2} \left(\nabla^2 X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)' \ \overline{a}_u(\overline{X}_{s,u}(x)) \ du + dM_u^{s,t}(x)\right]$$

In the above display, $u \in [s,t] \mapsto M_u^{s,t}(x)$ stands for the multivariate martingale given by the formulae

$$dM_u^{s,t}(x) := \left(\nabla X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)' \ d\overline{M}_{s,u}(x) \quad \text{with} \quad d\overline{M}_{s,u}(x) := \overline{\sigma}_u\left(X_{s,u}(x)\right) \ dW_u$$

We further assume that $\overline{X}_{s,t}(x)$ is the stochastic flow associated with a jump diffusion process adapted to some filtration $\mathcal{F}_t$ and given by an stochastic differential equation of the following form

$$d\overline{X}_{s,t}(x) = b_t\left(X_{s,t}(x)\right) \ dt + d\overline{M}_{s,t}(x) + d\overline{N}_{s,t}(x)$$

In the above display, $\overline{N}_{s,t}(x)$ stands for some martingale associated with the jump of the process; that is, we have that

$$d\overline{N}_{s,t}(x) = \Delta_t\overline{X}_{s,t}(x) - \mathbb{E}\left(\Delta_t\overline{X}_{s,t}(x) \mid \mathcal{F}_{t-}\right) \quad \text{with} \quad \Delta_t\overline{X}_{s,t}(x) := \overline{X}_{s,t}(x) - \overline{X}_{s,t-}(x)$$

At some given jump rate $\iota_t(\overline{X}_{s,t}(x))$ the càdlàg stochastic flow $\overline{X}_{s,t-}(x)$ jumps to a new location $\overline{X}_{s,t}(x) = y$ randomly chosen with some distribution denoted by $J_t(\overline{X}_{s,t}(x), dy)$.

In this context, the jumps of the interpolating process $Z_u^{s,t}$ are given by

$$\Delta_u Z_u^{s,t}(x) = X_{u,t}(\overline{X}_{s,u-}(x) + \Delta_u\overline{X}_{s,u}(x)) - X_{u,t}(\overline{X}_{s,u-}(x))$$

We denote by $u \in [s,t] \mapsto N_u^{s,t}(x)$ the multivariate martingale given by the formulae

$$dN_u^{s,t}(x) := \Delta_u Z_u^{s,t}(x) - \mathbb{E}\left(\Delta_u Z_u^{s,t}(x) \mid \mathcal{F}_{u-}\right)$$

In this notation, following the same lines of arguments as in (1.7) we check the following proposition.

Proposition 4.2. *For any $s \leqslant t$ and $x \in \mathbb{R}^d$ we have backward-forward interpolation formula*

$$\overline{X}_{s,t}(x) - X_{s,t}(x) = \int_s^t \left[\frac{1}{2}\ \left(\nabla^2 X_{u,t}\right)(\overline{X}_{s,u}(x))'\ \overline{a}_u(\overline{X}_{s,u}(x))\ du + dM_u^{s,t}(x)\right]$$

$$+\int_s^t\ \left[\partial_u\,\mathbb{E}\left(\Delta_u Z_u^{s,t}(x) - (\nabla X_{u,t})(\overline{X}_{s,u}(x))'\ \Delta_u\overline{X}_{s,u}(x) \mid \mathcal{F}_{u-}\right)\ du + dN_u^{s,t}(x)\right]$$

Using the first order Taylor expansion

$$\Delta_u Z_u^{s,t}(x) - (\nabla X_{u,t})(\overline{X}_{s,u-}(x))'\Delta_u\overline{X}_{s,u}(x)$$

$$= \int_0^1 (1-\epsilon)\ (\nabla^2 X_{u,t})(\overline{X}_{s,u-}(x) + \epsilon\ \Delta_u\overline{X}_{s,u}(x))'\ \left(\Delta_u\overline{X}_{s,u}(x)\right)^{\otimes 2}\ d\epsilon$$

we also check the decomposition

$$\overline{X}_{s,t}(x) - X_{s,t}(x) = (M_t^{s,t}(x) - M_s^{s,t}(x)) + (N_t^{s,t}(x) - N_s^{s,t}(x))$$

$$+\frac{1}{2}\ \int_s^t \left(\nabla^2 X_{u,t}\right)(\overline{X}_{s,u}(x))'\ \overline{a}_u(\overline{X}_{s,u}(x))\ du$$

$$+\int_s^t\ \int_0^1 (1-\epsilon)\ \partial_u\mathbb{E}\left((\nabla^2 X_{u,t})(\overline{X}_{s,u-}(x) + \epsilon\ \Delta_u\overline{X}_{s,u}(x))'\ \left(\Delta_u\overline{X}_{s,u}(x)\right)^{\otimes 2}\ \mid \mathcal{F}_{u-}\right)\ d\epsilon\ du$$

We further assume that $(\mathcal{T})_2$ is met. In this situation, using (1.31) the norm of the last term in the above display is almost surely bounded by

$$\kappa\ \int_s^t\ e^{-\lambda_A(t-u)}\ \iota_{u-}(\overline{X}_{s,u-}(x))\ \mathbb{E}\left(\ \|\Delta_u\overline{X}_{s,u}(x)\|^2 \mid \mathcal{F}_{u-}\right)\ du$$

In the same vein, using (1.30) we have

$$\mathbb{E}(\|N_t^{s,t}(x) - N_s^{s,t}(x)\|^2)$$

$$= \int_s^t\ \mathbb{E}\left[\iota_{u-}(\overline{X}_{s,u-}(x))\ \|X_{u,t}(\overline{X}_{s,u-}(x) + \Delta_u\overline{X}_{s,u}(x)) - X_{u,t}(\overline{X}_{s,u-}(x))\|^2\right]\ du$$

$$\leqslant \int_s^t\ e^{-2\lambda_A(t-u)}\ \mathbb{E}\left[\iota_{u-}(\overline{X}_{s,u-}(x))\ \|\ \Delta_u\overline{X}_{s,u}(x)\|^2\right]\ du$$

This yields the estimate

$$\mathbb{E}\left(\|\overline{X}_{s,t}(x)-X_{s,t}(x)\|^2\right)^{1/2} \leqslant \kappa_1\ \left(\|\|\overline{\sigma}(x)\|\|_2+\|\|\overline{a}(x)\|\|_2\right)$$

$$+\kappa_2\ \left(\int_s^t\ e^{-2\lambda_A(t-u)}\ \mathbb{E}\left(\iota_{u-}(\overline{X}_{s,u-}(x))\ \|\ \Delta_u\overline{X}_{s,u}(x)\|^2\right)\ du\right)^{1/2}$$

$$+\kappa_3\ \int_s^t\ e^{-\lambda_A(t-u)}\ \mathbb{E}\left(\iota_{u-}(\overline{X}_{s,u-}(x))^2\ \mathbb{E}\left(\ \|\Delta_u\overline{X}_{s,u}(x)\|^2\mid\mathcal{F}_{u-}\right)^2\right)^{1/2}\ du$$

For instance, whenever the jump rates and amplitudes are uniformly bounded

$$\iota_{u-}(\overline{X}_{s,u-}(x))\leqslant\kappa_0\quad\text{and}\quad\mathbb{E}\left(\ \|\Delta_u\overline{X}_{s,u}(x)\|^2\mid\mathcal{F}_{u-}\right)\leqslant\delta^2$$

for some parameters κ_0 and δ we have the uniform estimates

$$\mathbb{E}\left(\|\overline{X}_{s,t}(x)-X_{s,t}(x)\|^2\right)^{1/2}\leqslant\kappa\ \left(\delta\ (1+\delta)+\|\|\overline{\sigma}(x)\|\|_2+\|\|\overline{a}(x)\|\|_2\right)$$

5 Skorohod fluctuation processes

5.1 A variance formula

Let $\varsigma_t(x)$ be some differentiable $(d\times r)$-matrix valued function on $\mathbb{R}^d$ such that

$$\|\nabla\varsigma\|<\infty\quad\text{and}\quad\|\varsigma(0)\|:=\sup_t\|\varsigma_t(0)\|<\infty$$

We denote by $S_{s,t}(\varsigma)(x)$ the Skorohod stochastic integral defined by

$$S_{s,t}(\varsigma)(x):=\int_s^t\ \left[(\nabla X_{u,t})'\circ\overline{X}_{s,u}\right](x)\ \varsigma_u(\overline{X}_{s,u}(x))\ dW_u$$

As in (3.1), the Skorohod stochastic integral is defined by the $\mathbb{L}_2$-convergence formula

$$S_{s,t}(\varsigma)(x):=\lim_{h\to0}\sum_{u\in[s,t]_h}(\nabla X_{u+h,t})\,(\overline{X}_{s,u}(x))'\ \varsigma_u(\overline{X}_{s,u}(x))\ (W_{u+h}-W_u)\tag{5.1}$$

Observe that $(W_{u+h}-W_u)$ is independent of the flows $\overline{X}_{s,u}$ and $\nabla X_{u+h,t}$. This already implies that Skorohod stochastic integral is centered; that is, we have that $\mathbb{E}(S_{s,t}(\varsigma)(x))=0$.

As in (3.1), the variance can be computed using the following approximation formula

$$\mathbb{E}\left[\|S_{s,t}(\varsigma)(x)\|^2\right]=\lim_{h\to0}\sum_{u,v\in[s,t]_h}\ \sum_{i,j,k}$$

$$\mathbb{E}\left\{\left[(\nabla X_{u+h,t})\,(\overline{X}_{s,u}(x))'\ \varsigma_u(\overline{X}_{s,u}(x))\right]_{i,j}\ \left[(\nabla X_{v+h,t})\,(\overline{X}_{s,v}(x))'\ \varsigma_v(\overline{X}_{s,v}(x))\right]_{i,k}\right.$$

$$\left.(W^j_{u+h}-W^j_u)(W^k_{v+h}-W^k_v)\ \right\}$$

We consider the matrix valued function

$$\Sigma_{s,u,t}(x):=\left[(\nabla X_{u,t})'\circ\overline{X}_{s,u}\right](x)\ \varsigma_u(\overline{X}_{s,u}(x))\tag{5.2}$$

In this notation, the limiting diagonal term $u = v$ in the r.h.s. of (5.1) is clearly equal to

$$\int_s^t \mathbb{E}\left[\sum_{i,j} \Sigma_{s,u,t}(x)_{i,j}\ \Sigma_{s,u,t}(x)_{i,j}\right]\ du = \int_s^t \mathbb{E}\left[\|\Sigma_{s,u,t}(x)\|_F^2\right]\ du$$

In addition, whenever condition $(\mathcal{T})_2$ is met and ς is bounded, (2.2) readily yields the estimate

$$\left[\int_s^t \mathbb{E}\left[\|\Sigma_{s,u,t}(x)\|_F^2\right]\ du\right]^{1/2} \leqslant \|\varsigma\|_2\ \sqrt{d/(2\lambda_A)} \tag{5.3}$$

More generally, using (2.7) whenever $(\overline{\mathcal{M}})_{2/\delta}$ and $(\mathcal{T})_{2/(1-\delta)}$ are met for some $\delta \in]0,1[$ we have the estimate

$$\mathbb{E}\left[\|\Sigma_{s,u,t}(x)\|^2\right] \leqslant c_{1,\delta}\ \left[\|\varsigma(0)\|^2 + \|\nabla\varsigma\|^2\ (1+\|x\|)^2\right]\ e^{-2\lambda_A(2/(1-\delta))(t-u)}$$

This implies that

$$\left[\int_s^t \mathbb{E}\left[\|\Sigma_{s,u,t}(x)\|_F^2\right]\ du\right]^{1/2} \leqslant\ c_{2,\delta}\ \left[\|\varsigma(0)\| + \|\nabla\varsigma\|\ (1+\|x\|)\right]/\sqrt{\lambda_A} \tag{5.4}$$

The non-diagonal term can be computed in a more direct way using Malliavin derivatives. It is clearly out of the scope of this article to review the analytical construction of Malliavin differential calculus. Formally, when $u \leqslant v$ one can think the Malliavin derivatives $D_v^i X_{u,v}(x)$ as way to extract from the random variable $X_{u,v}(x)$ the integrand of Brownian increment dW_v^i. To simplify the presentation we further assume that $d = r$ and $\varsigma_t(x)$ are symmetric matrices.

