Convergence rate of strong approximations of compound random maps

Emmanuel Gobet, Mohamed Mrad

Convergence rate of strong approximations of compound random maps

Emmanuel Gobet* Mohamed Mrad †

April 27, 2016

Abstract

We consider a random map $x \mapsto F(\omega, x)$ and a random variable $\Theta(\omega)$, and we denote by $F^N(\omega, x)$ and $\Theta^N(\omega)$ their approximations: We establish a strong convergence result, in $\mathbf{L}_p$-norms, of the compound approximation $F^N(\omega, \Theta^N(\omega))$ to the compound variable $F(\omega, \Theta(\omega))$, in terms of the approximations of F and Θ. We then apply this result to the composition of two Stochastic Differential Equations through their initial conditions, which can give a way to solve some Stochastic Partial Differential Equations.

Keywords: strong approximation, Garsia-Rodemich-Rumsey lemma, Euler schemes, stochastic flow.

MSC: 60Hxx, 60Gxx

1 Introduction

Since the seventies, the numerical analysis of stochastic systems is a research field on its own and it has tremendous applications in engineering sciences. This work enriches this vast area by addressing the following natural questions. Consider a continuous random map $x \mapsto F(\omega, x)$ and a random variable $\Theta(\omega)$, and their numerical approximations $F^N(\omega, x)$ and $\Theta^N(\omega)$ for some convergence parameter $N \to +\infty$:

- Under which assumptions does the compound approximation $\omega \mapsto F^N(\omega, \Theta^N(\omega))$ converge in $\mathbf{L}_p$ to the compound map $\omega \mapsto F(\omega, \Theta(\omega))$?
- What is the convergence rate and how does it depend on those related to the approximations F^N to F and Θ^N to Θ?

It is easy to guess that the analysis would be straightforward if (F, F^N) were independent of (Θ, Θ^N), by using a conditioning argument. On the contrary, here

*Centre de Mathématiques Appliquées, Ecole Polytechnique and CNRS, Route de Saclay, 91128 Palaiseau Cedex, France. Email: `emmanuel.gobet@polytechnique.edu`. Corresponding author.

†Laboratoire Analyse, Géométrie et Applications (UMR CNRS 7539), Institut Galilée, Université Paris 13, France. Email: `mrad@cmap.polytechnique.fr`.

our aim is to allow arbitrary dependencies and study the strong convergence in this general setting (convergence in $\mathbf{L}_p$-norms).

Among the applied probability community, there is an increasing interest for strong convergence rates because they constitute the corner stone for designing efficient Multi-Level Monte Carlo methods [Hei01, Gil08] (which significantly speeds-up Crude Monte Carlo methods). In this work, we provide generic results which pave the way for establishing strong convergence rates in complicated situations where results were not available so far. Hopefully, it will open the door for many other interesting issues.

The paper is organized as follows. In Section 2 we state a general convergence result (Theorem 1) estimating the $\mathbf{L}_p$-error $\left\|F^N(\Theta^N) - F(\Theta)\right\|_{\mathbf{L}_p}$, and then we prove it. For this we assume locally uniform approximations on $F^N - F$, and local-Hölder continuity on F. These assumptions being possibly difficult to check in practice, we then give much easier conditions that imply the first ones, using the Garsia-Rodemich-Rumsey lemma with precise quantitative controls. In Section 3, we study the error induced by compound Euler schemes related to Stochastic Differential Equations (SDEs for short), through their initial conditions. This question originates in the resolution of Stochastic PDEs using stochastic flows.

2 $\mathbf{L}_p$-approximation of compound random maps

The section is devoted to stating and proving a general result (Theorem 1). Applications are postponed to Section 3.

2.1 Assumptions

Let $(\mathcal{E}, |.|)$ be a separable Banach space and $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We are given

- a random field, i.e. a $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable mapping $(\omega, x) \in (\Omega, \mathbb{R}^d) \mapsto F(\omega, x) \in \mathcal{E}$, continuous in x for a.e. ω;
- a $\mathcal{F}$-random variable $\Theta : \Omega \mapsto \mathbb{R}^d$.

Let F^N and Θ^N be respectively approximations of F and Θ, where $N \to +\infty$ is a asymptotic parameter; we aim at controlling in $\mathbf{L}_p$ the random variable

$$\omega \in \Omega \mapsto F^N(\omega, \Theta^N(\omega)) - F(\omega, \Theta(\omega)) \in \mathcal{E}$$

which will be denoted by $F^N(\Theta^N) - F(\Theta)$ for the sake of simplicity. For $p > 0$ and for a random variable $Z : \Omega \mapsto \mathcal{E}$ or $\mathbb{R}^d$, we set $\|Z\|_{\mathbf{L}_p} = (\mathbb{E}\,|Z|^p)^{1/p}$: We say that $Z \in \mathbf{L}_p$ if $\|Z\|_{\mathbf{L}_p} < +\infty$. Despite $\|.\|_{\mathbf{L}_p}$ is not a norm for $p < 1$, we refer to it as $\mathbf{L}_p$-norm to simplify the discussion.

(H1) For any $p > 0$, there exist constants $\alpha_p^{(\mathbf{H1})} \in [0, +\infty)$ and $C_p^{(\mathbf{H1})} \in [0, +\infty)$ such that

$$\left\| \sup_{|x|\le\lambda} |F(\cdot, x)| \right\|_{\mathbf{L}_p} \le C_p^{(\mathbf{H1})} \lambda^{\alpha_p^{(\mathbf{H1})}}, \quad \forall \lambda \ge 1. \qquad (\mathbf{H1})$$

(H2) There is a $\kappa \in (0,1]$ such that for any $p > 0$, there exist constants $\alpha_p^{(\mathbf{H2})} \in [0,+\infty)$ and $C_p^{(\mathbf{H2})} \in [0,+\infty)$ such that

$$\left\| \sup_{x \neq y, |x| \leq \lambda, |y| \leq \lambda} \frac{|F(\cdot,y) - F(\cdot,x)|}{|y-x|^\kappa} \right\|_{\mathbf{L}_p} \leq C_p^{(\mathbf{H2})} \lambda^{\alpha_p^{(\mathbf{H2})}}, \quad \forall \lambda \geq 1. \qquad (\mathbf{H2})$$

(H3) For any $p > 0$, there exist a constant $\alpha_p^{(\mathbf{H3})} \in [0,+\infty)$ and a sequence $(C_p^{N,(\mathbf{H3})})_{N \geq 1}$ with $C_p^{N,(\mathbf{H3})} \in [0,+\infty)$ such that

$$\left\| \sup_{|x| \leq \lambda} |F^N(\cdot,x) - F(\cdot,x)| \right\|_{\mathbf{L}_p} \leq C_p^{N,(\mathbf{H3})} \lambda^{\alpha_p^{(\mathbf{H3})}}, \quad \forall \lambda \geq 1, \forall N \geq 1. \qquad (\mathbf{H3})$$

(H4) For any $p > 0$, there exist a constant $C_p^{(\mathbf{H4}\text{-a})} \in [0,+\infty)$ and a sequence $(C_p^{N,(\mathbf{H4}\text{-b})})_{N \geq 1}$ with $C_p^{N,(\mathbf{H4}\text{-b})} \in [0,+\infty)$ such that

$$\|\Theta\|_{\mathbf{L}_p} \vee \left\|\Theta^N\right\|_{\mathbf{L}_p} \leq C_p^{(\mathbf{H4}\text{-a})}, \quad \forall N \geq 1, \qquad (\mathbf{H4}\text{-a})$$

$$\left\|\Theta^N - \Theta\right\|_{\mathbf{L}_p} \leq C_p^{N,(\mathbf{H4}\text{-b})}, \quad \forall N \geq 1. \qquad (\mathbf{H4}\text{-b})$$

These conditions state that all random variables belong to any $\mathbf{L}_p$, with some locally uniform estimates w.r.t. the space dependance; the extension to belonging to some $\mathbf{L}_p$ only would be easy and is left to the reader.

2.2 Main results

Had the random variable Θ be bounded by a finite constant Λ, we would have directly obtained $\left\|F^N(\Theta) - F(\Theta)\right\|_{\mathbf{L}_p} \leq C_p^{N,(\mathbf{H3})} \Lambda^{\alpha_p^{(\mathbf{H3})}}$. The extension to non bounded r.v. Θ is non trivial and is being achieved in Theorem 1 and its proof. The following result (inspired by [KS97, Lemma 2.1]) is instrumental in our analysis. In particular, it enables to justify that the quantities of study are well defined as $\mathbf{L}_p$ random variables.

Proposition 1. *Let E be an Euclidean space. Let G be a $\mathcal{F} \otimes \mathcal{B}(E)$-measurable mapping taking values in $\mathcal{E}$ such that, for any $p > 0$ there exist constants $\alpha_p^{(G)} \in [0,+\infty)$ and $C_p^{(G)} \in [0,+\infty)$ for which*

$$\left\| \sup_{|x| \leq \lambda} |G(\cdot,x)| \right\|_{\mathbf{L}_p} \leq C_p^{(G)} \lambda^{\alpha_p^{(G)}}, \quad \forall \lambda \geq 1. \qquad (1)$$

Let ξ be a random variable taking values in E, with finite $\mathbf{L}_p$ norms for any $p > 0$. Then for any $p > 0$, $\omega \mapsto G(\omega, \xi(\omega)) \in \mathbf{L}_p$ and for any finite conjugate exponents r and s ($r^{-1} + s^{-1} = 1$), we have the estimate

$$\|G(\xi)\|_{\mathbf{L}_p} \leq C_{pr}^{(G)} \left(\zeta(r)\right)^{1/(pr)} 2^{\alpha_{pr}^{(G)}+1/p} \left(1 + \|\xi\|_{\mathbf{L}_{s(\alpha_{pr}^{(G)}p+1)}}^{\alpha_{pr}^{(G)}+1/p}\right)$$

where $\zeta(r) := \sum_{n \geq 1} n^{-r}$ is the Riemann zeta function.

The above result will be extended later in Proposition 2, when the polynomial growth (1) is replaced by an exponential one and when the random variable ξ has exponential moments.

