Stochastic Differential Equations Driven by Processes Generated by Divergence Form Operators II: Convergence results

Antoine Lejay

Stochastic Differential Equations Driven by Processes Generated by Divergence Form Operators II: Convergence Results

Antoine Lejay[†]

June 1, 2007

Abstract:

We have seen in a previous article how the theory of "rough paths" allows us to construct solutions of differential equations driven by processes generated by divergence form operators. In this article, we study a convergence criterion which implies that one can interchange the integral with the limit of a family of stochastic processes generated by divergence form operators. As a corollary, we identify stochastic integrals constructed with the theory of rough paths with Stratonovich or Itô integrals already constructed for stochastic processes generated by divergence form operators by using time-reversal techniques.

Keywords:

Rough paths, stochastic differential equations, stochastic process generated by divergence form operators, Condition UTD, convergence of stochastic integrals.

AMS Classification: 60H10, 60J60

[†]Projet TOSCA,
INRIA & Institut Élie Cartan UMR 7502, Nancy-Université,
CNRS, INRIA. Campus scientifique — BP 239
54506 Vandœuvre-lès-Nancy CEDEX France
Email: `Antoine.Lejay@iecn.u-nancy.fr`
URL: `www.iecn.u-nancy.fr/~lejay`

This work has been partially supported by European Union's TMR Stochastic Analysis Network (grant reference 960075).

1 Introduction

In [Lej06a], we have seen how the theory of rough paths developed in [Lyo98] (See also [Lej03a, LQ02, Lej06b]) could be used to prove the pathwise existence of stochastic integrals of type

$$Z_t = z + \sum_{i=1}^{N} \int_0^t g_i(X_r)\, \mathrm{d}X_r^i \tag{1}$$

and solutions of the Stochastic Differential Equations (SDE) of type

$$Y_t = y + \sum_{i=1}^{N} \int_0^t f_i(Y_r)\, \mathrm{d}X_r^i \tag{2}$$

provided that f and g are smooth enough, where $X = (X^1, \ldots, X^N)$ is the stochastic process generated by a divergence form operator of type

$$L = \sum_{i,j=1}^{N} \frac{1}{2} \frac{\partial}{\partial x_i} \left(a_{i,j} \frac{\partial}{\partial x_j} \right) + \sum_{i=1}^{N} b_i \frac{\partial}{\partial x_i} \tag{3}$$

for a measurable function a taking its values in the space of symmetric $N \times N$-matrix and a measurable function b from $\mathbb{R}^N$ to $\mathbb{R}$. Here, a and b are bounded and a is uniformly elliptic, but a and b are not assumed to be continuous.

Let us denote by $\mathfrak{K}$ the map giving Z from X in (1), and $\mathfrak{I}$ the map giving Y from X in (2). Since in our case X has the same regularity properties of the Brownian motion's trajectories, the maps $\mathfrak{K}$ and $\mathfrak{I}$ are not functions of X, but functions of a pair $\mathbf{X} = (\mathbf{X}^1, \mathbf{X}^2)$ called a *rough path* or a *multiplicative functional*. This path $\mathbf{X}$ is a function from $[0, T]$ with values in $\mathrm{T}_1^{(2)}(\mathbb{R}^N) = \mathbb{R}^N \oplus \mathbb{R}^N \otimes \mathbb{R}^N$ where $\mathbf{X}_t^1 = X_t$ and $\mathbf{X}^2$ takes its values in $\mathbb{R}^N \otimes \mathbb{R}^N$. On the space $\mathrm{T}_1^{(2)}(\mathbb{R}^N) = \mathbb{R}^N \oplus \mathbb{R}^N \otimes \mathbb{R}^N$ in which $\mathbf{X}$ lives, we introduce the gauge[1]

$$\|\mathbf{x}\| = \max\{|\mathbf{x}^1|, \sqrt{|\mathbf{x}^2|}\},\ \mathbf{x} = (\mathbf{x}^1, \mathbf{x}^2),\ \mathbf{x}^1 \in \mathbb{R}^N,\ \mathbf{x}^2 \in \mathbb{R}^N \otimes \mathbb{R}^N,$$

where the norm on $\mathbb{R}^N \otimes \mathbb{R}^N$ is such that $|x \otimes y| \leq |x| \cdot |y|$ for all $x, y \in \mathbb{R}^N$. This space $\mathrm{T}_1^{(2)}(\mathbb{R}^N)$ is a sub-group of the non-commutative Lie group $\mathrm{T}^{(2)}(\mathbb{R}^N)$ with the tensor product $\otimes$ as group operation. For a rough path

[1] It can be think as a norm, but does not satisfy $|\lambda \mathbf{x}| = |\lambda| \cdot |\mathbf{x}|$ for $\lambda \in \mathbb{R}$, $\mathbf{x} \in \mathrm{T}_1^{(2)}(\mathbb{R}^N)$. Instead, it satisfies $|\delta_\lambda(\mathbf{x})| = |\lambda| \cdot |\mathbf{x}|$, where δ_λ is the dilatation of parameter λ: $\delta_\lambda(\mathbf{x}) = (\lambda \mathbf{x}^1, \lambda^2 \mathbf{x}^2)$. In addition, $|\mathbf{x} \otimes \mathbf{y}| \leq 2(|\mathbf{x}| + |\mathbf{y}|)$ for all $\mathbf{x}$, $\mathbf{y}$ in $\mathrm{T}_1^{(2)}(\mathbb{R}^N)$.

$\mathbf{X}$, its increment between time s and t is $\mathbf{X}_{s,t} = \mathbf{X}_s^{-1} \otimes \mathbf{X}_t$ where $\mathbf{x} \otimes \mathbf{y} = (\mathbf{x}^1 + \mathbf{y}^1, \mathbf{x}^2 + \mathbf{y}^2 + \mathbf{x}^1 \otimes \mathbf{y}^1)$ and $\mathbf{x}^{-1} = (-\mathbf{x}^1, -\mathbf{x}^2 - \mathbf{x}^1 \otimes \mathbf{x}^1)$ for $\mathbf{x}, \mathbf{y}$ in $\mathrm{T}_1^{(2)}(\mathbb{R}^N)$.

The maps $\mathfrak{K}$ and $\mathfrak{I}$ are continuous on the set of rough paths when one uses, for $\alpha \in [2,3)$, the topology $\mathcal{V}^\alpha$ generated by the norm

$$\|\mathbf{X}\|_{\mathcal{V}^\alpha} = \sup_{t\in[0,T]} |\mathbf{X}_t| + \mathrm{Var}_{\alpha,[0,T]}(\mathbf{X})$$

where

$$\mathrm{Var}_{\alpha,[0,T]}(\mathbf{X}) = \left(\sup_{\text{partition } 0\le t_1<\ldots<t_k\le T} \sum_{i=1}^{k-1} |\mathbf{X}_{t_i,t_{i+1}}|^\alpha \right)^{1/\alpha}. \tag{4}$$

The quantity $\mathrm{Var}_{\alpha,[0,T]}\,\mathbf{X}$ is the α-variation of $\mathbf{X}$. The choice of the range of α depends on the regularity of the trajectories of X (a function which is β-Hölder continuous is of finite β^{-1}-variation).

In [Lej06a], we have shown that if $(X, \mathbb{P}_x)$ is the process generated by L given by (3), one can use for $\mathbf{X}$ the process $t \mapsto (X_t^i, K_{0,t}^{i,j}(X))_{i,j=1,\ldots,N}$, where $K_{s,t}^{i,j}(X)$ is either

$$K_{s,t}^{i,j}(X) = \int_s^t (X_r^i - X_s^i) \circ \mathrm{d}X_r^j \text{ or } K_{s,t}^{i,j}(X) = \int_s^t (X_r^i - X_s^i)\,\mathrm{d}X_r^j.$$

These integrals are defined using the forward-backward martingales decomposition (or Lyons-Zheng decomposition) of X, as constructed by A. Rozkosz [Roz96a] and T. Lyons and L. Stoica [LS99]. In addition to have proved that $\mathbf{X}$ is of finite α-variation under $\mathbb{P}_x$ for any starting point x, we also established a result of type Wong-Zakai. Let us note that in a recent article [FV06a], P. Friz and N. Victoir have dealt with the same problem in a completely different way, using Dirichlet forms (see also Section 3.5.2).

Let us consider a family $(X^\varepsilon, K(X^\varepsilon))_{\varepsilon>0}$ where X^ε is generated by a divergence form operator

$$L^\varepsilon = \sum_{i,j=1}^N \frac{1}{2} \frac{\partial}{\partial x_i} \left(a_{i,j}^\varepsilon \frac{\partial}{\partial x_j} \right) + \sum_{i=1}^N b_i^\varepsilon \frac{\partial}{\partial x_i},$$

for which the uniform ellipticity and boundedness constants of a^ε and b^ε are uniform in ε. If $(X^\varepsilon)_{\varepsilon>0}$ converges uniformly in distribution to X, what can be said on the convergence of $(X^\varepsilon, K(X^\varepsilon))_{\varepsilon>0}$ in $\mathcal{V}^\alpha$? In particular, does $(K(X^\varepsilon))_{\varepsilon>0}$ converge to $K(X)$? If true, the continuity of the maps $\mathfrak{K}$ and $\mathfrak{I}$ implies the convergence of the stochastic integrals and solutions of SDEs driven by X^ε to the same stochastic integral and solution of SDE where X^ε is replaced by X.

Using the results in [Lej02, LL06], one can construct examples — when the coefficients are smooth and X^ε is then solution to some SDE — giving a negative answer to this question if we drop the assumption that the drift b^ε is uniformly bounded in ε. For this, we use the *homogenization theory*, which allows us to study the asymptotic behavior as ε decreases to 0 of X^ε when $a^\varepsilon = a(\cdot/\varepsilon)$ and $b^\varepsilon = \frac{1}{\varepsilon} b(\cdot/\varepsilon)$ for some a and b that are periodic (See for example [BPL78, JKO94] among many references).

However, the results of [Lej02, LL06] were developed for processes generated by non-divergence operators. If we use these results on the processes X^ε generated by divergence form operators $\sum_{i,j=1}^N \frac{1}{2}\partial_{x_i}(a_{i,j}(\cdot/\varepsilon)\partial_{x_j})$ with a periodic, smooth function a, we get that $K(X^\varepsilon)$ converges in distribution to $K(\overline{X})$ when Stratonovich integrals are used, where $\overline{X}$ is the process generated by a divergence form operator $\sum_{i,j=1}^N \frac{1}{2}\partial_{x_i}(a^{\text{eff}}_{i,j}\partial_{x_j})$. The coefficient a^{eff} is a constant matrix that catches the large scale behavior of the process X^ε. Indeed, X^ε is equal in distribution to the rescaled process $\varepsilon X_{\cdot/\varepsilon^2}$ for X generated by $\sum_{i,j=1}^N \frac{1}{2}\partial_{x_i}(a\partial_{x_j})$, so that the homogenization theory consists in proving a functional Central Limit Theorem. Using the results in [FV06a] and some analytical convergence results of [BBJR95], we prove the convergence results of the rough path lying above X^ε even if the coefficients are not assumed to be continuous.

Still with the homogenization theory, we get that the brackets of the martingales parts of X^ε does not necessarily converge to the brackets of the martingale part of $\overline{X}$. Hence, when one uses for $K(X^\varepsilon)$ and $K(X)$ the Itô integrals, then $K(X^\varepsilon)$ does not converge to $K(X)$, but to $t \mapsto K_{0,t}(X) + ct$ for some constant, $N \times N$-symmetric matrix c.

In the case of divergence-form operators without drift, the homogenization theory does not give a counter-example of the convergence of the iterated integrals to the iterated integrals of the limit in the Stratonovich case. We have been unable to find a sequence $(X^\varepsilon)_{\varepsilon>0}$ such that $(K(X^\varepsilon))_{\varepsilon>0}$ does not converge to $K(X)$ in the Stratonovich case when X^ε converges in distribution to X. Yet we give a natural sufficient (but not necessary!) condition to ensure that $(K(X^\varepsilon))_{\varepsilon>0}$ converges to $K(X)$ (whatever the type of the integrals), which turns out to be *Condition UTD* introduced by F. Coquet, A. Rozkosz and L. Słomiński in [RS98, CS99]. This condition implies that, given a family of Dirichlet processes converging to a Dirichlet process, their martingale parts and their zero-quadratic variation parts also converge jointly the the corresponding parts of the limiting process. In the context of semi-martingales seen as rough paths, the equivalent criterion is Condition UT (see for example [KP96]), as seen in [CL03]. From the heuristic point of view, this may be understood with the notion of (p, q)-rough paths defined

in [LV06], although the results in this last article cannot be applied directly in our case. The basic idea is that the limit of $K(X^\varepsilon)$ may not correspond to $K(X)$ when the part of finite or zero-quadratic variation "contributes" to the martingale part of the limit. In our homogenization example, the sequence (X^ε) does not satisfy Condition UTD.

In Section 4, we prove the continuity of $\mathfrak{K}(\mathbf{X}) = \int g(X_s)\,\mathrm{d}\mathbf{X}_s$ with respect to g. Although this continuity follows easily from the very construction of $\mathfrak{K}$, it seems to have never been stated. The convergence of $\mathfrak{I}$ with respect to the function f is also given in [CFV05] under a slightly stronger assumption of convergence of the vector fields. In Section 5, we identify the integrals given by the theory of rough paths with the integrals of Itô or Stratonovich type in function of the choice of $K(X)$. This identification is done by smoothing the coefficients of L and using a similar identification for semi-martingales proved in [CL03].

2 Stochastic integral driven by process generated by divergence-form operators

For $0 < \lambda < \Lambda$, let $\Xi^{\mathrm{coeff}}(\lambda, \Lambda)$ denotes the set of all the functions (a, b) such that

$$
\begin{aligned}
&a \text{ and } b \text{ are measurable,}\\
&a(x) = (a_{i,j}(x))_{i,j=1,\dots,N} \text{ is a symmetric matrix,}\\
&\forall \xi \in \mathbb{R}^N,\ \forall x \in \mathbb{R}^N,\ \lambda|\xi|^2 \leq \sum_{i,j=1}^{N} a_{i,j}(x)\xi_i\xi_j \leq \Lambda|\xi|^2,\\
&b = (b_i)_{i=1,\dots,N} \text{ and } \forall x \in \mathbb{R}^N,\ \sup_{i=1,\dots,N} |b_i(x)| \leq \Lambda.
\end{aligned}
$$

For $(a, b) \in \Xi^{\mathrm{coeff}}(\lambda, \Lambda)$, we may define L by (3) as a closed operator on $\mathrm{L}^2(\mathbb{R}^N)$. This operator is then the infinitesimal generator of a continuous, strong Markov, stochastic process $(X, (\mathcal{F}_t)_{t>0}, (\mathbb{P}_x)_{x\in\mathbb{R}^N})$: See for example [Str88, Lej00].

We denote by $\Xi(\lambda, \Lambda)$ the class of processes generated by divergence form operators with coefficients in $\Xi^{\mathrm{coeff}}(\lambda, \Lambda)$.

Unless its diffusion coefficient a is smooth enough, $X \in \Xi(\lambda, \Lambda)$ is not a semi-martingale, but belongs to the more general class of Dirichlet processes under $\mathbb{P}_x$ for any starting point x [RS98, Theorem 2.2] (note that there are several possible definitions of a Dirichlet process, with slight variations. For example, the definition of a Dirichlet processes given by the theory of

Dirichlet forms [FOT04], which may be used to study the process X, is different).

