On the Importance of the Levy Area for Studying the Limits of Functions of Converging Stochastic Processes. Application to Homogenization

Antoine Lejay, Terry Lyons

On the Importance of the Lévy Area for Studying the Limits of Functions of Converging Stochastic Processes. Application to Homogenization

Antoine Lejay[1,†] — Projet OMEGA, INRIA/Institut Élie Cartan de Nancy

Terry Lyons[2,†] — Mathematical Institute, Oxford University

Abstract: Two concrete examples show us that the convergence of a family of stochastic processes "as controls", *i.e.*, as integrators of SDEs or differential forms, may require more information than simply the limit in the uniform norm of the processes. This may be particularly important when one deals with the homogenization theory. The theory of rough paths is then used to bring some new results about interchanging limits and functionals of stochastic processes.

Keywords: rough paths, homogenization, approximation of SDEs by ODEs, Lévy area, convergence of stochastic processes, p-variation

AMS Classification: 60J60 (secondary) 35B27, 60F17

Published in *Current Trends in Potential Theory Conference Proceedings, Bucharest, September 2002 and 2003.* D. Bakry, L. Beznea, Gh. Bucur, M. Röckner (eds). The Theta Foundation, 2005

Archives, links & reviews:

- ISBN: 9738543266

[1]Current address: Projet OMEGA (INRIA/IECN)
IECN
BP 239
54506 Vandœuvre-lès-Nancy Cedex (France)
E-mail: `Antoine.Lejay@iecn.u-nancy.fr`

[2]Current address: Mathematical Institute
Oxford University
24–29 St Giles'
Oxford OX1 3LB (United Kingdom)
E-mail: `tlyons@maths.ox.ac.uk`

[†]Supported by European Union's TMR *Stochastic Analysis* Network (Project ERBF MRX CT 960075A, grant reference 960075)

 29 pages

1 Introduction

This article illustrates the notion of convergence of stochastic processes "as controls". Suppose that $(X^\varepsilon)_{\varepsilon>0}$ is a family of stochastic processes. Let $\mathfrak{K}(X^\varepsilon)$ and $\mathfrak{I}(X^\varepsilon)$ be the stochastic processes obtained when one integrates a differential form along the trajectories of X^ε or considers the solution to some SDE driven by X^ε. We investigate conditions ensuring the functionals and the limit of X^ε may be interchanged.

For that, we study two examples. One comes from the homogenization theory and provides a coherent interpretation of some of the results presented in [Lej02]. The second example is a problem of interpolation. In appearance, these examples are quite different although both are of practical interest. However, these results are rather similar, as will be proved using the rough paths theory developed in [Lyo98] (See also [LQ02, Lej03]). In fact, a similar phenomenon appears in these two examples and this provides a strong intuitive support to understand what happens.

The problem of homogenization concerns the large scale behavior of functionals of stochastic processes X^ε processes whose infinitesimal generators have coefficients that oscillate more and more rapidly as ε goes to 0. For example, if X^ε has as infinitesimal generator

$$\frac{1}{2}e^{2V(\cdot/\varepsilon)}\frac{\partial}{\partial x_i}\left(a_{i,j}(\cdot/\varepsilon)e^{-2V(\cdot/\varepsilon)}\frac{\partial}{\partial x_i}\right),$$

it is standard that X^ε converges in distribution to $\overline{\sigma}B$, where B is a Brownian motion, and $\overline{\sigma}$ is a constant matrix characterizing the large-scale, or effective, behavior of the media (See [BLP78] for example). This result motivates using $\overline{\sigma}B$ instead of more complex diffusion models in applications.

One could consider now that $\mathfrak{K}_t(X^\varepsilon) = \int_0^t f(X^\varepsilon_r)\circ \mathrm{d}X^\varepsilon_r$, and that $\mathfrak{I}(X^\varepsilon)$ gives the solution Y^ε of the SDE $\mathrm{d}Y^\varepsilon_t = g(Y^\varepsilon_t)\circ \mathrm{d}X^\varepsilon_t$, where f and g are smooth enough. A natural question is to know whether $\lim_{\varepsilon\to 0}\mathfrak{K}(X^\varepsilon) = \mathfrak{K}(\lim_{\varepsilon\to 0} X^\varepsilon)$ and $\lim_{\varepsilon\to 0}\mathfrak{I}(X^\varepsilon) = \mathfrak{I}(\lim_{\varepsilon\to 0} X^\varepsilon)$. In other words, could one interchange $\mathfrak{K}$ or $\mathfrak{I}$ with the limit? Or does $\overline{\sigma}B$ provides enough information to know the asymptotic behavior of $\mathfrak{K}(X^\varepsilon)$ and $\mathfrak{I}(X^\varepsilon)$? This is false in general, and one also has to keep track of the various areas enclosed between the trajectories of X^ε and its chords.

The second example concerns the interpolation of trajectories of Brownian motion. This has attracted interest, since it allows effective approximated solutions $\mathfrak{I}(X)$ of stochastic differential equations driven by a Brownian motion X using ordinary differential equations. However, the approach is not completely straightforward. If the trajectories are not interpolated by carefully chosen straight lines, one may obtain as a limit solutions to various SDEs driven by the Brownian motion X, but with a drift reflecting the choice of the interpolation (see for example [McS72, IW89, Kun90, Sus91, KP91] or more recently [CE00]). There are strong parallels between this phenomenon and the form of the limiting SDE in our homogenization result and the use of rough paths theory brings a new light in this fact. Moreover, although this observation appears to be a pitfall in the theory of SDEs, it also provides us with a numerical method to compute a large scale approximation of $\mathfrak{K}(X^\varepsilon)$ or $\mathfrak{I}(X^\varepsilon)$ in the case of homogenization, without having to compute the first derivative of the vector fields or differential form we integrate.

In this article, we also explain what is the link between Stratonovich integrals and rough paths' integration theory for semi-martingales, and we give some convergence and tightness results.

This paper is organized as follows: In Section 2, we recall the results about the construction of interpolation function leading to a process with a corrected area. In Section 3, we identify the limit of the SDEs driven by a family of stochastic processes that converges with the homogenization property. We link the corrective drift with the correction that appears in the limit of the Lévy areas of $(X^\varepsilon)_{\varepsilon>0}$. In Section 4, we put these results in the context of rough paths. In Sections 4.2 and 4.3, we explain how the construction of different geometric rough paths with different areas leads to different solutions of differential equations, which could be identified. Section 4.5, is devoted to state some tightness and convergence results in the topology generated by the norm in p-variation. In Section 4.4 and Section 5, we prove that the convergence of the process and its Lévy area studied in Sections 2 and 3 also holds in this topology.

2 The interpolation problem

In 1965, E. Wong and M. Zakai proved in [WZ65] that the solutions of differential equations driven by piecewise linear approximations of trajectories of a Brownian motion converge uniformly in probability to the solution of a Stratonovich SDE driven by a Brownian motion[1].

One may consider more general piecewise smooth approximations $B^\delta(\omega)$ of the trajectory $B(\omega)$ of a Brownian motion than piecewise linear approximations. This leads to a slightly different result. Here, we recall the main fact, without stating the hypotheses. These results, initially treated by McShane [McS72], are developed in Chapter 5.7 of [Kun90, p. 274] and in Section VI-7,

[1]They did it for 1-dimensional SDEs. The results was quickly and substantially refined by others, including Clark, Stroock, Varadhan, ...

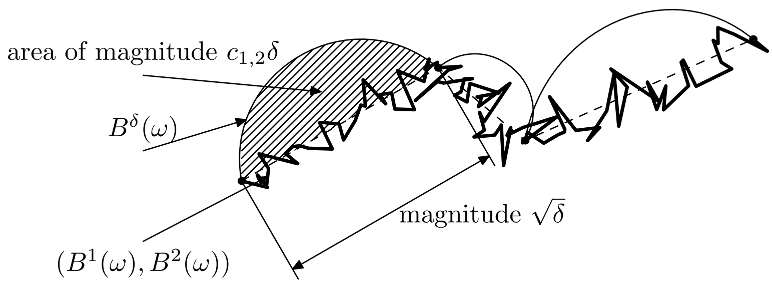

Figure 1: Meaning of $c_{1,2}$ in the interpolation problem

p. 392 of [IW89]. The reader is refereed to these books for details. If

$$\sigma(x) = (\sigma^i_n(x))^{i=1,\dots,d}_{n=1,\dots,N} \in d \times N\text{-matrices and } b(x) = (b^i(x))_{i=1,\dots,d}$$

are smooth functions, then the solution Y^δ of

$$Y^\delta_t = x + \int_0^t \sigma_n(Y^\delta_s)\,\mathrm{d}B^{n,\delta}_s + \int_0^t b(Y^\delta_s)\,\mathrm{d}s,$$

converges uniformly on $[0,1]$ in $\mathrm{L}^1(\mathbb{P})$ to the solution Y of

$$Y_t = x + \int_0^t \sigma_n(Y_s) \circ \mathrm{d}B^n_s + \int_0^t b(Y_s)\,\mathrm{d}s + \frac{1}{2} c_{n,m} \int_0^t [\sigma_n, \sigma_m](Y_s)\,\mathrm{d}s, \qquad (1)$$

where $[\cdot,\cdot]$ is the Lie bracket of two vector fields:

$$[\sigma_n, \sigma_m] = \left(\left(\sigma^j_n \frac{\partial \sigma^i_m}{\partial x_j} - \sigma^j_m \frac{\partial \sigma^i_n}{\partial x_j}\right)\right)_{i=1,\dots,d}$$

and the anti-symmetric matrix $c = (c_{i,j})_{i,j=1,\dots,N}$ is computed by

$$c_{i,j} = \lim_{\delta\to 0} \frac{1}{2\delta} \mathbb{E}\left[\int_0^\delta B^{i,\delta}_s\,\mathrm{d}B^{j,\delta}_s - \int_0^\delta B^{j,\delta}_s\,\mathrm{d}B^{i,\delta}_s\right]. \qquad (2)$$

And for integration of a smooth differential form f from $\mathbb{R}^N$ into $\mathbb{R}^d$, the result is

$$\lim_{\delta\to 0} \mathbb{E}\left[\sup_{t\in[0,T]} \left|\int_0^t f_i(B^\delta_s)\mathrm{d}B^{i,\delta}_s - \int_0^t f_i(B_s) \circ \mathrm{d}B^i_s - \int_0^t c_{i,j} \frac{\partial f_i}{\partial x_j}(B_s)\,\mathrm{d}s\right|^2\right] = 0. \qquad (3)$$

As we see it from (2), the area of the interpolation function between two sample points provides a necessary and sufficient information to compute the limit (1).

3 Homogenization and convergence of solutions of SDEs

3.1 An illustrative example

We start with a simple example of homogenization, where the drift is only time-dependent. This example allows us to understand the role the drift plays.

