Available online at www.sciencedirect.com
SCIENCE DIRECT®
Ann. I. H. Poincaré – PR 41 (2005) 523–558
ANNALES DE L'INSTITUT HENRI POINCARÉ PROBABILITÉS ET STATISTIQUES
www.elsevier.com/locate/anihpb

The approximate Euler method for Lévy driven stochastic differential equations

Jean Jacod [a,*], Thomas G. Kurtz [b,1], Sylvie Méléard [c], Philip Protter [d,2]

[a] *Laboratoire de Probabilités et Modèles Aléatoires, CNRS UMR 7599 and Université P. et M. Curie, 4, Place Jussieu, 75252 Paris cedex 05, France*
[b] *Departments of Mathematics and Statistics, University of Wisconsin — Madison, 480 Lincoln Drive, Madison, WI 53706-1388, USA*
[c] *MODAL'X, Université Paris-10, 200 Avenue de la République, 92000 Nanterre, and Laboratoire de Probabilités et Modèles Aléatoires, France*
[d] *ORIE – 219 Rhodes Hall, Cornell University, Ithaca, NY 14853-3801, USA*

Received 11 July 2003; received in revised form 28 January 2004; accepted 30 January 2004

Available online 1 April 2005

Abstract

This paper is concerned with the numerical approximation of the expected value $\mathbb{E}(g(X_t))$, where g is a suitable test function and X is the solution of a stochastic differential equation driven by a Lévy process Y. More precisely we consider an Euler scheme or an "approximate" Euler scheme with stepsize $1/n$, giving rise to a simulable variable X_t^n, and we study the error $\delta_n(g) = \mathbb{E}(g(X_t^n)) - \mathbb{E}(g(X_t))$.

For a genuine Euler scheme we typically get that $\delta_n(g)$ is of order $1/n$, and we even have an expansion of this error in successive powers of $1/n$, and the assumptions are some integrability condition on the driving process and appropriate smoothness of the coefficient of the equation and of the test function g.

For an approximate Euler scheme, that is we replace the non-simulable increments of X by a simulable variable close enough to the desired increment, the order of magnitude of $\delta_n(g)$ is the supremum of $1/N$ and a kind of "distance" between the increments of Y and the actually simulated variable. In this situation, a second order expansion is also available.

Résumé

Cet article est consacré à l'approximation numérique de l'espérance $\mathbb{E}(g(X_t))$, où g est une fonction test convenable et X est la solution d'une équation différentielle stochastique dirigée par un processus de Lévy Y. Plus précisément, on considère un

* Corresponding author.
E-mail address: jj@ccr.jussieu.fr (J. Jacod).

[1] Supported in part by the U. S. Army Research Laboratory and the U. S. Army Research Office under contract number DAAD19-01-1-0502 and by NSF grant DMS 02-05034.

[2] Supported in part by NSF grant DMS-0202958 and NSA grant MDA-904-00-1-0035.

0246-0203/$ – see front matter
doi:10.1016/j.anihpb.2004.01.007

schéma d'Euler ou un schéma d'Euler « approximatif » de pas $1/n$, produisant une variable X_t^n simulable, et on étudie l'erreur $\delta_n(g) = \mathbb{E}(g(X_t^n)) - \mathbb{E}(g(X_t))$.

Pour le schéma d'Euler classique on obtient typiquement que $\delta_n(g)$ est d'ordre $1/n$, et on exhibe un développement de cette erreur en puissances successives de $1/n$; les hypothèses sont d'une part une régularité suffisante de la fonction test g et des coefficients de l'équation différentielle, d'autre part l'existence de momemnts appropriés pour le processus directeur.

Pour le schéma d'Euler approximatif, on remplace les accroissements de Y, en général non simulables, par des variables simulables assez proches de ces accroissements. L'ordre de grandeur de $\delta_n(g)$ devient ainsi le maximum de $1/N$ et d'une sorte de « distance » entre les accroissements de Y et les variables effectivement simulées. Dans ce cadre, nous donnons également un développement à l'ordre 2.

MSC: 60H10; 65U05; 60J30

Keywords: Euler scheme; Stochastic differential equations; Simulations; Approximate simulations

1. Introduction

1) *Approximating Markov process expectations.* In applications of Markov processes, it is frequently necessary to compute $\mathbb{E}(g(X_t))$, where X is the process modelling the system of interest. While this expectation can sometimes be obtained by direct numerical computation, for example, by applying numerical schemes for partial differential equations, Monte Carlo methods may provide the only effective approach. In the simplest Monte Carlo approach, the desired expectation is approximated as

$$\mathbb{E}\big(g(X_t)\big) \approx \frac{1}{N}\sum_{i=1}^{N} g(\widehat{X}_t^i),$$

where the $\widehat{X}_t^i$ are simulated, independent copies of X_t. In practice, it may not be possible to simulate draws from the distribution of X_t exactly, so the Monte Carlo approximation may introduce a bias

$$\Delta g = \mathbb{E}g(\widehat{X}_t) - \mathbb{E}g(X_t),$$

where $\hat{X}$ is a simulatable approximation of X. We are interested in developing methods for estimating this bias for a large class of Markov processes and corresponding approximations.

The simulated process used in the Monte Carlo approximation will typically be a discrete time Markov chain in which the discrete time-step is identified with a small interval on the real-time axis. To simplify notation, we take the length of this interval to be $1/n$ for some integer n and assume that $t = 1$. We will denote the approximating process by X^n to emphasize the dependence on the time-step in the simulation. We also note that the bias will depend on the initial state x, particularly if g is unbounded, so we want to estimate the bias

$$\Delta_n g(x) = \mathbb{E}g(X_1^n) - \mathbb{E}g(X_1) \tag{1.1}$$

as a function of x and n.

Let $(P_t)_{t\geqslant 0}$ denote the transition semigroup for the process X, and also $(P_t^n)_{t\geqslant 0}$ be defined as

$$P_t^n = P_{[nt]/n}^n, \quad \text{where, for } j \text{ an integer,} \quad P_{j/n}^n h(x) = \mathbb{E}\big[h(X_{j/n}^n)|X_0^n = x\big]$$

($[nt]$ is the integer part of the number nt). Then we can write the bias as

$$\Delta_n g(x) = P_1^n g(x) - P_1 g(x) = \sum_{j=1}^{n} P_{1-j/n}^n (P_{1/n}^n - P_{1/n}) P_{(j-1)/n} g(x)$$

$$= \int_0^1 P_{1-\eta_n(s)-1/n}^n E^n P_{\eta_n(s)} g(x)\,\mathrm{d}s, \tag{1.2}$$

where $\eta_n(s) = [ns]/n$ and

$$E^n h(x) = n(Q_{1/n}^n - P_{1/n})h(x).$$

If X^n is a good approximation of X, then the operator P_t^n is "close" to the operator P_t. Hence the identity (1.2) suggests that the rate of convergence of $\Delta_n g(x)$ to 0 is governed by the rate at which the error operator E^n goes to 0. More precisely, one might expect to obtain estimates of the form

$$\left|E^n h(x)\right| \leqslant \varepsilon_n \rho(x) \|h\|_E,$$

for some sequence $\varepsilon_n \to 0$ and function $\rho \geqslant 0$ and for h in some collection of functions D_E with $\|\cdot\|_E$ a norm on D_E. If further P_t maps D_E into itself (a requirement which, for diffusion processes, is sometimes implied by regularity results for partial differential equations but, as we shall show, can also be obtained by direct probabilistic calculations), then we may expect a bound of the form

$$\left|\Delta_n g(x)\right| \leqslant \varepsilon_n \int_0^1 P_{1-\eta_n(s)}^n \rho(x) \|P_{\eta_n(s)} g\|_E\,\mathrm{d}s.$$

This analysis suggests the possibility of an exact asymptotic limit of the form

$$\Gamma^{(1)} g(x) \equiv \lim_{n\to\infty} \varepsilon_n^{-1} \Delta_n g(x) = \int_0^1 P_{1-s} E P_s g(x)\,\mathrm{d}s,$$

where

$$Eh(x) = \lim_{n\to\infty} \varepsilon_n^{-1} E^n h(x),$$

for $h \in D_E$. One can also consider the rate of convergence in this limit or attempt to derive higher order expansions.