Definition 1 ([Föl81b]**).** Let $(\Pi^n)_{n\in\mathbb{N}}$ be a family of partitions of $[0,T]$ whose meshes decrease to 0 as $n\to\infty$. For a continuous function A, we introduce

$$Q(A,\Pi^n)=\sum_{i=0}^{\ell^n-1}|A_{t_{i+1}^n}-A_{t_i^n}|^2 \text{ if } \Pi^n=\{t_0^n\le t_1^n\le\cdots\le t_{\ell^n}^n\}$$

and we say that a process A is of *zero quadratic variation* (along $(\Pi^n)_{n\in\mathbb{N}}$) if $Q(A,\Pi^n)$ converges to 0 in probability as $n\to\infty$.

A *Dirichlet process* (along $(\Pi^n)_{n\in\mathbb{N}}$) is a $\mathcal{F}_\cdot$-adapted process the sum of a local martingale and a term of zero quadratic variation which are both $\mathcal{F}_\cdot$-adapted.

Let $\Upsilon^\infty(\mathbb{R}^N)$ be the class of functions $g\in \mathrm{W}^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^N)$ for which $\nabla g\in \mathrm{L}^\infty(\mathbb{R}^N)$.

Let $g\in\Upsilon^\infty(\mathbb{R}^N)$ and $T>0$, as well as $X\in\Xi(\lambda,\Lambda)$. Let us denote by $(\overline{\mathcal{F}}_t)_{t\in[0,T]}$ the backward filtration of X, that is $\overline{\mathcal{F}}_t=\sigma(X_s;s\in[t,T])$. From [LS99, Roz96a, RS98], the process $g(X)$ may be decomposed under $\mathbb{P}_x$ for $x\in\mathbb{R}^N$ as

$$g(X_t)=g(X_0)+\frac{1}{2}M_t^g+\frac{1}{2}(\overline{M}^g_{T-t}-\overline{M}^g_T)+V_t^g,\ t\in[0,T], \tag{5}$$

where M^g is a $\mathcal{F}_\cdot$-martingale, $\overline{M}^g$ is a $\overline{\mathcal{F}}_\cdot$-martingale and

$$V_t^g=\int_0^t b(X_r)\nabla g(X_r)\,\mathrm{d}r+\int_0^t\frac{1}{p(r,x,X_r)}a(X_r)\nabla p(r,x,X_r)\nabla g(X_r)\,\mathrm{d}r.$$

Such a decomposition is called a *Lyons-Zheng decomposition*. In addition, V^g is a term of integrable variation under $\mathbb{P}_x$ and M^g and $\overline{M}^g$ are square-integrable martingales with brackets

$$\langle M^g\rangle_t=\int_0^t a\nabla g\cdot\nabla g(X_s)\,\mathrm{d}s \text{ and } \langle\overline{M}^g\rangle_t=\int_0^t a\nabla g\cdot\nabla g(X_{T-s})\,\mathrm{d}s.$$

Finally, the process $g(X)$ is a Dirichlet process under $\mathbb{P}_x$ for any starting point x (provided we choose a version of g which is continuous around x, whose existence follows from the Morrey theorem [A75, Theorem 5.4, p. 97]) whose term of zero quadratic variation A^g is

$$A_t^g=-\frac{1}{2}M_t^g+\frac{1}{2}(\overline{M}^g_{T-t}-\overline{M}^g_T)+V_t^g,$$

and M^g is its martingale part [Roz96a, RS98]. The backward martingale $\overline{M}^g$ is the martingale part of the reversed process $g(\overline{X}_\cdot) = g(X_{T-\cdot})$, which is also a Dirichlet process.

We set for $g, \varphi \in \Upsilon^\infty$,

$$\mathfrak{L}(t; X, g, \varphi) = \int_0^t g(X_s) \circ \mathrm{d}\varphi(X_s) \tag{6}$$

$$\text{or } \mathfrak{L}(t; X, g, \varphi) = \int_0^t g(X_s)\, \mathrm{d}\varphi(X_s) \tag{7}$$

in the way these integrals are defined in [Roz96a] (see also [LS99]). There are two ways to characterize them: In the Stratonovich case, where $\mathfrak{L}$ is defined by (6),

$$\mathfrak{L}(t; X, g, \varphi) = \lim_{n\in\mathbb{N}} \mathfrak{L}^n(t; X, g, \varphi)$$
$$\text{with } \mathfrak{L}^n(t; X, g, \varphi) = \frac{1}{2} \sum_{j=0}^{\ell^n(t)-1} (g(X_{t_{j+1}^n}) + g(X_{t_j^n}))(\varphi(X_{t_{j+1}^n}) - \varphi(X_{t_j^n})))$$

In the Itô case, where $\mathfrak{L}$ is defined by (7),

$$\mathfrak{L}(t; X, g, \varphi) = \lim_{n\in\mathbb{N}} \mathfrak{L}^n(t; X, g, \varphi)$$
$$\text{with } \mathfrak{L}^n(t; X, g, \varphi) = \sum_{j=0}^{\ell^n(t)-1} g(X_{t_j^n})(\varphi(X_{t_{j+1}^n}) - \varphi(X_{t_j^n})).$$

In both cases, the limits hold in probability under $\mathbb{P}_x$ for any $x \in \mathbb{R}^N$. Of course, the difference between the Itô and the Stratonovich case lies in the quadratic variation, since

$$\begin{aligned}
&\frac{1}{2} \sum_{j=0}^{\ell^n(t)-1} (g(X_{t_{j+1}^n}) + g(X_{t_j^n}))(\varphi(X_{t_{j+1}^n}) - \varphi(X_{t_j^n}))) \\
&\qquad = \frac{1}{2} \sum_{j=0}^{\ell^n(t)-1} (g(X_{t_{j+1}^n}) - g(X_{t_j^n}))(\varphi(X_{t_{j+1}^n}) - \varphi(X_{t_j^n})) \\
&\qquad\qquad + \sum_{j=0}^{\ell^n(t)-1} g(X_{t_j^n})(\varphi(X_{t_{j+1}^n}) - \varphi(X_{t_j^n})).
\end{aligned} \tag{8}$$

With (5), we also get that in the Stratonovich case, $\mathbb{P}_x$-almost surely,

$$\begin{aligned}\mathfrak{L}(t;X,g,\varphi) = \frac{1}{2}\int_0^t (g(X_r)-g(x))\,\mathrm{d}M_r^\varphi + \frac{1}{2}\int_0^t (g(\overline{X}_r)-g(\overline{X}_0))\,\mathrm{d}\overline{M}_r^\varphi \\ + \frac{1}{2}(g(X_t)-g(x))\overline{M}_t^\varphi + \int_0^t (g(X_r)-g(x))\,\mathrm{d}V_r^\varphi \\ - \frac{1}{2}\langle g(\overline{X}),\overline{M}^\varphi\rangle_t + \frac{1}{2}\langle g(X),M^\varphi\rangle_t\end{aligned}$$

with $\overline{X}_t = X_{T-t}$ for $t\in[0,T]$. In the Itô case, $\mathbb{P}_x$-almost surely,

$$\begin{aligned}\mathfrak{L}(t;X,g,\varphi) = \frac{1}{2}\int_0^t (g(X_r)-g(x))\,\mathrm{d}M_r^\varphi + \frac{1}{2}\int_0^t (g(\overline{X}_r)-g(\overline{X}_0))\,\mathrm{d}\overline{M}_r^\varphi \\ + \frac{1}{2}(g(X_t)-g(x))\overline{M}_t^\varphi + \int_0^t (g(X_r)-g(x))\,\mathrm{d}V_r^\varphi - \frac{1}{2}\langle g(\overline{X}),\overline{M}^\varphi\rangle_t.\end{aligned}$$

We end this section with a convergence result of the integral $\mathfrak{L}(X,g,\varphi)$ with respect to g and φ.

Proposition 1. *Let $(g_n)_{n\in\mathbb{N}}$ and $(\varphi_n)_{n\in\mathbb{N}}$ be sequences of functions in $\Upsilon^\infty(\mathbb{R}^N)$ and $g_0,\varphi_0\in\Upsilon^\infty(\mathbb{R}^N)$. We assume that g_n, φ_n are bounded uniformly for $n\geq 0$ by a constant K. We assume moreover, that g_n (resp. φ_n) and ∇g_n (resp. $\nabla\varphi_n$) converge uniformly to g_0 (resp. φ_0) and ∇g_0 (resp. $\nabla\varphi_0$). Then for any $x\in\mathbb{R}^N$, $t\mapsto\mathfrak{L}(t;X,g_n,\varphi_n)$ converges in probability under $\mathbb{P}_x$ to $t\mapsto\mathfrak{L}(t;X,g,\varphi)$ with respect to the uniform norm. This is true both for the Stratonovich and the Itô case.*

Proof. Let us note first that $\varphi\mapsto M^\varphi$ and $\varphi\mapsto V^\varphi$ are linear maps, as well as $(g,\varphi)\mapsto\mathfrak{L}(X,g,\varphi)$. Hence,

$$\mathfrak{L}(X,g_n,\varphi_n)-\mathfrak{L}(X,g_0,\varphi_0) = \mathfrak{L}(X,g_n-g_0,\varphi_n)+\mathfrak{L}(X,g_n,\varphi_n-\varphi_0).$$

We could then assume that $g_0=\varphi_0=0$.

Hence, for any $C>0$, the Burkholder-Davis-Gundy inequality (see for example Theorem 3.28 in [KS91, p. 166]),

$$\begin{aligned}\mathbb{P}_x\left[\sup_{t\in[0,T]}\left|\int_0^t g_n(X_r)\,\mathrm{d}M_r^{\varphi_n}\right|\geq C\right] \\ \leq \frac{1}{C^2}\mathbb{E}_x\left[\Lambda\int_0^T|\nabla\varphi_n(X_r)|^2 g_n(X_r)^2\,\mathrm{d}r\right]\xrightarrow[n\to\infty]{} 0. \quad (9)\end{aligned}$$

Similarly,

$$\mathbb{P}_x\left[\sup_{t\in[0,T]}\left|\int_0^t g_n(\overline{X}_r)\,\mathrm{d}\overline{M}_r^{\varphi_n}\right|\geq C\right] \leq \frac{1}{C^2}\mathbb{E}_x\left[\Lambda\int_0^T|\nabla\varphi_n(X_r)|^2 g_n(X_r)^2\,\mathrm{d}r\right]\xrightarrow[n\to\infty]{}0. \quad (10)$$

In addition, since $\mathbb{E}_x\left[\int_0^T|\mathrm{d}V_r^{\varphi_n}|\right]$ is bounded by $\|\nabla\varphi_n\|_\infty$ times a constant that depends only on λ, Λ and T (see [Lej06a]),

$$\mathbb{E}_x\left[\sup_{t\in[0,T]}\left|\int_0^t(g_n(X_r)-g_n(x))\,\mathrm{d}V_r^{\varphi_n}\right|\right]\leq 2\|g_n\|_\infty\mathbb{E}_x\left[\int_0^T|\mathrm{d}V_r^{\varphi_n}|\right]\xrightarrow[n\to\infty]{}0. \quad (11)$$

This proves that, in the Stratonovich case,

$$\sup_{t\in[0,T]}|\mathfrak{L}(t;X,g_n,\varphi_n)-\mathfrak{L}(t;X,0,\varphi_n)|\xrightarrow[n\to\infty]{}0,$$

since $\mathfrak{L}(t;X,0,\varphi_n)=0$.

From Remark 2.5 in [Roz96a, p. 107],

$$\langle g_n(X),M^{\varphi_n}\rangle_t=\int_0^t a\nabla g_n\cdot\nabla\varphi_n(X_r)\,\mathrm{d}r$$
$$\text{and }\langle g_n(\overline{X}),\overline{M}^{\varphi_n}\rangle_t=\int_{T-t}^T a\nabla g_n\cdot\nabla\varphi_n(X_r)\,\mathrm{d}r.$$

It follows easily that $\langle g_n(X),M^{\varphi_n}\rangle$ and $\langle g_n(\overline{X}),\overline{M}^{\varphi_n}\rangle$ converges uniformly in $t\in[0,T]$ to 0 as $n\to\infty$. This proves that, in the Itô case, that $\sup_{t\in[0,T]}|\mathfrak{L}(t;X,g_n,\varphi_n)-\mathfrak{L}(t;X,0,\varphi_n)|$ converges to 0.

To prove the convergence of $\mathfrak{L}(\cdot,X,g_n,\varphi_n)$ to 0 when $n\to\infty$ (since $\mathfrak{L}(\cdot,X,g_n,0)=0$), it is sufficient to use the previous inequalities (9), (10) and (11) as well as the convergence of $\langle g_n(X),M^{\varphi_n}\rangle$ and $\langle g_n(\overline{X}),\overline{M}^{\varphi_n}\rangle$ to 0. □

3 Convergence of processes generated by a divergence form operator

Given a process $X\in\Xi(\lambda,\Lambda)$, we have seen in [Lej06a] how to construct a rough path $\mathbf{X}$ lying above X. For this, we have constructed $\mathbf{X}$ with $\mathbf{X}_{s,t}^1=$

$X_t - X_s$ and $\mathbf{X}_{s,t}^{i,j,2} = K_{s,t}^{i,j}(X)$, where $K_{s,t}^{i,j}(X)$ is either $\int_s^t (X_r^i - X_s^i) \circ \mathrm{d}X_r^j$ (Stratonovich case) or $\int_s^t (X_r^i - X_s^i)\mathrm{d}X_r^j$ (Itô case).

According to Lemma Lemma 5.4.1 in [LQ02], all the rough paths lying above X may be constructed by adding to $\mathbf{X}$ a function $(s,t) \in \Delta_+ \mapsto \psi_t - \psi_s$ where ψ takes its values in $\mathbb{R}^N \otimes \mathbb{R}^N$ and is of finite $\alpha/2$-variation. In this article, we consider only the Stratonovich and the Itô cases, for they correspond to natural choices.

We have seen in [Lej06a] that for any $\alpha > 2$, any X in $\Xi(\lambda, \Lambda)$, any starting point x and any $\eta > 0$, there exists C large enough such that

$$\begin{cases} \mathbb{P}_x \left[\, \mathrm{Var}_{\alpha,[0,T]}(\mathbf{X}) > C \, \right] < \eta, \\ C \text{ depends only on } (\alpha, \lambda, \Lambda, T), \end{cases} \tag{12}$$

and that for any $C > 0$ and any $\eta > 0$, there exists δ small enough such that

$$\begin{cases} \mathbb{P}_x \left[\, \sup_{|t-s|<\delta} |\mathbf{X}| > C \, \right] \leq \eta, \\ \delta \text{ depends only on } (\alpha, \lambda, \Lambda, T). \end{cases} \tag{13}$$

3.1 Tightness results

Our first results on sequences of processes in $\Xi(\lambda, \Lambda)$ concern their tightness. For $X \in \Xi(\lambda, \Lambda)$, $K(X)$ denotes either the second-order iterated integrals of X in the Stratonovich or the Itô case.

We now consider a sequence $(X^\varepsilon)_{\varepsilon>0}$ of processes in $\Xi(\lambda, \Lambda)$. We denote by $\mathbb{P}_x$ the distribution of X^ε to denote that $\mathbb{P}_x \left[\, X^\varepsilon = x \, \right] = 1$.

Proposition 2. *For any $\varepsilon > 0$, let X^ε be in $\Xi(\lambda, \Lambda)$. Under $\mathbb{P}_x$ for any $x \in \mathbb{R}^N$, the sequence $(X^\varepsilon, K(X^\varepsilon))_{\varepsilon>0}$ is tight in $\mathcal{V}^\alpha$ for any $\alpha > 2$.*

Proof. With (12) and (13), the proof is an immediate consequence from the tightness criterion presented in [Lej03a, Lej06a].