Convergence of the process and convergence of the Lévy area to the "wrong" value. Let $B = (B^1, B^2)$ be a 2-dimensional Brownian motion, and b, $\widehat{b}$ be the functions, with complex notations where $\mathrm{i}^2 = -1$,

$$b(t) = e^{\mathrm{i}t} \text{ and } \widehat{b}(t) = b'(t) = \mathrm{i}e^{\mathrm{i}t}.$$

We set

$$X_t^\varepsilon = \varepsilon B_{t/\varepsilon^2} + \varepsilon b(t/\varepsilon^2) = \varepsilon B_{t/\varepsilon^2} + \frac{1}{\varepsilon}\int_0^t \widehat{b}(s/\varepsilon^2)\,\mathrm{d}s.$$

As b is bounded and $\varepsilon B_{\cdot/\varepsilon^2}$ remains a Brownian motion in distribution, X^ε converges to a Brownian motion $\overline{B}$ as ε goes to 0. But its infinitesimal generator is $\frac{1}{2}\triangle + \frac{1}{\varepsilon}\widehat{b}(t/\varepsilon^2)\nabla$ and has a highly-oscillating first order differential term.

Remark 1. In fact, the function $e^{\mathrm{i}t}$ provided a simple counter-example to the continuity with respect to the uniform norm of the application giving the area under the chord of a trajectory, since $A_{0,t}(\varepsilon e^{\mathrm{i}t/\varepsilon^2}) = t$ and $\varepsilon e^{\mathrm{i}t/\varepsilon^2}$ converges uniformly to 0 [Lyo98, Example 1.1.1, p. 217].

The Lévy area of $X = X^\varepsilon$ with $\varepsilon = 1$ between 0 and 2π is

$$A_{0,2\pi}(X) = A_{0,2\pi}(B) + A_{0,2\pi}(b) + \frac{1}{2}\int_0^{2\pi} B_t^1 b_t^2\,\mathrm{d}t - \frac{1}{2}\int_0^{2\pi} B_t^2 b_t^2\,\mathrm{d}t + \frac{1}{2}\int_0^{2\pi} \widehat{b}_t^1 \circ \mathrm{d}B_t^2 - \frac{1}{2}\int_0^{2\pi} \widehat{b}_t^2 \circ \mathrm{d}B_t^1$$

and then

$$\mathbb{E}\,[\,A_{0,2\pi}(X)\,] = \pi \text{ and } \mathbb{E}\,[\,A_{0,2\pi}(X^\varepsilon)\,] \simeq \pi \text{ for any } \varepsilon > 0.$$

In fact, using an integration by parts and proceeding as in [Lej02], it is easily proved that $A_{s,t}(X^\varepsilon)$ converges in distribution to $A_{s,t}(B) + (t-s)/2$. Here, $A_{s,t}(B) = \frac{1}{2}\int_s^t (B_r^1 - B_s^1)\circ \mathrm{d}B_r^2 - \frac{1}{2}\int_s^t (B_r^2 - B_s^2)\circ \mathrm{d}B_r^1$.

The drift term b induces some loops in the trajectories, and even after the renormalization, the area created by these loops keeps the same order of magnitude. At large scale, we see trajectories which become more and more closer to that of the Brownian motion. On the other hand, a look at the small scale shows us particles spinning around their "mean" trajectory.

Influence on SDEs. We consider now the SDE:

$$\mathrm{d}Y_t^\varepsilon = f(Y_t^\varepsilon)\circ \mathrm{d}X_t^\varepsilon = f_j(Y_t^\varepsilon)\circ \mathrm{d}B_t^{j,\varepsilon} - \frac{1}{\varepsilon}\widehat{b}^j(t/\varepsilon^2)f_j(Y_t^\varepsilon)\,\mathrm{d}t,$$

where $f = (f_j^k)_{j=1,2}^{k=1,\dots,d}$ is a smooth function from $\mathbb{R}^d$ into $\mathbb{R}^{d\times 2}$ and $B^\varepsilon = (B^{1,\varepsilon}, B^{2,\varepsilon})$ is the Brownian motion $\varepsilon B_{\cdot/\varepsilon^2}$. As in [Lej02], for $i=1,\dots,N$,

$$\begin{aligned}Y_t^{i,\varepsilon} - \varepsilon b^j(t/\varepsilon^2)f_j^i(Y_t^\varepsilon) - Y_0^{i,\varepsilon} - \varepsilon b^j(0) f_j^i(Y_0^\varepsilon) &= \int_0^t f_j^i(Y_s^\varepsilon)\circ \mathrm{d}B_s^{j,\varepsilon}\\ &- \int_0^t \widehat{b}^\ell(s/\varepsilon^2) b^j(s/\varepsilon^2)\frac{\partial f_j^i}{\partial y_k} f_\ell^k(Y_s^\varepsilon)\,\mathrm{d}s + V_t^\varepsilon,\end{aligned}$$

where V^ε contains all the terms of order ε and converges to zero in probability. The sequence $(Y^\varepsilon)_{\varepsilon>0}$ is tight, and has a unique limit Y solution to

$$\begin{aligned}\mathrm{d}Y_t^i &= f_j^i(Y_t)\circ \mathrm{d}\widehat{B}_t^j - \overline{b}^{\ell,j}\frac{\partial f_j^i}{\partial y_k} f_\ell^k(Y_t)\,\mathrm{d}t\\ &= f_j^i(Y_t)\circ \mathrm{d}\widehat{B}_t^j + \frac{1}{2}c_{j,\ell}[f_j,f_\ell](Y_t)\,\mathrm{d}t,\end{aligned}$$

$$\text{with } \overline{b}^{\ell,j} = \int_0^{2\pi} \widehat{b}_s^\ell b_s^j\,\mathrm{d}s \text{ and } c_{j,\ell} = \frac{1}{2}(\overline{b}^{\ell,j} - \overline{b}^{j,\ell}) = A_{0,2\pi}^{j,\ell}(b).$$

In fact, $c_{1,2}=1/2$, $c_{2,1}=-1/2$ and $c_{1,1}=c_{2,2}=0$. Although X^ε is very close to a Brownian motion, the drift b has a strong influence on the behavior of Y^ε when ε is very small.

3.2 Good sequence of semi-martingales

The problem of interchanging stochastic integrals or SDE and limit of semi-martingales has given rise to an abundant work around the notion of *good sequence of semi-martingales*. The article [KP96] contains a review of results about these notions.

Definition 1 (Good sequence of semi-martingales). A sequence $(Y^n)_{n\in\mathbb{N}}$ of semi-martingales is a *good sequence* if for any sequence of stochastic processes $(H^n)_{n\in\mathbb{N}}$ converging jointly with $(Y^n)_{n\in\mathbb{N}}$ to some process (H,Y) and such that the stochastic integrals $\int_0^\cdot H_s^n\,\mathrm{d}Y_s^n$ and $\int_0^\cdot H_s\,\mathrm{d}Y_s$ are well defined (which yields that H^n and Y^n are adapted with respect to the same filtration,...), then Y is a semi-martingale and $\int_0^\cdot H_s^n\,\mathrm{d}Y_s^n$ converges in distribution to $\int_0^\cdot H_s\,\mathrm{d}Y_s$.

It is immediate that if $(Y^n)_{n\in\mathbb{N}}$ is a good sequence of semi-martingales converging to Y, then $\langle Y^{i,n}, Y^{j,n}\rangle$ converges in distribution to $\langle Y^i, Y^j\rangle$. We give now a simple criterion ensuring that a sequence of semi-martingales is a good sequence.

Proposition 1. *A sequence $(Y^n)_{n\in\mathbb{N}}$ of continuous semi-martingales on $[0,T]$ decomposed as the sum of a local martingale M^n and a term of finite variation A^n is a good sequence if and only if it satisfies the* condition UCV *(for Uniformly Controlled Variations), that is*

$$(\langle M^n\rangle_T)_{n\in\mathbb{N}} \text{ and } (\mathrm{Var}_{1,[0,T]} V^n)_{n\in\mathbb{N}} \text{ are tight.} \tag{4}$$

This criterion could be refined when one works with discontinuous martingales, (in which case, the jumps shall be taken into account) or when the time interval $[0,T]$ is $\mathbb{R}_+$.

3.3 Homogenization of stochastic process with highly-oscillating drift

In this section, we consider the solutions

$$\mathrm{d}X_t^\varepsilon = \sigma(X_t^\varepsilon/\varepsilon)\,\mathrm{d}B_t + \frac{1}{\varepsilon}b(X_t^\varepsilon/\varepsilon)\,\mathrm{d}t,$$

generated by the operators

$$L^\varepsilon = \frac{1}{2}a_{i,j}(\cdot/\varepsilon)\frac{\partial^2}{\partial x_i\partial x_j} + \frac{1}{\varepsilon}b_i(\cdot/\varepsilon)\frac{\partial}{\partial x_i} \tag{5}$$

where $a = \sigma\sigma^{\mathrm{T}}$ is measurable function in the space of $N\times N$-symmetric matrix and is uniformly elliptic and bounded. The vector b is measurable and bounded.

This type of diffusion covers many physical problems, since it contains the cases where the diffusion operators has the form

$$\frac{1}{2}e^{2V(\cdot/\varepsilon)}\frac{\partial}{\partial x_i}\left(a_{i,j}(\cdot/\varepsilon)e^{-2V(\cdot/\varepsilon)}\frac{\partial}{\partial x_j}\right)$$

so that $b = -a\nabla V$. We assume that the coefficients are smooth enough and periodic. It is well known that the processes X^ε converge in distribution to some Brownian motion $\overline{\sigma}B$, where $\overline{\sigma}$ is a constant, non degenerate matrix (see for example [BLP78]). However, it is easily seen that the presence of the highly-oscillating term b implies that $(X^\varepsilon)_{\varepsilon>0}$ does *not* satisfy Condition (4). In fact, as it was shown in [Lej02], this sequence is *not* a good sequence of semi-martingales in general.

One may directly apply the Central Limit Theorem of the martingale part of X^ε using the ergodic theorem on its cross-variation, since the coefficients are periodic, and the projection of X^ε on the torus is an ergodic process with respect to an invariant measure $m(x)\,\mathrm{d}x$ whose density is solution to:

$$\frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j}(a_{i,j}(x)m(x)) - \frac{\partial}{\partial x_i}(b_i(x)m(x)) = 0 \text{ on the torus } \mathbb{T}^N,$$

where $\mathbb{T}^N = \mathbb{R}^N/\mathbb{Z}^N$.

However, the difficulty is to deal with the drift term. For that, we will introduce some functions called *correctors*, *i.e.*, the solutions of

$$L^1 u_i = -b_i \text{ on the torus } \mathbb{T}^N \tag{6}$$

for $i = 1, \dots, N$. Equation (6) may be solved if and only if

$$\int_{\mathbb{T}^N} b_i(x)m(x)\,\mathrm{d}x = 0. \tag{7}$$

From now, we assume that (7) is satisfied. Otherwise, it is easily shown, by substracting to the drift its average, that X^ε does not converge.

We remark that $u_i^\varepsilon = \varepsilon u_i(\cdot/\varepsilon)$ is solution to $L^\varepsilon u_i^\varepsilon = -\varepsilon^{-1}b_i(\cdot/\varepsilon)$. Furthermore, $L^\varepsilon v_i^\varepsilon(x) = 0$ on $\mathbb{R}^N$, where $v_i(x) = x_i + u_i(x)$ and $v_i^\varepsilon(x) = \varepsilon v_i(x/\varepsilon)$.