2) *Euler approximations for solutions of stochastic differential equations.* We develop the desired estimates for solutions of a stochastic differential equation

$$X_t = x + \int_0^t f(X_{s-})\,\mathrm{d}Y_s \tag{1.3}$$

driven by a Lévy process Y. The process X is d-dimensional, while Y is d'-dimensional, so f takes its values in $\mathbb{R}^d \otimes \mathbb{R}^{d'}$, and we systematically use (column) vector and matrix notation. The initial value is some given $x \in \mathbb{R}^d$. The precise assumptions are stated later.

The process X^n will be given by an Euler approximation with stepsize $1/n$ defined recursively at times i/n by

$$X_0^n = x, \qquad X_{(i+1)/n}^n = X_{i/n}^n + f(X_{i/n}^n)(Y_{(i+1)/n} - Y_{i/n}), \tag{1.4}$$

or since the true increment $Y_{(i+1)/n} - Y_{1/n}$ may be difficult to simulate, in practice, we may substitute the i.i.d. random variables ζ_i^n which are close enough to the true increments, and are exactly simulatable. That is, instead of the *genuine Euler scheme* given by (1.4), we consider the *approximate Euler scheme*, still denoted by X^n, given by

$$X_0^n = x, \qquad X_{(i+1)/n}^n = X_{i/n}^n + f(X_{i/n}^n)\zeta_{i+1}^n. \tag{1.5}$$

In this approximation, we have two sources of error, the discretization error and the error due to the approximation of the increment. The latter error does not have a natural generic characterization, so we will simply assume that we have estimates for

$$\delta_n(h) = \mathbb{E}\big(h(\zeta_1^n)\big) - \mathbb{E}\big(h(Y_{1/n})\big) \tag{1.6}$$

of the form

$$\big|\delta_n(h)\big| = \big|\mathbb{E}\big(h(\zeta_1^n)\big) - \mathbb{E}\big(h(Y_{1/n})\big)\big| \leqslant \frac{Ku_n}{n}\|h\|_E,$$

for h in a sufficiently large space D_E with an appropriate norm and a sequence u_n going to 0. For higher order results, we need an expansion of $\delta_n(h)$ around 0.

The space D_E will be of the form $C^k_\alpha(\mathbb{R}^d)$ for some integer $k \geqslant 0$ and some function $\alpha \geqslant 1$, where $C^k_\alpha(\mathbb{R}^d)$ is the space of k-times continuously differentiable functions on $\mathbb{R}^d$ with norm

$$\|h\|_{\alpha,k} = \inf\big\{a > 0\colon \big|\nabla^i h(x)\big| \leqslant a\alpha(x) \text{ for } i = 0, \dots, k\big\}.$$

For all $p \geqslant 0$ we introduce the function

$$\alpha_p(x) = \begin{cases} 1 + |x|^p & \text{if } p > 0, \\ 1 & \text{if } p = 0, \end{cases} \tag{1.7}$$

and we write $C^k_p(\mathbb{R}^d)$ and $\|h\|_{p,k}$ instead of $C^k_{\alpha_p}(\mathbb{R}^d)$ and $\|h\|_{\alpha_p,k}$.

3) *The basic problems.* We are interested in results of the following types:

(A) An estimate of the bias: for example, if f is in $C^4_0(\mathbb{R}^d)$ and Y_t has finite moments up to order 8 and

$$h \in C^4_0(\mathbb{R}^{d'}) \quad \Longrightarrow \quad \big|\delta_n(h)\big| \leqslant \frac{Ku_n}{n}\,\|h\|_{0,4} \tag{1.8}$$

for some constant K and some sequence (u_n) of positive numbers, then we have

$$g \in C^4_0(\mathbb{R}^d) \quad \Longrightarrow \quad \big|\Delta_n(g)\big| \leqslant K'\left(u_n \vee \frac{1}{n}\right)\|g\|_{0,4} \tag{1.9}$$

for another constant K'. More general results are provided in Theorem 2.1 below.

(B) A first order expansion: assume for example that f is in $C^{10}_0(\mathbb{R}^d)$, that Y_t has finite moments up to order 20, that (1.8) with $u_n = \frac{1}{n}$ holds, and that

$$h \in C^6_0(\mathbb{R}^{d'}) \quad \Longrightarrow \quad \left|\delta_n(h) - \frac{1}{n^2}\beta(h)\right| \leqslant \frac{Ku'_n}{n}\,\|h\|_{0,6} \tag{1.10}$$

for some constant K, some linear map β on $C^6_0(\mathbb{R}^{d'})$ and some sequence (u'_n) with $nu'_n \to 0$. Then there is a linear map γ_x on $C^6_0(\mathbb{R}^d)$ (where x is the starting point), such that

$$g \in C^6_0(\mathbb{R}^d) \quad \Longrightarrow \quad \left|\Delta_n(g) - \frac{1}{n}\gamma_x(g)\right| \leqslant K'\left(u'_n \vee \frac{1}{n^2}\right)\|g\|_{0,6} \tag{1.11}$$

for another constant K'. Here again, see Theorem 2.2 below for more general results.

(C) Higher order expansions, that is, the existence of operators $\Gamma^{(k)}$ for $k = 1, \dots, m$, such that

$$\left|\Delta_n g(x) - \sum_{k=1}^{m} \frac{1}{n^k}\Gamma^{(k)}g(x)\right| \leqslant \frac{C_m(g,x)}{n^{m+1}}. \tag{1.12}$$

Remark 1.1. For the genuine Euler scheme $\zeta_1^n = Y_{1/n}$, (1.9) and (1.10) with $u_n = u'_n = 0$ are obviously satisfied, hence we recover the results of Protter and Talay [12]: the assumptions made above are a bit stronger, but in Theorems 2.1 and 2.2 they will be (unsignificantly) weaker.

Remark 1.2. The comparison between the rate $1/n$ due to the Euler scheme and the rate u_n due to the approximation of $Y_{1/n}$ by ζ_1^n (in law) is instructive: since the time needed for the simulation is proportional to n and is also usually an increasing function of u_n, from a practical point of view it is best (asymptotically) to choose $u_n = 1/n$. This is why we assume $u_n = 1/n$ in the problem (B) above, which allows for a Romberg-type method of simulation.

It is noteworthy to observe that, except for genuine Euler schemes, we have in general (1.10) with $n^2 u'_n \to \infty$, unfortunately, as we will see in the examples below.