However, we give a short proof, restricted to the first level for the sake of simplicity, to endow the role of (12) and (13).

For some constant C, let $K(C)$ be the set of continuous functions x from $[0,T]$ to $\mathbb{R}^N$ such that $\mathrm{Var}_{\alpha,[0,T]}(x) \leq C$, where $\mathrm{Var}_{\alpha,[0,T]}(x)$ is defined by (4) with $\mathbf{X}_{s,t}$ replaced by $x_t - x_s$ and $|\cdot|$ is the Euclidean norm on $\mathbb{R}^N$. Let also K' be a set of continuous functions from $[0,T]$ to $\mathbb{R}^N$ which is relatively compact for the uniform convergence. If $(x^\varepsilon)_{\varepsilon>0}$ is a sequence of functions in $K' \cap K(C)$ that converges to a continuous function x^0 of finite α-variation, then $(x^\varepsilon)_{\varepsilon>0}$ also converges to x^0 in β-variation for any $\beta > \alpha$. This follows from

$$\mathrm{Var}_{\beta,[0,T]}(x^\varepsilon - x^0)^\beta \leq 2^{\alpha-1} \|x^\varepsilon - x^0\|_\infty^{\beta-\alpha} (\mathrm{Var}_{\alpha,[0,T]}(x^\varepsilon)^\alpha + \mathrm{Var}_{\alpha,[0,T]}(x^0)).$$

This also implies that $K' \cap K(C)$ is relatively compact with respect to the space of continuous functions of finite α-variation with the norm $\|\cdot\|_{\mathcal{V}^\beta}$ for all $\beta > \alpha$.

Now, since the choice of C and η in (12) depend only on λ and Λ and $(X^\varepsilon)_{\varepsilon>0}$ is tight for the uniform norm, we deduce that for any $\eta > 0$, there exists a constant C such that $\sup_{\varepsilon>0} \mathbb{P}_x [\, X^\varepsilon \in K(C) \cap K' \,] \leq \eta$, where $K' = (X^\varepsilon)_{\varepsilon>0}$, which is relatively compact for the uniform norm according to (13). Thus $K(C) \cap K'$ is relatively compact for the space of continuous functions of finite α-variation with the $\|\cdot\|_{\mathcal{V}^\beta}$-norm for any $\beta > \alpha > 2$.

The similar reasoning is easily carried to deal with the second level. □

Hypothesis 1. *The elements of the sequence $(X^\varepsilon)_{\varepsilon>0}$ belong to $\Xi(\lambda, \Lambda)$ and converge in distribution to X^0 in $\Xi(\lambda, \Lambda)$ in $(\mathcal{C}([0,T];\mathbb{R}^N), \|\cdot\|_\infty)$ under $\mathbb{P}_x$ for any starting point x.*

Remark 1. Under Hypothesis 1 and Proposition 2, it follows immediately from the Prohorov theorem [Bil68] that $(X^\varepsilon)_{\varepsilon>0}$ converges to X^0 in $\mathcal{V}^\alpha$ for any $\alpha > 2$: As the convergence in $\mathcal{V}^\alpha$ implies the uniform convergence, any possible limit of $(X^\varepsilon)_{\varepsilon>0}$ in $\mathcal{V}^\alpha$ is necessarily X^0. Note that however, $\mathcal{V}^\alpha$ is not separable and thus the convergence of $(X^\varepsilon)_{\varepsilon>0}$ in $\mathcal{V}^\alpha$ does not necessarily imply its tightness.

Remark 2. Given a sequence $(X^\varepsilon)_{\varepsilon>0}$ of processes in $\Xi(\lambda, \Lambda)$ corresponding to a sequence of coefficients $(a^\varepsilon, b^\varepsilon)$, it is generally true that any cluster point X^0 belongs to $\Xi(\lambda', \Lambda')$ for some $0 < \lambda' < \Lambda'$. At least, this is always true if $b^\varepsilon = 0$ for any $\varepsilon > 0$. For that, one has to combine the results of [Mar79] and [Roz96b] for example. This means that Hypothesis 1 is not stringent at all.

Under Hypothesis 1 and in view of Proposition 2, a natural question is:

$$\text{Is the limit of } (K(X^\varepsilon))_{\varepsilon>0} \text{ equal to } K(X^0)? \qquad \text{(Q)}$$

As will we show it later in Section 3.5, the answer may depend whether we consider $K(X^\varepsilon)$ as Itô or Stratonovich integrals. In addition, we give in Section 3.3 a sufficient condition that allows one to give a positive answer to this question, but we show in Section 3.5 that it is not a necessary condition.

3.2 Rough paths and geometric rough paths

The iterated integrals $K(X^\varepsilon)$ denote either Itô or Stratonovich type integrals. We will see that our answer to (Q) may depend on the type of integral we consider: for this, we need to explain the difference between the rough paths that are geometric and those which are not.

For this, we follow the way of seeing rough paths as paths with values in a non-commutative group, as introduced first in [FV06a] (see also the introductory article [Lej06b]). Basically, a *rough path* is a path with values in the subspace $\mathrm{T}_1^{(2)}(\mathbb{R}^N)$ of the truncated tensor space

$$\mathrm{T}^{(2)}(\mathbb{R}^N) = \mathbb{R} \oplus (\mathbb{R}^N) \oplus (\mathbb{R}^N \otimes \mathbb{R}^N)$$

whose projection on $\mathbb{R}$ is equal to 1. The space $\mathrm{T}^{(2)}(\mathbb{R}^N)$ is a Lie group with respect to the tensor product $\otimes$ (note that all the terms of type $x \otimes y \otimes z$ with $x, y, z \in \mathbb{R}^N$ vanish and $0 \otimes x = x$ by convention) and an associative algebra with the addition $+$ and the tensor product $\otimes$.

Given a norm on $\mathbb{R}^N \otimes \mathbb{R}^N$ such that $|x \otimes y| \leq |x| \cdot |y|$ for $x, y \in \mathbb{R}^N$, we introduce $\|\mathbf{x}\| = \max\{|\mathbf{x}^1|, \sqrt{|\mathbf{x}^2|}\}$ for $\mathbf{x} \in \mathrm{T}_1^{(2)}(\mathbb{R}^N)$.

Let us consider a non-commutative group $(\mathbb{H}^N, \boxplus)$ group of dimension $N(N+1)/2$ with basis $\{e_1, \ldots, e_d, e_{i,j}\}_{1=i<j=N}$ and the operation

$$\left(\sum_{1\leq i<j\leq N} x^i e_i + x^{i,j} e_{i,j}\right) \boxplus \left(\sum_{1\leq i<j\leq N} y^i e_i + y^{i,j} e_{i,j}\right)$$
$$= \sum_{1\leq i<j\leq N} (x^i + y^i) e_i + \left(x^{i,j} + y^{i,j} + \frac{1}{2}(x^i y^j - x^j x^i)\right) e_{i,j}.$$

This group is a Carnot group of step 2, as it may be decomposed as $\mathbb{H}^N = \mathbb{R}^N \oplus [\mathbb{R}^N, \mathbb{R}^N]$ with $[e_i, e_j] = e_{i,j}$ and $[e_k, e_{i,j}] = [e_{k,\ell}, e_{i,j}] = 0$ for the Lie bracket $[x, y] = x \boxplus y - y \boxplus x$ (see for example [Bau04]). If $N = 2$, then this group is the Heisenberg group. The distance we use on $\mathbb{H}^N$ is the sub-Riemannian metric (see [Mon02, Bau04, CD+07] for example), which is

$$d(x, y) = \inf_{\substack{\gamma:[0,1]\to\mathbb{R}^N \\ \Phi(\gamma)(0)=0,\ \Phi(\gamma)(1)=(-x)\boxplus y \\ \gamma \text{ Lipschitz}}} \int_0^1 |\dot\gamma(s)|\,\mathrm{d}s$$

$$\text{with } \Phi(\gamma)(t) = \sum_{i=1}^N (\gamma^i(t) - \gamma^i(0)) e_i$$
$$+ \sum_{1\leq i<j\leq N} \left(\int_0^t (\gamma^i(s) - \gamma^i(0))\,\mathrm{d}\gamma^j(s) - \int_0^t (\gamma^j(s) - \gamma^j(0))\,\mathrm{d}\gamma^i(s)\right) e_{i,j}.$$

Using the sub-Riemannian metric in the context of rough paths was introduced first by P. Friz and N. Victoir in [FV06a] (see also [Lej06b] for a presentation of this point of view).

A *geometric rough path* is a path with values in the sub-group $\mathrm{G}(\mathbb{R}^N)$ of $(\mathrm{T}_1^{(2)}(\mathbb{R}^N), \otimes)$, which is the image of $\mathbb{H}^N \subset \mathrm{T}^{(2)}(\mathbb{R}^N)$ by the (non-commutative) exponential mapping $\exp(x) = 1 + x + \frac{1}{2} x \otimes x$.

Any geometric rough path $\mathbf{X} = (\mathbf{X}^1, \mathbf{X}^2)$ with $\mathbf{X}_t^1 \in \mathbb{R}^N$ and $\mathbf{X}_t^2 \in \mathbb{R}^N \otimes \mathbb{R}^N$ can be identified as a path $\mathbf{Y}$ with values in $\mathbb{H}^N$ by setting $\mathbf{Y} = \log(\mathbf{X})$ with

$$\log(\mathbf{x}) = \sum_{i=1}^{N} \mathbf{x}^{1,i} e_i + \frac{1}{2} \sum_{1 \leq i < j \leq N} (\mathbf{x}^{2,i,j} - \mathbf{x}^{2,j,i}) e_{i,j} \text{ for } \mathbf{x} = (1, \mathbf{x}^1, \mathbf{x}^2) \in \mathrm{T}_1^{(2)}(\mathbb{R}^N).$$

Of course, log is the inverse of the exponential application exp.

A *smooth rough path* $\mathbf{X} = (\mathbf{X}^1, \mathbf{X}^2)$ is made of a piecewise smooth path $\mathbf{X}^1 = X$ living in $\mathbb{R}^N$ and $\mathbf{X}_t^{2,i,j} = \int_0^t (X_s^i - X_0^i) \, \mathrm{d}X_s^j$. Any geometric rough path of finite α-variation can be approximated by a sequence of smooth rough paths with respect to the topology generated by the β-variation, $\beta > \alpha$ (see [FV06a, Lej06b]).

The difference between two geometric rough paths of finite α-variation lying above the same path necessarily comes from an anti-symmetric path of finite $\alpha/2$-variation.

We then introduce the following sub-spaces of $\mathbb{R}^N \otimes \mathbb{R}^N$:

$$\mathrm{Anti}(\mathbb{R}^N) = \left\{ x \in \mathbb{R}^N \otimes \mathbb{R}^N \,\middle|\, x^{i,j} = -x^{j,i}, \ i, j = 1, \ldots, N \right\}$$

and

$$\mathrm{Sym}(\mathbb{R}^N) = \left\{ x \in \mathbb{R}^N \otimes \mathbb{R}^N \,\middle|\, x^{i,j} = x^{j,i}, \ i, j = 1, \ldots, N \right\}.$$

Lemma 1 (Lemma 5.4.1 in [LQ02])**.** *If $\mathbf{Y}$ is a geometric rough path of finite α-variation and $\psi : [0, T] \to \mathbb{R}^N \otimes \mathbb{R}^N$ is a path of finite $\alpha/2$-variation with values in* $\mathrm{Anti}(\mathbb{R}^N)$*, then $\mathbf{X} = \mathbf{Y} + \psi$ is a geometric rough path of finite α-variation. Conversely, if $\mathbf{X}$ and $\mathbf{Y}$ are geometric rough paths lying above a path X, then there exists a such a function ψ for which $\mathbf{X} = \mathbf{Y} + \psi$.*

As noted first in [LV06], the difference between a geometric rough path and a non-geometric one lies in the presence of a symmetric term.

Lemma 2 ([LV06])**.** *Any rough path $\mathbf{X}$ of finite α-variation, $\alpha \in [2, 3)$, can be decomposed as $\mathbf{X}_t = \mathbf{Y}_t + \psi_t$, where $\mathbf{Y}$ is a geometric rough path of finite α-variation and and ψ is a path of finite $\alpha/2$-variation with values in the space* $\mathrm{Sym}(\mathbb{R}^N \otimes \mathbb{R}^N)$.

Let us note that in the previous Lemmas 2 and 1, $\mathbf{X}_s^{-1} \otimes \mathbf{X}_t = \mathbf{Y}_s^{-1} \otimes \mathbf{Y}_t + \psi_t - \psi_s$ for any $0 \leq s \leq t \leq T$.

Combining Lemma 2 and 1 for a given rough path $\mathbf{X}$ of finite α-variation lying above a path $X : [0,T] \to \mathbb{R}^N$, we get a complete description of any rough path $\mathbf{Y}$ lying above X: there exist two paths ψ and φ of finite $\alpha/2$-variations that are respectively in $\mathrm{Anti}(\mathbb{R}^N)$ and $\mathrm{Sym}(\mathbb{R}^N)$ and such that $\mathbf{Y} = \mathbf{X} + \psi + \varphi$.

In the case of a rough path $\mathbf{X} = (X, K(X))$ generated by a divergence form operator, we have seen that $\mathbf{X}$ is a geometric rough path if we choose for $K(X)$ the Stratonovich integrals. But this is not a geometric one if we choose $K(X)$ to be of Itô type. The difference between these two cases lies in the symmetric term $-\frac{1}{2}\langle M \rangle$, since

$$\int_0^t (X_s^i - X_0^i)\,\mathrm{d}X_s^j = \int_0^t (X_s^i - X_0^i) \circ \mathrm{d}X_s^j - \frac{1}{2}\langle M^i, M^j \rangle_t,\ t \geq 0, \qquad (14)$$

where M is the martingale part for X.

Note that however,

$$\log(\mathbf{X}_t) = X_t^i e_i + \frac{1}{2} \sum_{1 \leq i < j \leq N} (K_{0,t}^{i,j}(X) - K_{0,t}^{j,i}(X)) e_{i,j}$$

so that it does not make any difference to construct the geometric rough path $\log(\mathbf{X})$ as a path living in the group $\mathbb{H}^N$ from the choice of Itô or Stratonovich integrals for the $K^{i,j}(X)$'s.

Let us come back to our question (Q). First, let us note that in view of our previous decomposition results, any limit $\mathbf{Y}$ of $(X^\varepsilon, K(X^\varepsilon))_{\varepsilon>0}$ can be written $\mathbf{Y} = \mathbf{X}^0 + \psi + \varphi$, where $\mathbf{X}^0 = (X^0, K(X^0))$, ψ (resp. φ) is an anti-symmetric (resp. symmetric) path from $[0,T]$ to $\mathbb{R}^N \otimes \mathbb{R}^N$ of finite $\alpha/2$-variation. The effect of φ and ψ on differential equations and integrals driven by $\mathbf{X}^0$ is to add a drift term (see [LQ02, Lej06b]).

If we use for $K(X^\varepsilon)$ ($\varepsilon \geq 0$) the Stratonovich integrals, then necessarily, in view of Lemma 1 and the results from [LV06], the symmetric path φ vanishes since the limit of geometric rough paths is necessarily a geometric rough path.

In order to deal with the Itô case, we may study the Stratonovich case and also the convergence of the brackets $\langle M^\varepsilon \rangle$ of the martingale part M^ε of X^ε.