Proof. Using twice Hölder inequalities, we obtain

$$\begin{aligned}
\mathbb{E}\left(|G(\cdot,\xi)|^p\right) &\leq \sum_{n\geq 1}\mathbb{E}\left(\sup_{|x|\leq n}|G(\cdot,x)|^p 1_{n-1\leq|\xi|<n}\right)\\
&\leq \sum_{n\geq 1}\left(\mathbb{E}\left(\sup_{|x|\leq n}|G(\cdot,x)|^{pr}\right)\right)^{1/r}\mathbb{P}\left(n-1\leq|\xi|<n\right)^{1/s}\\
&\leq [C_{pr}^{(G)}]^p\sum_{n\geq 1}\frac{1}{n}\, n^{\alpha_{pr}^{(G)}p+1}\mathbb{P}\left(n-1\leq|\xi|<n\right)^{1/s}\\
&\leq [C_{pr}^{(G)}]^p\left(\sum_{n\geq 1}\frac{1}{n^r}\right)^{1/r}\left(\sum_{n\geq 1}n^{s(\alpha_{pr}^{(G)}p+1)}\mathbb{P}\left(n\leq|\xi|+1<n+1\right)\right)^{1/s}\\
&\leq [C_{pr}^{(G)}]^p\left(\zeta(r)\right)^{1/r}\left(\mathbb{E}\left((|\xi|+1)^{s(\alpha_{pr}^{(G)}p+1)}\right)\right)^{1/s}.
\end{aligned}$$

Therefore, $\|G(\xi)\|_{\mathbf{L}_p}\leq C_{pr}^{(G)}\left(\zeta(r)\right)^{1/(pr)}\left(1+\|\xi\|_{\mathbf{L}_{s(\alpha_{pr}^{(G)}p+1)}}\right)^{\alpha_{pr}^{(G)}+1/p}$ where we have used the Minkowsky inequality. We complete our statement by using

$$(a+b)^\gamma\leq 2^{(\gamma-1)_+}\left(a^\gamma+b^\gamma\right)\leq 2^\gamma\left(a^\gamma+b^\gamma\right)\qquad(2)$$

for any non-negative a,b,γ. □

As a direct consequence of the above result, we deduce that $F(\Theta)$ is any $\mathbf{L}_p$ (owing to (**H1**) and (**H4**-a)). Moreover we can also apply it to $G=F^N$ and $\xi=\Theta^N$ in view of (**H4**-a) and since (1) is satisfied (owing to (**H1**) and (**H3**)): Thus, $F^N(\Theta^N)$ also belongs to any $\mathbf{L}_p$.

Our main result below states an error estimate on the approximation of $F(\Theta)$ by $F^N(\Theta^N)$, as a function of N, through the sequences $(C_{\cdot}^{N,(\mathbf{H3})})_{N\geq 1}$ and $(C_{\cdot}^{N,(\mathbf{H4}\text{-b})})_{N\geq 1}$.

Theorem 1. *Assume* (**H1**)-(**H2**)-(**H3**)-(**H4**-a)-(**H4**-b). *Then for any* $p>0$ *and any* $p_2>p$*, there is a constant* $c_{(3)}$ *independent on* N *such that*

$$\left\|F^N(\Theta^N)-F(\Theta)\right\|_{\mathbf{L}_p}\leq c_{(3)}\left(C_{2p}^{N,(\mathbf{H3})}+[C_{\kappa p_2}^{N,(\mathbf{H4}\text{-b})}]^\kappa\right),\quad\forall N\geq 1.\qquad(3)$$

Quite intuitively, the global approximation error inherits from that on F and that on Θ modified by the local Hölder regularity of $x\mapsto F(\omega,x)$.

Proof. Write $F^N(\Theta^N)-F(\Theta)=\left[F^N(\Theta^N)-F(\Theta^N)\right]+\left[F(\Theta^N)-F(\Theta)\right]$. First, a direct application of Proposition 1 (for $r=s=2$) with (**H3**) and (**H4**-a) yields

$$\begin{aligned}
\left\|F^N(\Theta^N)-F(\Theta^N)\right\|_{\mathbf{L}_p} &\leq C_{2p}^{N,(\mathbf{H3})}\left(\zeta(2)\right)^{1/(2p)}2^{\alpha_{2p}^{(\mathbf{H3})}+1/p}\left(1+\left\|\Theta^N\right\|_{\mathbf{L}_{2(\alpha_{2p}^{(\mathbf{H3})}p+1)}}^{\alpha_{2p}^{(\mathbf{H3})}+1/p}\right)\\
&\leq C_{2p}^{N,(\mathbf{H3})}\left(\zeta(2)\right)^{1/(2p)}2^{\alpha_{2p}^{(\mathbf{H3})}+1/p}\left(1+[C_{2(\alpha_{2p}^{(\mathbf{H3})}p+1)}^{(\mathbf{H4}\text{-a})}]^{\alpha_{2p}^{(\mathbf{H3})}+1/p}\right).
\end{aligned}$$

Consider now the second term $F(\Theta^N)-F(\Theta)$: Set

$$H_\kappa(\omega,\lambda):=\sup_{x\neq y,|x|\leq\lambda,|y|\leq\lambda}\frac{|F(\omega,y)-F(\omega,x)|}{|y-x|^\kappa}$$

and write $|F(\Theta^N)-F(\Theta)| \le H_\kappa(|\Theta^N|\vee|\Theta|)|\Theta^N-\Theta|^\kappa$. Then the Hölder inequality with p-conjugate numbers (p_1,p_2) (i.e. $p_1^{-1}+p_2^{-1}=p^{-1}$) gives

$$\left\|F(\Theta^N)-F(\Theta)\right\|_{\mathbf{L}_p} \le \left\|H_\kappa(|\Theta^N|\vee|\Theta|)\right\|_{\mathbf{L}_{p_1}} \left\|\Theta^N-\Theta\right\|^\kappa_{\mathbf{L}_{\kappa p_2}}.$$

The first factor is upper bound using Proposition 1 (for $r=s=2$) with (**H2**) and (**H4**-b), it readily leads to

$$\begin{aligned}
&\left\|F(\Theta^N)-F(\Theta)\right\|_{\mathbf{L}_p}\\
&\le \left\|H_\kappa(|\Theta^N|\vee|\Theta|)\right\|_{\mathbf{L}_{p_1}} \left\|\Theta^N-\Theta\right\|^\kappa_{\mathbf{L}_{\kappa p_2}}\\
&\le C_{2p_1}^{(\mathbf{H2})}\,(\zeta(2))^{1/(2p_1)}\,2^{\alpha_{2p_1}^{(\mathbf{H2})}+1/p_1}\left(1+\left\||\Theta^N|\vee|\Theta|\right\|^{\alpha_{2p_1}^{(\mathbf{H2})}+1/p_1}_{\mathbf{L}_{2(\alpha_{2p_1}^{(\mathbf{H2})}p_1+1)}}\right)[C_{\kappa p_2}^{N,(\mathbf{H4}\text{-b})}]^\kappa\\
&\le C_{2p_1}^{(\mathbf{H2})}\,(\zeta(2))^{1/(2p_1)}\,2^{\alpha_{2p_1}^{(\mathbf{H2})}+1/p_1}\left(1+[2C^{(\mathbf{H4}\text{-a})}_{2(\alpha_{2p_1}^{(\mathbf{H2})}p_1+1)}]^{\alpha_{2p_1}^{(\mathbf{H2})}+1/p_1}\right)[C_{\kappa p_2}^{N,(\mathbf{H4}\text{-b})}]^\kappa.
\end{aligned}$$

We are done. □

A variant of the main result: Instead of (**H3**) above, in some situations, especially when the variable x plays the role of time variable, computing the error approximation $(F^N(.,x)-F(.,x))$ leads to the following form of estimates.

(H3') For any $p>0$, there exist a constant $\alpha_p^{(\mathbf{H3'})}\in[0,+\infty)$ and a sequence $(C_p^{N,(\mathbf{H3'})})_{N\ge1}$ with $C_p^{N,(\mathbf{H3'})}\in[0,+\infty)$ such that

$$\left\|\sup_{|x|\le\lambda}|F^N(\cdot,x)-F(\cdot,x)|\right\|_{\mathbf{L}_p} \le C_p^{N,(\mathbf{H3'})}e^{\alpha_p^{(\mathbf{H3'})}\lambda},\quad \forall\lambda\ge1,\forall N\ge1. \qquad (\mathbf{H3'})$$

As the reader may guess, the exponential term w.r.t. time comes from an application of Gronwall's lemma in the derivation of error estimates. Note that, under the assumption that Θ admits finite exponential moments, this exponential factor does not enable us to deduce that $(F^N(.,x)-F(.,x))|_{x=\Theta}$ belongs to any $\mathbf{L}_p$, with good $\mathbf{L}_p$-estimates, by applying Proposition 1 since the latter requires a polynomial growth in the stochastic argument. But actually, we can modify Proposition 1 by taking advantage of the finite exponential moments of Θ. This is the next statement.

Proposition 2. *Let E be an Euclidean space and G be a $\mathcal{F}\otimes\mathcal{B}(E)$-measurable mapping taking values in $\mathcal{E}$ such that, for any $p>0$, there exist constants $\alpha_p^{(G)}\in[0,+\infty)$ and $C_p^{(G)}\in[0,+\infty)$ for which*

$$\left\|\sup_{|x|\le\lambda}|G(\cdot,x)|\right\|_{\mathbf{L}_p} \le C_p^{(G)}\exp(\alpha_p^{(G)}\lambda),\quad\forall\lambda\ge0.$$

Then, for any E-valued random variable ξ with exponential moments, $\omega\mapsto G(\omega,\xi(\omega))\in\mathbf{L}_p$ for any p, and for any $\rho>0$ and any finite conjugate exponents r and s, we have

$$\|G(\xi)\|_{\mathbf{L}_p} \le C_{pr}^{(G)}\,(e^{\rho r}-1)^{-1/(pr)}\left(\mathbb{E}\left(e^{(p\alpha_{pr}^{(G)}+\rho)s|\xi|}\right)\right)^{1/(ps)}.$$

Combine (**H3'**) with the above (for $\rho = 1$, $r = s = 2$), this proves that

$$\left\|F^N(.,\Theta^N) - F(.,\Theta^N))\right\|_{\mathbf{L}_p} \le C_{2p}^{N,(\mathbf{H3'})} \left(e^2-1\right)^{-1/(2p)} \left(\mathbb{E}\left(e^{2(p\alpha_{2p}^{(\mathbf{H3'})}+1)|\Theta^N|}\right)\right)^{1/(2p)}.$$

Since we assume that Θ and Θ^N admit finite exponential moments, we can adapt the arguments in the proof of Theorem 1, and finally, we demonstrate the following result.