For instance, we shall adopt the following definition

$$X_{u,v}(x) = x + \int_u^v b_\tau(X_{u,\tau}(x))\ d\tau + \sum_i\ \int_u^v \sigma_{\tau,i}(X_{u,\tau}(x))\ dW_\tau^i$$

$$\Longrightarrow D_v^i X_{u,v}(x) := \sigma_{v,i}(X_{u,v}(x))$$

In the same vein, we have

$$\nabla X_{u,v}(x) = I + \int_u^v \nabla X_{u,\tau}(x)\ \nabla b_\tau\left(X_{u,\tau}(x)\right)\ d\tau + \sum_i \int_u^v \nabla X_{u,\tau}(x)\ \nabla\sigma_{\tau,i}\left(X_{u,\tau}(x)\right)\ dW_\tau^i$$

$$\Longrightarrow D_v^i \nabla X_{u,v}(x) := \nabla X_{u,v}(x)\ \nabla\sigma_{v,i}(X_{u,v}(x))$$

In a more synthetic way, the D_v^i-Malliavin derivatives of $X_{u,v}$ and $\nabla X_{u,v}$ are given by the random tensor functions

$$\begin{aligned}(D_v X_{u,v})_{i,j} &= (D_v^i X_{u,v})_j = D_v^i X_{u,v}^j := \sigma_{i,v}^j \circ X_{u,v} \\ (D_v \nabla X_{u,v})_{i,j,k} &= D_v^i\left(\nabla X_{u,v}\right)_{j,k} \\ &:= \left(\nabla X_{u,v}\ \left((\nabla\sigma_{i,v}) \circ X_{u,v}\right)\right)_{j,k} = D_v^i\left(\nabla X_{u,v}\right)'_{k,j} = \left(D_v\left(\nabla X_{u,v}\right)'\right)_{i,k,j}\end{aligned} \tag{5.5}$$

As conventional differentials, Malliavin derivatives satisfy the chain rule properties

$$\begin{aligned}D_u\left(X_{u,v} \circ X_{s,u}\right) &:= (D_u X_{s,u})\ \left[(\nabla X_{u,v}) \circ X_{s,u}\right] \\ D_u\left(\varsigma_v \circ X_{s,v}\right) &= (D_u X_{s,v})\ \left[(\nabla\varsigma_v) \circ X_{s,v}\right]\end{aligned} \tag{5.6}$$

In the same vein, we have the chain rule formulae

$$\begin{aligned} &D_v\left(\nabla X_{u,v}\ \left[(\nabla X_{v,t})\circ X_{u,v}\right]\right)\\ &=(D_v\nabla X_{u,v})\left[(\nabla X_{v,t})\circ X_{u,v}\right]+(D_vX_{u,v}\otimes\nabla X_{u,v})\left[\left(\nabla^2X_{v,t}\right)\circ X_{u,v}\right]\end{aligned}\tag{5.7}$$

as well as

$$D_v\left\{\left[(\nabla X_{u,t})'\circ\overline{X}_{s,u}\right]\ \left[\varsigma_u\circ\overline{X}_{s,u}\right]\right\}=\left[\left(D_v\left(\nabla X_{u,t}\right)'\right)\circ\overline{X}_{s,u}\right]\ \left[\varsigma_u\circ\overline{X}_{s,u}\right]\tag{5.8}$$

More detailed proofs of (5.6) and (5.7) are provided in the appendix, on page 32. Observe that

$$\nabla\sigma=0\quad\Longrightarrow\quad D_v\Sigma_{s,u,t}(x)=0$$

In the reverse angle, whenever $v\leqslant u$ we have the chain rule formula

$$\begin{aligned} &D_v\left(\left[\varsigma_u\circ\overline{X}_{s,u}\right]\ \left[(\nabla X_{u,t})\circ\overline{X}_{s,u}\right]\right)\\ &:=\left[D_v\left(\varsigma_u\circ\overline{X}_{s,u}\right)\right]\left[(\nabla X_{u,t})\circ\overline{X}_{s,u}\right]+\left[D_v\overline{X}_{s,u}\otimes\left(\varsigma_u\circ\overline{X}_{s,u}\right)\right]\left[\left(\nabla^2X_{u,t}\right)\circ\overline{X}_{s,u}\right]\end{aligned}\tag{5.9}$$

As above, Malliavin differentials $D_v\left(\varsigma_u\circ\overline{X}_{s,u}\right)$ and $D_v\overline{X}_{s,u}$ can be computed using the chain rule formulae (5.6). A more detailed and constructive proof of (5.9) is provided in the appendix, on page 33. Observe that

$$\nabla\varsigma=0\quad\Longrightarrow\quad D_v\left[\Sigma'_{s,u,t}\right]=\left[D_v\overline{X}_{s,u}\otimes\left(\varsigma_u\circ\overline{X}_{s,u}\right)\right]\left[\left(\nabla^2X_{u,t}\right)\circ\overline{X}_{s,u}\right]$$

We consider the inner product

$$\langle D_u\Sigma_{s,v,t}(x),D_v\Sigma_{s,u,t}(x)\rangle\ :=\ \sum_{i,j,k}\left(D_v\Sigma_{s,u,t}(x)\right)_{k,i,j}\ \left(D_u\Sigma_{s,v,t}(x)\right)_{j,i,k}$$

In this notation, an explicit description of the $\mathbb{L}_2$-norm of the Skorohod integral in terms of Malliavin derivatives is given below.

Lemma 5.1. *The $\mathbb{L}_2$-norm of the Skorohod integral $S_{s,t}(\varsigma)(x)$ introduced in (5.1) is given for any $x\in\mathbb{R}^d$ and $s\leqslant t$ by the formulae*

$$\mathbb{E}\left[\|S_{s,t}(\varsigma)(x)\|^2\right]=\int_{[s,t]}\mathbb{E}\left[\|\Sigma_{s,u,t}(x)\|_F^2\right]\ du+\int_{[s,t]^2}\mathbb{E}\left[\langle D_v\Sigma_{s,u,t}(x),D_u\Sigma_{s,v,t}(x)\rangle\right]\ du\ dv$$

with the random matrix function $\Sigma_{s,u,t}$ defined in (5.2) and the Malliavin derivative $D_v\Sigma_{s,u,t}$ given in formulae (5.8) and (5.9). In addition, we have

$$\nabla\sigma=0\quad\Longrightarrow\quad\mathbb{E}\left[\|S_{s,t}(\varsigma)(x)\|^2\right]=\int_{[s,t]}\mathbb{E}\left[\|\Sigma_{s,u,t}(x)\|_F^2\right]\ du$$

The above lemma can be seen as an extended version of the Ito isometry to Skorohod integrals. The above formulation of the $\mathbb{L}_2$-norm of Skorohod integrals in terms of Malliavin derivatives of the integrands is rather well known in Malliavin theory literature, see for instance [26], as well as chapters 1.3 to 1.5 in the seminal book by Nualart [24]. A constructive proof of the above lemma based on the $\mathbb{L}_2$-approximation of two-sided stochastic integrals is provided in the appendix on page 34.

5.2 Some quantitative estimates

For any $p > 1$ and any tensor norms we also quote the rather well known $\mathbb{L}_p$-norm estimates

$$\begin{aligned}
&\mathbb{E}\left[\|S_{s,t}(\varsigma)(x)\|^p\right]^{2/p} \\
&\leqslant c_{1,p} \int_{[s,t]} \mathbb{E}\left[\|\Sigma_{s,u,t}(x)\|^2\right]\ du + c_{2,p}\ \mathbb{E}\left[\left(\left(\int_{[s,t]^2} \|D_v\Sigma_{s,u,t}(x)\|^2\ du\ dv\right)^{p/2}\right]^{2/p}
\end{aligned}$$

for some finite constants $c_{i,p}$ whose values only depends on p. A proof of these estimates can be found in [25, 30], see also [26] for multiple Skorohod integrals. By the generalized Minkowski inequality, for any $n \geqslant 2$ we also have the estimate

$$\begin{aligned}
&\mathbb{E}\left[\|S_{s,t}(\varsigma)(x)\|^n\right]^{2/n} \\
&\leqslant c_{1,n} \int_{[s,t]} \mathbb{E}\left[\|\Sigma_{s,u,t}(x)\|^2\right]\ du + c_{2,n} \int_{[s,t]^2} \mathbb{E}\left[\|D_v\Sigma_{s,u,t}(x)\|^n\right]^{2/n}\ du\ dv
\end{aligned} \tag{5.10}$$

Observe that for any $n \geqslant 2$ we have

$$(\overline{\mathcal{M}})_n \Longrightarrow \vert\!\vert\!\vert \varsigma(x) \vert\!\vert\!\vert_n \leqslant \kappa_n\ \left(\|\varsigma(0)\| + \|\nabla\varsigma\|\right)(1 \vee \|x\|)$$

The main objective of this section is to prove the following theorem.

Theorem 5.2. *Assume that $(M)_{2n/\delta}$ and $(T)_{2n/(1-\delta)}$ are satisfied for some parameter $n \geqslant 2$ and some $\delta \in]0,1[$. In this situation, we have the uniform estimate*

$$\mathbb{E}\left[\|S_{s,t}(\varsigma)(x)\|^n\right]^{1/n} \leqslant \kappa_{\delta,n}\ \vert\!\vert\!\vert \varsigma(x) \vert\!\vert\!\vert_{2n/\delta}\ (1 \vee \|x\|) \tag{5.11}$$

For uniformly bounded diffusion functions $(\varsigma, \sigma, \overline{\sigma})$ whenever $(T)_{2n}$ is met for some $n \geqslant 2$ we have

$$\mathbb{E}\left[\|S_{s,t}(\varsigma)(x)\|^n\right]^{1/n} \leqslant \kappa_n\ \left(\|\varsigma\| + \|\nabla\varsigma\|\right) \tag{5.12}$$

In addition, for constant diffusion functions $(\varsigma, \sigma, \overline{\sigma})$ whenever $(T)_2$ is met, for any $n \geqslant 2$ we have the uniform estimate

$$\mathbb{E}\left[\|S_{s,t}(\varsigma)(x)\|^n\right]^{1/n} \leqslant \kappa_n\ \|\varsigma\| \tag{5.13}$$

The proof of the above theorem, including a more detailed description of the parameters $\kappa_{\delta,n}$ and κ_n is provided below.