In the case of Stratonovich integrals, we believe that the answer to (Q) is negative in general. However, we have been unable to exhibit a counter-example. Yet in [Lej02, LL06], we have shown using the homogenization theory that the Lévy area of the limit of SDEs can be different from the Lévy area of the limit. As we will see it in Section 3.5, these results applied

to divergence form operators leads to the right convergence of the Lévy areas, unless we drop the assumption that the drifts are uniformly bounded.

In the Itô case, our homogenization results allows us to give a negative answer, as the limit of the brackets of the martingale parts of the X^ε's will be different from the brackets of the limit.

Thus, we give below in Theorem 1 a sufficient condition to ensure a positive answer to our question in both Itô and Stratonovich cases. Yet, the homogenization result proves that this is not a sufficient condition.

3.3 Condition UTD

Let $(\Pi^n)_{n\in\mathbb{N}}$ be a family of deterministic partitions of $[0,T]$ whose meshes decrease to 0 as $n \to \infty$.

Definition 2 (Condition UTD [RS98]**).** A family of Dirichlet processes $(X^\varepsilon, \mathbb{P})$ (along $(\Pi^n)_{n\in\mathbb{N}}$) with decomposition $X_t^\varepsilon = X_0^\varepsilon + M_t^\varepsilon + A_t^\varepsilon$ for $t \in [0,T]$ satisfies Condition UTD if $(\langle M^\varepsilon \rangle_T)_{\varepsilon>0}$ and $(\sup_{t\in[0,T]} |A_t^\varepsilon|)_{\varepsilon>0}$ are tight, and for any $C > 0$,

$$\sup_{\varepsilon>0} \mathbb{P}\left[\, Q(A^\varepsilon, \Pi^n) > C \,\right] \xrightarrow[n\to+\infty]{} 0.$$

Condition UTD is a natural generalization of Condition UCV for semimartingales (See [KP96] for a review of this notion). The following Proposition is a "specialization" of Theorem 1.1, p. 84 in [RS98] to processes generated by divergence form operators.

Proposition 3 ([RS98]**).** *Let $(X^\varepsilon)_{\varepsilon>0}$ and X^0 be as in Hypothesis 1. We assume that $(X^\varepsilon)_{\varepsilon>0}$ satisfies Condition UTD under $\mathbb{P}_x$. The decomposition of X^ε as a Dirichlet process is written $x + M^\varepsilon + A^\varepsilon$. Then $(M^\varepsilon, A^\varepsilon)_{\varepsilon>0}$ converges in distribution to (M^0, A^0), where $X^0 = x + M^0 + A^0$ is the decomposition of X^0 as a Dirichlet process under $\mathbb{P}_x$.*

Remark 3. It follows from Definition 1 that M^0 and A^0 are both $(\mathcal{F}_t^0)_{t\geq 0}$-adapted. Condition UTD (that does not necessarily concern processes generated by divergence form operators) was introduced in [CS99] where a similar result is given but without ensuring that M^0 and A^0 are $(\mathcal{F}_t^0)_{t\geq 0}$-adapted. It is shown in [RS98] that the limit (M^0, A^0) of $(M^\varepsilon, A^\varepsilon)_{\varepsilon>0}$ is really adapted to the filtration generated by X^0 and may then be identified with the decomposition of X^0 as a Dirichlet process given in Section 2.

Remark 4. Let us note that Condition UTD also implies that $(M^\varepsilon, \langle M^\varepsilon \rangle)_{\varepsilon>0}$ converges in distribution to $(M^0, \langle M^0 \rangle)$ (see for example [KP96, Theorem 7.12, p. 30]). If the limit of $(\langle M^\varepsilon \rangle)_{\varepsilon>0}$ is different from $\langle M^0 \rangle$, then $(X^\varepsilon)_{\varepsilon>0}$ does not satisfy Condition UTD. This can be an easy way to identify the sequences

$(X^\varepsilon)_{\varepsilon>0}$ for which more work has to be done just by studying the convergence of the brackets of their martingales parts.

Remark 5. From the proof of Theorem 2.2 in [RS98], we see that X^ε satisfies Condition UTD under $\mathbb{P}_x$ for any $x \in \mathbb{R}^N$ if for any compact K and φ_i is a function with compact support of $\mathbb{R}^N$ such that $\varphi_i(x) = x_i$ on K, $\alpha G^\varepsilon_\alpha \varphi_i$ converges uniformly in ε to φ_i in $\mathrm{H}^1(K)$ as $\alpha \to \infty$ for $i = 1, \ldots, N$, where G^ε_α is the resolvent operator $(\alpha - L^\varepsilon)^{-1}$ and L^ε is the infinitesimal generator of L^ε. This is equivalent to the uniform convergence of $L^\varepsilon G^\varepsilon_\alpha \varphi_i$ to 0 in $\mathrm{H}^1(K)$ uniformly in ε as $\alpha \to \infty$.

Lemma 3. *Let g and φ be some functions in $\Upsilon^\infty(\mathbb{R}^N)$ and $(X^\varepsilon)_{\varepsilon>0}$ a sequences of processes in $\Xi(\lambda, \Lambda)$. Then for all $x \in \mathbb{R}^N$ and all $\varepsilon > 0$, $\mathfrak{L}^n(t; X^\varepsilon, g, \varphi)$ converges in probability under $\mathbb{P}_x$ to $\mathfrak{L}(t; X^\varepsilon, g, \varphi)$ uniformly in $t \in [0, T]$. If in addition, $(g(X^\varepsilon))_{\varepsilon>0}$ satisfies Condition UTD, then for any $\delta > 0$ and any $C > 0$, there exists n_0 large enough such that*

$$\sup_{\varepsilon>0} \mathbb{P}_x \left[\sup_{t\in[0,T]} |\mathfrak{L}^n(t; X^\varepsilon, g, \varphi) - \mathfrak{L}^n(t; X^\varepsilon, g, \varphi)| \geq C \right] \leq \delta$$

for any $n \geq n_0$ and any $x \in \mathbb{R}^N$.

These statements are true both for the Stratonovich and the Itô integrals.

Proof. Let X be in $\Xi(\lambda, \Lambda)$. For $R > 0$, let us denote by $\Phi(x, R)$ the event $\{\sup_{t\in[0,T]} |X_t - x| \geq R\}$. One knows from Lemma II.1.12 in [Str88] that there exist some constants K and K' that depend only on λ, Λ, T and the dimension N such that

$$\mathbb{P}_x [\, \Phi(x, R) \,] \leq K \exp\left(-K' R^2\right). \tag{15}$$

Hence, for any $C > 0$, $t \in [0, T]$ and $R > 0$,

$$\begin{aligned}\mathbb{P}_x \left[\sup_{t\in[0,T]} |\mathfrak{L}^n(t; X, g, \varphi) - \mathfrak{L}(t; X, g, \varphi)| \geq C \right] &\leq \mathbb{P}_x [\, \Phi(x, R) \,] \\ &+ \mathbb{P}_x \left[\sup_{t\in[0,T]} |\mathfrak{L}^n(t; X, \widetilde{g}, \widetilde{\varphi}) - \mathfrak{L}(t; X, \widetilde{g}, \widetilde{\varphi})| \geq C; \Phi(x, R)^{\mathrm{c}} \right],\end{aligned}$$

where $\widetilde{g}$ (resp. $\widetilde{\varphi}$) is any continuous function equal to g (resp. φ) on the ball of radius R around x.

Thus, using this localization argument, one may assume that g and φ are bounded on $\mathbb{R}^N$ when x is fixed.

To estimate the speed of convergence of $\mathfrak{L}^n(\cdot; X, g, \varphi)$ to $\mathfrak{L}(\cdot; X, g, \varphi)$, and in view of (8), we have only to consider the speed of convergence to 0 of

$$I_1^n(t) = \int_0^t g(X_r)\,\mathrm{d}M_r^\varphi - \sum_{i=1}^{k^n(t)-1} g(X_{t_i^n})(M^\varphi_{t_{i+1}^n} - M^\varphi_{t_i^n}),$$

$$I_2^n(t) = \int_0^t g(\overline{X}_r)\,\mathrm{d}\overline{M}_r^\varphi - \sum_{i=1}^{k^n(t)-1} g(\overline{X}_{t-t_i^n})(\overline{M}^\varphi_{t-t_i} - \overline{M}^\varphi_{t-t_{i+1}^n}),$$

$$I_3^n(t) = \int_0^t g(X_r)\,\mathrm{d}V_r^\varphi - \sum_{i=1}^{k^n(t)-1} g(X_{t_i^n})(V^\varphi_{t_{i+1}^n} - V^\varphi_{t_i^n}),$$

$$I_4^n(t) = \langle g(\overline{X}), \overline{M}^\varphi\rangle_t - \sum_{i=1}^{k^n(t)-1} (g(\overline{X}_{t-t_{i+1}^n}) - g(\overline{X}_{t-t_i^n}))(\overline{M}^\varphi_{t-t_{i+1}^n} - \overline{M}^\varphi_{t-t_i^n})$$

and, in order to deal with the Itô integral,

$$I_5^n(t) = \langle g(X), M^\varphi\rangle_t - \sum_{i=1}^{k^n(t)-1} (g(X_{t_{i+1}^n}) - g(X_{t_i^n}))(M^\varphi_{t_{i+1}^n} - M^\varphi_{t_i^n}),$$

where $0 \le t_1^n \le \cdots \le t^n_{k^n(t)} \le t$ are the points of $\Pi^n \cap [0,t]$.

We have seen in [Lej06a, Lemma 3] that for any $\beta < 1/2$, there exists some random variable C_β such that $\mathbb{E}_x[C_\beta]$ is finite and depends only on $(\lambda, \Lambda, \beta)$ and such that X is β-Hölder continuous with constant C_β (and so is $\overline{X}$). In addition, since this relies on the Kolmogorov Lemma [KS91, Theorem 2.1, p. 25], it is easily obtained that $\mathbb{E}_x\left[C_\beta^2\right]$ is also finite and depends only on $(\lambda, \Lambda, \beta)$. Since g has a Lipschitz continuous version (this result follows from the Morrey theorem [A75, Theorem 5.4, p. 97]), we get that $g(X)$ and $g(\overline{X})$ are β-Hölder continuous with constant $C_\beta\|\nabla g\|_\infty$ when this version of g is used. As

$$\langle M^\varphi, M^\varphi\rangle_t = \int_0^t a\nabla\varphi\cdot\nabla\varphi(X_r)\,\mathrm{d}r \le \Lambda t\|\nabla\varphi\|_\infty^2 \text{ for any } t \ge 0,$$

the Burkholder-Davis-Gundy inequality implies that for some constant K,

$$\begin{aligned}\mathbb{E}_x\left[\sup_{t\in[0,T]} I_1^n(t)^2\right] &\le K\mathbb{E}_x\left[\int_0^T (g(X_r) - g(X_{t_i^n}))^2\mathbf{1}_{r\in[t_i^n, t_{i+1}^n)}\,\mathrm{d}\langle M^\varphi, M^\varphi\rangle_r\right]\\ &\le KT\Lambda\|\nabla\varphi\|_\infty^2\|\nabla g\|_\infty^2\mathbb{E}_x\left[C_\beta^2\right](\operatorname{mesh}\Pi^n)^\beta.\end{aligned}$$

Also $\langle\overline{M}^\varphi, \overline{M}^\varphi\rangle_t \le \Lambda t\|\nabla\varphi\|_\infty^2$ for any $t \ge 0$ and similarly,

$$\mathbb{E}_x\left[\sup_{t\in[0,T]} I_2^n(t)^2\right] \le KT\Lambda\|\nabla\varphi\|_\infty^2\|\nabla g\|_\infty^2\mathbb{E}_x\left[C_\beta^2\right](\operatorname{mesh}\Pi^n)^\beta.$$

As seen in [Lej06a], the term V^φ is of finite variation and $\mathbb{E}_x\left[\operatorname{Var}_{1,[0,T]} V^\varphi\right]$ is finite and bounded by a constant depending only on λ, Λ and $\|\nabla\varphi\|_\infty$. Besides, $I_3^n(t) = \int_0^t Z_r^n \,\mathrm{d}V_r^\varphi$ with $Z_r^n = g(X_r) - g(X_{t_i^n})$ when $r \in [t_i^n, t_{i+1}^n)$. Hence, $\mathbb{E}_x\left[\sup_{r\in[0,T]} |Z_r^n|\right] \le \mathbb{E}_x\left[C_\beta\right] \|\nabla g\|_\infty (\operatorname{mesh}\Pi^n)^\beta$. Hence, for any $C > 0$ and any $K > 0$,

$$\begin{aligned}\mathbb{P}_x\left[\sup_{t\in[0,T]} |I_3^n(t)| > C\right] &\le \mathbb{P}_x\left[\sup_{t\in[0,T]} |Z_r^n| \operatorname{Var}_{1,[0,T]} V^\varphi > C\right]\\ &\le \mathbb{P}_x\left[\sup_{t\in[0,T]} |Z_r^n| > K\right] + \mathbb{P}_x\left[\operatorname{Var}_{1,[0,T]} V^\varphi > \frac{C}{K}\right]\\ &\le \|\nabla g\|_\infty \frac{\mathbb{E}_x\left[C_\beta\right]}{K} (\operatorname{mesh}\Pi^n)^\beta + \frac{K}{C}\mathbb{E}_x\left[\operatorname{Var}_{1,[0,T]} V^\varphi\right].\end{aligned}$$

For any $\varepsilon > 0$ and any $C > 0$, one may first choose K small enough such that $KC^{-1}\mathbb{E}_x\left[\operatorname{Var}_{1,[0,T]} V^\varphi\right] < \varepsilon/2$. Then, we choose n_0 large enough such that for any $n \ge n_0$, $\|\nabla g\|_\infty K^{-1}\mathbb{E}_x\left[C_\beta\right](\operatorname{mesh}\Pi^n)^\beta \le \varepsilon/2$ for any $n \ge n_0$. It yields that $\mathbb{P}_x\left[\sup_{t\in[0,T]} |I_3^n(t)| > C\right] \le \varepsilon$ for any $n \ge n_0$ and and n_0 may be chosen in function of $\operatorname{mesh}(\Pi^n)$, β, $\|\nabla g\|_\infty$, $\|\nabla\varphi\|_\infty$, λ, Λ, and of course, ε and C.

We now estimate of the speed of convergence of I_4^n to 0 in probability.

It is easily seen that for $s \in [0,T]$, $g(\overline{X}_s) = g(\overline{X}_0) + \frac{1}{2}\overline{M}_s^g + \frac{1}{2}(M_{T-s}^g - M_t^g) + V_{T-s}^g - V_T^g$. Since V^g is of integrable variation and $g(X_t)$ is a Dirichlet process,

$$\overline{B}_t^g = g(\overline{X}_0) + V_{T-t}^g - V_T^g - \frac{1}{2}\overline{M}_t^g + \frac{1}{2}M_{T-t}^g - \frac{1}{2}M_T^g$$

is of zero quadratic variation along $(\overline{\Pi}^n)_{n\in\mathbb{N}}$ where $\overline{\Pi}^n = \{\, T - t \,|\, t \in \Pi^n \,\}$. Indeed, it is the term $t \in [0,T] \mapsto M_t^g - \overline{M}_{T-t}^g$ which is of zero-quadratic variation along $(\Pi^n)_{n\in\mathbb{N}}$.

To simplify the notations, we set for two processes X and Y,

$$Q_t^n(X,Y) = \sum_{i=1}^{k^n(t)-1} (X_{T-t_{i+1}^n} - X_{T-t_i^n})(Y_{T-t_{i+1}^n} - Y_{T-t_i^n}).$$

Using the previous decomposition of $g(\overline{X})$,

$$Q_t^n(g(\overline{X}), \overline{M}^\varphi) = Q_t^n(\overline{M}^g, \overline{M}^\varphi) + Q_t^n(\overline{B}^g, \overline{M}^\varphi),$$

and $\langle g(\overline{X}), \overline{M}^\varphi\rangle = \langle \overline{M}^g, \overline{M}^\varphi\rangle$ almost surely.