So, $v^\varepsilon(X_t^\varepsilon)$ is a local martingale M^ε, whose cross variations are, for $i, j = 1, \dots, N$,

$$\langle M^{i,\varepsilon}, M^{j,\varepsilon}\rangle_t \overset{\text{dist.}}{=} \varepsilon^2 \int_0^{t/\varepsilon^2} a\nabla v_i \cdot \nabla v_j(X_s)\,\mathrm{d}s, \tag{8}$$

where $X = X^\varepsilon$ with $\varepsilon = 1$. So, these cross-variation converges according to the ergodic theorem to $t\overline{a}_{i,j}$ with $\overline{a}_{i,j} = \int_{\mathbb{T}^N} a\nabla v_i \cdot \nabla v_j(x)m(x)\,\mathrm{d}x$. Both M^ε and X^ε (since u_i is bounded on $\mathbb{R}^N$) converge to $\overline{\sigma}B$, where B is a Brownian motion, and $\overline{\sigma}$ is matrix such that $\overline{\sigma}\overline{\sigma}^{\mathrm{T}} = \overline{a}$.

Remark 2. In all this article, one may add a first-order differential term $c(\cdot/\varepsilon)\frac{\partial}{\partial x_i}$. This will add to the limit of X^ε a drift of the form $(\int_{\mathbb{T}^N} c(x)\nabla u_i m(x)\,\mathrm{d}x)_{i=1,\dots,N}$, which may be treated using the Girsanov theorem. There is no difficulty to extend all our results to this case. However, to simplify the computations, we do not consider such a term (see [Lej02]).

Consequence for Lévy areas and SDEs. It was proved in [Lej02] that in presence of a highly-oscillating first order term, the Lévy areas $A_{s,t}(X^\varepsilon)$ does not converge in general to $A_{s,t}(\overline{\sigma}B)$, but to $A_{s,t}(\overline{\sigma}B) + c(t-s)$, where c is the anti-symmetric $N \times N$-matrix defined by

$$c_{i,j} = \frac{1}{2}\left(\left\langle a_{i,k}\frac{\partial u_j}{\partial x_k}\right\rangle_m - \left\langle a_{j,k}\frac{\partial u_i}{\partial x_k}\right\rangle_m + \langle b_i u_j\rangle_m - \langle b_j u_i\rangle_m\right). \tag{9}$$

Here,

$$\langle f \rangle_m = \int_{\mathbb{T}^N} f(x)m(x)\,\mathrm{d}x.$$

denotes the averaging of periodic functions with respect to the invariant measure m of the projection of X on the torus.

The solution to $\mathrm{d}Y_t^\varepsilon = h(X_t^\varepsilon/\varepsilon, X_t^\varepsilon, Y_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon$ does not converge in general to the SDE driven by the limit of X^ε (See [Lej02]).

In Section 3.4, we identify of the limit of the solution of $\mathrm{d}Y_t^\varepsilon = h(Y_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon$ in a way which is coherent with the interpolation problem of Section 2.

What happens for Lévy area of martingales or for SDE driven by martingales? The local martingales $M^\varepsilon = v^\varepsilon(X^\varepsilon)$ form a good sequence of semi-martingales, and no corrective term appears when one interchanges the limit of M^ε and the functional giving the Lévy area or the solution of some SDE.

By transforming processes into martingales, the correctors allow to suppress "pathological" behavior of the particle, such a spinning statistically more clockwise than anticlockwise. For a clear picture of the situation, we have to remember that the average of the drift term b with respect to the invariant measure of X is equal to 0.

So, for the sake of the convergence, we approximate trajectories with a lot of loops and foldings by the trajectories of martingales. However, the solutions of SDEs are sensitive to the loops and the foldings through the area, and reacts in consequence. As our examples proves it, two processes may be very close one from each other, but they may drive distinct (through a of a drift term) SDEs. as for the interpolation problem in Section 2.

3.4 Identification of the limit of solutions of SDEs driven by X^ε

Let us consider the solution of the following SDE written in the Stratonovich form:

$$\mathrm{d}Y_t^\varepsilon = f(Y_t^\varepsilon) \circ \mathrm{d}X_t^\varepsilon, \tag{10}$$

where, to keep things simple, we assume that $f = (f_1, \dots, f_N)$ maps $\mathbb{R}$ to $\mathbb{R}^N$, that is Y_t belongs to $\mathbb{R}$. The SDE (10) could be transformed into an Itô SDE:

$$\mathrm{d}Y_t^\varepsilon = f(Y_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon + \frac{1}{2} f_i(Y_t^\varepsilon) \frac{\partial f_j}{\partial y}(Y_t^\varepsilon) a_{i,j}(X_t^\varepsilon/\varepsilon)\,\mathrm{d}t.$$

Here, the process X^ε is generated by the operator L^ε defined in (5). The scalar product with respect to the invariant measure $m(x)\,\mathrm{d}x$ of L acting on the space of periodic functions is denoted by

$$\langle f, g\rangle_m = \int_{\mathbb{T}^N} f(x)g(x)m(x)\,\mathrm{d}x.$$

Without any special notification, all the involved functions are considered to be smooth enough.

Lemma 1. *Let f and g be two periodic functions. Then*

$$\left\langle \frac{L+L^*}{2} f, g\right\rangle_m = \langle Lf, g\rangle_m + \langle Lg, f\rangle_m = \langle a\nabla f\cdot\nabla g\rangle_m\,, \tag{11}$$

where L^* is the adjoint of L with respect to the scalar product $\langle\cdot,\cdot\rangle_m$.

Proof. Equality (11) follows from two integrations by parts and the fact that $L^* m = 0$. □

The homogenized diffusion coefficient is $\overline{a} = (\overline{a}_{i,j})_{i,j}$ with

$$\begin{aligned}\overline{a}_{i,j} &= \langle a(I+\nabla u_i)\cdot(I+\nabla u_j)\rangle_m \\ &= \langle a_{i,j}\rangle_m + \left\langle a_{k,j}\frac{\partial u_i}{\partial x_k}\right\rangle_m + \left\langle a_{i,k}\frac{\partial u_j}{\partial x_k}\right\rangle_m + \langle a\nabla u_i\cdot\nabla u_j\rangle_m\,.\end{aligned} \tag{12}$$

As a is symmetric and $Lu_i = -b_i$, we deduce that

$$\begin{aligned}&\frac{1}{2}\int_{\mathbb{T}^N}(b_i(x)u_j(x) - b_j(x)u_i(x))m(x)\,\mathrm{d}x \\ &\qquad = \frac{1}{2}(\langle Lu_j, u_i\rangle_m - \langle Lu_i, u_j\rangle_m) = -\left\langle \frac{L-L^*}{2}u_i, u_j\right\rangle_m.\end{aligned} \tag{13}$$

Here, it is the anti-symmetric part of L with respect to $\langle\cdot,\cdot\rangle_m$ which is involved.

As in [Lej02], the limit Y of Y^ε, is obtained by adding to Y^ε the quantity $\varepsilon u_i(X_t^\varepsilon/\varepsilon)f_i(Y_t^\varepsilon)$ computed with the Itô formula for a product. So, the limit Y is solution to

$$\begin{aligned}\mathrm{d}Y_t = f(Y_t)\,\mathrm{d}\overline{X}_t + f_i\frac{\partial f_j}{\partial y}(Y_t)\left\langle a_{i,k}\frac{\partial u_j}{\partial x_k}\right\rangle_m \mathrm{d}t + f_i\frac{\partial f_j}{\partial y}(Y_t)\,\langle u_i b_j\rangle_m\,\mathrm{d}t \\ + \frac{1}{2}f_i\frac{\partial f_j}{\partial y}(Y_t)\,\langle a_{i,j}\rangle_m\,\mathrm{d}t.\end{aligned}$$

With (12), this equation becomes

$$\begin{aligned}\mathrm{d}Y_t = f(Y_t)\,\mathrm{d}\overline{X}_t + f_i\frac{\partial f_j}{\partial y}(Y_t)\,\langle u_i b_j\rangle_m\,\mathrm{d}t + f_i\frac{\partial f_j}{\partial y}(Y_t)\left\langle a_{i,k}\frac{\partial u_j}{\partial x_k}\right\rangle_m \mathrm{d}t \\ + \frac{1}{2}f_i\frac{\partial f_j}{\partial y}(Y_t)\left(\overline{a}_{i,j} - \left\langle a_{k,j}\frac{\partial u_i}{\partial x_k}\right\rangle_m - \left\langle a_{i,k}\frac{\partial u_j}{\partial x_k}\right\rangle_m - \langle a\nabla u_i\cdot\nabla u_j\rangle_m\right)\mathrm{d}t.\end{aligned}$$

Using (11) and (13), Y is solution to

$$\mathrm{d}Y_t = f(Y_t)\,\mathrm{d}\overline{X}_t + \frac{1}{2}\overline{a}_{i,j} f_i \frac{\partial f_j}{\partial y}(Y_t)\,\mathrm{d}t + \frac{1}{2}c_{i,j}\,[f_i, f_j]\,(Y_t)\,\mathrm{d}t, \tag{14}$$

or in the Stratonovich sense,

$$\mathrm{d}Y_t = f(Y_t) \circ \mathrm{d}\overline{X}_t + \frac{1}{2}c_{i,j}\,[f_i, f_j]\,(Y_t)\,\mathrm{d}t, \tag{15}$$

where $c_{i,j}$ is the corrective term appearing in the limit (9) of the Lévy area of $(X^{i,\varepsilon}, X^{j,\varepsilon})$. This SDE (15) has to be compared with (1).

Hence, the parallel with the result of Section 2 is now complete.

Integration of one-forms. If instead of considering solutions of SDEs, one considers the integral of a smooth differential form f with coordinates $(f_i)_{i=1,\dots,N}$ along the trajectories of X^ε, that is, $Y_t^\varepsilon = \int_0^t f(X_s^\varepsilon) \circ \mathrm{d}X_s^\varepsilon$, then similar computations prove that Y^ε converges to

$$\begin{aligned} Y_t &= \int_0^t f(\overline{X}_s) \circ \mathrm{d}\overline{X}_s + \int_0^t \frac{1}{2}c_{i,j}\left(\frac{\partial f_i}{\partial x_j} - \frac{\partial f_j}{\partial x_i}\right)(\overline{X}_s)\,\mathrm{d}s \\ &= \int_0^t f(\overline{X}_s) \circ \mathrm{d}\overline{X}_s + \int_0^t c_{i,j}\frac{\partial f_i}{\partial x_j}(\overline{X}_s)\,\mathrm{d}s. \end{aligned}$$

This limit is similar to the one in (3).

4 Convergence of the solution of differential equations controlled by rough paths

4.1 Background information on the theory of rough paths

We explain in this section how the theory of rough paths allows one to unify the previous results on interpolation and homogenization. The concept of rough path allows one to regard differential equations of "SDE" type:

$$\mathrm{d}Y_t = f_i(Y_t) \circ \mathrm{d}X_t^i \tag{16}$$

as solutions of deterministic differential equations where X is a randomly chosen "rough path[2]". These are now a well developed concept, and there

[2]Rough Path is a technical term.

are many publications setting out their basic properties (See for example [Lyo98, LQ02, Lej03]). We do not feel we can justify doing it again here. For our purposes it will be sufficient to recall that such functions are like smooth functions to the extent that to solve differential equations such as (16), one can define iterated integrals of all orders, and these integrals satisfy algebraic and analytic conditions. However, the analysis is richer and allows one to consider paths that are far from smooth. The price one pays is that the rough paths have an extra structure (coming from iterated integrals) that needs to be accomodated.