Theorem 2. *Assume, in addition to* (**H1**)-(**H2**)-(**H3'**)-(**H4**-a)-(**H4**-b)*, that the random variables* Θ, Θ^N *admit finite exponential moments (uniformly in* N*). Then for any* $p > 0$ *and any* $p_2 > p$*, there is a constant* $c_{(4)}$ *independent on* N *such that*

$$\left\|F^N(\Theta^N) - F(\Theta)\right\|_{\mathbf{L}_p} \le c_{(4)} \left(C_{2p}^{N,(\mathbf{H3'})} + [C_{\kappa p_2}^{N,(\mathbf{H4}\text{-b})}]^{\kappa}\right), \quad \forall N \ge 1. \tag{4}$$

This result can be applied directly to diffusion process in diffusion time $Z_t = X_{|Y_t|}$ (see for instance [AZ01]) and study its strong approximation.

Proof of Proposition 2. We adjust the proof of Proposition 1 to the current exponential controls. Briefly, we have

$$\begin{aligned}
\mathbb{E}\left(|G(\cdot,\xi)|^p\right) &\le [C_{pr}^{(G)}]^p \sum_{n\ge1} e^{-\rho n}\ e^{p\alpha_{pr}^{(G)} n} e^{\rho n}\ \mathbb{P}\left(n-1 \le |\xi| < n\right)^{1/s} \\
&\le [C_{pr}^{(G)}]^p \left(\sum_{n\ge1} e^{-\rho r n}\right)^{1/r} \left(\sum_{n\ge1} e^{(p\alpha_{pr}^{(G)}+\rho)sn}\mathbb{P}\left(n-1 \le |\xi| < n\right)\right)^{1/s} \\
&\le [C_{pr}^{(G)}]^p \left(\frac{e^{-\rho r}}{1-e^{-\rho r}}\right)^{1/r} \left(\mathbb{E}\left(e^{(\alpha_{pr}^{(G)}+\rho)s|\xi|}\right)\right)^{1/s}.
\end{aligned}$$

□

2.3 Simplified assumptions

In some situations, checking the assumptions (**H1-H2-H3**) may be difficult since we evaluate the $\mathbf{L}_p$-norms of a maximum. When x is a time variable, we may rely on Doob inequalities and other martingale estimates to achieve this. In other situations, it becomes much more complicated. One can apply the general Kolmogorov continuity criterion for random fields [Kun97, Theorem 1.4.1 p.31], but it does not yield the quantitative estimates we are looking for, in particular regarding the polynomial growth factor in (**H1-H2-H3**). Alternatively, here we use the Garsia-Rodemich-Rumsey lemma [GRR70] (see for instance [Nua06, p.353–354]) which gives refinement compared to the Kolmogorov criterion. This approach has been extensively developed in [BY82] for studying regularity of local times of continuous martingales w.r.t. the space variable.

Lemma 1 (Garsia-Rodemich-Rumsey lemma, control of modulus of continuity)**.** *Let* ρ*,* $\Psi : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ *be continuous and strictly increasing functions vanishing at zero and such that* $\lim_{t\to+\infty}\Psi(t) = +\infty$*. Suppose that* $\phi : \mathbb{R}^d \longrightarrow \mathcal{E}$ *is a continuous* *function with values on the separable Banach space $(\mathcal{E}, |.|)$. Denote by B_r the open ball in $\mathbb{R}^d$ centered at 0 with radius r. Then, provided*

$$\Gamma = \int_{\mathrm{B}_r}\int_{\mathrm{B}_r} \Psi\Big(\frac{|\phi(x)-\phi(y)|}{\rho(|x-y|)}\Big)\mathrm{d}x\ \mathrm{d}y < +\infty \tag{5}$$

it holds, for all $x, y \in \mathrm{B}_r$,

$$|\phi(x)-\phi(y)| \le 8\int_0^{2|x-y|} \Psi^{-1}\Big(\frac{4^{d+1}\Gamma}{\lambda_d u^{2d}}\Big)\rho(\mathrm{d}u) \tag{6}$$

where λ_d is a universal constant depending only on d.

We now aim at proving the following result, which allows to go from pointwise estimates to locally uniform estimates, by assuming Hölder regularity in $\mathbf{L}_p$. It will help to check (**H2**) using much easier conditions.

Theorem 3. *Let $p > d$. Assume that G is $\mathcal{F}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable mapping $(\omega, x) \in (\Omega, \mathbb{R}^d) \mapsto G(\omega, x) \in \mathcal{E}$, continuous in x for a.e. ω. Assume that $G(x)$ is in $\mathbf{L}_p$ for any x and that there exist constants $C^{(G)} \in [0, +\infty)$, $\beta^{(G)} \in (d/p, 1]$ and $\tau^{(G)} \in [0, +\infty)$ such that*

$$\|G(x)-G(y)\|_{\mathbf{L}_p} \le C^{(G)}|x-y|^{\beta^{(G)}}(1+|x|+|y|)^{\tau^{(G)}}, \quad \forall (x,y) \in \mathbb{R}^d\times\mathbb{R}^d.$$

Then, for any $\beta \in (0, \beta^{(G)} - d/p)$, we have

$$\Bigg\|\sup_{x\neq y, |x|\le\lambda, |y|\le\lambda} \frac{|G(y)-G(x)|}{|y-x|^{\beta}}\Bigg\|_{\mathbf{L}_p} \le c_{(7)}C^{(G)}\lambda^{\tau^{(G)}+\beta^{(G)}-\beta}, \quad \forall\lambda \ge 1, \tag{7}$$

where $c_{(7)}$ is a constant depending only on $d, p, \beta, \beta^{(G)}, \tau^{(G)}$.

A similar result is proved in [RY99, Theorem 2.1, p.26] using the Kolmogorov criterion, with x and y in a compact set, i.e. with $\tau^{(G)} = 0$; the quoted result is not sufficient for our study.

Proof. Since $x \mapsto G(x)$ is a.s. continuous, we can apply Lemma 1 by taking $\Psi(t) := t^p$ and $\rho(u) := u^{\gamma_2}$ with $\gamma_2 := \beta + 2d/p$: Defining Γ as in (5) with G instead of ϕ, we obtain

$$\begin{aligned}
\mathbb{E}(\Gamma) &= \int_{\mathrm{B}_r}\int_{\mathrm{B}_r} \mathbb{E}\left(\frac{|G(x)-G(y)|^p}{|x-y|^{p\gamma_2}}\right)\mathrm{d}x\ \mathrm{d}y \\
&\le [C^{(G)}]^p(1+2|r|)^{p\tau^{(G)}}\int_{\mathrm{B}_r}\int_{\mathrm{B}_r}|x-y|^{p\beta^{(G)}-p\gamma_2}\ \mathrm{d}x\ \mathrm{d}y \\
&= [C^{(G)}]^p(1+2|r|)^{p\tau^{(G)}} r^{p\beta^{(G)}-p\gamma_2+2d}V_1 \\
&= [C^{(G)}]^p(1+2|r|)^{p\tau^{(G)}} r^{p(\beta^{(G)}-\beta)}V_1
\end{aligned} \tag{8}$$

where $V_1 := \int_{\mathrm{B}_1}\int_{\mathrm{B}_1}|x-y|^{p\beta^{(G)}-p\gamma_2}\ \mathrm{d}x\ \mathrm{d}y$ is a finite integral since $p\beta^{(G)} - p\gamma_2 = p(\beta^{(G)}-\beta) - 2d > -d \Leftrightarrow \beta < \beta^{(G)} - d/p$. This proves that $\mathbb{E}(\Gamma) < +\infty$ thus Γ is finite a.s.

Moreover, a direct computation shows that

$$\int_0^r \Psi^{-1}\Big(\frac{4^{d+1}\Gamma}{\lambda_d u^{2d}}\Big)\rho(\mathrm{d}u) = \int_0^r \Big(\frac{4^{d+1}\Gamma}{\lambda_d u^{2d}}\Big)^{1/p}\gamma_2 u^{\gamma_2 - 1}\mathrm{d}u = \Big(\frac{4^{d+1}\Gamma}{\lambda_d}\Big)^{1/p}\frac{\beta + 2d/p}{\beta} r^{\beta}, \quad r \geq 0.$$

Therefore, from the above and (6) we derive

$$|G(x) - G(y)| \leq 8\Big(\frac{4^{d+1}}{\lambda_d}\Big)^{1/p}\frac{2^{\beta}(\beta + 2d/p)}{\beta}\Gamma^{1/p}|y - x|^{\beta} \tag{9}$$

for any x, y with $|x| \leq r$ and $|y| \leq r$. Owing to (8) this implies

$$\begin{aligned}
&\mathbb{E}\left(\left|\sup_{x\neq y, |x|\leq r, |y|\leq r}\frac{|G(y) - G(x)|}{|y - x|^{\beta}}\right|^p\right) \\
&\leq \left[8\Big(\frac{4^{d+1}}{\lambda_d}\Big)^{1/p}\frac{2^{\beta}(\beta + 2d/p)}{\beta}\right]^p \mathbb{E}(\Gamma) \\
&\leq \left[8\Big(\frac{4^{d+1}}{\lambda_d}\Big)^{1/p}\frac{2^{\beta}(\beta + 2d/p)}{\beta} C^{(G)}(1 + 2|r|)^{\tau^{(G)}} r^{(\beta^{(G)} - \beta)}\right]^p V_1.
\end{aligned}$$

The proof is complete. Observe that in the proof (see inequality (9)) we more precisely show the a.s. Hölder estimate on $\sup_{x\neq y, |x|\leq r, |y|\leq r}\frac{|G(y)-G(x)|}{|y-x|^{\beta}}$: This is interesting on its own. □

As a consequence, we obtain the following result that may serve to easily check (**H1**).