As seen in Lemma 6 in Appendix A, the family $(t \mapsto \langle\overline{M}^\varphi\rangle_t - \sum_{i=1}^{k^n(t)-1}(\overline{M}_{t_{i+1}^n}^\varphi - \overline{M}_{t_i^n}^\varphi)^2)_{n\in\mathbb{N}}$ converges in probability to 0 uniformly in $t \in [0,T]$ at a speed that

depends only mesh Π^n, $\|\nabla\varphi\|_\infty$ and Λ, since $\langle\overline{M}^\varphi\rangle_t - \langle\overline{M}^\varphi\rangle_s \leq \Lambda(t-s)\|\nabla\varphi\|_\infty^2$ for any $0 \leq s \leq t$.

Hence, it remains to study the speed of convergence of $Q_t^n(\overline{B}^g, \overline{M}^\varphi)$ to 0.

From the Cauchy-Schwarz inequality and since $t \mapsto Q_t^n(\overline{B}^g)$ and $t \mapsto Q_t^n(\overline{M}^\varphi)$ are increasing,

$$\sup_{t\in[0,T]} |Q_t^n(\overline{B}^g, \overline{M}^\varphi)| \leq Q_T^n(\overline{B}^g)^{1/2} Q_T^n(\overline{M}^\varphi)^{1/2}$$

with $Q_t^n(\overline{B}^g) = Q_t^n(\overline{B}^g, \overline{B}^g)$ and $Q_t^n(\overline{M}^\varphi) = Q_t^n(\overline{M}^\varphi, \overline{M}^\varphi)$.

As for any $a, b, \eta > 0$, $\sqrt{ab} \leq \eta a + \eta^{-1} b$, for any $C > 0$,

$$\begin{aligned}\mathbb{P}_x\left[\sup_{t\in[0,T]} Q_t^n(\overline{B}^g, \overline{M}^\varphi) > C\right] &\leq \mathbb{P}_x\left[\, Q_T^n(\overline{B}^g) \geq C/\eta \,\right] + \mathbb{P}_x\left[\, Q_T^n(\overline{M}^\varphi) \geq \eta C \,\right] \\ &\leq \mathbb{P}_x\left[\, Q_T^n(\overline{B}^g) \geq C/\eta \,\right] + \mathbb{P}_x\left[\, |Q_T^n(\overline{M}^\varphi) - \langle\overline{M}^\varphi\rangle_T| \geq \eta C/2 \,\right] \\ &\qquad + \mathbb{P}_x\left[\, \langle\overline{M}^\varphi\rangle_T \geq \eta C/2 \,\right].\end{aligned}$$

In addition $\mathbb{P}_x\left[\, \langle\overline{M}^\varphi\rangle_T \geq \eta C/2 \,\right] \leq \mathbb{P}_x\left[\, T\Lambda\|\nabla\varphi\|_\infty^2 \geq \eta C/2 \,\right]$. We choose first η and then n_0 large enough so that $\mathbb{P}_x\left[\, |Q_T^n(\overline{M}^\varphi) - \langle\overline{M}^\varphi\rangle_T| \geq \eta C/2 \,\right] \leq \varepsilon/2$ and

$$\mathbb{P}_x\left[\, Q_T^n(\overline{B}^g) \geq C/\eta \,\right] \leq \varepsilon/2 \tag{16}$$

for any $n \geq n_0$. The last convergence holds since $\overline{B}^g$ is of zero quadratic variation along $(\overline{\Pi}^n)_{n\in\mathbb{N}}$.

The convergence of $I_5^n(t)$ to 0 is proved using similar computations by replacing $\overline{B}^g$ with $B_t^g = g(x) - \frac{1}{2}M_t^g + \frac{1}{2}(\overline{M}_{T-t}^g - \overline{M}_T^g) + V_t^g$ of zero quadratic variation along $(\Pi^n)_{n\in\mathbb{N}}$.

We deduce that $\mathfrak{L}^n(t; X, g, \varphi)$ converges to $\mathfrak{L}(t; X, g, \varphi)$ in probability uniformly in $t \in [0, T]$. Moreover, except *a priori* for the convergence of $Q_t^n(\overline{B}^g)$ and $Q_t^n(B^g)$ to 0, the rate of convergence depends only on λ, Λ, T, $\|\nabla g\|_\infty$, $\|\nabla\varphi\|_\infty$ and mesh(Π^n).

We now apply this result on X^ε. As $(g(X^\varepsilon))_{\varepsilon>0}$ satisfies Condition UTD, it is easily obtained that $(g(\overline{X}^\varepsilon))_{\varepsilon>0}$ also satisfies Condition UTD, but along the family of partitions $(\overline{\Pi}^n)_{n\in\mathbb{N}}$, where $\overline{\Pi}^n$ is constructed from Π^n by changing any point t in Π^n into $T - t$. Using Condition UTD in (16) with $\overline{B}^g$ replaced by the zero-quadratic variation term $\overline{B}^{\varepsilon,g}$ of $g(\overline{X}^\varepsilon)$ and $B^{\varepsilon,g}$ of $g(X^\varepsilon)$, one sees that the speed of convergence of $\mathfrak{L}^n(\cdot; X^\varepsilon, g, \varphi)$ to $\mathfrak{L}(\cdot; X^\varepsilon, g, \varphi)$. is uniform in ε. □

3.4 A criterion ensuring the convergence

The goal of this section is to prove the following convergence theorem. Of course, its interest lies in the continuity result of the integrals and the solu-

tions of the differential equations driven by rough paths: See Theorem 5.5.2, p. 143 and Corollary 6.3.2, p. 179 in [LQ02].

Theorem 1. *If $(X^\varepsilon)_{\varepsilon>0}$ satisfies Hypothesis 1 and Condition UTD, then $(X^\varepsilon, K(X^\varepsilon))_{\varepsilon>0}$ converges in distribution under $\mathbb{P}_x$ to $(X^0, K(X^0))$ in $\mathcal{V}^\alpha$ for any $\alpha > 2$ for any starting point x.*

Remark 6. We will see below in Section 3.5 that Condition UTD is sufficient but not necessary.

Before proving this theorem, let us state a useful application.

Corollary 1. *Let $(a^0, b^0) \in \Xi^{\text{coeff}}(\lambda, \Lambda)$ and $(a^\varepsilon, b^\varepsilon) \in \Xi^{\text{coeff}}(\lambda, \Lambda)$ such that $(a^\varepsilon, b^\varepsilon)$ converges almost everywhere to (a, b). For $\varepsilon \geq 0$, let X^ε be the process corresponding to $(a^\varepsilon, b^\varepsilon)$. Then $(X^\varepsilon, K(X^\varepsilon))$ converges in distribution under $\mathbb{P}_x$ to $(X^0, K(X^0))$ in $\mathcal{V}^\alpha$ for any $\alpha \geq 2$ for any starting point x.*

Proof. It is well known that the convergence of $(a^\varepsilon, b^\varepsilon)$ to (a^0, b^0) ensures the convergence in distribution of X^ε to X^0 under $\mathbb{P}_x$ for any $x \in \mathbb{R}^N$ (see [Str88, Roz96b] for example).

If φ is a smooth function with compact support,then the core of the proof of Theorem 2.2 in [RS98] is to establish that $(\varphi(X^\varepsilon))_{\varepsilon>0}$ satisfies Condition UTD (see in particular (2.33) in [RS98, p. 103]).

With a localization argument, this proves that $(X^\varepsilon)_{\varepsilon>0}$ also satisfies Condition UTD.

For this, let us note first that, if $M^\varepsilon = (M^{1,\varepsilon}, \ldots, M^{N,\varepsilon})$ is the martingale part of X^ε, then $\langle M^{i,\varepsilon}\rangle_T \leq \Lambda T$ for $i = 1, \ldots, N$ and then $(\langle M^\varepsilon\rangle_T)_{\varepsilon>0}$ is tight.

Now, as in the proof of Lemma 3, let us set $\Phi^\varepsilon(R) = \{\sup_{t\in[0,T]} |X_t^\varepsilon - x| \geq R\}$ for $R > 0$, and choose $\varphi(x)$ such that $\varphi(x) = x_i$ on the ball $\{y \in \mathbb{R}^N | |y - x| \leq R\}$. Let us denote $A^{R,\varepsilon}$ (resp. A^ε) the term of zero-quadratic variation of $\varphi(X^\varepsilon)$ (resp. X^ε). Then

$$\mathbb{P}_x\left[\sup_{t\in[0,T]} A^\varepsilon > C\right] \leq \mathbb{P}_x\left[\sup_{t\in[0,T]} A^\varepsilon > C; \Phi^\varepsilon(R)^{\text{c}}\right] + \mathbb{P}_x\left[\,\Phi^\varepsilon(R)\,\right]$$

On $\Phi^\varepsilon(R)$, $A^\varepsilon = A^{R,\varepsilon}$, so that

$$\mathbb{P}_x\left[\sup_{t\in[0,T]} A^\varepsilon > C\right] \leq \mathbb{P}_x\left[\sup_{t\in[0,T]} A^{R,\varepsilon} > C\right] + \mathbb{P}_x\left[\,\Phi^\varepsilon(R)\,\right]$$

With (15) (which is uniform in ε), it is easily deduced that $(\sup_{t\in[0,T]} A^\varepsilon)_{\varepsilon>0}$ is tight. The same computation with $\sup_{t\in[0,T]}(A^\varepsilon)$ replaced by $Q(A^\varepsilon, \Pi^n)$ proves that $(X^\varepsilon)_{\varepsilon>0}$ satisfies Condition UTD.

Theorem 1 allows us to conclude. □

To prove Theorem 1, let us remark first that in the Stratonovich or the Itô case, for $i, j = 1, \ldots, N$,

$$K_{0,t}^{i,j}(X^\varepsilon) = \mathfrak{L}(t; X^\varepsilon, \chi_i, \chi_j) \text{ with } \chi_i(x) = x_i.$$

In addition, for all $\varepsilon \geq 0$,

$$K_{s,t}^{i,j}(X^\varepsilon) = K_{0,t}^{i,j}(X^\varepsilon) - K_{0,s}^{i,j}(X^\varepsilon) - (X_s^i - X_0^i)(X_t^j - X_s^j), \ \forall (s,t) \in \Delta_+, \quad (17)$$

where $\Delta_+ = \{ (s,t) \in [0,T]^2 \,|\, 0 \leq s \leq t \leq T \}$.

Theorem 1 is proved using Lemma 3 and the following Lemma.

Lemma 4. *If for $x \in \mathbb{R}^N$, any $t \in [0,T]$, $i, j = 1, \ldots, N$ and any constant $C > 0$,*

$$\limsup_{n \to \infty} \sup_{\varepsilon > 0} \mathbb{P}_x \left[\, |\mathfrak{L}^n(t; X^\varepsilon, \chi_i, \chi_j) - \mathfrak{L}(t; X^\varepsilon, \chi_i, \chi_j)| \geq C \, \right] = 0 \qquad (18)$$

then $(X^\varepsilon, K(X^\varepsilon))$ converges in distribution under $\mathbb{P}_x$ to $(X^0, K(X^0))$ as $\varepsilon \to 0$ in $\mathcal{V}^\alpha$ for any $\alpha > 2$.

Proof. As for each integer n and each $t \in [0,T]$, $(X^\varepsilon, \mathfrak{L}^n(t; X^\varepsilon, \chi_i, \chi_j))$ converges in distribution to $(X^0, \mathfrak{L}^n(t; X^0, \chi_i, \chi_j))$, Condition (18) together with Theorem 4.2 in [Bil68, p. 25] allow us to assert that $(X^\varepsilon, K_{0,t}^{i,j}(X^\varepsilon))$ with $K_{0,t}^{i,j}(X^\varepsilon) = \mathfrak{L}(t; X^\varepsilon, \chi_i, \chi_j)$ converges in distribution to $(X^0, K_{0,t}^{i,j}(X^0))$. With (17), $(X^\varepsilon, K_{s,t}^{i,j}(X^\varepsilon))$ converges in distribution to $(X^0, K_{s,t}^{i,j}(X^0))$ for all $(s,t) \in \Delta_+$. The tightness of $(X^\varepsilon, K(X^\varepsilon))_{\varepsilon > 0}$ in $\mathcal{V}^\alpha$ (see Proposition 2) and the continuity of $(s,t) \mapsto K_{s,t}(X^0)$ allows us to uniquely identify any possible limit of $(X^\varepsilon, K(X^\varepsilon))_{\varepsilon > 0}$ with $(X^0, K(X^0))$. □

3.5 Examples from the homogenization theory

Let us consider (a, b) in $\Xi^{\text{coeff}}(\lambda, \Lambda)$ such that $b = 0$ and the diffusion coefficient a is smooth and 1-periodic. Set $a^\varepsilon(x) = a(x/\varepsilon)$ and $b^\varepsilon = 0$. For any $\varepsilon > 0$, $(a^\varepsilon, b^\varepsilon)$ belongs to $\Xi^{\text{coeff}}(\lambda, \Lambda)$ and gives rise to a stochastic process X^ε which is solution to the SDE

$$X_t^\varepsilon = x + \int_0^t \sigma(X_s^\varepsilon/\varepsilon) \, \mathrm{d}B_s^\varepsilon + \frac{1}{2\varepsilon} \int_0^t \nabla a(X_s^\varepsilon/\varepsilon) \, \mathrm{d}s, \qquad (19)$$

$\mathbb{P}_x$-almost surely, where B^ε is a Brownian motion. Using a scaling argument, one gets that X^ε is equal to distribution to $\varepsilon X_{t/\varepsilon}$ under $\mathbb{P}_{x/\varepsilon}$, where X is process corresponding to $(a, 0)$.

3.5.1 Convergence of the process to a non-standard Brownian motion

The homogenization theory aims to study the asymptotic limit of X^ε. In the case of smooth coefficients, it is standard that X^ε converges to some non-standard Brownian motion, whose diffusion coefficient "catches" the large-scale properties of the behavior of X: see for example the books [BPL78, JKO94]. The case of processes generated by a divergence form operators, which we present here with smooth coefficients, can be carried without any regularity assumptions, as shown in [Lej00, Lej01].

Let us characterize the constant matrix σ that is such that $(X^\varepsilon)_{\varepsilon>0}$ converges to $\sigma\overline{B}$ for a N-dimensional Brownian motion $\overline{B}$.