In case of X is smooth and $f_i(y) = A_i y$ is linear then Y_t may be expressed in terms of iterated integrals:

$$Y_t = \left(\sum_{k \geq 0} \left(\prod_{j=1}^{j=k} A_{i_j} \right) \int_{0 \leq s_1 \leq \cdots \leq s_k \leq t} \mathrm{d}X_{s_1}^{i_1} \cdots \mathrm{d}X_{s_k}^{i_k} \right) Y_0$$

and in the general case there are asymptotic expansions for solutions involving the iterated Lie brackets of the vector field $f = (f_1, \ldots, f_N)$.

Rough paths can be thought of as coming from a closure procedure. The equations such as (16) can be thought of as functions on smooth path space. But they are not continuous functions (and they are very very non-linear functions). It is well understood that one might usefully close an unbounded *linear* operator and Sobolev spaces can be thought of in this spirit. The appropriate analogue in this non-linear context starts with the smooth paths and uses metrics (built out of the iterated integrals) to identify the space of rough paths and so close the function (16).

The value of this approach is the fortunate fact that the space of rough paths of a given p-variation is not particularly complicated.

If X is a smooth path in a Banach space V then we could define the i-th iterated integral to be $\mathbf{X}^i \in \mathrm{V}^{\otimes i}$:

$$X_{s,t}^i := \int_{s<u_1<\cdots<u_i<t} \cdots \int \mathrm{d}X_{u_1} \otimes \cdots \otimes \mathrm{d}X_{u_i} \in \mathrm{V}^{\otimes i}$$

and the truncated signature in the truncated tensor algebra

$$\mathbf{X}_{s,t} = \left(1, \mathbf{X}_{s,t}^1, \cdots, \mathbf{X}_{s,t}^k\right) \in T^{(k)}(\mathrm{V})$$
$$T^{(k)}(\mathrm{V}) = \mathbb{R} \oplus \mathrm{V} \oplus \mathrm{V} \otimes \mathrm{V} \oplus \cdots \mathrm{V}^{\otimes k}$$

and in this way, lift the path to the truncated tensor algebra. There are natural homogeneous metrics on the tensor algebra, and one may consider the p-variation of the lifted path in this space and with one of these metrics by considering

$$d_{p,[s,t]}(\mathbf{X},\mathbf{Y}) = \|\mathbf{X}-\mathbf{Y}\|_{\infty,[s,t]} + \sum_{i=1}^{k} \mathrm{Var}_{p/i,[s,t]}(\mathbf{X}^i - \mathbf{Y}^i),$$

where we define

$$\mathrm{Var}_{p/i,[s,t]}(\mathbf{X}^i) = \sup_{\substack{\text{partitions } s\leq t_1\leq\cdots\leq t_k\leq t\\ \text{for any integer } k}} \left(\sum_{j=1}^{k} |\mathbf{X}^i_{t_j,t_{j+1}}|^{p/i}\right)^{i/p}.$$

We may consider the completion of this space of paths in this metric. In particular we could take $k=1$ and get classical paths of finite p-variation. However, such a choice would not close the Itô map (16) if $p\geq 2$ as the second iterated integral is easily seen to be discontinuous in this metric. However, if $k\geq\lfloor p\rfloor$ the one can prove that the Itô map is closed and is a continuous function on the completion (and even on those paths with finite p-variation in $T^{(k)}(\mathrm{V})$ which are in the closure of the smooth paths in finite p'-variation for every $p'>p$.) and we call this closure the space of geometric rough paths on V.

The extension of the graph of the Itô map takes a p-rough path on V to a p-rough path on $\mathrm{V}\oplus\mathrm{W}$ and by projection to a p-rough path in W, where W is the original target (Banach) space where Y took its values in the equation (16).

4.2 Changing the area

A smooth path such as X is also continuous, but X has many additional features. Although we can represent X as a curve in $\mathbb{R}^N$, we could equally consider it through its lift into the tensor algebra, as described in the last section. This option is not available for a general continuous function as there can be no canonical choice for even the second iterated integral. To see this observe that $X\to X^2_{s,t}$ is not even continuous in the conventional 2-variation norm on the associated Banach space of continuous paths.

A geometric rough path can always be associated to a path in V as it is obvious from the definition of $d_{p,[0,T]}$ that if a sequence of smooth paths is Cauchy in $d_{p,[0,T]}$ then the paths are Cauchy in the classical p-variation norm defined for paths in V. However, a sequence that is Cauchy in $d_{p,[0,T]}$ has to have convergence of all higher order integrals with $k\leq\lfloor p\rfloor$. As we remarked, this does not follow from the convergence of the paths in p-variation if $p\geq 2$. By the same token, one expects different families of approximating smooth paths converging to a given path in V to have higher integrals converging to distinct limits. There are always infinitely many p-rough paths which lay above a given path of p-variation over a path in V if there is one.

For an example of this phenomena look back at examples. Suppose that $X(\varepsilon)$ are a Cauchy family of smooth paths converging to some rough path X, and that $x \in \mathrm{V}$ is the associated path of finite p-variation under X. Then

$$X_t(\varepsilon) + \varepsilon b\left(t/\varepsilon^2\right)$$

(where $b(t) = \exp(\mathrm{i}t) \in \mathbb{C}$) is also a Cauchy sequence in $d_{p,[0,T]}$ for $p > 2$.

In particular, one should note that a smooth path, regarded as a rough path with $p < 2$ has uniquely defined iterated integrals of all orders (and indeed these are used in the definition we have taken for geometric rough paths of any order) however, when $p \geq 2$, there are always many p-rough paths over a given smooth path in V which do not have the standard values for the higher iterated integrals[3]. One sometimes refers to these as *smooth rough paths* to distinguish them!

So a rough path carries extra information which is over and above a function carrying $[0,T] \to \mathrm{V}$ and which cannot be derived from it. In effect one is saying that, to approximate differential equations such as (16), it is enough that we know a chordal approximation to X in the case where X is of p-variation less than 2 but in the case where it is greater than two it is no longer sufficient as the errors that occur do not disappear as one refines the chordal partition. One needs a better statistic to describe the path X over an interval of time than its chord. The extra iterated integrals for $k \leq p$ provide an adequate description and if they are collectively used to describe the evolution of Y in (16) approximately one can develop schemes that converge as one refines the partitions in a way that would be impossible with chordal descriptions. (This was used for example to develop variable step size algorithms for SDEs in [GL97]).

The extra information contained in the first k iterated integrals is not totally simple to extract as there are lots of algebraic identities between them because the signature always takes its values in a certain nilpotent group G^k embedded in the tensor algebra in a non-linear way. A simple example of this can be seen when $k = 2$ and $p < 3$. In this case we can regard the second iterated integral as a matrix

$$\mathbf{X}^{2,i,j}_{s,t} := \iint_{s<u<v<t} \mathrm{d}X^i_u \,\mathrm{d}X^j_v$$

[3] Although this might seem counter intuitive to some one unfamiliar with the theory, and a cause of mistakes, its is in fact quite natural.

and can consider the anti-symmetric and symmetric parts separately. Now

$$\mathfrak{S}^{i,j}_{s,t}(\mathbf{X}) = \frac{1}{2}(\mathbf{X}^{2,i,j}_{s,t} + \mathbf{X}^{2,j,i}_{s,t})$$

can be identified as $\mathfrak{S}^{i,j}_{s,t}(\mathbf{X}) = \frac{1}{2}(X^i_t - X^i_s)(X^j_t - X^j_s)$ and one readily sees that it is a function of $X^1_{s,t}$. On the other hand the anti-symmetric part:

$$\mathfrak{A}^{i,j}_{s,t}(\mathbf{X}) = \frac{1}{2}(\mathbf{X}^{2,i,j}_{s,t} - \mathbf{X}^{2,j,i}_{s,t})$$

is completely fresh and represents the area between the chord and the path. It is simple to see from the above examples that we can find $d_{p,[0,T]}$ Cauchy sequences of paths converging to rough paths where $X^1_{s,t}$ and $\mathfrak{S}^{i,j}_{s,t}(\mathbf{X})$ agree but where $\mathfrak{A}^{i,j}_{s,t}(\mathbf{X})$ is only determined up to a function $\psi(t) - \psi(s)$ and ψ is a function of finite $\frac{p}{2}$-variation.

To improve the intuition, consider a smooth path X and its canonical iterated integrals (the unique extension to a finite 1-variation path in $T^{(k)}(\mathrm{V})$). Now consider the paths $X_t(\varepsilon) = X_t + \varepsilon \exp(2\mathrm{i}\psi(t)/\varepsilon^2)$ in the complex plane in a way similar to the example above. Then define

$$\mathrm{d}Y^\varepsilon_t = f_i(Y^\varepsilon_t) \circ \mathrm{d}X^i_t(\varepsilon).$$

Since the $(X(\varepsilon))_{\varepsilon>0}$ form a Cauchy sequence, the Y^ε will converge to some limit $\widetilde{Y}$ (assuming the f are smooth enough). Can we understand the limit — and describe $\widetilde{Y}$ as a classical solution to a classical differential equation? In fact the answer is yes:

$$\mathrm{d}\widetilde{Y}_t = f_i(\widetilde{Y}_t) \circ \mathrm{d}X^i_t + \frac{1}{2}[f_1, f_2](\widetilde{Y}_t)\,\mathrm{d}\psi_t$$

so we see that the difference between geometric rough paths over a fixed path is explicitly reflected in the response of the system and can be calculated quite precisely in terms of a new vector field built out of the Lie bracket of the original vector fields defining the original differential equation.