Corollary 1. *Let consider the assumptions and notations of Theorem 3. Then we have*

$$\left\|\sup_{|x|\leq\lambda}|G(x)|\right\|_{\mathbf{L}_p} \leq c_{(10)}\lambda^{\tau^{(G)} + \beta^{(G)}}, \quad \forall \lambda \geq 1, \tag{10}$$

where $c_{(10)} := \|G(0)\|_{\mathbf{L}_p} + c_{(7)}C^{(G)}$ *where* $c_{(7)}$ *is defined in Theorem 3 with* $\beta = (\beta^{(G)} - d/p)/2$*. In particular, the constant* $c_{(7)}$ *depends only on* $d, p, \beta^{(G)}, \tau^{(G)}$*.*

Proof. Using easy inequalities and applying (7) with $y = 0$ and β as announced, it readily follows

$$\begin{aligned}
\left\|\sup_{|x|\leq\lambda}|G(x)|\right\|_{\mathbf{L}_p} &\leq \|G(0)\|_{\mathbf{L}_p} + \left\|\sup_{0<|x|\leq\lambda}\frac{|G(x) - G(0)|}{|x|^{\beta}}\right\|_{\mathbf{L}_p}\lambda^{\beta} \\
&\leq \|G(0)\|_{\mathbf{L}_p} + c_{(7)}C^{(G)}\lambda^{\tau^{(G)} + \beta^{(G)}}.
\end{aligned}$$

Since $\lambda \geq 1$ and $\tau^{(G)} + \beta^{(G)} \geq 0$, the proof is complete. □

3 Application to compound Euler schemes

In this section, let T be a positive and finite time horizon and let us consider a standard filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$ supporting two q-dimensional standard Brownian motions $W = (W^1, \ldots, W^q)$ and $B = (B^1, \ldots, B^q)$ on $[0, T]$. We consider two $\mathbb{R}^d$-valued stochastic processes X and Y, solutions of the following stochastic differential equations (SDE for short)

$$\text{SDE}(\mu,\sigma,W)\text{: } \mathrm{d}X_t(x) = \mu(t, X_t(x))\mathrm{d}t + \sum_{i=1}^{q} \sigma_i(t, X_t(x))\mathrm{d}W_t^i, \quad X_0(x) = x, \tag{11}$$

$$\text{SDE}(b,\gamma,B)\text{: } \mathrm{d}Y_t(y) = b(t, Y_t(y))\mathrm{d}t + \sum_{i=1}^{q} \gamma_i(t, Y_t(y))\mathrm{d}B_t^i, \quad Y_0(y) = y, \tag{12}$$

where $\mu, b, \sigma_i, \gamma_i$ are functions from $[0,T] \times \mathbb{R}^d$ into $\mathbb{R}^d$, globally Lipschitz in space to ensure the existence of a unique strong solution. Depending on the potential applications, we may require that B and W are the same, or different. Denote by $X_T^N(x)$ (resp. $Y_T^N(y)$) the Euler scheme with time step T/N of $X_T(x)$ (resp. $Y_T(y)$): Using previous results, we aim at establishing a new convergence result of the compound scheme $X_t^N(Y_t^N(y))$ to the compound SDE $X_t(Y_t(y))$ as N goes to infinity, under the form

$$\left\| X_t^N(Y_t^N(y)) - X_t(Y_t(y)) \right\|_{\mathbf{L}_p} = O(N^{-1/2})$$

for any $p > 0$. For a rigorous statement under precise assumptions, see Theorem 4. This approximation issue, interesting on its own, is actually motivated by other potential applications we briefly expose and that will be subject of future and deeper investigations.

Relation with approximation of stochastic partial differential equations (SPDEs). This work constitutes a first step in a subject that until now has not been addressed to our knowledge, that is to approach solutions of SPDEs by approximating compound SDEs. Relating compound SDEs to SPDEs is, in a sense, obvious since it is sufficient to apply the Itô-Ventzel formula [Kun97, Section 3.3] (under good regularity assumptions on (μ, σ)) to the compound process $U(t,y) := X_t(Y_t(y))$ to show that $(t,y) \mapsto X_t(Y_t(y))$ solves the second order SPDE, with stochastic coefficients, given by (to simplify we take $d = q = 1$ and $W = B$)

$$\begin{aligned}
\mathrm{d}U(t,y) =& \bigg(\partial_y U(t,y) \frac{b(t, Y_t(y))}{\partial_y Y(t,y)} + \frac{1}{2}\Big(\partial_y^2 U(t,y) - \partial_y U(t,y) \frac{\partial_y^2 Y_t(y)}{\partial_y Y_t(y)} \Big) \frac{\gamma^2(t, Y_t(y)}{(\partial_y Y_t(y))^2} \\
&+ \mu(t, U(t,y)) + \partial_x \sigma(t, U(t,y))\gamma(t, Y_t(y)) \bigg)\mathrm{d}t \\
&+ \bigg(\partial_y U(t,y) \frac{\gamma(t, Y_t(y))}{\partial_y Y(t,y)} + \sigma(t, U(t,y)) \bigg)\mathrm{d}W_t.
\end{aligned}$$

In the reverse direction, i.e. starting from a SPDE, it is more delicate to establish a link with SDEs. But in the recent work [EM13] based on the theory of stochastic flows, El Karoui and Mrad have established a direct connection between a certain utility SPDE and two SDEs. Indeed, being concerned with progressive stochastic utilities $(U(t,x) : t \geq 0, x \in \mathbb{R}^d)$ (a.k.a. Forward Utilities or performance processes, see [MZ10]), the authors show that U (under some regularity assumptions) are inevitably solution of a second order fully nonlinear SPDE. Moreover the marginal utility $\partial_x U$ is characterized by two SDEs X and Y under the form $U_x = X(Y^{-1})$.

Here Y^{-1} is the inverse flow of Y and can be interpreted as another SDE, see the above reference for details. The current work paves the way to the derivation of convergent approximation of SPDEs of this form.

3.1 Hypotheses

We first study approximations on X and for this, we state related assumptions on the $\mathbb{R}^d$-valued drift coefficient $\mu = \{\mu(t,x); t \in [0,T], x \in \mathbb{R}^d\}$ and the $\mathbb{R}^d \times \mathbb{R}^q$-valued diffusion coefficient $\sigma = \{\sigma_i(t,x); 1 \le i \le q, t \in [0,T], x \in \mathbb{R}^d\}$ which we suppose to be regular enough in time and space. When we will discuss on approximation of $X(Y)$, similar assumptions will be made on the coefficients b and γ_i of Equation (12) for Y.

(HP1) The coefficients μ and σ are Lipschitz continuous in space uniformly in time. More precisely, there exists a finite constant C^X such that for any $t \in [0,T]$ and $x, y \in \mathbb{R}^d$

$$\begin{cases} |\mu(t,x) - \mu(t,y)| \le C^X |x-y|, & |\mu(t,0)| \le C^X, \\ |\sigma(t,x) - \sigma(t,y)| \le C^X |x-y|, & |\sigma(t,0]| \le C^X. \end{cases} \tag{\textbf{HP1}}$$

(HP2) μ and σ are continuously space-differentiable functions such that their derivatives $\nabla_x \mu := \{\nabla_x \mu(t,x); t \in [0,T], x \in \mathbb{R}^d\}$ and $\nabla_x \sigma = \{\nabla_x \sigma_i(t,x); 1 \le i \le q, t \in [0,T], x \in \mathbb{R}^d\}$ are δ-Hölder for a certain exponent $\delta \in (0,1]$. Namely, there exists a finite constant $C^{X,\nabla}$ such that for any $t \in [0,T]$ and $x, y \in \mathbb{R}^d$

$$\begin{cases} |\nabla_x \mu(t,x) - \nabla_x \mu(t,y)| \le C^{X,\nabla} |x-y|^\delta, & |\nabla_x \mu(t,x)| \le C^{X,\nabla}, \\ |\nabla_x \sigma(t,x) - \nabla_x \sigma(t,y)| \le C^{X,\nabla} |x-y|^\delta, & |\nabla_x \sigma(t,x)| \le C^{X,\nabla}. \end{cases} \tag{\textbf{HP2}}$$

(HP3) μ and σ are Hölder continuous in time, locally in space, i.e. there exists an exponent $\alpha \in (0,1]$ and a finite constant C^X, such that for any $x \in \mathbb{R}^d$ and $s, t \in [0,T]$

$$|\mu(t,x) - \mu(s,x)| + |\sigma(t,x) - \sigma(s,x)| \le C^X (1 + |x|)|t-s|^\alpha. \tag{\textbf{HP3}}$$

(HP4) μ and σ are continuously space-differentiable functions such that their derivatives are Hölder continuous in time, locally in space, i.e. there exists an exponent $\alpha \in (0,1]$ and a finite constant $C^{X,\nabla}$, such that for any $x \in \mathbb{R}^d$ and $s, t \in [0,T]$

$$|\nabla_x \mu(t,x) - \nabla_x \mu(s,x)| + |\nabla_x \sigma(t,x) - \nabla_x \sigma(s,x)| \le C^{X,\nabla}(1+|x|)|t-s|^\alpha. \tag{\textbf{HP4}}$$

Denoting in the same way the constants of (**HP1**) and (**HP3**) (resp. (**HP2**) and (**HP4**)) by C^X (resp. $C^{X,\nabla}$) is made for the sake of simplicity.
Assumption (**HP1**) ensures the existence of a strong continuous solution to the SDE(μ, σ, W), which is adapted to the natural filtration of W completed by the $\mathbb{P}$-null sets: (**HP1**) plays a crucial role to establish a $\mathbf{L}_p$-estimates. It is also well-known [Kun97, Theorem 4.5.1] that the map $(t,x) \mapsto X_t(\omega, x)$ has a modification

which is continuous a.s., we shall systematically work with this modification from now on. Assumption (**HP2**) is a sufficient condition (see [Kun97, Theorem 3.3.3]) under which the above map is C^1 in x. Assumptions (**HP3**) and (**HP4**) enable us, essentially, to establish convergence results of the Euler discretization scheme within the paper setting.

3.2 Compound Euler schemes: Main result

Under (**HP1**), let us consider the strong solution to (11): its Euler scheme with $N \geq 1$ discretization times and step-size $\frac{T}{N}$ is defined as usually as follows.