For $i = 1, \dots, N$, let v_i be the variational solution to

$$\sum_{k,j=1,\dots,N} \frac{\partial}{\partial x_k}\left(a_{k,j}(x)\frac{\partial v_i(x)}{\partial x_j}\right) = -\sum_{j=1}^N \frac{a_{i,j}(x)}{\partial x_j},\ v_i \text{ periodic.}$$

Thanks to the Fredholm alternative, this problem has a unique solution which is, since a is smooth, a classical solution. The functions v_i are called *correctors*. With the Itô formula applied to $x \mapsto x + v_i(x)$ and an ergodic theorem applied to the projection of X on the torus (whose invariant measure is the Lebesgue measure on $[0,1]^N$), one gets that for any $t \geq 0$,

$$X_t^\varepsilon + (v_1^\varepsilon(X_t^\varepsilon), \dots, v_N^\varepsilon(X_t^\varepsilon)) = x + (v_1^\varepsilon(x), \dots, v_N^\varepsilon(x)) + R_t^\varepsilon,$$

where, for $i = 1, \dots, N$, $v_i^\varepsilon(x) = \varepsilon v_i(x/\varepsilon)$ (which decreases uniformly to 0) and R^ε is a martingale such that for $i, j = 1, \dots, N$,

$$\begin{aligned}\langle R^{i,\varepsilon}, R^{j,\varepsilon}\rangle_t &\\ &\overset{\text{dist.}}{=} \sum_{k,\ell=1}^N \varepsilon^2 \int_0^{t/\varepsilon^2} a_{k,\ell}(X_s)\left(\delta_{i,k} + \frac{\partial v_i(X_s)}{\partial x_k}\right)\left(\delta_{j,\ell} + \frac{\partial v_j(X_s)}{\partial x_\ell}\right) \mathrm{d}s.\end{aligned}$$

With the ergodic theorem, for any $t \geq 0$, $\langle R^{i,\varepsilon}, R^{j,\varepsilon}\rangle_t$ converges almost surely under $\mathbb{P}_x$ to ta^{eff} with

$$a_{i,j}^{\text{eff}} \overset{\text{def.}}{=} \sum_{k,\ell=1}^N \int_{[0,1]^N} a_{k,\ell}(x)\left(\delta_{i,k} + \frac{\partial v_i(x)}{\partial x_k}\right)\left(\delta_{j,\ell} + \frac{\partial v_j(x)}{\partial x_\ell}\right) \mathrm{d}x \tag{20}$$

for $i, j = 1, \dots, N$. It follows from a Central Limit Theorem on the martingales [EK86] that there exists a N-dimensional Brownian motion $\overline{B}$ such that R^ε converges in distribution to $\sigma\overline{B}$ with $\sigma\sigma^{\mathrm{T}} = a^{\text{eff}}$. The coefficient a^{eff} is called an *effective coefficient*. It follows that X^ε converges in distribution in the space of continuous functions to $\overline{X} = \sigma\overline{B}$.

3.5.2 Convergence of the iterated integrals in the Stratonovich case

In [Lej02] and [LL06], we have studied the convergence of $(X^\varepsilon, K(X^\varepsilon))_{\varepsilon>0}$ for X^ε equal to $\varepsilon X_{./\varepsilon^2}$, where X is the solution of some SDE with periodic coefficients. There, we have proved that the Lévy area of the limit may be different from the limits of the Lévy area. The difference lies in a function of type $t \mapsto ct$, where c is an anti-symmetric matrix seen as an element of $\mathbb{R}^N \otimes \mathbb{R}^N$.

In our context, when we use for X^ε the solution to (19), it follows from a short computation performed on the formula given in Proposition 5 in [Lej02] that in the Stratonovich case, $(X^\varepsilon, K(X^\varepsilon))$ converges in distribution in $\mathcal{V}^\alpha$ to $(\overline{X}, K(\overline{X}))$, where $\overline{X}$ is given above (let us recall that the Lévy area of X^ε is just the anti-symmetric part of matrix values functions $(K^{i,j}(X))_{i,j=1,\dots,N}$).

Remark 7. There is an error in the statement of Proposition 5 in [Lej02], although the proof is correct. The article [LL06] presents a correct statement.

Now, we drop the assumption that a is smooth, so that X^ε is the process generated by $\sum_{i,j=1}^N \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{i,j}(\cdot/\varepsilon)\frac{\partial}{\partial x_j}\right)$ where a is measurable, uniformly elliptic, bounded and 1-periodic. We denote by $\overline{X}$ its limit which equal to $\sigma\overline{B}$ as above.

Although the homogenization results for X^ε holds without any smoothness assumption on a, we have assumed that a is smooth in order to apply the results in [Lej02]. The proof given in this article relies on the Itô stochastic calculus and the ergodic theorem, as in Section 3.5.1. Using the recent article [FV06b] from P. Friz and N. Victoir as well as analytical tools, we can extend the previous convergence results to the case of discontinuous coefficients.

Theorem 2. *Let $K(X^\varepsilon)$ (resp. $K(\overline{X})$) denotes the iterated integrals of X^ε (resp. $\overline{X}$) constructed with the Stratonovich integrals. Then $(X^\varepsilon, K(X^\varepsilon))$ converges in α-variation to $(\overline{X}, K(\overline{X}))$ for any $\alpha > 2$ under $\mathbb{P}_x$ for any $x \in \mathbb{R}^N$.*

Proof. In this proof, we use the notations introduced in Section 3.2.

Let D_i be the differential operator

$$D_i = \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{1\le j<i\le N} x^j \frac{\partial}{\partial x_{j,i}} - \frac{1}{2}\sum_{1\le i<j\le N} x^j \frac{\partial}{\partial x_{i,j}},$$

where $\frac{\partial}{\partial x_{i,j}}$ is the derivative in the direction $e_{i,j}$. We denote by $\mathrm{H}^1(\mathbb{H}^N)$ the completion of the space of smooth functions with compact support on $\mathbb{H}^N$

with respect to the norm

$$\left(\int_{\mathbb{H}^N} |f(x)|^2 m(\mathrm{d}x) + \sum_{i=1}^{N} \int_{\mathbb{H}^N} |D_i f(x)|^2 m(\mathrm{d}x) \right)^{1/2},$$

where $m(\mathrm{d}x)$ is the Haar measure on $\mathbb{H}^N$.

In [FV06b], P. Friz and N. Victoir consider the bilinear form

$$\mathcal{E}(f,g) = \sum_{i,j=1}^{N} \int_{\mathbb{H}^N} a_{i,j}(x) D_i f(x) D_j g(x) m(\mathrm{d}x), \;\; f,g \in \mathrm{H}^1(\mathbb{H}^N),$$

when the coefficient a is a measurable function from $\mathbb{H}^N$ to the space of symmetric matrices that are uniformly elliptic and bounded. To this bilinear form corresponds a semi-group $(P_t)_{t>0}$ that maps $\mathrm{L}^1(\mathbb{H}^N)$ into $\mathrm{L}^\infty(\mathbb{H}^N)$. The semi-group $(P_t)_{t>0}$ has a density $p(t,x,y)$ that satisfies a Gaussian type estimate:

$$p(t,x,y) \leq C_1 \exp\left(-\frac{d(x,y)^2}{C_2 t} \right), \tag{21}$$

where C_1 and C_2 depend only on λ and Λ and $d(x,y)$ is the sub-Riemannian distance between x and y defined in Section 3.2. From this, it is easily obtained that the semi-group $(P_t)_{t>0}$ is a Feller semi-group, and then that it generates a diffusion process $\mathbf{X}$ with values in $\mathbb{H}^N$.

As they proved a theorem of type Wong-Zakai for $\mathbf{X}$ along the dyadic partition using the piecewise linear interpolation, the rough path generated by the Dirichlet form $\mathcal{E}$ is the same as $(X, K(X))$ — up to the application of the exponential map from $\mathbb{H}^N$ to $\mathrm{T}_1^{(2)}(\mathbb{R}^N)$ — where we use the Stratonovich integral for $K(X)$.

Now, let $(P_t^\varepsilon)_{t>0}$ be the semi-groups associated to the Dirichlet forms $\mathcal{E}^\varepsilon(f,g) = \sum_{i,j=1}^N \int_{\mathbb{H}^N} a_{i,j}(x/\varepsilon) D_i f(x) D_j g(x) m(\mathrm{d}x)$, where a is 1-periodic. It has been proved in [BBJR95, Corollary 3.2] that for any f in $\mathrm{L}^2(\mathbb{H}^N)$ and any $t \geq 0$, $P_t^\varepsilon f$ converges in $\mathrm{L}^2(\mathbb{H}^N)$ to $\overline{P}_t f$, where $(\overline{P}_t)_{t>0}$ is the semi-group generated by $\overline{\mathcal{E}}(f,g) = \sum_{i,j=1}^N \int_{\mathbb{H}^N} a_{i,j}^{\mathrm{eff}} D_i f(x) D_j g(x) m(\mathrm{d}x)$, for a^{eff} defined in (20) when the coefficient a is only a function of $\mathbb{R}^N$ and not of $\mathbb{H}^N$, which is our case.

Using the Gaussian bound on (21), we obtain easily that for a function f on $\mathbb{H}^N$ that vanishes at infinity, then $P_t^\varepsilon f$ converges uniformly to $\overline{P}_t f$, when using integral representations of the semi-groups $(P_t^\varepsilon)_{t>0}$ and $(\overline{P}_t)_{t>0}$ with their transition density functions. It follows then from standard results (see for example Theorem 2.5 in [EK86] or [Roz96b] for the case of divergence-form operators) that the process generated by $(P_t^\varepsilon)_{t>0}$ converges uniformly in distribution to the process generated by $(\overline{P}_t)_{t>0}$ under $\mathbb{P}_x$ for any $x \in \mathbb{R}^N$. $\square$

Remark 8. In [FV06b], P. Friz and N. Victoir actually deal with the $(1/\alpha)$-Hölder norm, instead of the α-variation norm, which gives more precise results. In a forthcoming article, we will study the relationship between the constructions and results presented here and in [Lej06a] and those in [FV06b].

Let us remark that if we drop the assumption that the drift b is uniformly bounded, it is easy to consider a sequence of processes with a different Lévy area. As noted at the end of Introduction of [Lej06a], our theory may be applied to the case of time-homogeneous coefficients. For example, one has only to consider the family of the two-dimensional SDEs

$$X_t^\varepsilon = B_t + \varepsilon \begin{bmatrix} \cos(t/\varepsilon^2) \\ \sin(t/\varepsilon^2) \end{bmatrix}.$$

The infinitesimal generator of X^ε is then, with the complex notations, $\frac{1}{2}\triangle + \mathrm{i}\varepsilon^{-1} e^{\mathrm{i}t/\varepsilon^2}\nabla\cdot$ and $c = \left[\begin{smallmatrix} 0 & 1/2 \\ -1/2 & 0 \end{smallmatrix}\right]$ (See [LL06]).

3.5.3 Convergence of the brackets and convergence of the iterated integrals in the Itô case

Although the homogenization result does not allow us to give a negative answer to our question (Q) in the case of Stratonovich integrals, this example is interesting because it shows that Condition UTD is not a necessary condition. In addition, it may give a negative answer to (Q) in the case where Itô integrals are used for $K(X^\varepsilon)$.

Lemma 5. *The martingale part M^ε of X^ε converges to $\rho\overline{B}'$, where $\overline{B}'$ is a N-dimensional Brownian motion and ρ is a symmetric $N \times N$-matrix with*

$$\rho\rho^{\mathrm{T}} = \overline{a} = \int_{[0,1]^N} a(x)\,\mathrm{d}x,$$

which is in general different from a^{eff} given by (20)*. In this case, then X^ε does not satisfy condition UTD, and $\langle M^\varepsilon \rangle$ does not converge to $\langle \overline{X} \rangle$.*

Proof. Let us remark that the brackets $\langle M^\varepsilon \rangle$ of M^ε are equal in distribution to $t \mapsto \varepsilon^2 \int_0^{t/\varepsilon^2} a(X_s)\,\mathrm{d}s$. As the projection of X^1 on $[0,1]^N$ is ergodic with respect to the Lebesgue measure and a is 1-periodic, $\langle M^\varepsilon \rangle$ converge almost surely uniformly on $[0,T]$ to $t \mapsto t\overline{a}$. The details of the convergence may be found for example in [Lej01]. The latter quantity differs in general from a^{eff}, and then $(\langle M^\varepsilon \rangle)_{\varepsilon>0}$ does not necessarily converge to the bracket of $\sigma\overline{B}$ which is equal to $t \mapsto ta^{\mathrm{eff}}$.

If $a^{\mathrm{eff}} \neq \overline{a}$, then Remark 4 allows us to conclude. □

Theorem 3. *Let $K(X^\varepsilon)$ (resp. $K(\overline{X})$) denotes the iterated integrals of X^ε (resp. $\overline{X}$) constructed with the Stratonovich integrals. Then for any $x \in \mathbb{R}^N$, $(X^\varepsilon, K(X^\varepsilon))$ converges in distribution under $\mathbb{P}_x$ in α-variation for any $\alpha > 2$ to $(\overline{X}, \widehat{K}(\overline{X}))$ under $\mathbb{P}_x$ for any $x \in \mathbb{R}^d$, where*

$$\widehat{K}_{0,t}(\overline{X}) = K_{0,t}(\overline{X}) + \frac{t}{2}(a^{\text{eff}} - \overline{a}),\ t \geq 0.$$

Proof. As that $K(X^\varepsilon)$ converges to $K(\overline{X})$ in the Stratonovich case and $\langle M^\varepsilon \rangle_t$ converges in probability to $\overline{a}t$ for any $t \geq 0$, this proves with (14) that in the Itô case, $K_{0,t}(X^\varepsilon)$ converges to $\widehat{K}_{0,t}(\overline{X})$ for any $t \geq 0$. The convergence of $\langle M^\varepsilon \rangle$ to $t \mapsto t\overline{a}$ also holds in α-variation in $\mathbb{R}^N \times \mathbb{R}^N$ for any $\alpha > 1$, since the finite variation of $(\langle M^\varepsilon \rangle)_{\varepsilon > 0}$ is uniformly bounded. □

4 Continuity of integrals along rough paths

We give now a general result of convergence that allows us to ensure that the rough path integral $\int f_\delta(X_s)\,\mathrm{d}\mathbf{X}_s$ converges to $\int f(X_s)\,\mathrm{d}\mathbf{X}_s$ when f_δ converges to f in an appropriate sense. This result will be used in the next section to identify the integrals constructed in [Lej06a] using the rough paths theory with the one constructed using time-reversal techniques presented in Section 2.

We denote by $\mathbf{X} = (\mathbf{X}^1, \mathbf{X}^2)$ a rough path of finite α-variation with $\alpha \in [2, 3)$. We assume that $\mathbf{X}$ lies above X with $X_0 = x$ for a fixed x, which means that $\mathbf{X}^1_{s,t} = X_t - X_s$.

For $\gamma \in (0, 1)$ such that $2 + \gamma > \alpha$, we denote by $\mathrm{Lip}(\gamma)$ the set of continuous, bounded functions $g : \mathbb{R}^N \to \mathbb{R}^m$ with a bounded first derivative which is γ-Hölder continuous. On this space, we define the norm

$$\|g\|_{\mathrm{Lip}} = \|g\|_\infty + \sum_{i=1}^N \|\partial_{x_i} g\|_\infty + \sum_{i=1}^N \sup_{x,y \in \mathbb{R}^N,\ x \neq y} \frac{|\partial_{x_i} g(x) - \partial_{x_i} g(y)|}{|x - y|^\gamma}.$$

The theory of rough path allows us to construct a new rough path $\mathbf{Z}(f) = (\mathbf{Z}^1(f), \mathbf{Z}^2(f))$ from $(f_1, \ldots, f_N) \in \mathrm{Lip}(\gamma)^N$ and $\mathbf{X}$ that corresponds to the integral of f — identified as a differential form — along $\mathbf{X}$. We set

$$\mathbf{Z}_{0,t}(f) = \int_0^t f(X_s)\,\mathrm{d}\mathbf{X}_s.$$

For such a function $f = (f_1, \ldots, f_N)$, we set $\|f\|_{\mathrm{Lip}} = \max_{i=1,\ldots,N} \|f_i\|_{\mathrm{Lip}}$.

Proposition 4. *Let $f \in \mathrm{Lip}(\gamma)^N$ and $(f_\delta)_{\delta>0}$ a sequence of functions in $\mathrm{Lip}(\gamma)^N$ such that $\|f_\delta\|_{\mathrm{Lip}}$ remains bounded in δ, f_δ converges uniformly to f and ∇f_δ converges uniformly to ∇f as $\delta \to 0$.*

Then $\mathbf{Z}(f_\delta)$ converges to $\mathbf{Z}(f)$ in α-variation.