For the rest of this section we restrict to the situation where $\mathbf{X}$ is a geometric rough path $\mathbf{X}$ of finite p-variation with $2 \leq p < 3$, and we assume that $\mathbf{X}$ lies above a path X. The map $\mathfrak{K}$, integrating a one form, is defined on geometric rough paths in $T^{(2)}(\mathrm{V})$ and takes its values in $T^{(2)}(\mathrm{W})$. However, if $\mathbf{Y} = (1, \mathbf{Y}^1, \mathbf{Y}^2) = \mathfrak{K}(\mathbf{X})$, and if f is a one form of class $\mathcal{C}^q$ for $q > p - 1$, then $\mathbf{Y}^1$ and $\mathbf{Y}^2$ may be computed by:

$$\mathbf{Y}^1_{s,t} = \lim_{\delta\to 0} \sum_{i=1,\dots,k^\delta,\ t^\delta_i, t^\delta_{i+1}\in[s,t]} \widetilde{\mathbf{X}}^1_{t^\delta_i, t^\delta_{i+1}} \tag{17}$$

$$\text{with } \widetilde{\mathbf{X}}^1_{s,t} = \sum_{k=1}^N f_k(X_s)\mathbf{X}^{1,k}_{s,t} + \sum_{k,j=1}^N \frac{\partial f_k}{\partial x_j}(X_s)\mathbf{X}^{2,j,k}_{s,t},$$

$$\text{and } \mathbf{Y}^2_{s,t} = \lim_{\delta\to 0} \sum_{\substack{i=1,\dots,k^\delta\\ t^\delta_i, t^\delta_{i+1}\in[s,t]}} \sum_{k,\ell=1}^N f_\ell(X_{t^\delta_i}) f_k(X_{t^\delta_i}) \mathbf{X}^{2,\ell,k}_{t^\delta_i, t^\delta_{i+1}} \tag{18}$$

$$+ \sum_{\substack{i,j=1,\dots,k^\delta\\ t^\delta_i, t^\delta_{i+1}\in[s,t]\\ t^\delta_j, t^\delta_{j+1}\in[s,t]}} \widetilde{\mathbf{X}}^1_{t^\delta_i, t^\delta_{i+1}} \otimes \widetilde{\mathbf{X}}^1_{t^\delta_j, t^\delta_{j+1}}, \tag{19}$$

where $\Pi^\delta = \left\{ t^\delta_i \,\middle|\, i = 1,\dots,k^\delta \right\}$ is a partition of $[0,T]$ whose mesh decreases to 0 with δ. (The existence of the limit relies on an algebraic property of these iterated integrals inherited by geometric rough paths from their smooth cousins). Note that, even if $\mathbf{X}$ is a stochastic process with rough sample paths then this integral $\mathbf{Y}$ may be computed pathwise for each rough sample paths and all f.

We denote by $\mathbf{X}$ and $\widehat{\mathbf{X}}$ two geometric rough paths of finite p-variation with $2 \leq p < 3$, and such that $\widehat{\mathbf{X}}^1 = \mathbf{X}^1$. From Lemma 2.2.3 in [Lyo98, p. 250], we know that there exists a function $\psi = (\psi_{i,j})_{i,j=1}^N$ defined on $[0,1]$ and such that for any $0 \leq s \leq t \leq 1$,

$$\widehat{\mathbf{X}}^{2,i,j}_{s,t} = \mathbf{X}^{2,i,j}_{s,t} + \psi_{i,j}(t) - \psi_{i,j}(s)$$

where ψ is of finite $p/2$-variation.

The following proposition is immediate from (17) and (18).

Proposition 2. *Let $\mathbf{X}$, $\widehat{\mathbf{X}}$ and ψ as above. We set $\mathbf{Y} = \mathfrak{K}(\mathbf{X})$ and $\widehat{\mathbf{Y}} = \mathfrak{K}(\widehat{\mathbf{X}})$. Then for $i,j = 1,\dots,N$,*

$$\widehat{\mathbf{Y}}^{1,i}_{s,t} = \mathbf{Y}^{1,i}_{s,t} + \sum_{j=1}^N \int_s^t \frac{\partial f_i}{\partial x_j}(X_r)\,\mathrm{d}\psi_{i,j}(r)$$

$$\text{and } \widehat{\mathbf{Y}}^{2,i,j}_{s,t} = \mathbf{Y}^{2,i,j}_{s,t} + \sum_{k,\ell=1}^N \int_s^t f^i_k(X_r) f^j_\ell(X_r)\,\mathrm{d}\psi_{k,\ell}(r)$$

$$+ \sum_{k,\ell=1}^N \int_s^t \frac{\partial f_i}{\partial x_k}(X_r)\,\mathrm{d}\psi_{i,k}(r) \otimes \int_s^t \frac{\partial f_j}{\partial x_\ell}(X_r)\,\mathrm{d}\psi_{j,\ell}(r).$$

Now consider the more sophisticated functional: the Itô map $\mathfrak{I}$. Of course superficially, we simply want to construct a "solution" geometric rough path $\mathbf{Y} = \mathfrak{I}(\mathbf{X})$ in $T^{(2)}(\mathrm{V})$, but in fact it is necessary to construct a geometric rough path $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ in $T^{(2)}(\mathrm{V} \oplus \mathrm{W})$ in order to interpret the integral. If $\mathrm{V} = \mathbb{R}^N$ and $\mathrm{W} = \mathbb{R}^d$, then $\mathbf{Z}$ is solution to $\mathbf{Z} = \mathfrak{K}'(\mathbf{Z})$, where $\mathfrak{K}'$ is the application corresponding to the integration of the differential form $z = (z, y) \mapsto f_i(y)\,\mathrm{d}x^i$. Hence, $\mathbf{Z}^2 = (\mathbf{Z}^{2,i,j})_{(i,j)\in I}$, where I is the set of indexes such that i and j belong to $\{1, \dots, N\}$ or to $\{1, \dots, d\}$. Moreover, if (i, j) belongs to $\{1, \dots, N\} \times \{1, \dots, d\}$, then

$$\mathbf{Z}^{2,i,j}_{s,t} = \sum_{k=1}^{N} \int_s^t f_k^i(Y_r)\,\mathrm{d}\mathbf{X}^{2,k,j}_{s,r}. \tag{20}$$

A similar relation holds for $(i, j) \in \{1, \dots, m\} \times \{1, \dots, N\}$, and (20) is easily proved for smooth X, and then by passing to the limit using the continuity of $\mathfrak{I}$.

The following corollary is immediate from of (20) and Proposition 2.

Corollary 1. *Let $\mathbf{X}$, $\widehat{\mathbf{X}}$ and ψ as above. We set $\widehat{\mathbf{Y}} = \mathfrak{I}(\widehat{\mathbf{X}})$. Then $\widehat{\mathbf{Y}}$ lies above a path $\widehat{Y}$ which is solution to*

$$\widehat{Y}_t^i = \widehat{Y}_0^i + \left(\int_0^t f(\widehat{Y}_r)\,\mathrm{d}\mathbf{X}_r\right)^{1,i} + \sum_{j,k=1}^{d} \int_0^t \frac{\partial f_k^i}{\partial y_j}(\widehat{Y}_r) f_k(\widehat{Y}_r)\,\mathrm{d}\psi_{k,j}(r)$$

for $i = 1, \dots, d$. Here, $\left(\int_s^t f(\widehat{Y}_r)\,\mathrm{d}\mathbf{X}_r\right)_{0\leq s\leq t\leq 1}$ denotes the geometric rough path $\mathfrak{K}'((\mathbf{X}, \widehat{\mathbf{Y}}))$.

4.3 Rough paths and Stratonovich integrals of semi-martingales

We assume that $\mathrm{V} = \mathbb{R}^N$. Let X be a continuous semi-martingale. If one considers again (18) then in the expression

$$\lim_{\delta\to 0} \sum_{i=1,\dots,k^\delta,\ t_i^\delta\in[s,t]} \left(\sum_{k=1}^{N} f_k(X_{t_i^\delta}) \mathbf{X}^{1,k}_{t_i^\delta, t_{i+1}^\delta} + \sum_{k,j=1}^{N} \frac{\partial f_k}{\partial x_j}(X_{t_i^\delta}) \mathbf{X}^{2,j,k}_{t_i^\delta, t_{i+1}^\delta} \right)$$

the first term added to the symmetric term in the expression $\sum_{k,j=1}^{N} \frac{\partial f_k}{\partial x_j}(X_{t_i^\delta}) \mathbf{X}^{2,j,k}_{t_i^\delta, t_{i+1}^\delta}$ can easily be seen, using Taylors Formula, to be what most people call the Stratonovich integral. It is shown to converge in probability under fairly broad assumptions on X providing the times of the partitions are kept deterministic. As a consequence, we can identify the Stratonovich integral with the rough path integral if we can prove that the anti-symmetric part of

$$\lim_{\delta\to 0} \sum_{i=1,\dots,k^\delta,\ t_i^\delta\in[s,t]} \sum_{k,j=1}^{N} \frac{\partial f_k}{\partial x_j}(X_{t_i^\delta})\mathbf{X}^{2,j,k}_{t_i^\delta,t_{i+1}^\delta} = 0. \tag{21}$$

Note that (21) is always true in dimension 1, or if $f_k = \frac{\partial F}{\partial x_k}$ for some function F smooth enough.

Hence we see that this is essentially equivalent to proving that the polygonal approximations to a Brownian path are Cauchy in the sense of rough paths. This convergence can easily be checked independently and provides extra insight into the Wong-Zakai theorem: see for example [NY78, Pro85], ...

4.4 Construction of a Brownian motion with arbitrary area

In this section, we transpose to rough paths the results of Section 2.

Let B be a Brownian motion on a probability space $(\Omega,\mathcal{F},\mathbb{P})$, and let $\mathbf{B}$ be the "natural stochastic" geometric functional lying above the trajectories of the Brownian motion B:

$$\mathbf{B}^{1,i}_{s,t} = B^i_t - B^i_s \text{ and } \mathbf{B}^{2,i,j}_{s,t} = \int_s^t (B^i_r - B^i_s)\circ \mathrm{d}B^j_r.$$

Thus, the anti-symmetric part $A^{i,j}_{s,t}(B) = \frac{1}{2}(\mathbf{B}^{2,i,j}_{s,t} - \mathbf{B}^{2,j,i}_{s,t})$ of $\mathbf{B}^{2,i,j}_{s,t}$ is just the Lévy area of (B^i, B^j).

For some integer m, we set $\delta = 1/2^m$ and $t^m_j = j/2^m$ for $j = 0,\dots,2^m$. For $i = 1,\dots,N$ and almost every realization $\omega\in\Omega$,

$$B^{i,\delta}_t(\omega) = B^{i,\delta}_{t^m_j}(\omega) + \frac{t - t^m_j}{t^j_{m+1} - t^j_m}(B^{i,\delta}_{t^m_j}(\omega) - B^{i,\delta}_{t^m_{j+1}}(\omega)) \text{ for } t\in[t^m_j, t^m_{j+1}].$$

We already know that $(B^\delta, A(B^\delta))$ converges in probability to $(B, A(B))$ in $\mathbb{V}^p$ as $\delta\to 0$ for any $p > 2$ (in fact, there are other approximation of the trajectories of B leading to the same result: see [IW89, Section VI-7]).

Let $\psi = (\psi_{i,j})_{i,j=1,\dots,N}$ be a function on $[0,1]$ taking its values in the space of anti-symmetric $N\times N$-matrices. In Section 4.2, we have seen that $\widehat{\mathbf{B}}$ defined by $\widehat{\mathbf{B}}^1_{s,t} = \mathbf{B}^1_{s,t}$ and $\widehat{\mathbf{B}}^{2,i,j}_{s,t} = \mathbf{B}^{2,i,j}_{s,t} + \psi_{i,j}(t) - \psi_{i,j}(s)$ is also a geometric rough path. Moreover, we have seen that a drift is added when one consider $\mathfrak{K}(\widehat{\mathbf{B}})$ and $\mathfrak{I}(\widehat{\mathbf{B}})$ instead of $\mathfrak{K}(\mathbf{B})$ and $\mathfrak{I}(\mathbf{B})$.

Moreover, we know that $\widehat{\mathbf{B}}$ is the limit in $\mathbb{V}^p$ of rough paths lying above a piecewise smooth trajectory $\widehat{B}^\delta$ (see Lemma 2.3.1 in [Lyo98, p. 259]). The problem is to find some explicit expression for $\widehat{B}^\delta$. We do it below in a simple case.