Next, we estimate the $\mathbb{L}_n$-norm of the Malliavin differential $D_v\Sigma_{s,u,t}(x)$ in the two cases $(s \leqslant u \leqslant v \leqslant t)$ and $(s \leqslant v \leqslant u \leqslant t)$.

Case $(s \leqslant u \leqslant v \leqslant t)$:

Using (5.8) we have

$$\|D_v\Sigma_{s,u,t}(x)\| \leqslant c\ \|\varsigma_u(\overline{X}_{s,u}(x))\|\ \|(D_v\nabla X_{u,t})(\overline{X}_{s,u}(x))\|$$

Using (5.7) this yields the estimate

$$\|D_v\Sigma_{s,u,t}(x)\| \leqslant c_1\ \mathbb{I}_{s,u,t}(x) + c_2\ \mathbb{J}_{s,u,t}(x)$$

with the functions

$$\begin{aligned}
\mathbb{I}_{s,u,t}(x) &:= \|\nabla\sigma\|\ \|\varsigma_u(\overline{X}_{s,u}(x))\|\ \|(\nabla X_{u,v})(\overline{X}_{s,u}(x))\|\ \|(\nabla X_{v,t})(Z_u^{s,v}(x))\| \\
\mathbb{J}_{s,u,t}(x) &:= \|\sigma_v(Z_u^{s,v}(x))\|\ \|\varsigma_u(\overline{X}_{s,u}(x))\|\ \|(\nabla X_{u,v})(\overline{X}_{s,u}(x))\|\ \|(\nabla^2 X_{v,t})(Z_u^{s,v}(x))\|
\end{aligned}$$

- Firstly assume that $\|\varsigma\|\vee\|\sigma\|<\infty$ and $(\mathcal{T})_{2n}$ is satisfied for some parameter $n\geqslant 1$. In this situation, applying proposition 2.2 and proposition 2.3, for any $\epsilon\in]0,1[$ we have the uniform estimates

$$\mathbb{E}\left(\|D_v\Sigma_{s,u,t}(x)\|^n\right)^{1/n}\leqslant\|\varsigma\|\ \chi_{n,\epsilon}(b,\sigma)\ \exp\left(-(1-\epsilon)\lambda_A(2n)(t-u)\right)$$

with the parameter $\chi_{n,\epsilon}(b,\sigma)$ given by

$$\chi_{n,\epsilon}(b,\sigma):=c\ [\|\sigma\|\vee\|\nabla\sigma\|]\left[1+\frac{1}{\epsilon}\ \frac{n}{\lambda_A(2n)}\ \chi(b,\sigma)\right]\quad\text{ with }\chi(b,\sigma)\text{ given in (1.20).}$$

- More generally, when $\|\nabla\varsigma\|\vee\|\nabla\sigma\|<\infty$ the functions $\varsigma_t(x)$ and $\sigma_t(x)$ may grow at the most linearly with respect to $\|x\|$. Assume that conditions $(M)_{2n/\delta}$ and condition $(\mathcal{T})_{2n/(1-\delta)}$ are satisfied for some parameters $n\geqslant 1$ and $\delta\in]0,1[$. In this situation, applying Hölder inequality we check that

$$\mathbb{E}\left(\|\mathbb{I}_{s,u,t}(x)\|^n\right)^{1/n}\leqslant c\ \|\nabla\sigma\|\ \mathbb{E}\left(\|\varsigma_u(\overline{X}_{s,u}(x))\|^{n/\delta}\right)^{\delta/n}$$

$$\times\mathbb{E}\left(\|(\nabla X_{u,v})(\overline{X}_{s,u}(x))\|^{2n/(1-\delta)}\right)^{(1-\delta)/(2n)}\ \mathbb{E}\left(\|(\nabla X_{v,t})(Z_u^{s,v}(x))\|^{2n/(1-\delta)}\right)^{(1-\delta)/(2n)}$$

Applying proposition 2.2 we check that

$$\mathbb{E}\left(\|\mathbb{I}_{s,u,t}(x)\|^n\right)^{1/n}\leqslant c_{n,\delta}\ \|\nabla\sigma\|\ \left|\!\left|\!\left|\varsigma(x)\right|\!\right|\!\right|_{n/\delta}\ e^{-\lambda_A(2n/(1-\delta))(t-u)}$$

In the same vein, combining proposition 2.2 and proposition 2.3 with the uniform moment estimates (1.25) we check that

$$\begin{aligned}\mathbb{E}\left(\|\mathbb{J}_{s,u,t}(x)\|^n\right)^{1/n}\leqslant c_{n,\delta}\ [\|\sigma(0)\|+\|\nabla\sigma\|]\ \frac{1}{\epsilon}\ \frac{\chi(b,\sigma)}{\lambda_A(2n/(1-\delta))}\\ \left|\!\left|\!\left|\varsigma(x)\right|\!\right|\!\right|_{2n/\delta}\ \ [1+\|x\|]\ \ e^{-(1-\epsilon)\lambda_A(2n/(1-\delta))(t-u)}\end{aligned}$$

We conclude that

$$\mathbb{E}\left(\|D_v\Sigma_{s,u,t}(x)\|^n\right)^{1/n}\leqslant\chi_{n,\delta,\epsilon}(b,\sigma)\ \left|\!\left|\!\left|\varsigma(x)\right|\!\right|\!\right|_{2n/\delta}\ [1+\|x\|]\ \ e^{-(1-\epsilon)\lambda_A(2n/(1-\delta))(t-u)}$$

with the parameter

$$\chi_{n,\delta,\epsilon}(b,\sigma):=c_{n,\delta}\ [\|\sigma(0)\|+\|\nabla\sigma\|]\ \left(1+\frac{1}{\epsilon}\ \frac{\chi(b,\sigma)}{\lambda_A(2n/(1-\delta))}\right)$$

Case $(s\leqslant v\leqslant u\leqslant t)$**:**

We use (5.9) to check that

$$\begin{aligned}\|D_v\Sigma_{s,u,t}(x)\|\leqslant\|[D_v\left(\varsigma_u\circ\overline{X}_{s,u}\right)](x)\|\ \|\left(\nabla X_{u,t}\right)(\overline{X}_{s,u}(x))\|\\ +\|[D_v\overline{X}_{s,u}](x)\otimes\varsigma_u(\overline{X}_{s,u})(x)\|\ \|\left(\nabla^2X_{u,t}\right)(\overline{X}_{s,u}(x))\|\end{aligned}$$

On the other hand, using the chain rules (5.6) we have

$$\begin{array}{rcl}D_v\overline{X}_{s,u}&:=&\left(D_v\overline{X}_{s,v}\right)\ \left[\left(\nabla\overline{X}_{v,u}\right)\circ\overline{X}_{s,v}\right]\\ D_v\left(\varsigma_u\circ\overline{X}_{s,u}\right)&=&\left(D_v\overline{X}_{s,u}\right)\ \left[\left(\nabla\varsigma_u\right)\circ\overline{X}_{s,u}\right]\end{array}$$

This yields the estimate

$$\|D_v\Sigma_{s,u,t}(x)\| \leqslant c_1\ \|\overline{\sigma}_v(\overline{X}_{s,v}(x))\|\ \|\nabla\varsigma\|\ \|(\nabla\overline{X}_{v,u})(\overline{X}_{s,v}(x))\|\ \|\left(\nabla X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)\|$$

$$+c_2\ \|\overline{\sigma}_v(\overline{X}_{s,v}(x))\|\ \|\varsigma_u(\overline{X}_{s,u}(x))\|\ \|(\nabla\overline{X}_{v,u})(\overline{X}_{s,v}(x))\|\ \|\left(\nabla^2 X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)\|$$

- Firstly assume that $\|\varsigma\| \vee \|\overline{\sigma}\| < \infty$ and condition $(T)_{2n}$ is satisfied for some $n \geqslant 1$. In this situation, arguing as above for any $\epsilon \in]0,1[$ we have the uniform estimates

$$\mathbb{E}\left(\|D_v\Sigma_{s,u,t}(x)\|^n\right)^{1/n} \leqslant \left(\|\varsigma\| + \|\nabla\varsigma\|\right)\ \overline{\chi}_{n,\epsilon}(b,\sigma)\ \exp\left(-(1-\epsilon)\lambda_{A,\overline{A}}(2n)(t-v)\right)$$

  for some universal constant c and the parameter $\overline{\chi}_{n,\epsilon}(b,\sigma)$ given by

$$\overline{\chi}_{n,\epsilon}(b,\sigma) := c\ \|\overline{\sigma}\|\ \left[1 + \frac{1}{\epsilon}\ \frac{n}{\lambda_{A,\overline{A}}(2n)}\ \chi(b,\sigma)\right] \quad \text{with } \chi(b,\sigma) \text{ given in (1.20).}$$

- More generally assume that $\|\nabla\varsigma\| \vee \|\nabla\overline{\sigma}\| < \infty$. Also assume that conditions $(M)_{2n/\delta}$ and $(T)_{2n/(1-\delta)}$ are satisfied for some parameters $n \geqslant 1$ and $\delta \in]0,1[$. In this situation, we have

$$\mathbb{E}\left(\|D_v\Sigma_{s,u,t}(x)\|^n\right)^{1/n}$$

$$\leqslant \chi_{n,\delta,\epsilon}(b,\sigma,\overline{\sigma})\ \|\|\varsigma(x)\|\|_{2n/\delta}\ \left[1+\|x\|\right]\ \ e^{-(1-\epsilon)\lambda_{A,\overline{A}}(2n/(1-\delta))(t-v)}$$

  with the parameter

$$\chi_{n,\delta,\epsilon}(b,\sigma,\overline{\sigma}) := c_{n,\delta}\ \left[\|\overline{\sigma}(0)\| + \|\nabla\overline{\sigma}\|\right]\ \left(1 + \frac{1}{\epsilon}\ \frac{\chi(b,\sigma)}{\lambda_{A,\overline{A}}(2n/(1-\delta))}\right)$$

The end of the proof of theorem 5.2 is a direct consequence of the estimates discussed above combined with (5.10) and the diagonal estimates presented in (5.3). ∎

5.3 Some extensions

This section is concerned with the Skorohod stochastic integral (3.4). Using the gradient formula in (3.5) the Skorohod stochastic integral in (3.4) takes the form

$$\mathbb{S}_{s,t}(f,\Delta\sigma)(x) = \int_s^t\ \Sigma_{s,u,t}(f)(x)\ dW_u$$

with the Skorohod integrands

$$\Sigma_{s,u,t}(f)(x) := \nabla f(Z_u^{s,t}(x))'\ \Sigma_{s,u,t}(x) \quad \text{and} \quad \Sigma_{s,u,t}(x) := \left[(\nabla X_{u,t})' \circ \overline{X}_{s,u}\right]\ \left[\Delta\sigma_u \circ \overline{X}_{s,u}\right]$$

As in (5.7), using the chain rule properties of Malliavin derivatives we check that

$$D_v^i\Sigma_{s,u,t}(f) = \left(D_v^i\nabla f(Z_u^{s,t})'\right)\ \Sigma_{s,u,t} + \nabla f(Z_u^{s,t})'\ D_v^i\Sigma_{s,u,t}$$

as well as

$$D_v^i\nabla f(Z_u^{s,t})' = \nabla^2 f(Z_u^{s,t})'\ D_v^i Z_u^{s,t}$$

This yields the differential formula

$$D^i_v \Sigma_{s,u,t}(f) = \nabla f(Z^{s,t}_u)' \; D^i_v \Sigma_{s,u,t} + \nabla^2 f(Z^{s,t}_u)' \; (D^i_v Z^{s,t}_u) \; \Sigma_{s,u,t}$$

The Malliavin derivatives $D^i_v \Sigma_{s,u,t}$ are computed using formulae (5.8) and (5.9); thus, it remains to compute the Malliavin derivatives $D_v Z^{s,t}_u$ of the interpolating path.