- Set $X_0^N(x) = x$.
- For $k = 0, \dots, N-1$ and $t \in (k\frac{T}{N}, (k+1)\frac{T}{N}]$, set

$$X_t^N(x) = X^N_{k\frac{T}{N}}(x) + \mu(k\frac{T}{N}, X^N_{k\frac{T}{N}}(x))(t - \frac{kT}{n}) + \sum_{i=1}^q \sigma_i(k\frac{T}{N}, X^N_{k\frac{T}{N}}(x))(W_t^i - W^i_{k\frac{T}{N}}).$$

It can be equivalently written as a continuous Itô process: Denoting by $\tau_t := [\frac{Nt}{T}]\frac{T}{N}$ the last discretization-time before t, we have

$$X_t^N(x) = x + \int_0^t \mu(\tau_s, X^N_{\tau_s}(x))\mathrm{d}s + \sum_{i=1}^q \int_0^t \sigma_i(\tau_s, X^N_{\tau_s}(x))\mathrm{d}W_s^i. \tag{13}$$

Similarly, assume that b and γ fulfills (**HP1**), so that the strong solution Y to (12) is well defined, together with its Euler scheme Y^N.

The section is devoted to establish the following main result.

Theorem 4. *Assume that μ and σ satisfy Assumptions* (**HP1**), (**HP2**), (**HP3**) *and* (**HP4**) *(which α-parameter is denoted by α^X) and that b and γ satisfy Assumptions* (**HP1**) *and* (**HP3**) *(which α-parameter is denoted by α^Y).*

Then the compound Euler scheme $X^N_\cdot(Y^N_\cdot)$ converges to $X_\cdot(Y_\cdot)$ in any $\mathbf{L}_p$-norm, at the order $\beta := \min(\alpha^X, \alpha^Y, \frac{1}{2})$ w.r.t. N: For any $p > 0$, there is a finite constant C_p such that for any $s, t \in [0, T]$

$$\left\| X_t^N(Y_s^N) - X_t(Y_s) \right\|_{\mathbf{L}_p} \leq C_p N^{-\beta}, \quad \forall N \geq 1.$$

The rest of this section is devoted to its long proof, which requires intermediate estimates on the SDE and its Euler scheme, some of them being completely new (Theorems 5 and 8).

3.3 Proof of Theorem 4

In this subsection, we will make use of different constants that may depend on the integer p of $\mathbf{L}_p$-norm, on the dimensions d and q, on the time horizon T and on the constants from the assumptions: These constants will be called *generic constant* and will be denoted by the same notation C_p even if their values change from line to line. They will not depend on N.

We denote by C_p^{BDG} the constant of the upper Burkholder-Davis-Gundy inequality with $\mathbf{L}_p$-norm (see the right-hand side of [RY99, Theorem 4.1, p.160]).

3.3.1 SDE: differentiability, local and uniform estimates

To analyze the approximation of the compound SDE $X(Y)$, precise estimates on the maps $x \mapsto X_t(\omega, x)$ are needed: Such random fields are also called stochastic flows and are the main subject of Kunita's book [Kun97]. As aforementioned, under (**HP1**), the map $(t, x) \mapsto X_t(\omega, x)$ has a continuous modification we are working with. The additional space regularity is connected to the regularity of the coefficients (μ, σ), owing to (**HP2**), which can be described as follows.

Proposition 3 ([Kun97, Theorem 4.6.5]). *Under Assumptions* (**HP1**) *and* (**HP2**), *the strong solution $X_t(x)$ to* (11) *is continuously differentiable in space and its derivative denoted by $\nabla X_t(x)$ is locally ε-Hölder*[1] *for any $\varepsilon < \delta$. Furthermore, it is a semimartingale solution of a linear equation, with bounded stochastic parameters $(\nabla_x \mu(t, X_t(x)), \nabla_x \sigma(t, X_t(x)))$ given by*

$$\nabla X_0(x) = \mathrm{Id},$$
$$\mathrm{d}\nabla X_t(x) = \nabla_x \mu(t, X_t(x)) \nabla X_t(x) \mathrm{d}t + \sum_{i=1}^{q} \nabla_x \sigma_i(t, X_t(x)) \nabla X_t(x) \mathrm{d}W_t^i. \qquad (14)$$

We now proceed to L_p-estimates of $X_t(x)$ and its sensitivity w.r.t. x. We collect several useful results in the following Proposition.

Proposition 4. *Assume* (**HP1**). *For any $p > 0$, there exist generic constants $C_{p,(15)}$ and $C_{p,(16)}$ such that*

$$\|X_t(x)\|_{\mathbf{L}_p} \le C_{p,(15)}(1 + |x|), \qquad (15)$$
$$\|X_t(x) - X_t(y)\|_{\mathbf{L}_p} \le C_{p,(16)}|x - y| \qquad (16)$$

for any $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$. In addition under (**HP2**), *for any $p > 0$ there exist generic constants $C_{p,(17)}$ and $C_{p,(18)}$ such that*

$$\|\nabla X_t(x)\|_{\mathbf{L}_p} \le C_{p,(17)}, \qquad (17)$$
$$\|\nabla X_t(x) - \nabla X_t(y)\|_{\mathbf{L}_p} \le C_{p,(18)}|x - y|^{\delta} \qquad (18)$$

for any $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$.

Proof. The proofs of inequalities (15) and (16) are standard, see [Kun97, Lemmas 4.5.3 and 4.5.5]. The uniform estimate (17) is also easy to obtain, in view of (14) and owing to the boundedness of $\nabla_x \mu$ and $\nabla_x \sigma_i$, we leave the details to the reader.

It remains to show (18) under (**HP2**). To alleviate the notation, we provide the proof when $d = q = 1$, the general case being similar. Also, we can focus on the case $p \ge 2$ since we can deduce the result for $p < 2$ using the stability of $\mathbf{L}_p$-norm combined with the result for $p = 2$. First, from (14) write

$$\nabla X_t(x) - \nabla X_t(y) = \int_0^t \nabla_x \mu(s, X_s(x)) \Big(\nabla X_s(x) - \nabla X_s(y) \Big) \mathrm{d}s$$

[1] That is for any compact K of $\mathbb{R}^d$ there exists a finite positive random variable $C(K)$ such that for any $x, y \in K$ we have $|\nabla X_t(x, \omega) - \nabla X_t(y, \omega)| \le C(K, \omega)|x - y|^{\varepsilon}$ a.s., see [Kun97, Chapters 3 and 4] for details .

$$
\begin{aligned}
&+\int_0^t \Big(\nabla_x\mu(s,X_s(s))-\nabla_x\mu(s,X_s(y))\Big)\nabla X_s(y)\mathrm{d}s\\
&+\int_0^t \nabla_x\sigma(s,X_s(x))\Big(\nabla X_s(x)-\nabla X_s(y)\Big)\mathrm{d}W_s\\
&+\int_0^t \Big(\nabla_x\sigma(s,X_s(s))-\nabla_x\sigma(s,X_s(y))\Big)\nabla X_s(y)\mathrm{d}W_s.
\end{aligned}
$$

Take the power p and the expectation, then apply the Burkholder-Davis-Gundy inequality, the Jensen equality ($p \geq 2$) and the Cauchy-Schwarz inequality; it leads to

$$
\begin{aligned}
&\mathbb{E}\left(|\nabla X_t(x)-\nabla X_t(y)|^p\right)\\
&\leq 4^{p-1}t^{p-1}\int_0^t \mathbb{E}\left(|\nabla_x\mu(s,X_s(x))(\nabla X_s(x)-\nabla X_s(y))|^p\right)\mathrm{d}s\\
&+4^{p-1}t^{p-1}\int_0^t \sqrt{\mathbb{E}\left(|\nabla_x\mu(s,X_s(s))-\nabla_x\mu(s,X_s(y))|^{2p}\right)}\sqrt{\mathbb{E}\left(|\nabla X_s(y)|^{2p}\right)}\mathrm{d}s\\
&+4^{p-1}[C_p^{\mathrm{BDG}}]^p t^{p/2-1}\int_0^t \mathbb{E}\left(|\nabla_x\sigma(s,X_s(x))(\nabla X_s(x)-\nabla X_s(y))|^p\right)\mathrm{d}s\\
&+4^{p-1}[C_p^{\mathrm{BDG}}]^p t^{p/2-1}\int_0^t \sqrt{\mathbb{E}\left(|\nabla_x\sigma(s,X_s(s))-\nabla_x\sigma(s,X_s(y))|^{2p}\right)}\sqrt{\mathbb{E}\left(|\nabla X_s(y)|^{2p}\right)}\mathrm{d}s.
\end{aligned}
$$

Now, take advantage of the Assumptions (**HP1**) and (**HP2**), together with the estimates (16) and (17): it readily follows that $\iota(t) := \mathbb{E}\left(|\nabla X_t(x)-\nabla X_t(y)|^p\right)$ solves

$$
\begin{aligned}
\iota(t) \leq\ & 4^{p-1}[C^{X,\nabla}]^p\left(T^{p-1}+[C^{\mathrm{BDG}}]^pT^{p/2-1}\right)\int_0^t \iota(s)\mathrm{d}s\\
&+4^{p-1}[C^{X,\nabla}]^pC^p_{2p,(17)}C^{p\delta}_{2p\delta,(16)}\left(T^p+[C^{\mathrm{BDG}}]^pT^{p/2}\right)|x-y|^{p\delta}.
\end{aligned}
$$

The estimate (18) is then a direct consequence of Gronwall's lemma.

□

Thanks to the results of Section 2, we are now in a position to generalize Proposition 4 by putting the sup over the space variable inside the expectation. This is the following assertion, which is a new result to our knowledge.