4) *Relationship to other work.* The results above, starting with (A), are in deep contrast with "pathwise" rates of convergence obtained for example in [9,5,6], or [14], where one looks for sequences v_n increasing to infinity and such that $v_n(X^n - X)$ is tight (as processes, or at some time t) with non-zero limits. The rate v_n depends on the characteristics of the Lévy process Y, ranging from $v_n = \sqrt{n}$ when Y has a non-vanishing Wiener part to $v_n = (n/\log n)^\alpha$ when Y is a symmetric stable process with index α, and even to "$v_n = \infty$" (that is $X^n \equiv X$ for n big enough, depending on ω of course) when Y is a compound Poisson process. The mathematical reason for this discrepancy is a lack of uniform integrability which prevents exchanging limits of random variables and expectations. It is interesting to observe that $\Delta_n g(x)$ is *always* of order $1/n$, irrespective of the characteristics of Y, provided Y has some integrability. The reason for that is quite clear when Y is a compound Poisson process, and we will devote some space to that special case (although from a practical point of view there is a way to simulate X "exactly" in that case and one should not use an Euler scheme).

The identity (1.2) has been used by a number of authors to estimate the error in Markov process approximation. See for example [11,3], and Section 1.6 of [4].

The first expansions of type (B) or (C) are due to Talay and Tubaro [16] for diffusions, while Bally and Talay [2] have considerably weakened the smoothness assumptions on the test function g (using Malliavin calculus, and under non-degeneracy assumptions on the diffusion coefficient). The discontinuous case has been studied first by Protter and Talay [12], and more recently by Kurtz and Protter [10] and Kohatsu-Higa and Yoshida [8] for an equation driven by a Poisson random measure instead of a Lévy process.

5) *Implications for simulation.* The main motivations for these types of results are practical: we want to estimate $\mathbb{E}(g(X_1))$. We run the Euler scheme N times, giving rise to the simulated numbers $X_1^{n,1}, \ldots, X_1^{n,N}$, and we take the estimate $G_{N,n} = \frac{1}{N}\sum_{i=1}^N g(X_1^{n,i})$. This is where approximate Euler schemes come into the picture, because usually one does not know how to exactly simulate the increments of our Lévy process. Section 3 below is entirely devoted to these practical problems, and in particular to the evaluation of the time necessary to perform simulations.

The expansion (C) is mathematically interesting, but its practical relevance is more dubious: in principle it lays the foundation for studying higher order Romberg schemes, but these are probably quite unstable, even in the diffusion case.

6) *Organization of the paper.* In Section 2 we state the main results in full generality, and "practical" examples are expounded in Section 3. In Section 4 we state and prove a version of the first order expansion when the driving process Y is a compound Poisson process and for the genuine Euler scheme: this is easy to prove and serves as a good introduction to the general case. Section 5 is devoted to recalling some more or less known results on Eq. (1.3). In Section 6 we prove various technical lemmas, and the last three sections are devoted to the proof of the main results.

2. The main results

1) *Some notation.* We suppose that the time interval is bounded, and without loss of generality that it is $[0, 1]$. The starting point x plays a role in the results, and a crucial role in the proofs. So, instead of X_t we write X_t^x in (1.3). Similarly, we write $X_{i/n}^{n,x}$ for the solution (1.5) to the approximate Euler scheme. For more coherent notation, we also define $X_t^{n,x}$ for all $t \in [0, t]$, by setting $X_t^{n,x} = X_{i/n}^{n,x}$ when $i \leqslant nt < i + 1$.

We consider also the processes

$$Y_t^n = \sum_{i=1}^{[nt]} \zeta_i^n, \tag{2.1}$$

where $[s]$ denotes the integer part of s. For a genuine Euler scheme, we have of course $Y_t^n = Y_{[nt]/n}$. In general, Y^n is a non-homogeneous process with independent increments, however its increments over intervals of length a multiple of $1/n$ are stationary. If we use the notation $\varphi_n(t) = \frac{i}{n}$ when $i < nt \leqslant i + 1$, then we can rewrite (1.5) as

$$X_t^{n,x} = x + \int_0^t f\big(X_{\varphi_n(s)}^{n,x}\big)\,\mathrm{d}Y_s^n. \tag{2.2}$$

Therefore, by well known results on the stability of stochastic differential equations (plus the fact that a sequence of processes with independent increments converging in law to a Lévy process is "predictably uniformly tight" (PUT): see Słominski [15], or also Theorem IX.6.9 of [7]), we readily obtain that:

$$\left.\begin{array}{lll} Y^n \xrightarrow{\mathcal{L}} Y & \Rightarrow & X^{n,x} \xrightarrow{\mathcal{L}} X^x, \\ Y_t^n = Y_{[nt]/n} & \Rightarrow & \sup_{t\in[0,1]} |X_t^{n,x} - X_t^x| \xrightarrow{\mathbb{P}} 0 \end{array}\right\} \tag{2.3}$$

(the second case corresponds to the genuine Euler scheme).

From time to time we need a filtration. For the genuine Euler scheme we take for $(\mathcal{F}_t)$ the filtration generated by Y. Otherwise, we have convergence in law only (see (2.3)), so it is no restriction to assume that all processes are defined on the same probability space and that the Y^n's are mutually independent and independent of Y, and $(\mathcal{F}_t)$ is the filtration generated by all processes Y and Y^n for $n \geqslant 1$.

2) *Assumptions on f and Y.* We now state our assumptions, starting with those on the coefficient f:

Assumption H(l, N)**.** *The coefficient f is N times continuously differentiable and all partial derivatives of order $l, l + 1, \ldots, N$ are bounded.*

Here N is an integer, and l will be either 0 or 1. We usually assume **H**$(1, 1)$ at least, except when Y is a compound Poisson. Clearly **H**$(l, N + 1) \Rightarrow$ **H**(l, N).

Next, we denote by (b, c, F) the characteristics of the Lévy process Y, in the sense that

$$\mathbb{E}(\mathrm{e}^{i\langle u, Y_t\rangle}) = \exp t\left(\mathrm{i}\langle u, b\rangle - \frac{\langle u, cu\rangle}{2} + \int F(\mathrm{d}y)\big(\mathrm{e}^{\mathrm{i}\langle u,y\rangle} - 1 - \mathrm{i}\big\langle u, \tau(y)\big\rangle\big)\right),$$

where τ is a truncation function on $\mathbb{R}^{d'}$, that is a map from $\mathbb{R}^{d'}$ into itself, which is bounded and coincides with the identity near 0, and whose components can be assumed to be in $C_0^\infty(\mathbb{R}^{d'})$ without loss of generality. Then we need the following integrability assumption, where p is some nonnegative real:

Assumption F(p)**.** *We have $\int_{\{|y|>1\}} |y|^p F(\mathrm{d}y) < \infty$ (equivalently, $Y_t \in \mathbb{L}^p$ for all t).*