Proof. For $(s,t) \in \Delta_+$, set

$$\mathbf{Y}^1_{s,t} = \sum_{i=1}^N f(X_s)\mathbf{X}^{i,1}_{s,t} + \sum_{i,j=1}^N \frac{\partial f_i}{\partial x_j}(X_s)\mathbf{X}^{i,j,2}_{s,t}$$

and

$$\mathbf{Y}^2_{s,t}(f) = \sum_{i,j=1}^N f_i(X_s)f_j(X_s)\mathbf{X}^{i,j,2}_{s,t}.$$

We also assume that $\mathbf{X}$ is controlled by some function $\omega : \Delta_+ \to \mathbb{R}$, which means that

$$|\mathbf{X}_{s,t}| \leq \omega(s,t)^{1/\alpha} \text{ with } \omega(s,r) + \omega(r,t) \leq \omega(s,t)$$

for all $0 \leq s \leq r \leq t \leq T$ and ω is continuous near its diagonal. Then, $\mathbf{Y}(f)$ is an *almost rough path*, in the sense that

$$|\mathbf{Y}_{s,t}(f) - \mathbf{Y}_{s,r}(f) \otimes \mathbf{Y}_{r,t}(f)| \leq C\omega(s,t)^\theta, \ \forall 0 \leq s \leq r \leq t \leq T,$$

for $\theta = (\gamma+1)/\alpha > 1$ and a constant C that depends only on $\omega(0,T)$, α, γ and $\|f\|_{\mathrm{Lip}}$.

Given a family of partitions $(\Pi^n)_{n\in\mathbb{N}}$ whose meshes decreases to 0, we set

$$\mathbf{Y}^{\Pi^n}_{s,t}(f) = \mathbf{Y}_{s,t^n_{\ell(s)}}(f) \otimes \mathbf{Y}_{t^n_{\ell(s)},t_{\ell(s)^n+1}}(f) \otimes \cdots \otimes \mathbf{Y}_{t^n_{\ell(t)}-1,t_{\ell(t)^n}}(f) \otimes \mathbf{Y}_{t^n_{\ell(t)},t}(f)$$

where $\Pi^n = \{0 = t^n_0 \leq \cdots \leq t^n_{\ell^n} = T\}$ and $\ell(s)^n$ and $\ell(t)^n$ are such that $t^n_{\ell(s)-1} < s \leq t^n_{\ell(s)}$ and $t^n_{\ell(t)} \leq t < t^n_{\ell(t)}$. The rough path $\mathbf{Z}(f)$ is constructed from $\mathbf{Y}(f)$ as the limit of $\mathbf{Y}^{\Pi^n}(f)$. Indeed, by construction (see Theorem 3.2.1 in [LQ02, p. 41]), one gets that there exists a sequence $(K_n)_{n\in\mathbb{N}}$ decreasing to 0 such that

$$|\mathbf{Z}_{s,t}(f) - \mathbf{Y}^{\Pi^n}_{s,t}(f)| \leq K_n\omega(s,t)^\theta, \ \forall (s,t) \in \Delta_+. \tag{22}$$

These constants K_n depend on C and n, and thus on $\|f\|_{\mathrm{Lip}}$, $\omega(0,T)$, α and γ.

For each integer n, $\mathbf{Y}^{\Pi^n}_{s,t}(f_\delta)$ converges to $\mathbf{Y}^{\Pi^n}_{s,t}(f)$ as $\delta \to 0$. In addition, for each $\delta > 0$, $\mathbf{Y}^{\Pi^n}_{s,t}(f_\delta)$ converges to $\mathbf{Z}(f_\delta)$ at a speed that does not depend

on δ, since $(\|f_\delta\|_{\mathrm{Lip}})_{\delta>0}$ is bounded. This is sufficient to ensure that $\mathbf{Z}_{s,t}(f_\delta)$ converges to $\mathbf{Z}_{s,t}(f)$ for each $(s,t) \in \Delta_+$.

Moreover, it is easily checked that

$$|\mathbf{Y}^{\Pi^n}_{s,t}(f_\delta) - \mathbf{Y}^{\Pi^n}_{s,t}(f)| \leq C \max\{\sup_{\delta>0} \|f_\delta - f\|_\infty, \sup_{\delta>0} \|\nabla f_\delta - f\|_\infty\} \omega(s,t)^{1/\alpha}$$

for some constant C that depends only on $\sup_{\delta>0} \|f_\delta\|_{\mathrm{Lip}}$ and $\|f\|_{\mathrm{Lip}}$. With (22), one obtains that for any $\varepsilon > 0$, there exists δ_0 small enough such that for all $\delta < \delta_0$, $|\mathbf{Z}_{s,t}(f_\delta) - \mathbf{Z}_{s,t}(f)| \leq \varepsilon \omega(s,t)^{1/\alpha}$ for all $(s,t) \in \Delta_+$.

The result is then proved. □

5 Identification of the integrals

Let us note that a function in $\mathrm{Lip}(\gamma)$ also belongs to Υ^∞. Given $g = (g_1, \ldots, g_N) \in \mathrm{Lip}(\gamma)^N$, we are given two ways in considering the integrals of type $\sum_{i=1}^N \int_0^t g_i(X_s) \circ \mathrm{d}X^i_s$ and $\sum_{i=1}^N \int_0^t g_i(X_s)\, \mathrm{d}X^i_s$ for $X \in \Xi(\lambda, \Lambda)$: either by using the Lyons-Zheng decomposition (5) or by using the rough paths $(X, K(X))$, where $K(X)$ is either the Stratonovich or the Itô integral, as constructed in [Lej06a].

We have seen in [CL03] that the integrals given by the proper choice of $K(X)$ are the same as the Itô or Stratonovich integral when X is a semi-martingale. The identification relies on the Wong-Zakai theorem for semi-martingales.

In our case, we will approximate a process X by semi-martingales obtained by smoothing the coefficients of its infinitesimal generator, and then pass to the limit.

We denote by $\mathfrak{K}_{\mathrm{s}}(t; X, g)$ the integral $\mathbf{Z}^1_{0,t}(g)$ with $\mathbf{Z} = \int g(X_s)\, \mathrm{d}\mathbf{X}_s$, where $\mathbf{X}$ is the rough path $\mathbf{X} = (X, K(X))$ with $K(X)$ defined as a Stratonovich integral. Similarly, if we use an Itô integral for $K(X)$, then we denote $\mathbf{Z}^1_{0,t}(g)$ by $\mathfrak{K}_{\mathrm{i}}(t; X, g)$.

We also set $\mathfrak{L}_{\mathrm{s}}(t; X, g) = \sum_{i=1}^N \mathfrak{L}(t; X, g_i, \chi_i)$, when $\mathfrak{L}$ is defined in the Stratonovich, and $\mathfrak{L}_{\mathrm{i}}(t; X, g) = \sum_{i=1}^N \mathfrak{L}(t; X, g_i, \chi_i)$ when $\mathfrak{L}$ is defined in the Itô sense.

Remark 9. Here, we consider only the "first level", and not the iterated integrals of $\int g(X)\, \mathrm{d}X$ against itself since it is useless for our result.

Theorem 4. *For any $X \in \Xi(\lambda, \Lambda)$ and $g \in \mathrm{Lip}(\gamma)^N$ with $\gamma \in (0,1)$. Then $\mathfrak{K}_{\mathrm{s}}(X, g) = \mathfrak{L}_{\mathrm{s}}(X, g)$ and $\mathfrak{K}_{\mathrm{i}}(X, g) = \mathfrak{L}_{\mathrm{i}}(X, g)$ under $\mathbb{P}_x$ for any starting point $x \in \mathbb{R}^N$.*

We already know that this result is true for semi-martingales.

Proposition 5 ([CL03]). *Let $(X, \mathbb{P})$ be a continuous semi-martingale and $g \in \mathrm{Lip}(\gamma)^N$ with $\gamma \in (0,1)$. Then $\mathbb{P}$-almost surely, $\mathfrak{K}_{\mathrm{s}}(X,g) = \mathfrak{L}_{\mathrm{s}}(X,g)$ and $\mathfrak{K}_{\mathrm{i}}(X,g) = \mathfrak{L}_{\mathrm{i}}(X,g)$.*

Proof of Theorem 4. We drop any references to the subscripts s and i in $\mathfrak{L}$ and $\mathfrak{K}$, since the proof is the same in both cases.

Let $(a^\varepsilon, b^\varepsilon)$ a sequence of elements in $\Xi^{\mathrm{coeff}}(\lambda, \Lambda)$ such that a^ε and b^ε are smooth and converge almost everywhere to a and b. The corresponding process X^ε is then a semi-martingale, since $L^\varepsilon = \frac{1}{2}\nabla(a^\varepsilon \nabla \cdot) + b^\varepsilon \nabla$ may be transformed into a non-divergence form operator.

In addition, let us assume in a first time that g is smooth. From the proof of Theorem 2.2 in [Roz96a] (see also the proof of Corollary 1 on the localization argument), $(g(X^\varepsilon))_{\varepsilon>0}$ satisfies Conditions UTD. From Lemma 3, one knows that for $i = 1, \dots, N$, $\mathfrak{L}^n(t; X^\varepsilon, g, \chi_i)$ converges to $\mathfrak{L}(t; X^\varepsilon, g, \chi_i)$ in probability at a speed that is uniform in ε. As in the proof of Lemma 4, it is then easily deduced from Theorem 4.2 in [Bil68] that $\mathfrak{L}(t; X^\varepsilon, g)$ converges $\mathfrak{L}(t; X, g)$ in distribution. Yet we have seen in Corollary 1 that $(X^\varepsilon, K(X^\varepsilon))$ converges in $\mathcal{V}^\alpha$ to $(X, K(X))$. From the continuity of $\mathfrak{K}$ in $\mathcal{V}^\alpha$, we deduce that $\mathfrak{K}(X^\varepsilon, g)$ converges in distribution to $\mathfrak{K}(X, g)$. Necessarily, $\mathfrak{L}(X, g) = \mathfrak{K}(X, g)$ under the assumption that g is smooth.

Now, if g only belongs to $\mathrm{Lip}(\gamma)^N$, let us introduce a sequence of mollifiers $(\varphi_\delta)_{\delta>0}$. Then the convolution $g \star \varphi_\delta$ of g with φ_δ is a smooth functions satisfies $\|g \star \varphi_\delta\|_{\mathrm{Lip}} \leq \|g\|_{\mathrm{Lip}}$. In addition, $g \star \varphi_\delta$ and $\nabla g \star \varphi_\delta$ converge uniformly to g and ∇g.

The conclusion follows from Propositions 1 and 4. □

A On the convergence of the sum of squares of a martingale to its bracket

The following Lemma is used in the proof of Lemma 3. It corresponds to a classical result, but we write it explicitly to show the dependence of the rate of convergence with respect to the Lipschitz constant of the brackets.

Lemma 6. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_t)_{t\geq 0}$. Let M be a square integrable martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ such that for some constant K, $\langle M\rangle_t - \langle M\rangle_s \leq K(t-s)$ for any $0 \leq s \leq t \leq T$ for a fixed $T > 0$. For each $n \geq 1$, let $\{t_i^n\}_{i=1,\dots,n}$ be a partition of $[0,T]$ such that $\sup_{i=1,\dots,n-1}(t_{i+1}^n - t_i^n)$ decreases to 0 as $n \to \infty$. Let*

$$Q_t^n = \sum_{i=1}^{\ell^n(t)-1} (M_{t_{i+1}^n} - M_{t_i^n})^2,$$

where $\ell(t)$ is the integer such that $t^n_{\ell(t)} \leq t < t^n_{\ell(t)+1}$.

Then, for any $\kappa > 0$ and any $\varepsilon > 0$, there exists an integer n_0 such that

$$\mathbb{P}\left[\sup_{t\in[0,T]} |Q^n_t - \langle M\rangle_t| \geq \kappa\right] \leq \varepsilon \text{ for all } n \geq n_0$$

and the choice of n_0 depends only on K, κ, T and the mesh $\sup_{i=1,\dots,n-1}(t^n_{i+1} - t^n_i)$ of $\{t^n_i\}_{i=1,\dots,n}$.

Proof. The convergence in probability of Q^n_t to $\langle M\rangle_t$ for any $t \geq 0$ is classical (See for example [KS91, Theorem 1.5.8, p. 33]). However, we write it to show that the rate of convergence depends only on K. Let N be the martingale defined by $N = M^2 - \langle M\rangle$ (note that with our condition on the growth of $\langle M\rangle$, N is really a martingale and not a local martingale).

Let us set $\Delta^n_i M = M_{t^n_{i+1}} - M_{t^n_i}$, $\Delta^n_i N = N_{t^n_{i+1}} - N_{t^n_i}$, $\Delta^n_i\langle M\rangle = \langle M\rangle_{t^n_{i+1}} - \langle M\rangle_{t^n_i}$.

Then

$$\begin{aligned}\left(\sum_{i=1}^{\ell^n(t)-1} \Delta^n_i M - \langle M\rangle_t\right)^2 &\leq 2\left(\sum_{i=1}^{\ell^n(t)-1} \Delta^n_i M - \langle M\rangle_{t^n_\ell}\right)^2 + 2(\langle M\rangle_t - \langle M\rangle_{t^n_{\ell^n(t)}})^2 \\ &= 2\left(\sum_{i=1}^{\ell^n(t)-1} (\Delta^n_i M - \Delta^n_i\langle M\rangle)\right)^2 + 2(\langle M\rangle_t - \langle M\rangle_{t^n_{\ell^n(t)}})^2 \\ &= 2\sum_{i=1}^{\ell^n(t)-1} \Delta^n_i N^2 + 2\sum_{\substack{i,j=1,\dots,\ell^n(t)-1\\ i\neq j}} \Delta^n_j N \Delta^n_i N + 2(\langle M\rangle_t - \langle M\rangle_{t^n_{\ell^n(t)}})^2.\end{aligned}$$

If $j < i$, since N is a martingale,

$$\mathbb{E}\left[\,\Delta^n_j N \Delta^n_i N\,\right] = \mathbb{E}\left[\,\Delta^n_j N \mathbb{E}\left[\,\Delta^n_i N \,|\, \mathcal{F}_{t^N_i}\,\right]\right] = 0.$$

Thus,

$$\mathbb{E}\left[\left(\sum_{i=1}^{\ell^n(t)-1} \Delta^n_i M - \langle M\rangle_t\right)^2\right] \leq 2\sum_{i=1}^{\ell^n(t)-1} \mathbb{E}\left[\,\Delta^n_i N^2\,\right] + 2\mathbb{E}\left[\,(\langle M\rangle_t - \langle M\rangle_{t^n_{\ell^n(t)}})^2\,\right].$$

With the Burkholder-Davis-Gundy inequality, there exists a constant C' such that for $i = 1, \dots, \ell^n(t) - 1$,

$$\begin{aligned}\mathbb{E}\left[\,\Delta^n_i N^2\,\right] &= \mathbb{E}\left[\,\Delta^n_i M^4\,\right] - 2\mathbb{E}\left[\,\Delta^n_i M^2 \Delta^n_i\langle M\rangle^2\,\right] + \mathbb{E}\left[\,\Delta^n_i\langle M\rangle^2\,\right] \\ &\leq C'\mathbb{E}\left[\,\Delta^n_i\langle M\rangle^2\,\right] \leq KC'(t^n_{i+1} - t^n_i)^2,\end{aligned}$$

so that

$$\mathbb{E}\left[\left(\sum_{i=1}^{\ell^n(t)-1}\Delta_i^n M-\langle M\rangle_t\right)^2\right]\leq 2K(C'T+1)\sup_{i=1,\dots,n-1}(t_{i+1}^n-t_i^n).$$

This proves the $\mathrm{L}^2(\mathbb{P})$-convergence of the one-dimensional marginals of Q^n to $\langle M\rangle$.