- When $u \leqslant v$ we have

$$Z^{s,t}_u = (X_{v,t} \circ X_{u,v}) \circ \overline{X}_{s,u} = X_{v,t} \circ Z^{s,v}_u$$

In this situation, as in (5.6) using the chain rule properties of Malliavin derivatives we check that

$$D_v Z^{s,t}_u = D_v Z^{s,v}_u \; ((\nabla X_{v,t}) \circ Z^{s,v}_u) = ((D_v X_{u,v}) \circ \overline{X}_{s,u}) \; ((\nabla X_{v,t}) \circ Z^{s,v}_u)$$

By (5.5) we conclude that

$$D_v Z^{s,t}_u = (\sigma_v \circ Z^{s,v}_u) \; ((\nabla X_{v,t}) \circ Z^{s,v}_u)$$

- When $v \leqslant u$ we have

$$Z^{s,t}_u = X_{u,t} \circ (\overline{X}_{v,u} \circ \overline{X}_{s,v}) = Z^{v,t}_u \circ \overline{X}_{s,v}$$

In this situation, arguing as above we check that

$$D_v Z^{s,t}_u = D_v \overline{X}_{s,v} \; ((\nabla Z^{v,t}_u) \circ \overline{X}_{s,v}) = D_v \overline{X}_{s,v} \; ((\nabla \overline{X}_{v,u}) \circ \overline{X}_{s,v}) \; ((\nabla X_{u,t}) \circ \overline{X}_{s,u})$$

By (5.5) we conclude that

$$D_v Z^{s,t}_u = (\overline{\sigma}_v \circ \overline{X}_{s,v}) \; ((\nabla \overline{X}_{v,u}) \circ \overline{X}_{s,v}) \; ((\nabla X_{u,t}) \circ \overline{X}_{s,u})$$

6 Some illustrations

6.1 Perturbation analysis

Assume that $\overline{\sigma} = \sigma$ and the drift function $\overline{b}_t$ is given by a first order expansion

$$\overline{b}_t(x) = b_{\delta,t}(x) := b_t(x) + \delta \; b^{(1)}_{\delta,t}(x) \quad \text{with} \quad b^{(1)}_{\delta,t}(x) = b^{(1)}_t(x) + \frac{\delta}{2} \; b^{(2)}_{\delta,t}(x)$$

for some perturbation parameter $\delta \in [0,1]$ and some functions $b^{(i)}_{\delta,t}(x)$ with $i = 1,2$.

In this context, the stochastic flow $\overline{X}_{s,t}(x) := X^\delta_{s,t}(x)$ can be seen as a δ-perturbation of $X_{s,t}(x) := X^0_{s,t}(x)$.

We further assume that the unperturbed diffusion satisfies condition $(\mathcal{T})_2$.

To avoid unnecessary technical discussions on the existence of absolute moments of the flows we also assume that $b^{(i)}_{\delta,t}(x)$ are uniformly bounded w.r.t. the parameters (δ,t,x). In addition, $b^{(1)}_t(x)$ is differentiable w.r.t. the coordinate x and it has uniformly bounded gradients. In this situation, we set

$$\|b^{(i)}\| := \sup_{\delta,t,x} \|b^{(i)}_{\delta,t}(x)\| \quad \text{and} \quad \|\nabla b^{(1)}\| := \sup_{t,x} \|\nabla b^{(1)}_t(x)\|$$

With some additional work to estimate the absolute moments of the flows, the perturbation analysis presented below allows to handle more general models. The methodology described in this section can also be extended to expand the flow $X^\delta_{s,t}(x)$ at any order as soon as $\delta \mapsto b_{\delta,t}(x)$ is sufficiently smooth.

The first order approximation is given by the following theorem.

Theorem 6.1. *For any $s \leqslant t$, $x \in \mathbb{R}^d$ and $\delta \geqslant 0$ we have the first order expansion*

$$X^{\delta}_{s,t}(x) = X_{s,t}(x) + \delta\ \partial X_{s,t}(x) + \frac{\delta^2}{2}\ \partial^2_{\delta} X_{s,t}(x) \tag{6.1}$$

with the first order stochastic flow

$$\partial X_{s,t}(x) := \int_s^t\ (\nabla X_{u,t})\,(X_{s,u}(x))'\ b^{(1)}_u(X_{s,u}(x))\ du$$

The remainder second order term $\partial^2_{\delta} X_{s,t}(x)$ in the above display is such that for any $n \geqslant 2$ s.t. $\lambda_A(n) > 0$ we have the uniform estimate

$$\sup_{s,t,x} \mathbb{E}[\|\partial^2_{\delta} X_{s,t}(x)\|^n]^{1/n} \leqslant c_n$$

Observe that the first order stochastic flow satisfies the diffusion equation

$$d\ \partial X_{s,t}(x) = \left[b^{(1)}_t(X_{s,t}(x)) + \nabla b_t\,(X_{s,t}(x))'\,\partial X_{s,t}(x)\right]\ dt + \sum_{1\leqslant k\leqslant r} \nabla\sigma_{t,k}\,(X_{s,t}(x))'\,\partial X_{s,t}(x)\ dW^k_t$$

The rest of this section is dedicated to the proof of the above theorem.

Using (4.1) we readily check that

$$DX^{\delta}_{s,t}(x) \ := \ \delta^{-1}[X^{\delta}_{s,t}(x) - X_{s,t}(x)] = \int_s^t\ (\nabla X_{u,t})\,(X^{\delta}_{s,u}(x))'\ b^{(1)}_{\delta,u}(X^{\delta}_{s,u}(x))\ du$$

By proposition 2.2 for any $n \geqslant 2$ we have

$$\lambda^+_A(n) := \lambda_A - (n-2)\rho(\nabla\sigma)^2/2 > 0 \Longrightarrow \mathbb{E}\left(\|DX^{\delta}_{s,t}(x)\|^n\right)^{1/n} \leqslant c\ \|b^{(1)}\|/\lambda^+_A(n) \tag{6.2}$$

This yields the first order Taylor expansion (6.1) with

$$\partial^2_{\delta} X_{s,t}(x) := R^{(2,1)}_{\delta,s,t}(x) + R^{(2,2)}_{\delta,s,t}(x)$$

and the second order remainder terms

$$\begin{aligned}
\partial^{(2,2)}_{\delta} X_{s,t}(x) \ &:= \ \int_s^t\ (\nabla X_{u,t})\,(X^{\delta}_{s,u}(x))'\ b^{(2)}_{\delta,t}(X^{\delta}_{s,u}(x))\ du \\
\partial^{(2,1)}_{\delta} X_{s,t}(x) \ &:= \ 2\delta^{-1}\int_s^t\ \left[(\nabla X_{u,t})\,(X^{\delta}_{s,u}(x)) - (\nabla X_{u,t})\,(X_{s,u}(x))\right]'\ b^{(1)}_u(X^{\delta}_{s,u}(x))\ du \\
&\qquad + 2\delta^{-1}\int_s^t\ (\nabla X_{u,t})\,(X_{s,u}(x))'\ [b^{(1)}_u(X^{\delta}_{s,u}(x)) - b^{(1)}_u(X_{s,u}(x))]\ du
\end{aligned}$$

Arguing as above, for any $n \geqslant 2$ s.t. $\lambda^+_A(n) > 0$ we have the uniform estimate

$$\mathbb{E}\left(\|\partial^{(2,2)}_{\delta} X_{s,t}(x)\|^n\right)^{1/n} \leqslant c\ \|b^{(2)}\|/\lambda^+_A(n)$$

To estimate $\partial^{(2,1)}_{\delta} X_{s,t}(x)$ we need to consider the second order decompositions

$$\begin{aligned}
&2^{-1}\ \partial^{(2,1)}_{\delta} X_{s,t}(x) \\
&= \int_0^1 \int_s^t \left[\nabla^2 X_{u,t}\right] \left(X_{s,u}(x) + \epsilon(X^{\delta}_{s,u}(y) - X_{s,u}(x))\right)'\ \left[b^{(1)}_u(X^{\delta}_{s,u}(x)) \otimes DX^{\delta}_{s,u}(x)\right]\ du\ d\epsilon \\
&+ \int_0^1 \int_s^t\ (\nabla X_{u,t})\,(X_{s,u}(x))'\ \nabla b^{(1)}_u \left(X_{s,u}(x) + \epsilon(X^{\delta}_{s,u}(x) - X_{s,u}(x)), y\right)'\ DX^{\delta}_{s,u}(x)\ du\ d\epsilon
\end{aligned}$$

Combining proposition 2.3 with the estimate (6.2) for any $n \geqslant 2$ s.t. $\lambda_A(n) > 0$ we check that

$$\mathbb{E}[\|\partial_\delta^{(2,1)} X_{s,t}(x)\|^n]^{1/n} \leqslant c\ \left(1 + n\ \chi(b,\sigma)/\lambda_A(n)\right)\ \left(\|b^{(1)}\|/\lambda_A(n)\right)^2$$

for some universal constant $c < \infty$ and the parameter $\chi(b,\sigma)$ introduced in (1.20). This ends the proof of (6.1). ■

6.2 Interacting diffusions

Consider a system of N interacting and $\mathbb{R}^d$-valued diffusion flows $X^i_{s,t}(x)$, with $1 \leqslant i \leqslant N$ given by a stochastic differential equation of the form

$$dX^i_{s,t}(x) = B_t\left(X^i_{s,t}(x), \frac{1}{N}\sum_{1\leqslant i\leqslant N} X^j_{s,t}(x)\right)\ dt + \sigma_t\left(\frac{1}{N}\sum_{1\leqslant i\leqslant N} X^j_{s,t}(x)\right)\ dW^i_t$$

for some Lipschitz functions $B_t(x,y)$ and $\sigma_t(y)$ with appropriate dimensions. In the above display, W^i_t stands for a collection of independent copies of d-dimensional Brownian motion W_t. Assume that $B_t(x,y)$ linear w.r.t. the first coordinate.