3) *Assumptions on the variables* ζ_i^n. As said before, for each n the sequence of d'-dimensional variables ζ_i^n for $i = 1, \ldots, n$ is i.i.d., and some concrete ways of constructing appropriate ζ_i^n's are given in the next section. The discrepancy between ζ_1^n and $Y_{1/n}$ is measured by the quantities $\delta_n(g)$ of (1.6), and we make different assumptions according to the kind of result (of type (A), (B) or (C)) which we want to prove. Below, $p \in \mathbb{R}_+$ and $N \in \mathbb{N}^\star$ are arbitrary, and $(u_n)_{n\geqslant 1}$ and $(u'_n)_{n\geqslant 1}$ are arbitrary sequences of positive numbers tending to 0 and such that $u'_n/u_n \to 0$:

Assumption $\mathbf{G}(\{u_n\}, p)$**.** *We have* $\mathbf{F}(p)$*, and there is a constant* K *such that*

$$h \in C_p^4(\mathbb{R}^{d'}) \quad \Longrightarrow \quad \bigl|\delta_n(h)\bigr| \leqslant \frac{Ku_n}{n}\,\|h\|_{p,4}. \tag{2.4}$$

Observe that we put u_n/n and not u_n on the right: this is because the variable $Y_{1/n}$ is close to 0, and indeed "of order $1/n$" already as $n \to \infty$ in the sense that $\mathbb{E}(h(Y_{1/n})) = \mathrm{O}(1/n)$ for any $h \in C_p^0(\mathbb{R}^{d'})$ under $\mathbf{F}(p)$. We will see later that this assumption is enough to ensure the first convergence in law in (2.3).

Assumption $\mathbf{G}'(\{u_n\}, \{u'_n\}, p)$**.** *We have* $\mathbf{G}(\{u_n\}, p)$*, and there are a constant* K *and a linear map* ϕ *on the space* $C_p^6(\mathbb{R}^{d'})$ *such that*

$$h \in C_p^6(\mathbb{R}^{d'}) \quad \Longrightarrow \quad \left|\delta_n(h) - \frac{u_n}{n}\phi(h)\right| \leqslant \frac{Ku'_n}{n}\,\|h\|_{p,6}. \tag{2.5}$$

Assumption $\mathbf{G}''(N, p)$**.** *We have* $\mathbf{F}(p)$ *and there are a constant* K *and linear maps* ϕ_k *on the spaces* $C_p^{2k+4}(\mathbb{R}^{d'})$ *for* $k = 1, \ldots, N$*, such that* (*with an empty sum set equal to* 0):

$$k = 0, \ldots, N,\ h \in C_p^{2k+4}(\mathbb{R}^{d'}) \quad \Longrightarrow \quad \left|\delta_n(h) - \sum_{i=1}^{k} \frac{1}{n^{i+1}}\phi_i(h)\right| \leqslant \frac{K}{n^{k+2}}\,\|h\|_{p,2k+4}. \tag{2.6}$$

Clearly $\mathbf{G}''(N, p) \Rightarrow \mathbf{G}''(N-1, p)$, and $\mathbf{G}''(1, p) = \mathbf{G}'(\{1/n\}, \{1/n^2\}, p)$, and also $\mathbf{G}'(\{u_n\}, \{u'_n\}, p) \Rightarrow \mathbf{G}(\{u_n\}, p)$.

Finally, observe that for the genuine Euler scheme we have $\delta_n(h) = 0$ for all h, hence all the above assumptions are trivially fulfilled in this case.

4) *The main results.* Our aim is to evaluate the "error" involved by the – approximate or genuine – Euler scheme, and measured through the quantity

$$\Delta_{n,t}g(x) = \mathbb{E}\bigl(g(X^{n,x}_{[nt]/n})\bigr) - \mathbb{E}\bigl(g(X^{x}_{[nt]/n})\bigr) \tag{2.7}$$

for suitable test functions g.

Our first result is an estimate on $\Delta_{n,t}g(x)$, that is, it solves problem (A) for approximate Euler schemes:

Theorem 2.1. *Let* $p \geqslant 0$ *and* $l = 0$ *or* $l = 1$*. Assume* $\mathbf{H}(l, 4)$ *and* $\mathbf{G}(\{u_n\}, 4 + 4 \vee p)$ *for some sequence* u_n *decreasing to* 0*. Then there is a constant* K *depending on* p*,* f*,* Y *only, such that for any* $g \in C_p^4(\mathbb{R}^d)$ *we have, for all* $t \in [0, 1]$*,* $n \geqslant 1$*,* $x \in \mathbb{R}^d$*:*

$$\bigl|\Delta_{n,t}g(x)\bigr| \leqslant Kt\left(u_n \vee \frac{1}{n}\right)\|g\|_{p,4}\bigl(1 + |x|^{p+4l}\bigr). \tag{2.8}$$

As said before, for the genuine Euler scheme $\mathbf{F}(p)$ implies $\mathbf{G}(\{u_n\}, p)$ with $u_n = 0$, so we recover the estimates of Protter and Talay in [12]. For the approximate Euler scheme this result allows us to single out the contributions

of the error of the Euler scheme on the one side (that is, $1/n$), and of the simulation discrepancy on the other side (which is u_n).

The second main result is a first order expansion for $\Delta_{n,t}g(x)$. The result goes as follows, and we see that there are in fact *two* "first order terms" corresponding respectively to the Euler scheme error and to the simulation discrepancy.

Theorem 2.2. *Let $l = 0$ or $l = 1$. Assume $\mathbf{H}(l, 10)$ and $\mathbf{G}'(\{u_n\}, \{u'_n\}, 10 + 10 \vee p)$ for some $p \geqslant 0$ and some sequences u_n and u'_n with $u_n \to 0$ and $u'_n/u_n \to 0$. Then there is a constant K depending on p, f, Y only, and linear operators U_t and V_t on $C_p^{10}(\mathbb{R}^d)$, such that for any $g \in C_p^{10}(\mathbb{R}^d)$ the functions $U_t g$ and $V_t g$ belong to $C_{p+6l}^4(\mathbb{R}^d)$, and also that for all $t \in [0, 1]$, $n \geqslant 1$, $x \in \mathbb{R}^d$:*

$$\|U_t g\|_{p+6l,4} \leqslant Kt\|g\|_{p,10}, \qquad \|V_t g\|_{p+6l,4} \leqslant Kt\|g\|_{p,10}. \tag{2.9}$$

$$\left|\Delta_{n,t}g(x) - u_n U_t g(x) - \frac{1}{n}\, V_t g(x)\right| \leqslant Kt\left(u'_n \vee u_n^2 \vee \frac{1}{n^2}\right)\|g\|_{p,10}\big(1 + |x|^{p+8l}\big). \tag{2.10}$$

For the genuine Euler scheme and under $\mathbf{F}(10 + 10 \vee p)$ we have the previous hypotheses with $u_n = u'_n = 0$, so the first order term is $\frac{1}{n}V_t g(x)$ and the remainder is of order $1/n^2$.