Theorem 5. *Assume Assumption* (**HP1**). *For any $p > 0$ and any $\beta \in (0,1)$, there exist generic constants $C_{p,(19)}$ and $C_{p,(20)}$ such that, for any $t \in [0,T]$,*

$$
\left\|\sup_{|x|\leq\lambda}|X_t(x)|\right\|_{\mathbf{L}_p} \leq C_{p,(19)}\lambda, \quad \forall\lambda\geq 1, \tag{19}
$$

$$
\left\|\sup_{x\neq y,|x|\leq\lambda,|y|\leq\lambda}\frac{|X_t(x)-X_t(y)|}{|y-x|^\beta}\right\|_{\mathbf{L}_p} \leq C_{p,(20)}\lambda^{1-\beta}, \quad \forall\lambda\geq 1. \tag{20}
$$

Assume furthermore Assumption (**HP2**). *For any $p > 0$ and any $\beta \in (0,\delta)$, there exist generic constants $C_{p,(21)}$, $C_{p,(22)}$ and $C_{p,(23)}$ such that, for any $t \in [0,T]$,*

$$
\left\|\sup_{|x|\leq\lambda}|\nabla X_t(x)|\right\|_{\mathbf{L}_p} \leq C_{p,(21)}\lambda^\delta, \quad \forall\lambda\geq 1, \tag{21}
$$

$$\left\| \sup_{x\neq y, |x|\leq\lambda, |y|\leq\lambda} \frac{|\nabla X_t(x) - \nabla X_t(y)|}{|y-x|^\beta} \right\|_{\mathbf{L}_p} \leq C_{p,(22)}\lambda^{\delta-\beta}, \quad \forall \lambda \geq 1, \tag{22}$$

$$\left\| \sup_{x\neq y, |x|\leq\lambda, |y|\leq\lambda} \frac{|X_t(x) - X_t(y)|}{|y-x|} \right\|_{\mathbf{L}_p} \leq C_{p,(23)}\lambda^{\delta}, \quad \forall \lambda \geq 1. \tag{23}$$

Proof. Let $\beta \in (0,1)$: We first show (19) and (20) for any $p > d/(1-\beta) > d$. Owing to (16), we can apply Theorem 3 to $G(x) := X_t(x)$ with $\beta^{(G)} = 1 \in (d/p, 1]$ and $\tau^{(G)} = 0$, to conclude that (20) holds with the given index β since $\beta < 1 - d/p \Leftrightarrow p > d/(1-\beta)$. Moreover the application of Corollary 1 provides (19). Now it remains to relax the constraint on p: for $p \leq d/(1-\beta)$, set $\bar{p} = 2d/(1-\beta)$ for which (20) holds and take advantage of the stability property of $\mathbf{L}_p$-norm to write

$$\left\| \sup_{x\neq y, |x|\leq\lambda, |y|\leq\lambda} \frac{|X_t(x) - X_t(y)|}{|y-x|^\beta} \right\|_{\mathbf{L}_p} \leq \left\| \sup_{x\neq y, |x|\leq\lambda, |y|\leq\lambda} \frac{|X_t(x) - X_t(y)|}{|y-x|^\beta} \right\|_{\mathbf{L}_{\bar{p}}} \leq C_{\bar{p},(20)}\lambda^{1-\beta}.$$

The same arguments apply to prove that (19) holds for any $p > 0$.

The justification of (21) and (22) follows the same arguments as above, using (18) instead of (16): then Theorem 3 and Corollary 1 can be applied to $G(x) := \nabla X_t(x)$ with $\beta^{(G)} = \delta$ and $\tau^{(G)} = 0$. We leave the details to the reader.

Last, observe that for any x, y such that $|x| \leq \lambda$ and $|y| \leq \lambda$, we have $|X_t(x) - X_t(y)| \leq \sup_{|z|\leq\lambda} |\nabla X_t(z)| \, |y-x|$: thus, (23) readily follows from (21). □

Observe that the additional smoothness in (**HP2**) enables us to improve (20) (for $\beta < 1$) to (23) (i.e. $\beta = 1$): this improvement will play an important role in the derivation of Theorem 4.

3.3.2 Euler scheme: local and uniform estimates

Still as intermediate steps to prove Theorem 4, we partly generalize the previous results about the SDE to its Euler approximation. Some derivations are more subtle and require details at some places. Recall the definition of Euler scheme in (13).

First, as for the solution of the SDE(μ, σ), some estimates for its approximation scheme are needed. This is the analogue of Proposition 4.

Proposition 5. *Under* (**HP1**)*, for any $p > 0$ there exist generic constants $C_{p,(24)}$ and $C_{p,(25)}$ such that*

$$\left\| X_t^N(x) \right\|_{\mathbf{L}_p} \leq C_{p,(24)}(1 + |x|), \tag{24}$$

$$\left\| X_t^N(x) - X_t^N(y) \right\|_{\mathbf{L}_p} \leq C_{p,(25)}|x-y| \tag{25}$$

for any $(t, x, y) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^d$.

We omit the proof which is quite standard. Following the same arguments than for the SDE case (Theorem 5), we can put the sup over the space variable inside the $\mathbf{L}_p$-norm, it gives the following.

Proposition 6. *Under* (**HP1**)*, the estimates* (19) *and* (20) *where we replace X by X^N hold true, up to changing the generic constants.*

Let us now show the following estimates on local increments, it will be needed for the sequel.

Lemma 2. *Assume Assumption* (**HP1**) *and let* $p > 0$. *Then there exist generic constants* $C_{p,(26)}$ *and* $C_{p,(27)}$ *such that, for any* $x, y \in \mathbb{R}^d$ *and any* $t \in [0,T]$,

$$\left\| \sup_{\tau_t \le u \le t} |X_u^N(x) - X_{\tau_u}^N(x)| \right\|_{\mathbf{L}_p} \le C_{p,(26)} \frac{(1+|x|)}{N^{1/2}}, \tag{26}$$

$$\left\| \sup_{\tau_t \le u \le t} |X_u^N(x) - X_u^N(y) - X_{\tau_u}^N(x) + X_{\tau_u}^N(y)| \right\|_{\mathbf{L}_p} \le C_{p,(27)} \frac{|x-y|}{N^{1/2}}. \tag{27}$$

Proof. Here again, it is enough to prove the estimates for $p \ge 2$, which we assume from now on. Also we take $d = q = 1$ to simplify the exposure. Similarly to the proof of Proposition 4, Burkholder-Davis-Gundy's inequality combined with Jensen's inequality readily leads to

$$\mathbb{E}\left(\sup_{\tau_t \le u \le t} |X_u^N(x) - X_{\tau_u}^N(x)|^p \right) \le 2^{p-1}\Big((t-\tau_t)^{p-1} \int_{\tau_t}^t \mathbb{E}\left(|\mu(\tau_s, X_{\tau_s}^N(x))|^p \right) \mathrm{d}s$$
$$+ (t-\tau_t)^{p/2-1} [C_p^{\mathrm{BDG}}]^p \int_{\tau_t}^t \mathbb{E}\left(|\sigma(\tau_s, X_{\tau_s}^N(x))|^p \right) \mathrm{d}s \Big).$$

Finally, from Assumption (**HP1**), we have $|\mu(t,x)| + |\sigma(t,x)| \le C^X(1+|x|)$ for any $t \in [0,T]$; combined with (24), we deduce

$$\mathbb{E}\left(\sup_{\tau_t \le u \le t} |X_u^N(x) - X_{\tau_u}^N(x)|^p \right) \le 2^{p-1}(C^X)^p \Big((t-\tau_t)^p 2^{p-1}(1 + C_{p,(24)}^p (1+|x|)^p)$$
$$+ [C_p^{\mathrm{BDG}}]^p (t-\tau_t)^{p/2} 2^{p-1}(1 + C_{p,(24)}^p (1+|x|)^p) \Big)$$

which readily leads to the announced estimate (26).

Let us now turn to the second inequality: The same arguments combined with Assumption (**HP1**) and (25) lead to

$$\begin{aligned}
&\mathbb{E}\big(\sup_{\tau_t \le u \le t} |X_u^N(x) - X_u^N(y) - X_{\tau_u}^N(x) + X_{\tau_u}^N(y)|^p \big) \\
&\le 2^{p-1}(t-\tau_t)^{p-1} \int_{\tau_t}^t \mathbb{E}\left(|\mu(\tau_s, X_{\tau_s}^N(x)) - \mu(\tau_s, X_{\tau_s}^N(y))|^p \right) \mathrm{d}s \\
&+ 2^{p-1}[C_p^{\mathrm{BDG}}]^p (t-\tau_t)^{p/2-1} \int_{\tau_t}^t \mathbb{E}\left(|\sigma(\tau_s, X_{\tau_s}^N(x)) - \sigma(\tau_s, X_{\tau_s}^N(y))|^p \right) \mathrm{d}s \\
&\le 2^{p-1}(C^X)^p C_{p,(25)}^p \left((t-\tau_t)^p + [C_p^{\mathrm{BDG}}]^p (t-\tau_t)^{p/2} \right) |x-y|^p,
\end{aligned}$$

which completes the proof. □

Strong convergence (classical result). Since in the Euler scheme dynamics the coefficients μ and σ are computed at the left of each time interval, we need to account for their time regularity in order to derive a sharp convergence result: This is stated through Assumption (**HP3**). The proof of the following result can be found in [BL93, Theorem B.1.4 p. 276].

Theorem 6. *Assume Assumptions* (**HP1**) *and* (**HP3**) *and set* $\beta = \min(\alpha, \frac{1}{2})$. *Then, for any* $p > 0$ *there exists a generic constant* $C_{p,(28)}$ *such that for any* $x \in \mathbb{R}^d$

$$\left\| \sup_{t \le T} |X_t(x) - X_t^N(x)| \right\|_{\mathbf{L}_p} \le C_{p,(28)} \frac{(1+|x|)}{N^\beta}. \tag{28}$$

Furthermore, for any $\gamma < \beta$, *the random variables* $(N^\gamma \sup_{t \le T} |X_t - X_t^N|)_{N \ge 1}$ *converge almost surely to* 0 *as* N *tends to* $+\infty$.

Unfortunately, the classical estimate of Theorem 6 is not sufficient to analyze the error of compound Euler schemes: in view of Theorem 1 and its assumptions (in particular (**H3**)), one should have a sup over $|x| \le \lambda$ inside the $\mathbf{L}_p$-norm. This is the purpose of the next derivations.

Strong convergence (new results). To obtain locally uniform in space convergence results, the supplementary assumptions of regularity in space and time for $\nabla_x \mu$ and $\nabla_x \sigma_i$ (see (**HP2**) and (**HP4**)) are seemingly important. Thus Theorem 6 can be generalized to the following crucial one.

Theorem 7. *Assume* (**HP1**), (**HP2**), (**HP3**), (**HP4**) *and let* $\beta = \min(\alpha, \frac{1}{2})$. *For any* $p > 0$, *there exists a generic constant* $C_{p,(29)}$ *such that*

$$\left\| \sup_{u \le t} |X_u(x) - X_u^N(x) - X_u(y) + X_u^N(y)| \right\|_{\mathbf{L}_p} \le C_{p,(29)} (1+|x|+|y|) \frac{|x-y| + |x-y|^\delta}{N^\beta} \tag{29}$$

for all $x, y \in \mathbb{R}^d$ *and* $t \in [0, T]$.