In this situation, up to a change or probability space, the empirical mean of the process

$$\overline{X}_{s,t}(x) := \frac{1}{N}\sum_{1\leqslant i\leqslant N} Y^j_{s,t}(x)$$

satisfies the stochastic differential equation

$$d\overline{X}_{s,t}(x) = b_t\left(\overline{X}_{s,t}(x)\right)\ dt + \frac{1}{\sqrt{N}}\ \sigma_t\left(\overline{X}_{s,t}(x)\right)\ dW_t \quad \text{with} \quad b_t(x) := B_t(x,x)$$

Formally, the above diffusion converges as $N \to \infty$ to the flow of the dynamical system

$$\partial_t X_{s,t}(x) = b_t\left(X_{s,t}(x)\right)$$

Without further work, the forward-backward interpolation formula (1.8) yields the bias-variance error decomposition

$$\begin{aligned}\overline{X}_{s,t}(x) - X_{s,t}(x) &= \frac{1}{2N}\int_s^t \left(\nabla^2 X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)'\ a_u(\overline{X}_{s,u}(x))\ du\\ &\qquad + \frac{1}{\sqrt{N}}\int_s^t\ \left(\nabla X_{u,t}\right)\left(\overline{X}_{s,u}(x)\right)'\ \sigma_u(\overline{X}_{s,u}(x))\ dW_u\end{aligned}$$

After some elementary manipulations we check that

$$\lim_{N\to\infty} N\ \left[\mathbb{E}(\overline{X}_{s,t}(x)) - X_{s,t}(x)\right] = \frac{1}{2}\int_s^t \left(\nabla^2 X_{u,t}\right)\left(X_{s,u}(x)\right)'\ a_u(X_{s,u}(x))\ du$$

We also have the almost sure fluctuation theorem

$$\lim_{N\to\infty} \sqrt{N}\ \left[\overline{X}_{s,t}(x) - X_{s,t}(x)\right] = \int_s^t\ \left(\nabla X_{u,t}\right)\left(X_{s,u}(x)\right)'\ \sigma_u(X_{s,u}(x))\ dW_u$$

6.3 Time discretisation

We fix some parameter $h > 0$ and some $s \geqslant 0$ and for any $t \in [s + kh, s + (k+1)h[$ we set

$$dX^h_{s,t}(x) = Y^h_{s,t}(x)\ dt + \sigma\ dW_t \quad \text{with} \quad Y^h_{s,t}(x) := b\left(X^h_{s,s+kh}(x)\right)$$

for some fluctuation parameter $\sigma \geqslant 0$. For any $s + kh \leqslant u < s + (k+1)h$ we have

$$X^h_{s,u}(x) - X^h_{s,s+kh}(x) = Y^h_{s,u}(x)\ (u - (s+kh)) + \sigma\ (W_u - W_{s+kh})$$

Using (4.1) we readily check that

$$X^h_{s,t}(x) - X_{s,t}(x) = \int_s^t \ (\nabla X_{u,t})\,(X^h_{s,u}(x))' \ \left[Y^h_{s,u}(x) - b(X^h_{s,u}(x))\right]\ du$$

On the other hand, for any $s + kh \leqslant u < s + (k+1)h$ we have

$$\begin{aligned}
&b(X^h_{s,u}(x)) - Y^h_{s,u}(x)\\
&= \int_0^1 \nabla b\left(X^h_{s,s+kh}(x) + \epsilon(X^h_{s,u}(x) - X^h_{s,s+kh}(x))\right)' \ (X^h_{s,u}(x) - X^h_{s,s+kh}(x))\ d\epsilon\\
&= \left[\int_0^1 \nabla b\left(X^h_{s,s+kh}(x) + \epsilon(X^h_{s,u}(x) - X^h_{s,s+kh}(x))\right)' \ b\left(X^h_{s,s+kh}(x)\right)\ d\epsilon\right]\ (u - (s+kh))\\
&\qquad + \left[\int_0^1 \nabla b\left(X^h_{s,s+kh}(x) + \epsilon(X^h_{s,u}(x) - X^h_{s,s+kh}(x))\right)' \ d\epsilon\right]\ \sigma\ (W_u - W_{s+kh})
\end{aligned}$$

Let us assume that

$$\nabla b + (\nabla b)' \leqslant -2\lambda\ I \qquad \|\nabla b\| := \sup_x \|\nabla b(x)\| < \infty \quad \text{and} \quad \|b\| := \sup_x \|b(x)\| < \infty \tag{6.3}$$

for some $\lambda > 0$. In this case, for any $s + kh \leqslant u < s + (k+1)h$ and any $n \geqslant 1$ we have

$$\mathbb{E}\left(\|b(X^h_{s,u}(x)) - Y^h_{s,u}(x))\|^n\right)^{1/n} \leqslant \|\nabla b\| \ \left(\|b\|\ h + \sigma\ \sqrt{h}\right)$$

This implies that

$$\mathbb{E}\left(\|X^h_{s,t}(x) - X_{s,t}(x)\|^n\right)^{1/n} \leqslant \|\nabla b\| \ \left(\|b\|\ h + \sigma\ \sqrt{h}\right)\int_s^t \ e^{-\lambda(t-u)}\ du \leqslant \|\nabla b\| \ \left(\|b\|\ h + \sigma\ \sqrt{h}\right)/\lambda$$

from which we conclude that

$$\mathbb{E}\left(\|X^h_{s,t}(x) - X_{s,t}(x)\|^n\right)^{1/n} \leqslant \|\nabla b\| \ \left(\|b\|\ h + \sigma\ \sqrt{h}\right)/\lambda$$

The uniform boundedness drift condition stated in the r.h.s. of (6.3) is rather too strong as it is not met for conventional linear models. Instead of $\|b\| < \infty$ let us assume that

$$\langle x, b(x)\rangle \leqslant -\beta\ \|x\|^2 \quad \text{for some} \quad \beta > 0$$

In this situation, the stochastic flows $X_{s,t}(x)$ has uniform absolute moments of any order $n \geqslant 1$ w.r.t. the time horizon; that is, we have that

$$m_n(x) \leqslant \kappa_n \ (1 + \|x\|) \quad \text{with } m_n(x) \text{ defined in (1.24)}$$

Observe that for any $t \in [s + kh, s + (k+1)h[$ we have

$$\begin{aligned} & d\|X^h_{s,t}(x)\|^2 \\ & \leqslant \left[-2\lambda_0 \ \|X^h_{s,t}(x)\|^2 + 2 \ \langle X^h_{s,t}(x), b(X^h_{s,s+kh}(x)) - b(X^h_{s,t}(x))\rangle + \sigma^2 d\right] \ dt + 2\sigma \ X^h_{s,t}(x)' dW_t \end{aligned}$$

Thus, for any $\epsilon > 0$ we have

$$d\|X^h_{s,t}(x)\|^2 \leqslant \left[(-2\lambda_0 + \epsilon)\|X^h_{s,t}(x)\|^2 + \epsilon^{-1}\|\nabla b\| + \sigma^2 d\right] \ dt + 2\sigma \ X^h_{s,t}(x)' dW_t$$

Arguing as above we check that the stochastic flows $X^h_{s,t}(x)$ also has uniform moments w.r.t. the time horizon; that is, for any $n \geqslant 1$ we have that

$$\widehat{m}_n(x) := \sup_{h \geqslant 0} \sup_{t \geqslant s} \mathbb{E}\left[\|X^h_{s,t}(x)\|^n\right]^{1/n} \leqslant c_n \ (1 + \|x\|)$$

In this situation, for any $s + kh \leqslant u < s + (k+1)h$ and any $n \geqslant 1$ we have

$$\mathbb{E}\left(\|b(X^h_{s,u}(x)) - Y^h_{s,u}(x))\|^n\right)^{1/n} \leqslant \|\nabla b\| \ \left([\|b(0)\| + \widehat{m}_n(x) \ \|\nabla b\|] \ h + \sigma \ \sqrt{h}\right)$$

Arguing as above we conclude that

$$\mathbb{E}\left(\|X^h_{s,t}(x) - X_{s,t}(x)\|^n\right)^{1/n} \leqslant \|\nabla b\| \ \left([\|b(0)\| + \widehat{m}_n(x) \ \|\nabla b\|] \ h + \sigma \ \sqrt{h}\right)/\lambda$$

For the Langevin diffusion discussed in (1.33) and (2.6) the drift function is given by $b = -\nabla U$. We let P^h_t be the Markov transition of the diffusion flow $X^h_t(x)$.

When $\|\nabla U\| < \infty$ the above estimates yields for any $n \geqslant 1$ the uniform inequality

$$\mathbb{W}_n(\pi P^h_t, \pi) \leqslant \|\nabla^2 U\| \ \left(\|\nabla U\| \ h + \sigma \ \sqrt{h}\right)/\lambda$$

Letting $t \to \infty$ we check that the invariant measure π^h of the diffusion flow $X^h_t(x)$ is such that

$$\mathbb{W}_n(\pi^h, \pi) \leqslant \|\nabla^2 U\| \ \left(\|\nabla U\| \ h + \sigma \ \sqrt{h}\right)/\lambda$$

Using (2.12) for any f s.t. $\|f\| \leqslant 1$ we also check that

$$|[\pi P^h_t - \pi](f)| \leqslant \ \frac{c}{\epsilon\lambda} \ \|\nabla^2 U\| \ \left(\|\nabla U\| \ h + \sigma \ \sqrt{h}\right) \int_0^t \ \frac{1}{\sqrt{s}} \ e^{-\lambda_1(1-\epsilon)s} \ ds$$

This yields the total variation norm estimate

$$\|\pi^h - \pi\|_{tv} \leqslant c \ \|\nabla^2 U\| \ \left(\|\nabla U\| \ h + \sigma \ \sqrt{h}\right)/\lambda$$

Appendix

Proof of (1.24)

Whenever $(\mathcal{M})_n$ is satisfied, we have

$$2\langle x, b_t(x)\rangle + \|\sigma_t(x)\|_F^2 \leqslant \gamma_0 + \gamma_1\|x\| - \gamma_2\|x\|^2$$

with the parameters

$$\gamma_0 = \alpha_0 + 2\beta_0 \qquad \gamma_1 = \alpha_1 + 2\beta_1 \quad \text{and} \quad \gamma_2 = 2\beta_2 - \alpha_2$$

Observe that

$$\begin{aligned} & d\|X_{s,t}(x)\|^2 \\ & = \left[2\left\langle X_{s,t}(x), b_t(X_{s,t}(x))\right\rangle + \|\sigma_t(X_{s,t}(x))\|_F^2\right]\ dt + 2\textstyle\sum_k \langle X_{s,t}(x), \sigma_{k,t}(X_{s,t}(x))\rangle\ dW_t^k \end{aligned}$$