Remark 3. Let $(e_i)_{i=1,\dots,N}$ be an orthonormal basis of $\mathrm{V} = \mathbb{R}^N$. At any time t the expectation of $\mathbf{B}_t$ is an element of $T^{(2)}(\mathrm{V})$ equal to

$$\mathbb{E}\,[\,\mathbf{B}_t\,] = \exp\left(\frac{t}{2}(e_1 \otimes e_1 + \cdots + e_N \otimes e_N)\right),$$

where the exp, the inverse of the function log previously introduced, is defined by the projection on $T^{(2)}(\mathrm{V})$ of the non-commutative power series

$$\exp(a^1 + \cdots + a^N) = \frac{1}{k!} \sum_{k \geq 0} \sum_{\text{multi-index } (i_1,\dots,i_k)} a^{i_1} \dots a^{i_k}.$$

One the other side, one may embed the points $(S_k)_{k \in \mathbb{N}}$ of a simple random walk S in a trajectory $(S_t)_{t \geq 0}$ for which the area above the path is different from the area above the path when the points of the random walk S are simply linked by straight lines. The renormalization $S_t^n = S_{nt}/\sqrt{n}$ of $(S_t)_{t \geq 0}$ satisfies

$$\lim_{n \to \infty} \mathbb{E}\,[\,\mathbf{S}_t^n\,] = \exp\left(\frac{t}{2}\sum_{i=1}^N e_i \otimes e_i + \frac{t}{2}\sum_{i,j=1}^N c_{i,j}[e_i, e_j]\right),$$

where $\mathbf{S}^n$ is the geometric rough path lying above S^n, $[e_i, e_j] = e_i \otimes e_j - e_j \otimes e_i$ and $(c_{i,j})_{i,j=1,\dots,N}$ is an anti-symmetric matrix given by the expectation of the additional area between two points (See [Faw03, LV03]).

A theorem of McShane revisited. In [McS72] McShane gave an explicit example to see that the Wong-Zakai theorem was wrong when one used a generic interpolation. We can interpret his result as saying that in general, if one provides an interpolation of a Brownian path to a path that is smooth off a discrete set of times and continuous and agreeing in value with the Brownian path at deterministic times path, one expects that the family of paths will be tight in the metric of rough paths. Converging subsequences will provide converging solutions to the SDE. However, in general it will not be the case that the limit of the convergent sequences will have the "natural stochastic" lift to a rough path although it will often be the case that it does converge to some unique rough path if the interpolation has enough independence of the Brownian path.

For that, we explicitly construct, when the dimension N is 2, an approximation $X^\delta(\omega)$ of $B(\omega)$ such that the limit of $A_{s,t}(X^\delta)$ is $A_{s,t}(B)+c(t-s)$, where $c=(c_{i,j})_{i,j=1,2}$ is an anti-symmetric 2×2-matrix.

We assume that the dimension N of the space is 2. Let us introduce a smooth function $\phi=(\phi^1,\phi^2):[0,1]\to\mathbb{R}^2$ such that $\phi(0)=(0,0)$, $\phi(1)=(1,0)$ and

$$c_{1,2}=A^{1,2}_{0,1}(\phi)=\frac{1}{2}\left(\int_0^1\phi^1_s\,\mathrm{d}\phi^2_s-\int_0^1\phi^2_s\,\mathrm{d}\phi^1_s\right).$$

We identify $\mathbb{R}^2$ with the complex plane $\mathbb{C}$: The Brownian motion $B=(B^1,B^2)$ is now $B_t=B^1+\mathrm{i}B^2$ and $\phi=\phi^1+\mathrm{i}\phi^2$. Let us define the process $(X^\delta_t)_{t\in[0,1]}$ by setting

$$X^\delta_t=B_{t^m_j}+\phi((t-t^m_j)/\delta)(B_{t^m_{j+1}}-B_{t^m_{j+1}})$$

for $t\in[t^m_j,t^m_{j+1}]$. If $\phi(t)=(t,0)$, then X^δ is the piecewise linear approximation of B along the partition $\left\{t^m_j\right\}_{j=1,\dots,k}$.

Proposition 3. *If X^δ is constructed as previously, then with probability one:*

$$(X^\delta,A(X^\delta))\xrightarrow[\delta\to 0]{\mathbb{V}^p}(B_t-B_s,A_{s,t}(B)+2c(t-s))_{(s,t)\in\Delta^+}.$$

Proof. One knows that the piecewise linear path B^δ together with its area $A_{s,t}(B^\delta)$ converge a.s. in the rough path metric to Brownian motion B and the standard Lévy area $A(B)$ (See for example [Sip93, LLQ02],...). An easy calculation shows that the area associated to a second interpolation X^δ differs from $A_{s,t}\left(B^\delta\right)$ by the sum of the areas in the individual intervals of interpolation, that is, for dyadic times s,t and δ small enough,

$$A_{s,t}(X^\delta)-A_{s,t}(B^\delta)=\sum_{t_i,t_{i+1}\in[s,t]}A_{t_i,t_{i+1}}(X^\delta)=c\sum_{t_i,t_{i+1}\in[s,t]}\left|B_{t_{i+1}}-B_{t_i}\right|^2.$$

Hence, the result comes easily from the fact that $(s,t)\mapsto\sum_{t_i,t_{i+1}\in[s,t]}|B_{t_{i+1}}-B_{t_i}|^2$ converges almost surely both uniformly and in $p/2$-variation to $(s,t)\mapsto 2(t-s)$. □

Remark 4. Using the results from [CL05], this result is easily generalized to continuous semi-martingales.

4.5 Useful results to prove the convergence in p-variation

We end this section with lemmas that will be useful in the next Section, in order to compute prove that some convergence also holds in p-variation.

Let $(\mathbf{X}^\varepsilon)_{\varepsilon>0}$ be a family of rough paths in $T^{(k)}(\mathrm{V})$ of finite p-variation. We are interested in giving a simple relative compactness criteria for this family with respect to $d_{p,[0,1]}$. Let us denote by $\mathbb{V}^p$ the topology this distance $d_{p,[0,1]}$ generates on $T^{(k)}(\mathrm{V})$. Let also $(\mathcal{C}, \|\cdot\|_\infty)$ be the space of continuous functions from $[0,T]$ to the metric space V with the uniform norm $\|x\|_\infty = \sup_{t\in[0,T]} |x(t)|$.

We set $\Delta^+ = \{ (s,t) \,|\, 0 \leq s \leq t \leq T \}$. We remark first that if $(Y^\varepsilon)_{\varepsilon>0}$ is a family of functions from Δ^+ to $\mathbb{R}$ of finite q-variation such that $Y^\varepsilon_{s,t}$ converges to $Y_{s,t}$ for any $(s,t) \in \Delta^+$, then

$$\mathrm{Var}_{q,[0,T]}(Y) \leq \liminf_{\varepsilon\to 0} \mathrm{Var}_{q,[0,T]}(Y^\varepsilon). \tag{22}$$

We also remark that for any $\alpha > 0$,

$$\begin{aligned}\mathrm{Var}_{q+\alpha,[0,T]}(Y - Y^\varepsilon)^{q+\alpha} \\ \leq 2^{\alpha+q-1} \sup_{(s,t)\in\Delta^+} |Y_{s,t} - Y^\varepsilon_{s,t}|^\alpha \left(\mathrm{Var}_{q,[0,T]}(Y^\varepsilon)^q + \mathrm{Var}_{q,[0,T]}(Y)^q\right).\end{aligned} \tag{23}$$

Thus, if we combine (23) and (22), we see that the uniform convergence of Y^ε to Y and the condition $\sup_{\varepsilon>0} \mathrm{Var}_{q,[0,T]}(Y^\varepsilon) < +\infty$ implies that Y^ε converges to Y in $\mathbb{V}^{q+\alpha}$ for any $\alpha > 0$, and the limit is of finite q-variation.

Proposition 4. *Let $(\mathbf{X}^\varepsilon)_{\varepsilon>0}$ be a family of random rough paths in $T^{(k)}(\mathrm{V})$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We assume that for any $\kappa > 0$ and for $i = 1, \dots, k$,*

$$\lim_{\eta\to 0} \sup_{\varepsilon>0} \mathbb{P}\left[\sup_{|t-s|<\eta} |\mathbf{X}^{i,\varepsilon}_{s,t}| > \kappa\right] = 0,$$

and that for any $\delta > 0$, there exists some $\kappa > 0$ such that for $i = 1, \dots, k$,

$$\sup_{\varepsilon>0} \mathbb{P}\left[\mathrm{Var}_{q/i,[0,T]}(\mathbf{X}^{i,\varepsilon}) > \kappa\right] < \delta. \tag{24}$$

Then there exists some rough path $\mathbf{X}$ in $T^{(k)}(\mathrm{V})$ of finite p-variation such that, for any $q > p$, $\mathbf{X}^\varepsilon$ converges in distribution to $\mathbf{X}$ with respect to the topology $\mathbb{V}^q$ generated by $d_{q,[0,T]}$ along a subsequence.

If furthermore $\mathbf{X}^\varepsilon$ lies above a stochastic process X^ε and $(X^\varepsilon_0)_{\varepsilon>0}$ is tight, then $\mathbf{X}$ lies above some stochastic process X which is a limit of $(X^\varepsilon)_{\varepsilon>0}$ in the space of continuous functions.

Remark 5. As $\mathbb{V}^p$ is not separable, the convergence of $\mathbf{X}^\varepsilon$ to $\mathbf{X}$ in distribution in $\mathbb{V}^p$ does not necessarily implies that (24) is satisfied.

Let $\mathbf{X} = (1, \mathbf{X}^1, \mathbf{X}^2)$ be a rough path in $T^{(2)}(\mathrm{V})$. One of the practical problem with the rough paths theory is to prove that a candidate to a geometric rough path is really of finite p-variation. For that, one could use the following estimate. Let us set

$$\omega_{p,[s,t]}(\mathbf{X}) = \sup_{k\geq 1,\ \text{partition } (t_i)_{i=1}^k \text{ of } [s,t]} \sum_{i=1}^{k} (|\mathbf{X}^1_{t_i,t_{i+1}}|^p + |\mathbf{X}^2_{t_{i+1},t_i}|^{p/2}),$$

Proposition 5 (See [BHL02] or [LLQ02] for example)**.** *There exists some constant $C = C(p)$ such that for any $\gamma > p - 1$,*

$$\omega_{p,[s,t]}(\mathbf{X}) \leq C(\eta_{s,t}(\mathbf{X}^1, 1, p) + \eta_{s,t}(\mathbf{X}^2, 2, p)),$$

where

$$\eta_{s,t}(Y, m, p) = \sum_{k\geq 1} k^\gamma \sum_{i=1}^{2^k - 1} |Y_{t_i^k, t_{i+1}^k}|^{p/m},\ t_i^k = s + (t-s)i/2^k.$$

What this Proposition said is that one could estimate the p-variation of $\mathbf{X}$ if one has a good estimate $\mathbf{X}_{t_i,t_{i+1}}$ at the dyadics points of $[0, T]$.