Similarly to Theorem 5, we can now derive estimates locally uniformly in space.

Theorem 8. *Under Assumptions of Theorem 7, for any* $p > 0$ *there exists a finite generic constant* $C_{p,(30)}$ *such that, for any* $t \in [0, T]$,

$$\left\| \sup_{|x| \le \lambda} |X_t(x) - X_t^N(x)| \right\|_{\mathbf{L}_p} \le \frac{C_{p,(30)}}{N^\beta} \lambda^2, \quad \forall \lambda \ge 1. \tag{30}$$

Proof. We aim at applying Corollary 1 by checking the assumptions of Theorem 3 applied to $G(x) := X_t(x) - X_t^N(x)$. From (29) we have

$$\begin{aligned} \|G(x) - G(y)\|_{\mathbf{L}_p} &\le C_{p,(29)} (1 + |x| + |y|) \frac{|x-y| + |x-y|^\delta}{N^\beta} \\ &\le 2 C_{p,(29)} (1 + |x| + |y|)^{2-\delta} \frac{|x-y|^\delta}{N^\delta} \end{aligned}$$

using $|x-y| + |x-y|^\delta = |x-y|^\delta (1 + |x-y|^{1-\delta}) \le 2|x-y|^\delta (1+|x|+|y|)^{1-\delta}$. Thus, we can take $C^{(G)} = 2C_{p,(29)}/N^\delta$, $\tau^{(G)} = 2 - \delta$ and $\beta^{(G)} = \delta$ provided that $\delta \in (d/p, 1]$, which is true for p large enough. Therefore for such p, the estimate (10) holds true, which is the announced inequality of Theorem 8. The estimate for smaller values of p are automatically satisfied invoking once again the stability of $\mathbf{L}_p$ norms as p decreases. □

Proof of Theorem 7. As in the previous proofs, we argue that it is enough to assume $p \geq 2$. To alleviate the presentation, we additionally assume $d = q = 1$, the derivation in the general case being similar. From the dynamics of X and X^N, we write

$$\begin{aligned}
&X_t(x) - X_t^N(x) - X_t(y) + X_t^N(y) \\
&= \int_0^t \Big(\mu(s, X_s(x)) - \mu(\tau_s, X_{\tau_s}^N(x)) - \mu(s, X_s(y)) + \mu(\tau_s, X_{\tau_s}^N(y))\Big)\mathrm{d}s \\
&+ \int_0^t \Big(\sigma(s, X_s(x)) - \sigma(\tau_s, X_{\tau_s}^N(x)) - \sigma(s, X_s(y)) + \sigma(\tau_s, X_{\tau_s}^N(y))\Big)\mathrm{d}W_s.
\end{aligned}$$

Then, as in the proof of Proposition 4, we obtain

$$\begin{aligned}
&\mathbb{E}\left(\sup_{u\leq t} |X_u(x) - X_u^N(x) - X_u(y) + X_u^N(y)|^p\right) \qquad (31)\\
&\leq 2^{p-1}t^{p-1}\int_0^t \mathbb{E}\left(\Big|\mu(s, X_s(x)) - \mu(\tau_s, X_{\tau_s}^N(x)) - \mu(s, X_s(y)) + \mu(\tau_s, X_{\tau_s}^N(y))\Big|^p\right)\mathrm{d}s \\
&+ 2^{p-1}[C_p^{\mathrm{BDG}}]^p t^{p/2-1}\int_0^t \mathbb{E}\left(\Big|\sigma(s, X_s(x)) - \sigma(\tau_s, X_{\tau_s}^N(x)) - \sigma(s, X_s(y)) + \sigma(\tau_s, X_{\tau_s}^N(y))\Big|^p\right)\mathrm{d}s.
\end{aligned}$$

Actually, both terms of the right side of above inequality can be treated in the same way, thus we only detail the computations for the second integral. First write that

$$\begin{aligned}
&\sigma(s, X_s(x)) - \sigma(\tau_s, X_{\tau_s}^N(x)) - \sigma(s, X_s(y)) + \sigma(\tau_s, X_{\tau_s}^N(y)) \\
&= \sigma(s, X_s(x)) - \sigma(s, X_s^N(x)) - \sigma(s, X_s(y)) + \sigma(s, X_s^N(y)) \\
&+ \sigma(s, X_s^N(x)) - \sigma(\tau_s, X_{\tau_s}^N(x)) - \sigma(s, X_s^N(y)) + \sigma(\tau_s, X_{\tau_s}^N(y)). \qquad (32)
\end{aligned}$$

Now, we treat the two lines above separately.
<u>**Step 1.**</u> Denoting by $X_s^{N,\lambda,x} := X_s(x) + \lambda\big(X_s^N(x) - X_s(x)\big)$ for $\lambda \in [0,1]$, we have

$$\begin{aligned}
&\sigma(s, X_s(x)) - \sigma(s, X_s^N(x)) - \sigma(s, X_s(y)) + \sigma(s, X_s^N(y)) \\
&= \big(X_s(x) - X_s^N(x) - X_s(y) + X_s^N(y)\big)\int_0^1 \nabla_x\sigma\big(s, X_s^{N,\lambda,x}\big)\mathrm{d}\lambda \\
&+ \big(X_s(y) - X_s^N(y)\big)\int_0^1 \big(\nabla_x\sigma(s, X_s^{N,\lambda,x}) - \nabla_x\sigma(s, X_s^{N,\lambda,y})\big)\mathrm{d}\lambda.
\end{aligned}$$

Now we use the definition of the process $X^{N,\lambda,x}$, the fact that $|\nabla_x\sigma(t,x)| \leq C^{X,\nabla}$ and $|\nabla_x\sigma(t,x) - \nabla_x\sigma(t,y)| \leq C^{X,\nabla}|x-y|^\delta$; we then deduce (for a generic constant C_p which values may change from line to line)

$$\begin{aligned}
&|\sigma(s, X_s(x)) - \sigma(s, X_s^N(x)) - \sigma(s, X_s(y)) + \sigma(s, X_s^N(y))|^p \\
&\leq C_p\Big(|X_s(x) - X_s^N(x) - X_s(y) + X_s^N(y)|^p \\
&\quad + |X_s(y) - X_s^N(y)|^p \int_0^1 \big|(1-\lambda)\big(X_s(x) - X_s(y)\big) + \lambda\big(X_s^N(x) - X_s^N(y)\big)\big|^{\delta p}\mathrm{d}\lambda\Big) \\
&\leq C_p\Big[|X_s(x) - X_s^N(x) - X_s(y) + X_s^N(y)|^p \\
&\quad + |X_s(y) - X_s^N(y)|^p\Big(|X_s(x) - X_s(y)|^{\delta p} + |X_s^N(x) - X_s^N(y)|^{\delta p}\Big)\Big]
\end{aligned}$$

where we have invoked the Minkowsky inequality to handle the dλ-integral and also used (2). From this, integrating over (s,ω) and applying the Cauchy-Schwarz inequality, we obtain (with a larger constant C_p)

$$
\begin{aligned}
&\mathbb{E}\int_0^t |\sigma(s,X_s(x))-\sigma(s,X_s^N(x))-\sigma(s,X_s(y))+\sigma(s,X_s^N(y))|^p \mathrm{d}s \\
&\le C_p\Big[\int_0^t \mathbb{E}\left(|X_s(x)-X_s^N(x)-X_s(y)+X_s^N(y)|^p\right)\mathrm{d}s \\
&+\int_0^t \sqrt{\mathbb{E}\left(|X_s(y)-X_s^N(y)|^{2p}\right)}\sqrt{\mathbb{E}\left(|X_s(x)-X_s(y)|^{2\delta p}+|X_s^N(x)-X_s^N(y)|^{2\delta p}\right)}\mathrm{d}s\Big]
\end{aligned}
$$

which rewrites, owing to (16)-(25) and (28),

$$
\begin{aligned}
&\mathbb{E}\int_0^t |\sigma(s,X_s(x))-\sigma(s,X_s^N(x))-\sigma(s,X_s(y))+\sigma(s,X_s^N(y))|^p \mathrm{d}s \\
&\le C_p\Big(\int_0^t \mathbb{E}\left(|X_s(x))-X_s^N(x)-X_s(y)+X_s^N(y)|^p\right)\mathrm{d}s+\frac{(1+|y|)^p}{N^{\beta p}}|x-y|^{\delta p}\Big) \quad (33)
\end{aligned}
$$

for a new generic constant C_p.

Step 2. Now we are concerned by the second line of Identity (32). Similarly to before, we can write

$$
\begin{aligned}
&\sigma(s,X_s^N(x))-\sigma(\tau_s,X_{\tau_s}^N(x))-\sigma(s,X_s^N(y))+\sigma(\tau_s,X_{\tau_s}^N(y)) \\
&=\sigma(s,X_s^N(x))-\sigma(s,X_{\tau_s}^N(x))-\big(\sigma(s,X_s^N(y))-\sigma(s,X_{\tau_s}^N(y))\big) \\
&+\sigma(s,X_{\tau_s}^N(x))-\sigma(s,X_{\tau_s}^N(y))-\big(\sigma(\tau_s,X_{\tau_s}^N(x))-\sigma(\tau_s,X_{\tau_s}^N(y))\big) \\
&=\int_0^1 \nabla_x\sigma\big(s,X_s^N(x)+\lambda(X_{\tau_s}^N(x)-X_s^N(x))\big)\mathrm{d}\lambda\ (X_s^N(x)-X_{\tau_s}^N(x)) \\
&-\int_0^1 \nabla_x\sigma\big(s,X_s^N(y)+\lambda(X_{\tau_s}^N(y)-X_s^N(y))\big)\mathrm{d}\lambda\ (X_s^N(y)-X_{\tau_s}^N(y)) \\
&+\int_0^1 \nabla_x\sigma\big(s,X_{\tau_s}^N(x)+\lambda(X_{\tau_s}^N(y)-X_{\tau_s}^N(x))\big)\mathrm{d}\lambda\ (X_{\tau_s}^N(x)-X_{\tau_s}^N(y)) \\
&-\int_0^1 \nabla_x\sigma\big(\tau_s,X_{\tau_s}^N(x)+\lambda(X_{\tau_s}^N(y)-X_{\tau_s}^N(x))\big)\mathrm{d}\lambda\ (X_{\tau_s}^N(x)-X_{\tau_s}^N(y)) \\
&=\int_0^1 \Big(\nabla_x\sigma\big(s,X_s^N(x)+\lambda(X_{\tau_s}^N(x)-X_s^N(x))\big)-\nabla_x\sigma\big(s,X_s^N(y)+\lambda(X_{\tau_s}^N(y)-X_s^N(y))\big)\Big)\mathrm{d}\lambda \\
&\quad\times(X_s^N(x)-X_{\tau_s}^N(x)) \\
&+\int_0^1 \nabla_x\sigma\big(s,X_s^N(y)+\lambda(X_{\tau_s}^N(y)-X_s^N(y))\big)\mathrm{d}\lambda\ (X_s^N(x)-X_{\tau_s}^N(x)-X_s^N(y)+X_{\tau_s}^N(y)) \\
&+\int_0^1 \Big(\nabla_x\sigma\big(s,X_{\tau_s}^N(x)+\lambda(X_{\tau_s}^N(y)-X_{\tau_s}^N(x))\big)-\nabla_x\sigma\big(\tau_s,X_{\tau_s}^N(x)+\lambda(X_{\tau_s}^N(y)-X_{\tau_s}^N(x))\big)\Big)\mathrm{d}\lambda \\
&\quad\times(X_{\tau_s}^N(x)-X_{\tau_s}^N(y)).
\end{aligned}
$$