After some elementary computations, for any $n \geqslant 1$ we check that

$$\begin{aligned} n^{-1}\partial_t \mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right] &\leqslant -\left[\gamma_2 - 2(n-1)\alpha_2\right]\ \mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right] \\ & \quad + \left[\gamma_1 + 2(n-1)\alpha_1\right]\ \mathbb{E}\left[\|X_{s,t}(x)\|^{2n-1}\right] + \left[\gamma_0 + 2(n-1)\alpha_0\right]\ \mathbb{E}\left[\|X_{s,t}(x)\|^{2(n-1)}\right] \end{aligned}$$

This implies that

$$\begin{aligned} \partial_t \mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right]^{1/n} &\leqslant -\left[\gamma_2 - 2(n-1)\alpha_2\right]\ \mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right]^{1/n} \\ & \quad + \left[\gamma_1 + 2(n-1)\alpha_1\right]\ \mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right]^{1/(2n)} + \left[\gamma_0 + 2(n-1)\alpha_0\right] \end{aligned}$$

from which we check that for any $\epsilon > 0$ we have

$$\begin{aligned} & \partial_t \mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right]^{1/n} \\ & \leqslant -\left[\gamma_2 - 2(n-1)\alpha_2 - 2\epsilon\right]\ \mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right]^{1/n} + \frac{1}{8\epsilon}\left[\gamma_1 + 2(n-1)\alpha_1\right]^2 + \left[\gamma_0 + 2(n-1)\alpha_0\right] \end{aligned}$$

This implies that

$$\begin{aligned} & \partial_t \mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right]^{1/n} \\ & \leqslant -2\left[\beta_2 - (n-1/2)\alpha_2 - \epsilon\right]\ \mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right]^{1/n} + \frac{1}{8\epsilon}\left[\gamma_1 + 2(n-1)\alpha_1\right]^2 + \left[\gamma_0 + 2(n-1)\alpha_0\right] \end{aligned}$$

from which we check that

$$\mathbb{E}\left[\|X_{s,t}(x)\|^{2n}\right]^{1/n} \leqslant e^{-2[\beta_2 - (n-1/2)\alpha_2 - \epsilon](t-s)}\ \|x\|^2 + \frac{1}{8\epsilon}\ \frac{\left[\gamma_1 + 2(n-1)\alpha_1\right]^2 + \left[\gamma_0 + 2(n-1)\alpha_0\right]}{2\left[\beta_2 - (n-1/2)\alpha_2 - \epsilon\right]}$$

as soon as $\epsilon < \beta_2 - (n-1/2)\alpha_2$ and $n \geqslant 1$. Replacing ϵ by $\epsilon(\beta_2 - (n-1/2)\alpha_2)$ and then $(2n)$ by n we check that

$$\begin{aligned} & \mathbb{E}\left[\|X_{s,t}(x)\|^{n}\right]^{1/n} \\ & \leqslant e^{-(1-\epsilon)\beta_2(n)(t-s)}\ \|x\| + \frac{1}{4\sqrt{\epsilon(1-\epsilon)}}\ \frac{\gamma_1(n) + \gamma_0(n)^{1/2}}{\beta_2(n)^{1/2}} \quad \text{with} \quad \gamma_i(n) := \gamma_i + (n-2)\alpha_i \end{aligned}$$

This ends the proof of (1.24). ■

Proof of (5.7)

When $u+h \leqslant v$ we have

$$\begin{aligned}&(\nabla X_{u+h,t})(y)'\\&= (\nabla X_{v+h,t})(X_{v,v+h}(X_{u+h,v}(y)))'(\nabla X_{v,v+h})(X_{u+h,v}(y))'(\nabla X_{u+h,v})(y)'\end{aligned}$$

Observe that

$$X_{v,v+h}(X_{u+h,v}(y)) = X_{u+h,v}(y) + \sigma_{v,k}(X_{u+h,v}(y))\ (W^k_{v+h} - W^k_v) + \mathrm{O}(\mathcal{T})$$

This implies that

$$\begin{aligned}&(\nabla X_{v+h,t})(X_{v,v+h}(X_{u+h,v}(y)))'\\&= (\nabla X_{v+h,t})(X_{u+h,v}(y))' + \textstyle\sum_k \left(\nabla^2 X_{v+h,t}\right)(X_{u+h,v}(y))'\sigma_{v,k}(X_{u+h,v}(y))'\ (W^k_{v+h} - W^k_v)\end{aligned}$$

In the same vein, we have

$$(\nabla X_{v,v+h})(X_{u+h,v}(y))' = I + \sum_k \nabla\sigma_{v,k}(X_{u+h,v}(y))'\ (W^k_{v+h} - W^k_v) + \mathrm{O}(\mathcal{T})$$

This yields the formula

$$(\nabla X_{u+h,t})(y)' = \nabla(X_{v+h,t}\circ X_{u+h,v})(y)' + \sum_k D^k_v (\nabla X_{u+h,t})(y)'\ (W^k_{v+h} - W^k_v) + \mathrm{O}(\mathcal{T})$$

with

$$\begin{aligned}D^k_v (\nabla X_{u+h,t})(y)' &= \left(\nabla^2 X_{v+h,t}\right)(X_{u+h,v}(y))'\ \sigma_{v,k}(X_{u+h,v}(y))'(\nabla X_{u+h,v})(y)'\\&\quad + (\nabla X_{v+h,t})(X_{u+h,v}(y))'\ \nabla\sigma_{v,k}(X_{u+h,v}(y))'(\nabla X_{u+h,v})(y)'\end{aligned}$$

Rewritten in a slightly different form, we have

$$\begin{aligned}D^k_v (\nabla X_{u+h,t})(y)' &= (\nabla X_{v+h,t})(X_{u+h,v}(y))'\ D^k_v (\nabla X_{u+h,v})'(y)\\&\quad + \left(\nabla^2 X_{v+h,t}\right)(X_{u+h,v}(y))'\ (D^k_v X_{u+h,v})(y)'(\nabla X_{u+h,v})(y)'\end{aligned}$$

we also have

$$\begin{aligned}&\left(D_v (\nabla X_{u+h,t})'\right)_{k,i,j} := \left(D^k_v (\nabla X_{u+h,t})'\right)_{i,j} = \left(D^k_v \nabla X_{u+h,t}\right)_{j,i}\\&= \textstyle\sum_l\ D_v (\nabla X_{u+h,v})_{k,j,l}\ \ (\nabla X_{v+h,t})(X_{u+h,v})_{l,i}\\&\qquad + \displaystyle\sum_{l_1,l_2}\ (D_v X_{u+h,v})_{k,l_1}(\nabla X_{u+h,v})_{j,l_2}\ \ \left(\nabla^2 X_{v+h,t}\right)(X_{u+h,v}(y))_{(l_1,l_2),i}\end{aligned}$$

We obtain the tensor differential equation (5.7) by taking the limit as $h \to 0$. This ends the proof of (5.7). ∎

Proof of (5.6) and (5.9)

When $v+h\leqslant u$ we have

$$\begin{aligned}
&\left[(\nabla X_{u+h,t})\,(\overline{X}_{s,u})'\ \varsigma_u(\overline{X}_{s,u})\right]_{i,j}\\
&=\left[(\nabla X_{u+h,t})\,(\overline{X}_{v+h,u}\circ\overline{X}_{s,v})'\ \varsigma_u(\overline{X}_{v+h,u}\circ\overline{X}_{s,v})\right]_{i,j}\\
&\qquad\qquad+D^k_v\left[(\nabla X_{u+h,t})\,(\overline{X}_{s,u})'\ \varsigma_u(\overline{X}_{s,u})\right]_{i,j}\ (W^k_{v+h}-W^k_v)+\mathrm{O}(\mathcal{T})
\end{aligned}$$

with

$$\begin{aligned}
&D^k_v\left[(\nabla X_{u+h,t})\,(\overline{X}_{s,u})'\ \varsigma_u(\overline{X}_{s,u})\right]_{i,j}\\
&=\sum_{l_1,l_2,l_3}\left(\nabla^2 X_{u+h,t}\right)(\overline{X}_{v+h,u}\circ\overline{X}_{s,v})_{(l_1,l_2),i}\ \tau^{l_1}_{u,j}(\overline{X}_{v+h,u}\circ\overline{X}_{s,v})\ (\nabla\overline{X}_{v+h,u})(\overline{X}_{s,v})'_{l_2,l_3}\ \overline{\sigma}^{l_3}_{v,k}(\overline{X}_{s,v})\\
&+\sum_{l_1,l_2,l_3}(\nabla X_{u+h,t})\,(\overline{X}_{v+h,u}\circ\overline{X}_{s,v})_{l_1,i}\ (\nabla\tau_{u,j})(\overline{X}_{v+h,u}\circ\overline{X}_{s,v})_{l_2,l_1}\ (\nabla\overline{X}_{v+h,u})'_{l_2,l_3}\ \overline{\sigma}^{l_3}_{v,k}(\overline{X}_{s,v})
\end{aligned}$$

Rewritten in a slightly different form, we have

$$\begin{aligned}
&D_v\left[(\nabla X_{u+h,t})\,(\overline{X}_{s,u})'\ \varsigma_u(\overline{X}_{s,u})\right]_{k,i,j}=D^k_v\left[(\nabla X_{u+h,t})\,(\overline{X}_{s,u})'\ \varsigma_u(\overline{X}_{s,u})\right]_{i,j}\\
&=D^k_v\left[\varsigma_u(\overline{X}_{s,u})\,(\nabla X_{u+h,t})\,(\overline{X}_{s,u})\right]_{j,i}=D_v\left[\varsigma_u(\overline{X}_{s,u})\,(\nabla X_{u+h,t})\,(\overline{X}_{s,u})\right]_{k,j,i}\\
&=\textstyle\sum_{l_1}D_v\left[\varsigma_u\circ\overline{X}_{v+h,u}\circ\overline{X}_{s,v}\right]_{k,j,l_1}\ \ (\nabla X_{u+h,t})\,(\overline{X}_{v+h,u}\circ\overline{X}_{s,v})_{l_1,i}\\
&\qquad+\textstyle\sum_{l_1,l_2}(D_v(\overline{X}_{v+h,u}\circ\overline{X}_{s,v}))_{k,l_2}\ \varsigma_u(\overline{X}_{v+h,u}\circ\overline{X}_{s,v})_{j,l_1}\ \left(\nabla^2 X_{u+h,t}\right)(\overline{X}_{v+h,u}\circ\overline{X}_{s,v})_{(l_1,l_2),i}
\end{aligned}$$

with

$$\begin{aligned}
(D_v(\overline{X}_{v+h,u}\circ\overline{X}_{s,v}))_{k,l_2}&:=\sum_{l_3}(D_v\overline{X}_{s,v})_{k,l_3}\ (\nabla\overline{X}_{v+h,u})(\overline{X}_{s,v})_{l_3,l_2}\\
D_v\left[\varsigma_u\circ\overline{X}_{v+h,u}\circ\overline{X}_{s,v}\right]_{k,j,l_1}&:=\sum_{l_2}(D_v(\overline{X}_{v+h,u}\circ\overline{X}_{s,v}))_{k,l_2}\ \left[(\nabla\varsigma_u)\circ\overline{X}_{s,u}\right]_{l_2,j,l_1}
\end{aligned}$$