Finally we state the result about problem (C), that is expansions of arbitrary order, of the form:

$$\Delta_{n,t}g(x) = \sum_{k=1}^{N} \frac{1}{n^k}\, \Gamma^{(k)}_{[nt]/n} g(x) + \frac{1}{n^{N+1}}\, R_{N,n,t}g(x). \tag{2.11}$$

Theorem 2.3. *Let $p \geqslant 0$ and $l = 0$ or $l = 1$ and $N \geqslant 1$, and assume $\mathbf{H}(l, 6N + 4)$ and $\mathbf{G}''(N, 6N + 4 + (6N + 4) \vee p)$. Then there is a constant K depending only on p, f, Y and N, and linear operators $\Gamma_t^{(k)}$ on $C_p^{6k}(\mathbb{R}^d)$ for $k = 1, \ldots, N$, such that if $r = 6k, 6k + 1, \ldots, 6N + 4$ we have*

$$g \in C_p^r(\mathbb{R}^d) \quad \Rightarrow \quad \Gamma_t^{(k)} g \in C_{p+6lk}^{r-6k}(\mathbb{R}^d), \quad \|\Gamma_t^{(k)} g\|_{p+6lk,r-6k} \leqslant Kt\|g\|_{p,r}, \tag{2.12}$$

and moreover if $g \in C_p^{6N+4}(\mathbb{R}^d)$ we have the expansion (2.11) *with a remainder satisfying*

$$\big|R_{N,n,t}g(x)\big| \leqslant Kt\|g\|_{p,6N+4}\big(1 + |x|^{p+4(N+1)l}\big). \tag{2.13}$$

Remark 2.4. In a sense (2.11) is not a true expansion because the "coefficients" $\Gamma^{(k)}_{[nt]/n}g(x)$ depend on n, *except when $t = 1$* of course. But, except for $t = 1$ again, $\Delta_{n,t}g(x)$ is not really a function of t but rather of the discretized time $[nt]/n$, so having an expansion which depends on time through $[nt]/n$ is also natural.

For the genuine Euler scheme, one could also use Y instead of Y^n in (2.2) (the two Euler approximations coincide at all times i/n). Instead of (2.7) one naturally takes $\Delta_{n,t}g(x) = \mathbb{E}(g(X_t^{n,x})) - \mathbb{E}(g(X_t^x))$. Then we have (2.8), but not (2.10), essentially because $nt - [nt]$ oscillates (when $t \neq 1$) between 0 and 1 as n varies.

Remark 2.5. For the genuine Euler scheme one can slightly improve the result of Theorem 2.3: we only need $\mathbf{H}(l, 4N + 2)$ and $\mathbf{F}(4N + 2 + (4N + 2) \vee p)$, and $g \in C_p^{4N+2}(\mathbb{R}^d)$, and then $\Gamma_t^{(k)} g \in C_{p+4lk}^{4N+2-4k}(\mathbb{R}^d)$. Similarly, Theorem 2.2 holds under $\mathbf{H}(l, 6)$ and $\mathbf{F}(6 + 6 \vee p)$ and for $g \in C_p^6(\mathbb{R}^d)$ (and of course $U_t g$ does not show up in that case).

Remark 2.6. When $u_n = 1/n$ in Theorem 2.2, the sum $U_t + V_t$ is equal to the operator $\Gamma_t^{(1)}$ of Theorem 2.3.

Remark 2.7. At the end of the paper (Remark 9.1) we give an "explicit" form for $\Gamma^{(1)}$ in the 1-dimensional case. This operator is well defined under $\mathbf{F}(2)$ only, so it is likely that we cannot drop integrability assumptions, even

when g is bounded (but the assumption that Y_t has finite moments of order 20 (or 12 in the genuine Euler case, see Remark 2.5) is obviously too strong!).

Remark 2.8. From a practical point of view, only the first two theorems are interesting: the first one for the plain (approximate) Euler scheme, and the second one if one wishes to use the Romberg method. And, in the latter case, this method can be applied only when the two "first order" terms are comparable, that is when $u_n = 1/n$. This is not a true practical restriction since the probabilist or statistician can indeed choose the accuracy u_n (at the price of a more or less long time for the simulation of a single variable ζ_1^n): we explain in Section 3 below how this works on a particular example.

5) Let us end this section with another set of assumptions, in a particular case. Actually, checking Assumptions $\mathbf{G}(\{u_n\}, p)$ or $\mathbf{G}'(\{u_n\}, \{u'_n\}, p)$ if we have a procedure to approximate $Y_{1/n}$ by variables ζ_1^n when we know only the laws of the latter may be quite difficult (not to mention $\mathbf{G}''(N, p)$).

However there is a situation which occurs often in practice and for which we have simpler conditions: we will say that we have a *restricted approximate Euler scheme* if each ζ_1^n is (in law) the value at time $1/n$ of a Lévy process Y'^n (equivalently, the law of ζ_1^n is infinitely divisible). That is, for each n we have a Lévy process Y'^n, and we take $\zeta_i^n = Y'^n_{i/n} - Y'^n_{(i-1)/n}$. Then of course the process Y^n of (2.1) is the discretization of Y'^n, that is $Y_t^n = Y'^n_{[nt]/n}$.

In this situation it is usually the case that the characteristics of Y'^n are known, and they are denoted by (b'_n, c'_n, F'_n) (w.r.t. the same truncation function τ than Y). We also consider the second modified characteristics of Y and Y'^n, given by

$$\tilde{c} = c + F(\tau\tau^\star), \qquad \tilde{c}'_n = c'_n + F'_n(\tau\tau^\star) \tag{2.14}$$

(here and below, we write $F(g)$ instead of $\int g(y)F(\mathrm{d}y)$). With this notation, and if we further denote by $C'^k_p(\mathbb{R}^{d'})$ (for $k \geqslant 2$) the set of all $h \in C^k_p(\mathbb{R}^{d'})$ such that $\nabla^i h(0) = 0$ for $i = 0, 1, 2$, we can introduce the following two assumptions (we suppose, as above, that $u_n \to 0$ and $u'_n/u_n \to 0$):

Assumption $\widehat{\mathbf{G}}(\{u_n\}, p)$. *We have $\mathbf{F}(p)$, and there is a constant K such that*

$$\left.\begin{aligned} &|b'_n - b| \leqslant Ku_n, \quad |\tilde{c}'_n - \tilde{c}| \leqslant Ku_n, \\ &h \in C'^4_p(\mathbb{R}^{d'}) \quad \Rightarrow \quad \bigl|F'_n(h) - F(h)\bigr| \leqslant Ku_n\|h\|_{p,4}. \end{aligned}\right\} \tag{2.15}$$

Assumption $\widehat{\mathbf{G}}'(\{u_n\}, \{u'_n\}, p)$. *We have $\mathbf{F}(p)$, and there are a vector $\beta \in \mathbb{R}^{d'}$ and a $d' \times d'$ matrix σ and a linear map Φ on $C'^6_p(\mathbb{R}^{d'})$ and a constant K such that*

$$\left.\begin{aligned} &|b'_n - b - u_n\beta| \leqslant Ku'_n, \quad |\tilde{c}'_n - \tilde{c} - u_n\sigma| \leqslant Ku'_n, \\ &h \in C'^4_p(\mathbb{R}^{d'}) \quad \Rightarrow \quad \bigl|F'_n(h) - F(h)\bigr| \leqslant Ku_n\|h\|_{p,4}, \\ &h \in C'^6_p(\mathbb{R}^{d'}) \quad \Rightarrow \quad \bigl|F'_n(h) - F(h) - u_n\Phi(h)\bigr| \leqslant Ku'_n\|h\|_{p,6}. \end{aligned}\right\} \tag{2.16}$$

Proposition 2.9. *In the case of a restricted approximate Euler scheme, that is $\zeta_i^n = Y'^n_{i/n} - Y'^n_{(i-1)/n}$ for some Lévy process Y'^n with characteristics and second modified characteristic (b'_n, c'_n, F'_n) and $\tilde{c}'_n$, we have for any $p \geqslant 2$:*

(a) $\widehat{\mathbf{G}}(\{u_n\}, p)$ *implies* $\mathbf{G}(\{u_n \vee \frac{1}{n}\}, p)$.
(b) *If $u_n \geqslant 1/n$, then* $\widehat{\mathbf{G}}'(\{u_n\}, \{u'_n\}, p)$ *implies* $\mathbf{G}'(\{u_n\}, \{u'_n \vee \frac{u_n}{n}\}, p)$.