Now, by taking advantage of the boundedness and regularity assumptions on $\nabla_x\sigma$, it readily follows

$$
\begin{aligned}
&|\sigma(s,X_s^N(x))-\sigma(\tau_s,X_{\tau_s}^N(x))-\sigma(s,X_s^N(y))+\sigma(\tau_s,X_{\tau_s}^N(y))| \\
&\le C^{X,\nabla}\Big[\int_0^1 \Big|(1-\lambda)(X_s^N(x)-X_s^N(y))+\lambda(X_{\tau_s}^N(x)-X_{\tau_s}^N(y))\Big|^\delta \mathrm{d}\lambda\ |X_s^N(x)-X_{\tau_s}^N(x)|
\end{aligned}
$$

$$+ \left|X_s^N(x) - X_{\tau_s}^N(x) - X_s^N(y) + X_{\tau_s}^N(y)\right|$$
$$+ |s-\tau_s|^\alpha \int_0^1 \Big(1 + \left|X_{\tau_s}^N(x) + \lambda(X_{\tau_s}^N(y) - X_{\tau_s}^N(x))\right|\Big)\mathrm{d}\lambda \, \left|X_{\tau_s}^N(x) - X_{\tau_s}^N(y)\right|\Big].$$

By taking the power p and integrating w.r.t. (s,ω), we get, after standard computations,

$$\int_0^t \mathbb{E}\left(|\sigma(s,X_s^N(x)) - \sigma(\tau_s,X_{\tau_s}^N(x)) - \sigma(s,X_s^N(y)) + \sigma(\tau_s,X_{\tau_s}^N(y))|^p\right)\mathrm{d}s$$
$$\le C_p\Bigg[\int_0^t \left(\sqrt{\mathbb{E}\left(|X_s^N(x) - X_s^N(y)|^{2p\delta}\right)} + \sqrt{\mathbb{E}\left(|X_{\tau_s}^N(x) - X_{\tau_s}^N(y))|^{2p\delta}\right)}\right)$$
$$\times \sqrt{\mathbb{E}\left(|X_s^N(x) - X_{\tau_s}^N(x)|^{2p}\right)}\mathrm{d}s$$
$$+ \int_0^t \mathbb{E}\left(|X_s^N(x) - X_{\tau_s}^N(x) - X_s^N(y) + X_{\tau_s}^N(y)|^p\right)\mathrm{d}s$$
$$+ \frac{1}{N^{\alpha p}}\int_0^t \left(1 + \sqrt{\mathbb{E}\left(|X_{\tau_s}^N(x)|^{2p}\right)} + \sqrt{\mathbb{E}\left(|X_{\tau_s}^N(y)|^{2p}\right)}\right)\sqrt{\mathbb{E}\left(\left|X_{\tau_s}^N(x) - X_{\tau_s}^N(y)\right|^{2p}\right)}\mathrm{d}s\Bigg].$$

for some new generic constant C_p. Finally, by plugging into the above the results of Proposition 5 and Lemma 2, we obtain (for a new constant C_p)

$$\mathbb{E}\int_0^t |\sigma(s,X_s^N(x)) - \sigma(\tau_s,X_{\tau_s}^N(x)) - \sigma(s,X_s^N(y)) + \sigma(\tau_s,X_{\tau_s}^N(y))|^p\mathrm{d}s$$
$$\le C_p\Big(\frac{|x-y|^{p\delta}}{N^{p/2}}(1+|x|)^p + \frac{|x-y|^p}{N^{p/2}} + \frac{|x-y|^p}{N^{\alpha p}}(1+|x|^p+|y|^p)\Big)$$
$$\le C_p(1+|x|+|y|)^p\frac{|x-y|^p + |x-y|^{\delta p}}{N^{\beta p}}. \tag{34}$$

We then obtain, by combining (32), (33) and (34),

$$\mathbb{E}\int_0^t |\sigma(s,X_s(x)) - \sigma(\tau_s,X_{\tau_s}^N(x)) - \sigma(s,X_s(y)) + \sigma(\tau_s,X_{\tau_s}^N(y))|^p\mathrm{d}s$$
$$\le C_p\Big[\int_0^t \mathbb{E}\left(|X_s(x) - X_s^N(x) - X_s(y) + X_s^N(y)|^p\right)\mathrm{d}s$$
$$+ (1+|x|+|y|)^p\frac{|x-y|^p + |x-y|^{\delta p}}{N^{\beta p}}\Big],$$

for some new constant C_p. The same estimates hold for μ instead of σ. Hence, plugging the above into (31), we obtain the existence of generic constants C_p such that

$$\mathbb{E}\left(\sup_{u\le t}|X_u(x) - X_u^N(x) - X_u(y) + X_u^N(y)|^p\right)$$
$$\le C_p\Big[\int_0^t \mathbb{E}\left(\sup_{u\le s}|X_u(x) - X_u^N(x) - X_u(y) + X_u^N(y)|^p\right)\mathrm{d}s$$
$$+ (1+|x|+|y|)^p\frac{|x-y|^p + |x-y|^{\delta p}}{N^{\beta p}}\Big]$$
$$\le C_p(1+|x|+|y|)^p\frac{|x-y|^p + |x-y|^{\delta p}}{N^{\beta p}}$$

where the last inequality follows from Gronwall's Lemma; the proof is complete. □

3.3.3 Completion of the proof of Theorem 4

We now aim at applying Theorem 1 with $F(\omega, x) := X_t(\omega, x)$, $F^N(\omega, x) := X_t^N(\omega, x)$, $\Theta := Y_s(\omega, y)$ and $\Theta^N := Y_s^N(\omega, y)$.

Assumption (**H1**) is satisfied with $C_p^{(\mathbf{H1})} := C_{p,(19)}$ and $\alpha_p^{(\mathbf{H1})} := 1$ in view of Theorem 5.

Thanks to the inequality (23) of Theorem 5, Assumption (**H1**) holds true with $\kappa := 1$, $C_p^{(\mathbf{H2})} := C_{p,(23)}$ and $\alpha_p^{(\mathbf{H2})} := \delta$.

Moreover (**H3**) is valid owing to Theorem 8 where we take $C_p^{N,(\mathbf{H3})} := \frac{C_{p,(30)}}{N^{\beta^X}}$ (with $\beta^X := \min(\alpha^X, \frac{1}{2})$) and $\alpha_p^{(\mathbf{H3})} := 2$.

Last, (**H4**) is clearly true using

- Propositions 4 and 5 applied to Y instead of X, which yields $C_p^{(\mathbf{H4}\text{-a})} := \max(C_{p,(15)}, C_{p,(24)})(1 + |y|)$,
- Theorem 6, applied to Y and Y^N, which gives $C_p^{N,(\mathbf{H4}\text{-b})} := \frac{C_{p,(28)}}{N^{\beta^Y}}(1+|y|)$ with $\beta^Y := \min(\alpha^Y, \frac{1}{2})$.

We are done. □

References

[AZ01] H. Allouba and W. Zheng. Brownian-time processes: the PDE connection and the half-derivative generator. *Annals of probability*, pages 1780–1795, 2001.

[BL93] N. Bouleau and D. Lépingle. *Numerical Methods for Stochastic Processes.* Wiley series in probability and mathematical statistics. Wiley & Sons, Inc, 1993.

[BY82] M.T. Barlow and M. Yor. Semi-martingale inequalities via the Garsia-Rodemich-Rumsey lemma, and applications to local times. *Journal of Functional Analysis*, 49(2):198–229, 1982.

[EM13] N. El Karoui and M. Mrad. An exact connection between two solvable SDEs and a non linear Utility Stochastic PDEs. *SIAM Journal on Financial Mathematics*, 4(1):697–736, 2013.

[Gil08] M.B. Giles. Multilevel Monte Carlo path simulation. *Operation Research*, 56:607–617, 2008.

[GRR70] A.M. Garsia, E. Rodemich, and H.Jr. Rumsey. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana University Mathematics Journal*, 20(6):565–578, 1970.

[Hei01] S. Heinrich. Multilevel Monte Carlo Methods. In *LSSC '01 Proceedings of the Third International Conference on Large-Scale Scientific Computing*, volume 2179 of *Lecture Notes in Computer Science*, pages 58–67. Springer-Verlag, 2001.

[KS97] A. Kohatsu-Higa and M. Sanz-Solé. Existence and regularity of density for solutions to stochastic differential equations with boundary conditions. *Stochastics Stochastics Rep.*, 60(1-2):1–22, 1997.

[Kun97] H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1997.

[MZ10] M. Musiela and T. Zariphopoulou. Stochastic partial differential equations and portfolio choice. In *Contemporary Quantitative Finance*, pages 195–216. Springer, 2010.

[Nua06] D. Nualart. *Malliavin calculus and related topics*. Springer Verlag, second edition, 2006.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Comprehensive Studies in Mathematics. Berlin: Springer, third edition, 1999.