We obtain the tensor differential equation (5.9) by taking the limit as $h\rightarrow 0$ and recalling that

$$\begin{aligned}
D_v\left(\overline{X}_{v,u}\circ\overline{X}_{s,v}\right)&:=\left[D_v\overline{X}_{s,v}\right]\ \left[\left(\nabla\overline{X}_{v,u}\right)\circ\overline{X}_{s,v}\right]\\
D_v\left[\varsigma_u\circ\overline{X}_{s,u}\right]&=\left(D_v\overline{X}_{s,u}\right)\ \left[(\nabla\varsigma_u)\circ\overline{X}_{s,u}\right]
\end{aligned}$$

This ends the proof of (5.9). ∎

Proof of lemma 5.1

When $u < v$ we have

$$\mathbb{E}\Big\{\big[(\nabla X_{u+h,t})\,(\overline{X}_{s,u}(x))'\ \varsigma_u(\overline{X}_{s,u}(x))\big]_{i,j}\ \ \big[(\nabla X_{v+h,t})\,(\overline{X}_{s,v}(x))'\ \varsigma_v(\overline{X}_{s,v}(x))\big]_{i,k}$$

$$(W^j_{u+h}-W^j_u)(W^k_{v+h}-W^k_v)\ \Big\}$$

$$=\sum_{m,\overline{m}}\mathbb{E}\Big\{\big[D^m_v\left[(\nabla X_{u+h,t})\,(\overline{X}_{s,u}(x))'\varsigma_u(\overline{X}_{s,u}(x))\right]\big]_{i,j}\ \ \big[D^{\overline{m}}_u\left[(\nabla X_{v+h,t})\,(\overline{X}_{s,v})'\ \varsigma_v(\overline{X}_{s,v})\right]\big]_{i,k}$$

$$(W^{\overline{m}}_{u+h}-W^{\overline{m}}_u)(W^j_{u+h}-W^j_u)(W^k_{v+h}-W^k_v)(W^m_{v+h}-W^m_v)\ \Big\}+\mathrm{O}(h^2)$$

$$=\mathbb{E}\left\{\left[D^k_v\left[(\nabla X_{u+h,t})\,(\overline{X}_{s,u}(x))'\varsigma_u(\overline{X}_{s,u}(x))\right]\right]_{i,j}\ \ \left[D^j_u\left[(\nabla X_{v+h,t})\,(\overline{X}_{s,v})'\ \varsigma_v(\overline{X}_{s,v})\right]\right]_{i,k}\right\}+\mathrm{O}(h^2)$$

In summary, we have proved that

$$\mathbb{E}\Big\{\big[(\nabla X_{u+h,t})\,(\overline{X}_{s,u}(x))'\ \varsigma_u(\overline{X}_{s,u}(x))\big]_{i,j}\ \ \big[(\nabla X_{v+h,t})\,(\overline{X}_{s,v}(x))'\ \varsigma_v(\overline{X}_{s,v}(x))\big]_{i,k}$$

$$(W^j_{u+h}-W^j_u)(W^k_{v+h}-W^k_v)\ \Big\}$$

$$=\mathbb{E}\Big\{D_u\left[\varsigma_v(\overline{X}_{s,v})\ \ (\nabla X_{v+h,t})\,(\overline{X}_{s,v})\right]_{j,k,i}\ D_v\left[(\nabla X_{u+h,t})\,(\overline{X}_{s,u}(x))'\varsigma_u(\overline{X}_{s,u}(x))\right]_{k,i,j}\ \Big\}+\mathrm{O}(h^2)$$

This ends the proof of lemma 5.1. ■

Proof of proposition 2.3

The proof of the estimate (2.9) is mainly based on the following technical lemma of its own interest.

Lemma 6.2. *Let Z_t be a non negative diffusion process satisfying an inequality of the form*

$$dZ_t \leqslant (-\lambda Z_t+\alpha_t\ \sqrt{Z_t}+\beta_t)\ dt+dM_t\quad\text{with}\quad \partial_t\langle M\rangle_t\leqslant (u_t\sqrt{Z_t}+v_tZ_t)^2$$

for some parameters $\lambda > 0$ and $v_t \geqslant 0$, and some non negative processes (α_t,β_t,u_t). In this situation, for any $\epsilon > 0$ we have

$$\mathbb{E}(Z^n_t)^{1/n}\leqslant e^{\int_0^t\lambda_{n,s}(\epsilon)ds}\ \mathbb{E}(Z^n_0)^{1/n}+\int_0^t e^{\int_s^t\lambda_{n,u}(\epsilon)du}\ z^n_s(\epsilon)\ ds \tag{6.4}$$

with the parameters

$$\begin{aligned}\lambda_{n,t}(\epsilon)\ &:=\ -\lambda+\frac{n-1}{2}\ v_t^2+\frac{\epsilon}{2}\\ z^n_t(\epsilon)\ &:=\ \mathbb{E}\left[\beta^n_t\right]^{1/n}+\frac{n-1}{2}\ \mathbb{E}\left[u^{2n}_t\right]^{1/n}+\frac{1}{\epsilon}\ \left(\mathbb{E}\left[\alpha^{2n}_t\right]^{1/n}+(n-1)^2\ \mathbb{E}\left[(u_tv_t)^{2n}\right]^{1/n}\right)\end{aligned}$$

Proof. Applying Ito's formula, for any $n \geqslant 2$, we have

$$
\begin{aligned}
& n^{-1}\partial_t \mathbb{E}(Z_t^n) \\
& \leqslant \mathbb{E}\left[Z_t^{n-1}(-\lambda Z_t + \alpha_t\ \sqrt{Z_t} + \beta_t) + \frac{n-1}{2}\ (u_t\sqrt{Z_t} + v_t Z_t)^2\ Z_t^{n-2}\right] \\
& = \left(-\lambda + \frac{n-1}{2}\ v_t^2\right)\ \mathbb{E}(Z_t^n) + \mathbb{E}\left[\left(\beta_t + \frac{n-1}{2}\ u_t^2\right) Z_t^{n-1}\right] + \mathbb{E}\left([\alpha_t + (n-1)u_t v_t]\ Z_t^{n-1/2}\right)
\end{aligned}
$$

On the other hand, for any $\epsilon > 0$ we have the almost sure inequality

$$
[\alpha_t + (n-1)u_t v_t]\ Z_t^{(n-1)/2}\ Z_t^{n/2} \leqslant \frac{1}{2\epsilon}\ [\alpha_t + (n-1)u_t v_t]^2\ Z_t^{n-1} + \frac{\epsilon}{2}\ Z_t^n
$$

This implies that

$$
\begin{aligned}
& n^{-1}\partial_t \mathbb{E}(Z_t^n) \\
& \leqslant \lambda_{n,t}(\epsilon)\ \mathbb{E}(Z_t^n) + \mathbb{E}\left[\left(\beta_t + \frac{n-1}{2}\ u_t^2 + \frac{1}{2\epsilon}\ [\alpha_t + (n-1)u_t v_t]^2\right) Z_t^{n-1}\right]
\end{aligned}
$$

Applying Hölder inequality we check that

$$
\begin{aligned}
& \mathbb{E}\left[\left(\beta_t + \frac{n-1}{2}\ u_t^2 + \frac{1}{2\epsilon}\ [\alpha_t + (n-1)u_t v_t]^2\right)\ Z_t^{n-1}\right] \\
& \leqslant \mathbb{E}\left[\left(\beta_t + \frac{n-1}{2}\ u_t^2 + \frac{1}{2\epsilon}\ [\alpha_t + (n-1)u_t v_t]^2\right)^n\right]^{1/n}\ \mathbb{E}(Z_t^n)^{1-1/n} \leqslant z_t^n\ \mathbb{E}(Z_t^n)^{1-1/n}
\end{aligned}
$$

This yields the estimate

$$
\partial_t \mathbb{E}(Z_t^n)^{1/n} = \mathbb{E}(Z_t^n)^{-(1-1/n)}\ n^{-1}\partial_t \mathbb{E}(Z_t^n) \leqslant \lambda_{n,t}(\epsilon)\ \mathbb{E}(Z_t^n)^{1/n} + z_t^n
$$

This ends the proof of the lemma. ∎

We set

$$
Y_{s,t}(x) := \|\nabla^2 X_{s,t}(x)\|_F^2 \quad \text{and} \quad T_{s,t}(x) := \|\nabla X_{s,t}(x)\|_F
$$

and we also consider the collection of parameters

$$
\|\tau\|_F \ := \ \sup_{t,x} \|\tau_t(x)\|_F \qquad \rho(\upsilon) \ := \ \sup_{t,x} \lambda_{max}(\upsilon_t(x))
$$

with the tensor functions (τ_t, υ_t) introduced in (2.8). Observe that

$$
\|\tau\|_F \leqslant \|\nabla^2 b\|_F + d\ \|\nabla^2\sigma\|_F^2 \quad \text{and} \quad \rho(\upsilon) \leqslant d\ \|\nabla^2\sigma\|_2^2
$$

Whenever $(\mathcal{T})_2$ is met we have

$$
\mathrm{Tr}\left[\nabla^2 X_{s,t}(x)\ A_t(X_{s,t}(x))\ \nabla^2 X_{s,t}(x)'\right] \leqslant -2\lambda_A\ Y_{s,t}(x)
$$

Also observe that

$$
|\mathrm{Tr}\left[[\nabla X_{s,t}(x) \otimes \nabla X_{s,t}(x)]\ \ \tau_t(X_{s,t}(x))\ \nabla^2 X_{s,t}(x)'\right]| \leqslant \|\tau\|_F\ Y_{s,t}(x)^{1/2}\ T_{s,t}(x)^2
$$

and

$$\mathrm{Tr}\left[[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)]\,\upsilon_t(X_{s,t}(x))\,[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)]'\right]\leqslant\rho(\upsilon)\ T_{s,t}(x)^4$$

In the same vein, we have

$$\begin{aligned}&|\mathrm{Tr}\left\{[\nabla X_{s,t}(x)\otimes\nabla X_{s,t}(x)]\ \nabla^2\sigma_{t,k}(X_{s,t}(x))\ \nabla^2 X_{s,t}(x)'\right.\\ &\qquad\qquad\qquad\left.+\nabla^2 X_{s,t}(x)\ \nabla\sigma_{t,k}(X_{s,t}(x))\ \nabla^2 X_{s,t}(x)'\right\}|\\ &\leqslant\|\nabla^2\sigma_k\|_F\ T_{s,t}(x)^2\ Y_{s,t}(x)^{1/2}+\rho(\nabla\sigma_k)\ Y_{s,t}(x)\end{aligned}$$

We are now in position to prove proposition 2.3.