5 Rough paths and homogenization

We come back now to the homogenization problem: X^ε is the process generated by (5). Let us recall that $v_i^\varepsilon(x) = x_i + \varepsilon u_i(x/\varepsilon)$, where u_i's are the correctors, *i.e.*, the solutions of $Lu_i = -b_i$ on the space of periodic functions.

For a semi-martingale X, we denote by $A_{s,t}(X)$ its "Lévy area" between the times s and t, that is $A^{i,j}_{s,t}(X) = \int_s^t (X^i_r - X^i_s) \circ \mathrm{d}X^j_r$.

We know from the theory of homogenization that both X^ε and the martingale $v^\varepsilon(X^\varepsilon)$ converge to the same limit $\overline{X}$ on $(\mathcal{C}, \|\cdot\|_\infty)$. Moreover, from [Lej02],

$$A_{0,t}(X^\varepsilon) \xrightarrow[\varepsilon\to 0]{(\mathcal{C},\|\cdot\|_\infty)} A_{0,t}(\overline{X}) + ct \text{ and } A_{0,t}(v^\varepsilon(X^\varepsilon)) \xrightarrow[\varepsilon\to 0]{(\mathcal{C},\|\cdot\|_\infty)} A_{0,t}(\overline{X}),$$

where the coefficients $c_{i,j}$ of the anti-symmetric matrix c are given by (9).

In order to apply the results from the theory of rough paths, mainly the continuity of the maps $\mathfrak{K}$ and $\mathfrak{I}$, one has to prove that his convergence also holds in p-variation.

5.1 Convergence with the corrected driving process

Let us fix a real $p > 2$. As a and the gradients of the v_i's are bounded, it is clear that $v^\varepsilon(X^\varepsilon)$ satisfies the condition UCV of Proposition 1 (See [Lej02]). Then it follows from [CL05] that

$$(v^\varepsilon(X^\varepsilon), A(v^\varepsilon(X^\varepsilon))) \xrightarrow[\varepsilon\to 0]{\mathbb{V}^p} (\overline{X}, A(\overline{X})).$$

Consequently, it is immediate that if f is a function in $\mathcal{C}^{p+\eta}$ for some $\eta > 0$, then the solution of the SDE $\mathrm{d}Y_t^\varepsilon = f(Y_t^\varepsilon) \circ \mathrm{d}v^\varepsilon(X_t^\varepsilon)$ converges to the solution of the SDE $\mathrm{d}Y_t = f(Y_t) \circ \mathrm{d}\overline{X}_t$.

5.2 Convergence with the initial driving process

Again, p is a fixed real satisfying $p > 2$.

Proposition 6. *The following convergence holds:*

$$(X_{s,t}^\varepsilon, A_{s,t}(X^\varepsilon))_{(s,t)\in\Delta^+} \xrightarrow[\varepsilon\to 0]{\mathbb{V}^p} (\overline{X}_t, A_{s,t}(\overline{X}) + c(t-s))_{(s,t)\in\Delta^+}$$

in distribution.

With the results of Section 4.4, it is clear that the solution of the SDE $\mathrm{d}Y_t^\varepsilon = f(Y_t^\varepsilon) \circ \mathrm{d}X_t^\varepsilon$ converges to the solution of the SDE (14).

Although $v^\varepsilon(X^\varepsilon)$ is very close to X^ε when ε is small, the behavior of the solutions of the differential equation controlled by these processes are different.

As $V_t^\varepsilon = \frac{1}{\varepsilon}\int_0^t b(X_s^\varepsilon/\varepsilon)\,\mathrm{d}s$ converges in distribution to a martingale, one cannot expect that the sequence $(\mathrm{Var}_{1,[0,T]} V^\varepsilon)_{\varepsilon>0}$ is tight and then that $(X^\varepsilon)_{\varepsilon>0}$ satisfies the condition UCV, and the results from [CL05] cannot be applied.

Lemma 2. *There exists some constant C such that for any $0 \leq s \leq 1$ and any $p \geq 2$,*

$$\sup_{\varepsilon>0}\sup_{x\in\mathbb{R}^N}\mathbb{E}_x\left[\sup_{r\in[s,t]}|X_r^\varepsilon - X_s^\varepsilon|^p\right] \leq C|t-s|^{p/2}, \tag{25}$$

where X^ε generated by L^ε and starting at point x under $\mathbb{P}_x$.

Proof. In this proof, we denote the distribution of X^ε by $\mathbb{P}^\varepsilon$ and by X the canonical projection at time t: $X_t^\varepsilon(\omega) = \omega_t$ for any continuous function ω.

From the Markov property of X,

$$\sup_{x\in\mathbb{R}^N}\mathbb{E}_x^\varepsilon\left[\sup_{r\in[s,t]}|X_r-X_s|^p\right]=\sup_{x\in\mathbb{R}^N}\mathbb{E}_x^\varepsilon\left[\sup_{r\in[0,t-s]}|X_r-X_0|^p\right].$$

Let us consider an arbitrary positive real R, and let us denote by τ the first-exit time from the ball $B(x,R)$ of radius R and center x: $\tau=\inf\{\,t\geq 0\,|\,|X_t|\geq R\,\}$. Then, as X is a strong Markov process and τ is a stopping time,

$$\begin{aligned}\sup_{x\in\mathbb{R}^N}\mathbb{E}_x^\varepsilon\left[\sup_{r\in[0,t]}|X_r-X_0|^p\right]&\leq\sup_{x\in\mathbb{R}^N}\mathbb{E}_x^\varepsilon\left[\sup_{r\in[0,t]}|X_r-X_0|^p;t\leq\tau\right]\\&\quad+2^{p-1}\sup_{x\in\mathbb{R}^N}\mathbb{E}_x^\varepsilon\left[\sup_{r\in[0,t-\tau]}|X_r-X_\tau|^p;t>\tau\right]\\&\quad+2^{p-1}\sup_{x\in\mathbb{R}^N}\mathbb{E}_x^\varepsilon\left[\,|X_\tau-X_0|^p;t>\tau\,\right]\\&\leq(1+2^{p-1})R^p+2^{p-1}\sup_{x\in\mathbb{R}^N}\mathbb{E}_x^\varepsilon\left[\sup_{r\in[0,t]}|X_r-X_0|^p\right]\sup_{x\in\mathbb{R}^N}\mathbb{P}_x^\varepsilon\left[\,t>\tau\,\right].\end{aligned}$$

If we assume that $\sup_{\varepsilon>0}\sup_{x\in\mathbb{R}^N}\mathbb{P}_x^\varepsilon\left[\,t>\tau\,\right]\leq c$ for a constant c small enough and $R=\beta\sqrt{t}$, then

$$\sup_{x\in\mathbb{R}^N}\mathbb{E}_x^\varepsilon\left[\sup_{r\in[0,t]}|X_r-X_0|^p\right]\leq\frac{(1+2^{p-1})R^p}{1-2^{p-1}\sup_{x\in\mathbb{R}^N}\mathbb{P}_x^\varepsilon\left[\,t>\tau\,\right]}\leq Ct^{p/2}$$

and our Lemma is proved.

The constant β will be chosen later. We assume first that ε and t satisfy the relation

$$\varepsilon\,\|u\|_\infty\leq\frac{\beta\sqrt{t}}{4}=\frac{R}{4}.\tag{26}$$

Then

$$\begin{aligned}\mathbb{P}_x^\varepsilon\left[\,t>\tau\,\right]=\mathbb{P}_x&\left[\sup_{r\in[0,t]}|M_t^\varepsilon-\varepsilon u(X_t^\varepsilon/\varepsilon)+\varepsilon u(x/\varepsilon)|>R\right]\\&\leq\mathbb{P}_x\left[\sup_{r\in[0,t]}|N_t^\varepsilon|>\frac{R}{2}\right],\end{aligned}$$

where N^ε is the martingale defined by $M_t^\varepsilon=X_t^\varepsilon-X_0^\varepsilon+\varepsilon u(X_t^\varepsilon/\varepsilon)-\varepsilon u(X_0^\varepsilon/\varepsilon)$ and $N_0^\varepsilon=0$. In distribution, the cross-variation of N^ε are

$$\begin{aligned}&\mathcal{L}\left(\langle N^{i,\varepsilon},N^{j,\varepsilon}\rangle_t\,\middle|\,\mathbb{P}_x^\varepsilon\right)\\&\qquad=\mathcal{L}\left(\varepsilon^2\int_0^{t/\varepsilon^2}a_{k,\ell}\left(\delta_{i,k}+\frac{\partial u_i}{\partial x_k}\right)\left(\delta_{j,\ell}+\frac{\partial u_j}{\partial x_\ell}\right)(X_r)\,\mathrm{d}r\,\middle|\,\mathbb{P}_{x/\varepsilon}^1\right),\end{aligned}$$

so that $\sup_{\varepsilon>0}\mathbb{E}_x^\varepsilon\left[\,\langle N^{i,\varepsilon},N^{i,\varepsilon}\rangle_t\,\right]\leq C_0 t$ for some constant C_0 that depends only on the bounds of a, b and ∇u. Using the Bienaymé-Chebyshev inequality and the estimate on the supremum of a martingale, there exists some constant C_1 independent from ε and t such that, if (26) holds,

$$\mathbb{P}_x^\varepsilon\left[\,t>\tau\,\right]\leq C_1\frac{t}{R^2}\leq\frac{C_1}{\beta}.$$

Now, if we assume that t and ε satisfy

$$\varepsilon\left\|u\right\|_\infty>\frac{\beta\sqrt{t}}{4}=\frac{R}{4}, \tag{27}$$

then we set $\tau^\varepsilon=\inf\{\,t\geq 0\,|\,|X_t-X_0|\geq R/\varepsilon,\ X_0=x/\varepsilon\,\}$. It follows from the renormalization property of X^ε and a standard estimate on the exponential martingales which appears in the Girsanov theorem that

$$\mathbb{P}_x^\varepsilon\left[\,t>\tau\,\right]\leq\mathbb{P}_{x/\varepsilon}\left[\,t/\varepsilon^2>\tau^\varepsilon\,\right]\leq e^{C_3t/\varepsilon^2}\widehat{\mathbb{P}}_{x/\varepsilon}\left[\,t/\varepsilon^2>\tau^\varepsilon\,\right]^{1/2},$$

where C_3 depends only on the constant of uniform ellipticity of a and the bounds of b, and where $\widehat{\mathbb{P}}_{x/\varepsilon}$ is the distribution of the process generated by $\frac{1}{2}a_{i,j}\frac{\partial^2}{\partial x_i\partial x_j}$. This process is in fact a martingale, and as previously, $\widehat{\mathbb{P}}_{x/\varepsilon}\left[\,t/\varepsilon^2>\tau^\varepsilon\,\right]\leq C_4t/R^2=C_4/\beta$ for some constant C_4 which depends only on the bounds of a. On the other hand, with (27), $t/\varepsilon^2\leq 4\left\|u\right\|_\infty/\beta$. Then, if β is large enough, $\mathbb{P}_x\left[\,t>\tau\,\right]$ may be chosen smaller than a constant not depending on x, t and ε.