The proof of this proposition is given in Subsection 6.1.

3. Some examples

In this section we consider some practical examples and compute the time necessary to achieve a given precision in the computation of $\mathbb{E}(g(X_1))$ via a Monte Carlo method, as explained in the Introduction (and especially §5).

We draw N independent copies of the approximation X_1^n of X_1 and take the estimate $G_{N,n} = \frac{1}{N}\sum_{i=1}^N g(X_1^{n,i})$. If we use a "Romberg technique", we draw N copies of X_1^n and of X_1^{2n} and take the estimate $G_{N,n} = \frac{1}{N}\sum_{i=1}^N (2g(X_1^{2n,i})) - g(X_1^{n,i})))$. Assuming g bounded, the expected error $e(N,n) = \mathbb{E}(|G_{N,n} - \mathbb{E}(g(X_1))|)$ is the sum of the statistical error, of order $1/\sqrt{N}$, plus the bias $\Delta_n g(x)$ (or $2\Delta_{2n}g(x) - \Delta_n g(x)$ for the Romberg method). Therefore if Theorem 2.1 applies we get

$$e(N,n) = \mathrm{O}\left(\frac{1}{\sqrt{N}} + \frac{1}{n} \vee u_n\right), \tag{3.1}$$

while for the Romberg method and if we can apply Theorem 2.2, we get

$$e(N,n) = \mathrm{O}\left(\frac{1}{\sqrt{N}} + \frac{1}{n^2} \vee u'_n\right). \tag{3.2}$$

In both cases the time needed is $T(N,n) = \mathrm{O}(Nn\alpha_n)$, where α_n is the time necessary to calculate a single time step.

1) *Genuine Euler scheme.* If we can simulate exactly the increments of Y, the time α_n is $\alpha_n = \mathrm{O}(1)$, and we have (2.4) and (2.5) with $u_n = u'_n = 0$. Optimizing the choices of n and N in (3.1), subject to the condition $e(N,n) \leqslant \varepsilon$, leads to take $N = \mathrm{O}(n^2)$ and $n = \mathrm{O}(1/\varepsilon)$, and injecting into $T(N,n) = \mathrm{O}(Nn)$ gives us a time T_ε necessary to achieve a precision ε which satisfies:

$$T_\varepsilon = \mathrm{O}(\varepsilon^{-3}). \tag{3.3}$$

If we use the Romberg method, we apply (3.2) instead of (3.1): this leads to take $N = \mathrm{O}(n^4)$ and $n = \mathrm{O}(1/\sqrt{\varepsilon})$, and injecting into $T(N,n) = \mathrm{O}(Nn)$ gives us

$$T_\varepsilon = \mathrm{O}(\varepsilon^{-5/2}). \tag{3.4}$$

Observe that if we knew how to exactly simulate $g(X_1)$ without any bias, we would obtain a time $T_\varepsilon = \mathrm{O}(\varepsilon^{-2})$, to be compared with (3.3) or (3.4).

2) *Restricted approximate Euler scheme.* We can exactly simulate the drift (of course!) and the Wiener part of Y, but not the jump part (except when this jump part is compound Poisson, or is a stable process, but in the latter case the integrability assumptions of this paper are not fulfilled). Otherwise, we cannot exactly simulate the increments of Y. To approximate them, most methods resort to deleting in some way or another the "small jumps" of Y, so for the discontinuous part we are left with a compound Poisson process, which can usually be simulated. The reader can look at the papers [13] of Rosiński or [1] of Asmussen and Rosiński for various possibilities. Below we use the most simple-minded one, with a view towards minimizing the time needed. This method works if we can simulate a variable whose law is the (normalized) restriction of F to the complement of any neighborhood of 0: since F is often explicitly known, this is in general feasible.

So we truncate the jumps at some cut-off size v_n (going to 0 as $n \to \infty$). This amounts to a restricted approximate Euler scheme, the characteristics (b'_n, c'_n, F'_n) and $\tilde{c}'_n$ being chosen such that $\widehat{\mathbf{G}}(\{u_n\}, p)$ or $\widehat{\mathbf{G}}'(\{u_n\}, \{u'_n\}, p)$ holds for suitable sequences u_n and/or u'_n, and with $F'_n(\mathrm{d}x) = 1_{\{|x|>v_n\}}F(\mathrm{d}x)$. Then we can of course choose $b'_n = b_n$, and c_n in such a way that $\tilde{c}'_n = \tilde{c}$, so only the last parts of (2.15) or (2.16) have to be checked.

Observe that Y'^n is the sum of a drift, a Wiener process, and a compound Poisson process. So we can simulate exactly $\zeta_1^n = Y'^n_{1/n}$ by using a Gaussian variable, plus a Poisson variable Z (the number of jumps on the interval

$(0, 1/n])$, plus Z variables according to the law F'_n (normalized)). The time necessary to do that is random, with expectation

$$\alpha_n = \mathrm{O}\big(1 + \mathbb{E}(Z)\big) = \mathrm{O}\left(1 + \frac{1}{n} F'_n(\mathbb{R}^{d'})\right). \tag{3.5}$$

Now we introduce the assumptions. First, we suppose $\mathbf{F}(p)$ for all the values of p necessary to apply our theorems. Next, we set $\beta(t) = F(\{y\colon |y| > t\})$ for $t > 0$, and we assume that

$$t \leqslant 1 \quad \Rightarrow \quad \beta(t) \leqslant \frac{C}{t^\alpha} \tag{3.6}$$

for some constants $C > 0$ and $\alpha \in [0, 2]$: this assumption is always satisfied if $\alpha = 2$, because F integrates $x \mapsto |x|^2$ near 0, and if it holds for α_0 it also holds for any $\alpha > \alpha_0$. If this holds with $\alpha = 0$, then $\beta(0) < \infty$ and the purely discontinuous part of Y is compound Poisson. Note also that it holds for some $\alpha \in (0, 2)$ as soon as the Lévy measure F in a neighborhood of 0 is dominated by the Lévy measure of an α-stable process.

For $q \geqslant 2$ we also introduce the functions

$$\tilde\beta_q(t) = \int_{\{|y|\leqslant t\}} |y|^q F(\mathrm{d}y) = q \int_0^t s^{q-1}\big(\beta(s) - \beta(t)\big)\,\mathrm{d}s \leqslant \frac{Cq}{q-\alpha}\, t^{q-\alpha}, \tag{3.7}$$

where the last inequality holds under (3.6).

We say that we are in the *pseudo-symmetrical case* if $\int_{\{|y|\leqslant t\}} y_i y_j y_k F(\mathrm{d}y) = 0$ for all i, j, k when t is small enough (here y_i is the ith coordinate of $y \in \mathbb{R}^{d'}$; this holds e.g. when F is invariant by all rotations in $\mathbb{R}^{d'}$).