To now turn to prove the uniform convergence of Q^n to $\langle M\rangle$. For this, let us choose two integers $1\leq k<\ell\leq n$ and let us note that

$$\left(\sum_{i=k}^{\ell-1}\Delta_i^n M^2\right)^2=\sum_{i=k}^{\ell-1}\Delta_i M^4+2\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}}\Delta_i^n M^2\Delta_j^n M^2.$$

With the Burkholder-Davis-Gundy inequality, there exists some constant C' such that

$$\mathbb{E}\left[\sum_{i=k}^{\ell-1}\Delta_i^n M^4\right]\leq C'\mathbb{E}\left[\sum_{i=k}^{\ell-1}\Delta_i^n\langle M\rangle^2\right]\leq C'K(t_\ell-t_k)\sup_{i=k,\dots,\ell-1}(t_{i+1}^n-t_i^n). \tag{23}$$

On the other hand,

$$\begin{aligned}\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}}\Delta_i^n M^2\Delta_j^n M^2=&\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}}\Delta_i^n N\Delta_j^n\langle M\rangle+\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}}\Delta_i^n\langle M\rangle\Delta_j^n N\\ &+\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}}\Delta_i^n N\Delta_j^n N+\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}}\Delta_i^n\langle M\rangle\Delta_j^n\langle M\rangle.\end{aligned}$$

As N is martingale, one gets that

$$\mathbb{E}\left[\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}}a_i\Delta_j^n N\right]=\mathbb{E}\left[\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}}a_i\mathbb{E}\left[\,\Delta_j^n N\,|\,\mathcal{F}_{r_j}\,\right]\right]=0$$

for $a_i=\Delta_i^n N$ or $a_i=\Delta_i^n\langle M\rangle$. In addition, since $\Delta_i^n N^2=(\Delta_i^n M^2-$

$\Delta_i^n\langle M\rangle)^2 \le \Delta_i^n M^4 + \Delta_i^n\langle M\rangle^2$,

$$\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}} \Delta_i^n N \Delta_j^n \langle M\rangle = \sum_{i=k}^{\ell-1} \Delta_i^n N(\langle M\rangle_{t_\ell^n} - \langle M\rangle_{t_{i+1}^n})$$

$$\le (\langle M\rangle_{t_\ell^n} - \langle M\rangle_{t_k^n})^2 \sum_{i=k}^{\ell-1} \Delta_i^n N^2$$

$$\le (\langle M\rangle_{t_\ell^n} - \langle M\rangle_{t_k^n})^2 \left(\sum_{i=k}^{\ell-1} \Delta_i^n M^4 + \sum_{i=k}^{\ell-1} \Delta_i^n \langle M\rangle^2 \right).$$

With (23), $\mathbb{E}\left[\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}} \Delta_i^n N \Delta_j^n \langle M\rangle \right] \le C(t_\ell^n - t_k^n)^4$ for some constant C that depends only on T and K. On the other hand, still using the fact that $\langle M\rangle$ is increasing, $\sum_{\substack{i,j=k,\dots,\ell-1\\ i\neq j}} \Delta_i^n\langle M\rangle \Delta_j^n\langle M\rangle \le (\langle M\rangle_{t_\ell^n} - \langle M\rangle_{t_k^n})^2 \le K^2(t_\ell^n - t_k^n)^2$.

Hence, we finally get that

$$\mathbb{E}\left[\left(\sum_{i=k}^{\ell-1} \Delta_i^n M^2 \right)^2 \right] \le C(t_\ell^n - t_k^n)^2 \tag{24}$$

for some constant C that depends only on K and T.

Now, for a partition $\{s_i^m\}_{i=1,\dots,m}$ of $[0,T]$,

$$\sup_{t\in[0,T]} |Q_t^n - \langle M\rangle_t|$$
$$\le \sup_{i=1,\dots,m-1} \sup_{t\in[t_i^m, t_{i+1}^m]} \left(|Q_t^n - Q_{s_i^m}^n| + |Q_{s_i^m}^n - \langle M\rangle_{s_i^m}| + |\langle M\rangle_t - \langle M\rangle_{s_i^m}| \right). \tag{25}$$

Let us note that, since $t \mapsto Q_t^n$ is increasing, for $t \in [s_i^m, s_{i+1}^m]$

$$\sup_{t\in[s_i^m, s_{i+1}^m]} |Q_t^n - Q_{s_i^m}^n| \le \sum_{i=\ell^-(i,n)}^{\ell^+(i,n)-1} (M_{t_{i+1}^n} - M_{t_i^n})^2$$
$$- (M_{t^n_{\ell^-(i,n)-1}} - M_{t^n_{\ell^-(i,n)}})^2, \tag{26}$$

where $\ell^-(i,n)$ is such that $t^n_{\ell^-(i,n)-1} < s_i^m \le t^n_{\ell^-(i,n)}$ and $t^n_{\ell^+(i,n)} < s_{i+1}^m \le t^n_{\ell^+(i,n)+1}$.

With (25) and (26),

$$\mathbb{P}\left[\sup_{t\in[0,T]}|Q_t^n-\langle M\rangle_t|>4\kappa\right]\le\mathbb{P}\left[\sup_{i=1,\dots,m}\sup_{t\in[s_i^m,s_{i+1}^m]}|\langle M\rangle_t-\langle M\rangle_{s_i^m}|>\kappa\right]$$
$$+\mathbb{P}\left[\sup_{i=1,\dots,m}\sum_{i=\ell^-(i,n)}^{\ell^+(i,n)-1}(M_{t_{i+1}^n}-M_{t_i^n})^2>\kappa\right]$$
$$+\mathbb{P}\left[\sup_{i=1,\dots,m-1}|Q_{s_i^m}^n-\langle M\rangle_{s_i^m}|>\kappa\right]$$
$$+\mathbb{P}\left[\sup_{i=1,\dots,m-1}(M_{t^n_{\ell^-(i,n)-1}}-M_{t^n_{\ell^-(i,n)}})^2>\kappa\right]. \quad (27)$$

For any $\kappa>0$, it follows from (24) that

$$\mathbb{P}\left[\sup_{i=1,\dots,m-1}\sum_{j=\ell^-(i,n)}^{\ell^+(i,n)-1}\Delta_iM^2\ge\kappa\right]\le\frac{1}{\kappa^2}\sum_{i=1}^m\mathbb{E}\left[\left(\sum_{j=\ell^-(i,n)}^{\ell^+(i,n)-1}\Delta_iM^2\right)^2\right]$$
$$\le\frac{C}{\kappa^2}\sum_{i=1}^{m-1}(s_{i+1}^m-s_i^m)^2\le\frac{CT}{\kappa^2}\delta \quad (28)$$

with $\delta=\sup_{i=1,\dots,m-1}(s_{i+1}^m-s_i^m)$. Moreover,

$$\mathbb{P}\left[\sup_i\sup_{t\in[s_i^m,s_{i+1}^m]}|\langle M\rangle_t-\langle M\rangle_{s_i^m}|>\kappa\right]\le\mathbb{P}\left[\,\mathrm{osc}(\langle M\rangle,\delta)>\kappa\,\right]\le\frac{K\delta}{\kappa}, \quad (29)$$

where $\mathrm{osc}(f,\delta)=\sup_{s,t\in[0,T],\ |t-s|<\delta}|f(t)-f(s)|$ is the modulus of continuity of a continuous function f. With the Burkholder-Davis-Gundy, there exists a constant C' such that

$$\mathbb{P}\left[\sup_{i=1,\dots,m-1}(M_{t^n_{\ell^-(i,n)-1}}-M_{t^n_{\ell^-(i,n)}})^2>\kappa\right]$$
$$\le\sum_{i=1}^{m-1}\frac{1}{\kappa^2}\mathbb{E}\left[(M_{t^n_{\ell^-(i,n)-1}}-M_{t^n_{\ell^-(i,n)}})^2\right]\le\frac{mC'K}{\kappa^2}\sup_{j=1,\dots,n-1}(t_{j+1}^n-t_j^n). \quad (30)$$

Finally,

$$\mathbb{P}\left[\sup_{i=1,\dots,m-1}|Q_{s_i^m}^n-\langle M\rangle_{s_i^m}|>\kappa\right]\le\sum_{i=1}^{m-1}\mathbb{P}\left[|Q_{s_i^m}^n-\langle M\rangle_{s_i^m}|>\kappa\right]. \quad (31)$$

Given $\kappa, \varepsilon > 0$, the idea is now to fix first the mesh δ of the partition $\{s_i^m\}_{i=1,\dots,m}$, — which is arbitrary chosen — so that the quantities in (28) and (29) are smaller than $\varepsilon/4$. This fix m. We choose then n_0 large enough such that for any $n \geq n_0$, the quantities in (30) and (31) are also smaller than $\varepsilon/4$. We then obtain that with (27)

$$\mathbb{P}\left[\sup_{t\in[0,T]} |Q_t^n - \langle M\rangle_t| > 4\kappa\right] \leq \varepsilon \text{ for all } n \geq n_0$$

and we remark that the choice of δ depends only on K, T and κ, as well as the choice of n_0, as we claimed. □

Acknowledgments. This work was started while the author was a post-doctoral student in Oxford and granted by the TMR *Stochastic Analysis Network*. Thus, the author wishes to thank Prof. T.J. Lyons for his kind hospitality and for having taught him the theory of rough paths.

The author wishes also to thank the anonymous referee whose remarks have greatly improved the quality of this article.

References

[A75] R. Adams. *Sobolev spaces*. Academic Press, 1975.

[Bau04] F. Baudoin. *An introduction to the geometry of stochastic flows*. Imperial College Press, London, 2004.

[BPL78] A. Bensoussan, J.-L. Lions and G. Papanicolaou. *Asymptotic Analysis for Periodic Structures*. North-Holland, 1978.

[Bil68] P. Billingsley. *Convergence of Probability Measures*. Wiley, 1968.

[BBJR95] C.J.K. Batty, O. Bratteli P.E.T. Jørgensen, and D.W. Robinson. Asymptotics of periodic subelliptic operators. *J. Geom. Anal.*, **5**:4, 427–443, 1995.

[CD+07] L. Capogna, D. Danielli, S.D. Pauls and J.T. Tyson, *An Introduction to the Heisenberg Group and the Sub-Riemannian Isoperimetric Problem*, Progress in Mathematics, Vol. 259 Birkhäuser, 2007.

[CFV05] L. Coutin, P. Friz and N. Victoir. Good Rough Path Sequences and Applications to Anticipating & Fractional Stochastic Calculus, 2006. <arXiv:math.PR/0501197>.

[CL03] L. Coutin and A. Lejay. Semi-martingales and rough paths theory. *Electron. J. Probab.*, **10**:23, 761–785, 2005.

[CS99] F. Coquet and L. Słomiński. On the convergence of Dirichlet processes. *Bernoulli*, **5**:4, 615–639, 1999.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov Processes, Characterization and Convergence*. Wiley, 1986.

[Föl81b] H. Föllmer. Dirichlet processes. In *Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980)*, vol. 851 of *Lecture Notes in Math.*, 476–478. Springer, Berlin, 1981.

[FOT04] M. Fukushima, Y. Oshima and M. Takeda. *Dirichlet Forms and Symmetric Markov Process*, De Gruyter, 1994.

[FV06a] P. Friz and N. Victoir. A note on the notion of geometric rough path. *Probab. Theory Related Fields*, **136**:3, 395–416, 2006.

[FV06b] P. Friz and N. Victoir. On Uniformly Subelliptic Operators and Stochastic Area. Preprint Cambridge University, 2006. <arXiv:math.PR/0609007>.

[JKO94] V.V. Jikov, S.M. Kozlov and O.A. Oleinik. *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, 1994.

[KP96] T.G. Kurtz and P. Protter. Weak Convergence of Stochastic Integrals and Differential Equations. In *Probabilistic Models for Nonlinear Partial Differential Equations, Montecatini Terme, 1995*, edited by D. Talay and L. Tubaro, vol. 1629 of *Lecture Notes in Math.*, 1–41. Springer-Verlag, 1996.

[KS91] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, 2nd edition, 1991.

[Lej00] A. Lejay. *Méthodes probabilistes pour l'homogénéisation des opérateurs sous forme-divergence : cas linéaires et semi-linéaires*. PhD thesis, Université de Provence, Marseille, France, 2000. <URL: www.iecn.u-nancy.fr/~lejay/>.

[Lej01] A. Lejay. A Probabilistic Approach of the Homogenization of Divergence-Form Operators in Periodic Media. *Asymptot. Anal.*, **28**:2, 151–162, 2001.

[Lej02] A. Lejay. On the convergence of stochastic integrals driven by processes converging on account of a homogenization property. *Electron. J. Probab.*, **7**:18, 1–18, 2002.

[Lej03a] A. Lejay. An introduction to rough paths. In *Séminaire de probabilités, XXXVII*, vol. 1832 of *Lecture Notes in Math.*, 1–59, Springer, Berlin, 2003.

[Lej06a] A. Lejay. Stochastic Differential Equations driven by processes generated by divergence form operators I: a Wong-Zakai theorem. *ESAIM Probab. Stat.*, **10**, 356–379, 2006.

[Lej06b] A. Lejay. Yet another introduction to rough paths. Preprint Institut Élie Cartan, Nancy, 2006. <http://hal.inria.fr/inria-00107460>.

[LL06] A. Lejay and T.J. Lyons. On the Importance of the Lévy Area for Systems Controlled by Converging Stochastic Processes. Application to Homogenization. In *New Trend in Potential Theory*, D. Bakry, L. Beznea, Gh. Bucur and M. Röckner (eds), The Theta Foundation, 2006.

[LV06] A. Lejay and N. Victoir. On (p, q)-rough paths. *J. Differential Equations*, **225**:1, 103–133, 2006.

[LQ02] T. Lyons and Z. Qian. *System Control and Rough Paths*. Oxford Mathematical Monographs. Oxford University Press, 2002.

[LS99] T.J. Lyons and L. Stoica. The limits of stochastic integrals of differential forms. *Ann. Probab.*, **27**:1, 1–49, 1999.

[Lyo98] T.J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, **14**:2, 215–310, 1998.

[Mar79] P. Marcellini. Convergence of Second Order Linear Elliptic Operator. *Boll. Un. Mat. Ital. B (5)*, **16**:1, 278–290, 1979.

[Mon02] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications*. Mathematical Surveys and Monographs, 91. American Mathematical Society, Providence, RI, 2002.

[Roz96a] A. Rozkosz. Stochastic Representation of Diffusions Corresponding to Divergence Form Operators. *Stochastic Process. Appl.*, **63**:1, 11–33, 1996.

[Roz96b] A. Rozkosz. Weak Convergence of Diffusions Corresponding to Divergence Form Operator. *Stochastics Stochastics Rep.*, **57**, 129–157, 1996.

[RS98] A. Rozkosz and L. Słomiński. Extended Convergence of Dirichlet Processes. *Stochastics Stochastics Rep.*, **65**:1–2, 79–109, 1998.

[Str88] D.W. Stroock. Diffusion Semigroups Corresponding to Uniformly Elliptic Divergence Form Operator. In *Séminaire de Probabilités XXII*, vol. 1321 of *Lecture Notes in Math.*, 316–347. Springer-Verlag, 1988.