Lemma 2 is then proved by combining all these results. □

Proof of Proposition 6. Let us remark the following fact: Let V be an adapted process and M be a square-integrable martingale. We assume that there exists some constant C such that $\mathbb{E}\left[\,|V_t-V_s|\,\right]\leq C|t-s|$ and $\mathbb{E}\left[\,\langle M\rangle_t-\langle M\rangle_s\,\right]\leq C|t-s|$ for any $0\leq s\leq t\leq T$. Then for any $q\geq 2$,

$$\begin{aligned}\mathbb{E}\left[\left|\int_s^t(V_r-V_s)\,\mathrm{d}M_r\right|^{q/2}\right]&\leq\mathbb{E}\left[\left(\sup_{r\in[s,t]}|V_r-V_s|^2(\langle M\rangle_t-\langle M\rangle_s)\right)^{q/4}\right]\\&\leq\frac{1}{2}\mathbb{E}\left[\sup_{r\in[s,t]}|V_r-V_s|^q\right]+\frac{1}{2}\mathbb{E}\left[\,(\langle M\rangle_t-\langle M\rangle_s)^{q/2}\,\right]\leq C|t-s|^{q/2}.\end{aligned} \tag{28}$$

On the other hand,

$$|(V_t-V_s)(M_t-M_s)|^{q/2}\leq\frac{1}{2}|V_t-V_s|^q+\frac{1}{2}|M_t-M_s|^q.$$

Hence, with the Burkholder-Davis-Gundy inequality, there exists some universal constant K depending only on q such that

$$\mathbb{E}\left[\,|(V_t - V_s)(M_t - M_s)|^{q/2}\right] \leq KC|t-s|^{q/2}.$$

Combining this inequality with (28), one gets also that

$$\mathbb{E}\left[\left|\int_s^t (V_r - V_s) \circ \mathrm{d}M_r\right|^{q/2}\right] \leq K'|t-s|^{q/2}, \tag{29}$$

for some constant K' depending only on C and q. With the Burholder-Davis-Gundy inequality again, this result is also true when V is a square-integrable martingale such that $\mathbb{E}\left[\,\langle V\rangle_t - \langle V\rangle_s\,\right] \leq C|t-s|$ for any $0 \leq s \leq t \leq T$, or if M is a not a martingale, but a term of finite variation such that $\mathbb{E}\left[\,|M_t - M_s|\,\right] \leq C|t-s|$ for all $0 \leq s \leq t \leq T$.

Let us set $V_{s,t}^{i,\varepsilon} = \frac{1}{\varepsilon}\int_s^t b_i(X_\tau^\varepsilon/\varepsilon)\,\mathrm{d}\tau$, and

$$M_{s,t}^{i,\varepsilon} = X_t^{i,\varepsilon} - X_s^{i,\varepsilon} + \varepsilon u_i(X_t^\varepsilon/\varepsilon) - \varepsilon u_i(X_s^\varepsilon/\varepsilon).$$

One knows that M^ε is a martingale whose cross-variations are given by (8).

Then

$$\int_s^t (X_\tau^{i,\varepsilon} - X_s^{i,\varepsilon}) \circ \mathrm{d}X_\tau^{j,\varepsilon} = I_1^\varepsilon(s,t) + I_2^\varepsilon(s,t) + I_3^\varepsilon(s,t) + I_4^\varepsilon(s,t)$$

with

$$\begin{aligned}
I_1^\varepsilon(s,t) &= \int_s^t (M_\tau^{i,\varepsilon} - M_s^{i,\varepsilon}) \circ \mathrm{d}M_\tau^{j,X,\varepsilon},\\
I_2^\varepsilon(s,t) &= \int_s^t \varepsilon(u_i(X_\tau^\varepsilon/\varepsilon) - u_i(X_s^\varepsilon/\varepsilon)) \circ \mathrm{d}M_\tau^{j,X,\varepsilon},\\
I_3^\varepsilon(s,t) &= \int_s^t (M_\tau^{i,\varepsilon} - M_s^{i,\varepsilon})\,\mathrm{d}V_\tau^{j,\varepsilon},\\
I_4^\varepsilon(s,t) &= \int_s^t (u_i(X_\tau^\varepsilon/\varepsilon) - u_i(X_s^\varepsilon/\varepsilon))b_i(X_\tau^\varepsilon/\varepsilon)\,\mathrm{d}\tau.
\end{aligned}$$

As a is bounded, there exists some constant C such that $\langle M^{i,X,\varepsilon}\rangle_t - \langle M^{i,X,\varepsilon}\rangle_s \leq C|t-s|$ for any $0 \leq s \leq t \leq T$ and all $i \in \{\,1,\dots,N\,\}$. Similarly, as the coefficients a and b are smooth enough, the gradient ∇u_i of the correctors are bounded, so that $\langle M^{i,\varepsilon}\rangle_t - \langle M^{i,\varepsilon}\rangle_s \leq C|t-s|$ for any $0 \leq s \leq t \leq T$ and all $i \in \{\,1,\dots,N\,\}$. Using this and (25), one gets from Lemma 2 that for any $q > 2$, there exists a constant C such that

$$\sup_{\varepsilon>0}\sup_{x\in\mathbb{R}^N} \mathbb{E}_x\left[\sup_{\tau\in[s,t]} |\varepsilon u_i(X_\tau^\varepsilon/\varepsilon) - \varepsilon u_i(X_s^\varepsilon/\varepsilon)|^q\right] \leq C|t-s|^{q/2}.$$

Similarly, from Lemma 2 again, one has that

$$\sup_{\varepsilon>0}\sup_{x\in\mathbb{R}^N}\mathbb{E}_x\left[\sup_{\tau\in[s,t]}|V_t^{i,\varepsilon}-V_s^{i,\varepsilon}|^q\right]\leq C|t-s|^{q/2}.$$

It follows that applying (28) or (29) to all the terms I_1^ε, I_2^ε, I_3^ε and I_4^ε separately, there exists a constant C such that for any $0\leq s\leq t\leq T$,

$$\sup_{\varepsilon>0}\sup_{x\in\mathbb{R}^N}\mathbb{E}_x\left[\left|\int_s^t(X_\tau^{i,\varepsilon}-X_s^{i,\varepsilon})\circ\mathrm{d}X_\tau^{j,\varepsilon}\right|^{q/2}\right]\leq C|t-s|^{q/2}. \tag{30}$$

In Proposition 5, one could replace the second-order iterated integral $\mathbf{X}^2$ of $\mathbf{X}$ by its anti-symmetric part $((\mathbf{X}^{2,i,j}-\mathbf{X}^{2,j,i})/2)_{i,j=1,\dots,N}$. Thus, (30), Lemma 2 and Proposition 5 imply that

$$\sup_{\varepsilon>0}\sup_{x\in\mathbb{R}^N}\mathbb{E}_x\left[\operatorname{Var}_{q/2}A(X^\varepsilon)+\operatorname{Var}_q X^\varepsilon\right]<+\infty.$$

We have then proved that $(X^\varepsilon,A(X^\varepsilon))_{\varepsilon>0}$ is tight in $\mathbb{V}^p$ for any $p>q>2$. Since the limit in the uniform norm of $(X^\varepsilon,A(X^\varepsilon))$ is known, Proposition 6 is now proved. □

References

[BHL02] R.F. Bass, B. Hambly and T.J. Lyons. Extending the Wong-Zakai theorem to reversible Markov processes. *J. Eur. Math. Soc.*, **4**:3, pp. 237–269, 2002. <doi: `10.1007/s100970200040`>. 22

[BLP78] A. Bensoussan, J.L. Lions and G. Papanicolaou. *Asymptotic Analysis for Periodic Structures*. North-Holland, 1978. 2, 7

[CG98] V.V. Chistyakov and O.E. Galkin. On maps of bounded p-variations with $p>1$. *Positivity*, **2**:1, pp. 19–45, 1998.

[CE00] S. Cohen and A. Estrade. Non-symmetric approximations for manifold-valued semimartingales. *Ann. Inst. H. Poincaré Probab. Statist.*, **36**:1, pp. 45–70, 2000. 3

[CL05] L. Coutin and A. Lejay. Semi-martingales and rough paths theory. *Elect. J. Probab.*, **10**:23, pp. 761–785, 2005 <`www.math.washington.edu/~ejpecp`>. 20, 23

[Faw03] T. Fawcett. *Problems in Stochastic Analysis. Connections between Rough paths and Non-Commutative Harmonic Analysis.* PhD thesis, Oxford University, 2003. 19

[GL97] J. Gaines and T. Lyons. Variable step size control in the numerical solution of stochastic differential equations. *SIAM J. Appl. Math.*, **57**:5. pp. 1455–1484, 1997. 14

[IW89] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland, 2$^{\text{nd}}$ edition, 1989. 3, 4, 18

[KP91] T. G. Kurtz and P. Protter. Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. In *Stochastic analysis*, pp. 331–346. Academic Press, Boston, MA, 1991. 3

[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 Mathematics*, pp. 1–41. Springer-Verlag, 1996. 6

[Kun90] H. Kunita. *Stochastic flows and stochastic differential equations.* Cambridge University Press, 1990. 3

[Lej02] A. Lejay. On the convergence of stochastic integrals driven by processes converging on account of a homogenization property. *Electron. J. Probab.*, **7**:18, pp. 1–18, 2002 <www.math.washington.edu/~ejpecp>. 2, 5, 6, 7, 8, 9, 10, 22, 23

[Lej03] A. Lejay. An introduction to rough paths. *Séminaire de probabilités*, vol. XXXVII, vol. 1832 of *Lecture Notes in Math.*, Springer-Verlag, 2003. 2, 12

[LLQ02] M. Ledoux, T. Lyons and Z. Qian. Lévy area of Wiener processes in Banach spaces. *Ann. Probab.*, **30**:2, pp. 546–578, 2002. 20, 22

[Lyo98] T.J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, **14**:2, pp. 215–310, 1998. 2, 5, 12, 16, 19

[LQ02] T. Lyons and Z. Qian. *System Control and Rough Paths.* Oxford University Press, 2002. 2, 12

[LV03] T. Lyons and N. Victoir. Cubature on Wiener Space. *Proceedings of the Royal Society. Series A. Mathematical and Physical Sciences*, **460**:2041, pp. 169–198, 2004. 19

[McS72] E. J. McShane. Stochastic differential equations and models of random processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory*, pp. 263–294. Univ. California Press, 1972. 3, 19

[NY78] S. Nakao and Y. Yamato. Approximation theorem on stochastic differential equations. In *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*, pp. 283–296, New York, 1978. Wiley. 18

[Pro85] P. Protter. Approximations of solutions of stochastic differential equations driven by semimartingales. *Ann. Probab.*, **13**:3, pp. 716–743, 1985. 18

[RY90] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, 1990.

[Sip93] E.-M. Sipiläinen. *A pathwise view of solutions of stochastic differential equations*. PhD thesis, University of Edinburgh, 1993. 20

[Sus91] H. J. Sussmann. Limits of the Wong-Zakai type with a modified drift term. In *Stochastic analysis*, pp. 475–493. Academic Press, Boston, MA, 1991. 3

[WZ65] E. Wong and M. Zakai. On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist.*, **36**, pp. 1560–1564, 1965. 3