Note that, in view of (3.5), we have $\alpha_n = \mathrm{O}(1 + 1/nv_n^\alpha)$ under (3.6). Therefore the expected time necessary to perform the computation, namely $T(N, n) = \mathrm{O}(Nn\alpha_n)$, is

$$T(N, n) = \mathrm{O}\left(Nn + \frac{N}{v_n^\alpha}\right). \tag{3.8}$$

Now we want $\widehat{\mathbf{G}}(\{u_n\}, p)$. As said before, only the last part of (2.15) has to be checked. For any function h on $\mathbb{R}^{d'}$ we have $F'_n(h) - F(h) = -\int h(y) 1_{\{|y|\leqslant v_n\}} F(\mathrm{d}y)$. If $h \in C'^4_p(\mathbb{R}^{d'})$, by using a Taylor expansion of h around 0, up to order 4, we obtain

$$\big|F'_n(h) - F(h)\big| \leqslant \begin{cases} K\tilde\beta_4(v_n)\|h\|_{p,4} & \text{in the pseudo-symmetrical case,} \\ K\tilde\beta_3(v_n)\|h\|_{p,4} & \text{otherwise} \end{cases} \tag{3.9}$$

for some constant K. Hence (3.7) yields $\widehat{\mathbf{G}}(\{u_n\}, p)$ with $u_n = v_n^{4-\alpha}$ in the pseudo-symmetrical case, and $u_n = v_n^{3-\alpha}$ otherwise. Then Theorem 2.2 and Proposition 2.9 yield $e(N, n) = \mathrm{O}(1/\sqrt{N} + u_n \vee (1/n))$. So in view of minimizing (3.8) it is best to take $u_n = \mathrm{O}(1/n)$ and $N = \mathrm{O}(n^2)$. This leads to an expected time T_ε necessary to achieve an error smaller than ε which is

$$T_\varepsilon = \begin{cases} \mathrm{O}(\varepsilon^{-3}) & \text{in the pseudo-symmetrical case, or if } \alpha \leqslant \frac{3}{2}, \\ \mathrm{O}(\varepsilon^{-(6-\alpha)/(3-\alpha)}) & \text{otherwise.} \end{cases} \tag{3.10}$$

Moreover, it is noteworthy to observe that the expected number of jumps to simulate in a single interval is always smaller than 1 in the first case above.

If we want to use the Romberg method, based upon Theorem 2.2, we need the last two parts of (2.16) with $u_n = 1/n$, and for this an assumption like (3.6) is not enough, and we need an equivalent to $\beta(t)$ (or $\beta_\pm(t)$) as $t \to 0$. To keep things simple, we consider the very particular case where $d' = 1$ and the Lévy measure F satisfies

$$F(\mathrm{d}x)1_{[-v,v]} = \frac{A_+}{x^{1+\alpha}}\, 1_{(0,v)}(x) + \frac{A_-}{(-x)^{1+\alpha}}\, 1_{(-v,0)}(x), \tag{3.11}$$

where $\alpha \in (0,2)$ and $A_+, A_- \geqslant 0$ and for some number $v > 0$. This of course implies (3.6) with the same α. We already know that (2.15) holds with $u_n = v_n^{4-\alpha}$ in the pseudo-symmetrical case (corresponding to $A_+ = A_-$ here), and with $u_n = v_n^{3-\alpha}$ otherwise. Hence we take $v_n = n^{-1/(4-\alpha)}$ if $A_+ = A_-$ and $v_n = n^{-1/(3-\alpha)}$ otherwise. Then a simple calculation shows that the last assertion in (2.16) holds if $h \in C_p'^6(\mathbb{R}^{d'})$, with

$$\Phi(h) = \frac{A_+ + A_-}{24(4-\alpha)}\, h^{(4)}(0), \qquad u_n' = n^{-(6-\alpha)/(4-\alpha)}$$

if $A_+ = A_-$, and otherwise

$$\Phi(h) = \frac{A_+ - A_-}{6(3-\alpha)}\, h^{(3)}(0), \qquad u_n' = n^{-(4-\alpha)/(3-\alpha)}.$$

Then we can apply Theorem 2.2 and (3.2) to get that $e(N,n) = \mathrm{O}(1/\sqrt{N} + 1/n^{(6-\alpha)/(4-\alpha)})$ if $A_+ = A_-$, and $e(N,n) = \mathrm{O}(1/\sqrt{N} + 1/n^{(4-\alpha)/(3-\alpha)})$ otherwise. Then we take $N = \mathrm{O}(n^{(12-2\alpha)/(4-\alpha)})$ in the first case, and $N = \mathrm{O}(n^{(8-2\alpha)/(3-\alpha)})$ in the second case. This leads to an expected time T_ε necessary to achieve an error smaller than ε which is

$$T_\varepsilon = \begin{cases} \mathrm{O}(\varepsilon^{-(16-3\alpha)/(6-\alpha)}) & \text{if } A_+ = A_-, \\ \mathrm{O}(\varepsilon^{-(11-3\alpha)/(4-\alpha)}) & \text{if } A_+ \neq A_- \text{ and } \alpha \leqslant 3/2, \\ \mathrm{O}(\varepsilon^{-(8-\alpha)/(4-\alpha)}) & \text{if } A_+ \neq A_- \text{ and } \alpha > 3/2. \end{cases} \tag{3.12}$$

In all cases we have $T_\varepsilon = \varepsilon^{-\rho(\alpha)}$, and the smaller $\rho(\alpha)$ is, the better is the result. We can summarize all the results by stating the behavior of $\rho(\alpha)$ as a function of α, as follows:

α:	0		3/2		2
genuine simple	3	$\longrightarrow$	3	$\longrightarrow$	3
genuine Romberg	2.5	$\longrightarrow$	2.5	$\longrightarrow$	2.5
approximate simple, symm.	3	$\longrightarrow$	3	$\longrightarrow$	3
approximate simple, non-symm.	3	$\longrightarrow$	3	$\nearrow$	4
approximate Romberg, $A_+ = A_-$	2.66	$\searrow$	2.55	$\searrow$	2.5
approximate Romberg, $A_+ \neq A_-$	2.75	$\searrow$	2.6	$\nearrow$	3

The reader will observe that the rates of convergence are quite reasonable for the approximate scheme, compared with those for the genuine scheme. Also, the improvement of the Romberg method is not really significant.

4. The compound Poisson case

In this section we suppose that Y is a compound Poisson process. This can be expressed through its characteristics as follows: $c = 0$ and $F(\mathbb{R}^{d'}) < \infty$ and $b = \int \tau(y) F(\mathrm{d}y)$. We only consider the genuine Euler scheme, since we can simulate Y exactly in this situation. Actually, we can even simulate X^x exactly, so the result below is given only for the sake of comparison with the general result and because of the simplicity of its proof, and we restrict ourselves to the first order expansion.

In this situation, Eq. (1.3) has a unique solution X^x, with no assumption at all on the coefficient f. This is why we assume nothing like $\mathbf{H}(l,N)$ below. Observe also that there is no integrability assumption on the jumps of Y like $\mathbf{F}(p)$ below.