Proof of proposition 2.3:

Applying the above lemma to the processes

$$Z_t=Y_{s,t}(x)\qquad\lambda=2\lambda_A\qquad\alpha_t=2\|\tau\|_F\ T_{s,t}(x)^2\qquad\beta_t=\rho(\upsilon)\ T_{s,t}(x)^4$$

and the parameters

$$u_t=2\sqrt{d}\ \|\nabla^2\sigma\|_F\ T_{s,t}(x)^2\quad\text{and}\quad v_t=2\sqrt{d}\ \rho_\star(\nabla\sigma)$$

we obtain the estimate (6.4) with the parameters

$$\begin{aligned}\lambda_{n,t}(\epsilon)&:=&-2\left[\lambda_A-d(n-1)\rho_\star(\nabla\sigma)^2-\frac{\epsilon}{4}\right]\\ z_t^n(\epsilon)&:=&\left\{\rho(\upsilon)\ +2d(n-1)\ \|\nabla^2\sigma\|_F^2\ +\frac{4}{\epsilon}\ \left(\|\tau\|_F^2\ +4\ d^2(n-1)^2\ \rho_\star(\nabla\sigma)^2\ \|\nabla^2\sigma\|_F^2\ \right)\right\}\\ &&\times\mathbb{E}\left[\|\nabla X_{s,t}(x)\|_F^{4n}\right]^{1/n}\end{aligned}$$

Observe that

$$z_t^n(\epsilon)\leqslant cn^2\ (1\vee\epsilon^{-1})\ \chi(b,\sigma)^2\ \mathbb{E}\left[\|\nabla X_{s,t}(x)\|_F^{4n}\right]^{1/n}$$

for some universal constant $c<\infty$ and the parameter $\chi(b,\sigma)$ defined in (1.20). Using (2.7) we check that

$$\begin{aligned}&\mathbb{E}\left(\|\nabla^2 X_{s,t}(x)\|_F^{2n}\right)^{1/n}\\ &\leqslant cn^2\ (1\vee\epsilon^{-1})\ \chi(b,\sigma)^2\ \int_s^t e^{-2\left[\lambda_A-d(n-1)\rho_\star(\nabla\sigma)^2-\frac{\epsilon}{4}\right](t-u)}\ e^{-4\left[\lambda_A-(n-1)\rho(\nabla\sigma)^2\right](u-s)}\ du\\ &=cn^2\ (1\vee\epsilon^{-1})\ \chi(b,\sigma)^2\ e^{-2\left[\lambda_A-d(n-1)\rho_\star(\nabla\sigma)^2-\frac{\epsilon}{4}\right](t-s)}\\ &\qquad\int_s^t e^{-2\left[\lambda_A-(n-1)\rho(\nabla\sigma)^2+(n-1)[d\rho_\star(\nabla\sigma)^2-\rho(\nabla\sigma)^2]+\frac{\epsilon}{8}\right](u-s)}\ du\end{aligned}$$

Assume that

$$\lambda_A>d(n-1)\rho_\star(\nabla\sigma)^2$$

In this case there exists some $0<\epsilon_n\leqslant 1$ such that for any $0<\epsilon\leqslant\epsilon_n$ we have

$$\lambda_A-d(n-1)\rho_\star(\nabla\sigma)^2>\epsilon$$

and therefore

$$\mathbb{E}\left(\|\nabla^2 X_{s,t}(x)\|_F^{2n}\right)^{1/(2n)}\leqslant c\ n\ \epsilon^{-1}\ \chi(b,\sigma)\ \exp\left(-\left[\lambda_A-d(n-1)\rho_\star(\nabla\sigma)^2-\epsilon\right](t-s)\right)$$

This ends the proof of the proposition. ∎

Proof of (1.10)

Using (1.29) and (5.11) we check that

$$\begin{aligned}
&\mathbb{E}\left[\|X_{s,t}(x)-\overline{X}_{s,t}(x)\|^n\right]^{1/n}\\
&\leqslant \kappa_{(\delta_1,\delta_2),n}\ \left(\vert\!\vert\!\vert\Delta a(x)\vert\!\vert\!\vert_{n/(1-\delta_1)}+\vert\!\vert\!\vert\Delta b(x)\vert\!\vert\!\vert_{n/(1-\delta_1)}+\vert\!\vert\!\vert\Delta\sigma(x)\vert\!\vert\!\vert_{2n/\delta_2}\ (1\vee\|x\|)\right)
\end{aligned}$$

as soon as the regularity conditions $(\mathcal{T})_{n/\delta_1}$, $(M)_{2n/\delta_2}$ and $(T)_{2n/(1-\delta_2)}$ are satisfied for some parameter $n\geqslant 2$ and some $\delta_1,\delta_2\in]0,1[$. Choosing $\delta_1=(1-\delta_2)/2$ and setting $\delta=\delta_2$ we check that

$$\begin{aligned}
&\mathbb{E}\left[\|X_{s,t}(x)-\overline{X}_{s,t}(x)\|^n\right]^{1/n}\\
&\leqslant \kappa_{\delta,n}\ \left(\vert\!\vert\!\vert\Delta a(x)\vert\!\vert\!\vert_{2n/(1+\delta)}+\vert\!\vert\!\vert\Delta b(x)\vert\!\vert\!\vert_{2n/(1+\delta)}+\vert\!\vert\!\vert\Delta\sigma(x)\vert\!\vert\!\vert_{2n/\delta}\ (1\vee\|x\|)\right)
\end{aligned}$$

as soon as $(M)_{2n/\delta}$ and $(T)_{2n/(1-\delta)}$ are satisfied for some parameter $n\geqslant 2$ and some $\delta\in]0,1[$. For instance, $(\mathcal{M})_{2n/\delta}$ and $(\mathcal{T})_{2n/(1-\delta)}$ are satisfied as soon as

$$\beta_2-\alpha_2/2>(n/\delta-1)\ \alpha_2\quad\text{and}\quad \lambda_A>d(n/(1-\delta)-1)\ \rho_\star(\nabla\sigma)^2$$

This ends the proof of (1.10). ∎

References

[1] V. Alekseev. An estimate for the perturbations of the solution of ordinary differential equations. Vestn. Mosk.Univ., Ser. I, Math. Meh. vol. 2, (1961).

[2] T. Ando and J. L. van Hemmen. An inequality for trace ideals. Commun. Math. Phys., vol. 76, pp. 143–148 (1980).

[3] M. Arnaudon, P. Del Moral. A variational approach to nonlinear and interacting diffusions. ArXiv:1812.04269 (2018). Stochastic Analysis and Applications DOI: 10.1080/07362994.2019.1609985 (2019).

[4] M. Arnaudon, P. Del Moral. A duality formula and a particle Gibbs sampler for continuous time Feynman-Kac measures on path spaces. ArXiv preprint arXiv:1805.05044 (2018).

[5] M. Arnaudon, P. Del Moral. A second order analysis of McKean-Vlasov semigroups. hal-02151808 (2019).

[6] M. Arnaudon, H. Plank, A. Thalmaier. A Bismut type formula for the Hessian of heat semigroups. C. R. Math. Acad. Sci. Paris, vol. 336, no. 8, pp. 661–666 (2003).

[7] R. Bellman. Some inequalities for the square Root of a Positive Definite Matrix. Linear Algebra and its applications, vol. 1, no. 3, pp. 321–324 (1968).

[8] R. Bellman, Stability Theory of Differential Equations, McGraw Hill, New York, (1953).

[9] J.M. Bismut. Large deviations and the Malliavin calculus. Birkhauser Prog. Math. 45 (1984).

[10] W.A. Coppel. Dichotomies in Stability Theory. Springer (1978).

[11] G. Da Prato, J.L. Menaldi, L. Tubaro. Some results of backward Ito formula. Stochastic analysis and applications, vol. 25, no. 3, pp. 679–703 (2007).

[12] P. Del Moral. Feynman-Kac formulae. Genealogical and interacting particle systems with applications. Probability and its Applications (New York). (573p.) Springer-Verlag, New York (2004).

[13] P. Del Moral, A. Guionnet. On the stability of measure valued processes with applications to filtering. Comptes Rendus de l'Académie des Sciences-Series I-Mathematics, vol. 329, no. 5, pp. 429–434 (1999).

[14] P. Del Moral and A. Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. Ann. Inst. Henri Poincaré, vol. 37, no. 2, pp. 155–194 (2001).

[15] P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering. In *Séminaire de Probabilités, XXXIV*, volume 1729, *Lecture Notes in Math.*, pages 1–145. Springer, Berlin (2000).

[16] K.D. Elworthy, X.M. Li. Formulae for the Derivative of Heat Eemigroups. Journal of Functional Analysis 125, pp. 252–286 (1994).

[17] Gröbner, W. Die Lie-Reihen und ihre Anwendungen. VEB Deutscher Verlag der Wiss., Berlin (1960).

[18] T.H. Gronwall, Note on the derivatives with respect to a parameter of the solutions of a system of differential equations, Ann. Math., vol. 20, no. 2 , pp. 293–296 (1919).

[19] N. J. Higham. Functions of Matrices : Theory and Computation, SIAM, Philadelphia, PA (2008).

[20] A. Hudde, M. Hutzenthaler, A. Jentzen, S. Mazzonetto. On the Itô-Alekseev-Gröbner formula for stochastic differential equations. arXiv preprint arXiv:1812.09857 (2018).

[21] A. Iserles, G. Söderlind. Global bounds on numerical error for ordinary differential equations. Journal of Complexity, vol. 9, no. 1, pp. 97–112 (1993).

[22] A. Jentzen, F. Lindner, P. Pusnik. On the Alekseev-Gröbner formula in Banach spaces. arXiv preprint arXiv:1810.10030 (2018).

[23] X. M. Li. Doubly Damped Stochastic Parallel Translations and Hessian Formulas. International Conference on Stochastic Partial Differential Equations and Related Fields. Springer, Cham (2016).

[24] D. Nualart. The Malliavin calculus and related topics. Vol. 1995. Berlin: Springer (2006).

[25] D. Nualart, E. Pardoux. Stochastic calculus with anticipating integrands. Probability Theory and Related Fields, vol. 78, no. 4, pp. 535–581 (1988).

[26] D. Nualart, M. Zakai. Generalized multiple stochastic integrals and the representation of wiener functionals. Stochastics, vol. 23, no. 3, pp. 311–330 (1988).

[27] E. Pardoux, P. Protter. A two-sided stochastic integral and its calculus. Probability Theory and Related Fields, vol. 76, no. 1, pp. 15–49 (1987).

[28] M. Scheutzow. A stochastic Gronwall lemma. Infinite Dimensional Analysis, Quantum Probability and Related Topics, vol. 16, no. 2, p. 1350019 (2013).

[29] J. Thompson. Derivatives of Feynman-Kac semigroups. Journal of Theoretical Probability, vol. 32, no. 2, pp. 950–973 (2019).

[30] S. Watanabe. Lectures on stochastic differential equations and Malliavin Calculus. Tata Institute of Fundamental Research. Springer-Verlag (1984).

[31] M. Wu. A note on stability of linear time-varying systems. IEEE Transactions on Automatic Control. vol. 19, no. 2. pp. 162–162 (1974).