Theorem 4.1. *If Y is a compound Poisson process with Lévy measure F, and if $Y_t^n = Y_{\lfloor nt \rfloor / n}$ (the genuine Euler scheme), for any bounded measurable function g on $\mathbb{R}^d$ we have the expansion* (2.11) *for $N = 1$ with the operator $\Gamma_t^{(1)}$ given by*

$$\Gamma_t^{(1)}g(x) = \frac{1}{2}\int_0^t P_s H_{t-s} g(x)\,\mathrm{d}s, \tag{4.1}$$

$$H_s g(y) = \int F(\mathrm{d}u)F(\mathrm{d}v)\big(P_s g\big(y + f(y)(u+v)\big) - P_s g\big(y + f(y)u + f\big(y + f(y)u\big)v\big)\big), \tag{4.2}$$

where $(P_t)_{t\geqslant 0}$ is the transition semi-group of the process X^x, and for some constant K depending on F and f we have (below $\|g\|_\infty$ denotes the sup-norm).

$$\big|R_{1,n,t}g(x)\big| \leqslant tK\|g\|_\infty, \qquad \big|\Gamma_t^{(1)}g(x)\big| \leqslant tK\|g\|_\infty. \tag{4.3}$$

The reason such a result holds is simple enough: recall that in this case, we have $X^{n,x} \equiv X^x$ on the set A_n where there is at most one jump of Y on each interval $I_i^n = (\frac{i-1}{n}, \frac{i}{n}]$ and $\mathbb{P}(A_n) \to 1$. The set B_n on which there is exactly one interval I_i^n on which two jumps of Y occur and Y jumps at most once on all other intervals I_j^n has a probability of order $1/n$, and the complement of $A_n \cup B_n$ has a probability of order $1/n^2$. On B_n the values of $X_1^{n,x}$ and of X_1^x are possibly far apart, so $\mathbb{E}((g(X_1^{n,x}) - g(X_1^x))1_{C_n})$ is 0 when $C_n = A_n$ and of order $1/n$ when $C_n = B_n$, and of course of order $1/n^2$ if C_n is the complement of $A_n \cup B_n$ (when g is bounded).

Proof. We set $\lambda = F(\mathbb{R}^{d'})$ and $G = F/\lambda$ (a probability measure, which is the law of all jumps of Y). We denote by N_i^n the number of jumps of Y within the interval $((i-1)/n, i/n]$, and we set for $1 \leqslant i \leqslant k \leqslant n$:

$$C_{n,i} = \bigcap_{j=1}^{i}\{N_j^n \leqslant 1\}, \qquad D_{n,i,k} = \Bigg(\bigcap_{j:\ j\neq i,\ 1\leqslant j\leqslant k}\{N_j^n \leqslant 1\}\Bigg) \cap \{N_i^n = 2\}.$$

The sets $D_{n,i,k}$ for $i = 1, \ldots, k$ are pairwise disjoint, with a union denoted by $D_{n,k}$. By well-known properties of Poisson processes, $\mathbb{P}(D_{n,i,k})$ does not depend on i, and we have

$$\mathbb{P}(C_{n,n}) = 1 - \frac{\lambda^2}{2n} + \mathrm{O}\Big(\frac{1}{n^2}\Big), \quad \mathbb{P}\big((C_{n,n}\cup D_{n,k})^c\big) \leqslant \frac{K}{n^2}, \quad \mathbb{P}(D_{n,i,k}) = \frac{\lambda^2}{2n^2} + \mathrm{O}\Big(\frac{1}{n^3}\Big). \tag{4.4}$$

Set $Q_j^n g(x) = \mathbb{E}(g(X_{j/n}^x)1_{C_{n,j}})$ and

$$H_j^n g(y) = \int F(\mathrm{d}u)F(\mathrm{d}v)\big(Q_j^n g\big(y + f(y)(u+v)\big) - Q_j^n g\big(y + f(y)u + f\big(y + f(y)u\big)v\big)\big)$$

for $j = 0, 1, \ldots, n$. We deduce from the first part of (4.4) that

$$\big|P_{j/n}g(x) - Q_j^n g(x)\big| \leqslant \frac{K}{n}\|g\|_\infty, \tag{4.5}$$

hence if H_s is defined by (4.2) we get

$$\big|H_j^n g(x) - H_{j/n}g(x)\big| \leqslant \frac{K}{n}\|g\|_\infty. \tag{4.6}$$

Observe that $X_s^{n,x} = X_s^x$ for all $s \leqslant i$ on the set $C_{n,i}$, so the second part of (4.4) gives:

$$\Bigg|\Delta_{n,t}g(x) - \sum_{i=1}^{[nt]}\mathbb{E}\big(\big(g(X_{[nt]/n}^{n,x}) - g(X_{[nt]/n}^x)\big)1_{D_{n,i,[nt]}}\big)\Bigg| \leqslant \frac{K}{n^2}\|g\|_\infty. \tag{4.7}$$

Now, if we denote by W_1 and W_2 the sizes of the two jumps of Y on the interval $(\frac{i-1}{n}, \frac{i}{n}]$, when there are exactly two of them, we have on the set $D_{n,i,k}$ for any $k \geqslant i$:

$$X^x_{i/n} = X^x_{(i-1)/n} + f(X^x_{(i-1)/n})W_1 + f\big(X^x_{(i-1)/n} + f(X^x_{(i-1)/n})W_1\big)W_2,$$
$$X^{n,x}_{i/n} = X^x_{(i-1)/n} + f(X^x_{(i-1)/n})(W_1 + W_2).$$

Moreover, it is obvious that if $i < j$, then

$$\mathbb{E}\big(g(X^x_{[nt]/n})1_{D_{n,i,[nt]}}|\mathcal{F}_{i/n}\big) = Q^n_{[nt]-i}g(X^x_{i/n})1_{D_{n,i,i}},$$
$$\mathbb{E}\big(g(X^{n,x}_{[nt]/n})1_{D_{n,i,[nt]}}|\mathcal{F}_{i/n}\big) = Q^n_{[nt]-i}g(X^{n,x}_{i/n})1_{D_{n,i,i}}.$$

Therefore, taking into account the fact that conditionally on $D_{n,i,[nt]}$ the two variables W_1 and W_2 are i.i.d. with law $G = F/\lambda$, we get for $i \leqslant [nt]$:

$$\mathbb{E}\big(g(X^x_{[nt]/n})1_{D_{n,i}}\big) = \frac{\mathrm{e}^{-\lambda/n}}{2n^2}\int Q^n_{i-1}(x,\mathrm{d}y)F(\mathrm{d}u)F(\mathrm{d}v)Q^n_{[nt]-i}g\big(y + f(y)u + f\big(y + f(y)u\big)v\big),$$
$$\mathbb{E}\big(g(X^{n,x}_{[nt]/n})1_{D_{n,i}}\big) = \frac{\mathrm{e}^{-\lambda/n}}{2n^2}\int Q^n_{i-1}(x,\mathrm{d}y)F(\mathrm{d}u)F(\mathrm{d}v)Q^n_{[nt]-i}g\big(y + f(y)u + f(y)v\big).$$

Hence we have

$$\mathbb{E}\big(\big(g(X^{n,x}_{[nt]/n}) - g(X^x_{[nt]/n})\big)1_{D_{n,i}}\big) = \frac{\mathrm{e}^{-\lambda/n}}{2n^2}Q^n_{i-1}H^n_{[nt]-i}g(x). \tag{4.8}$$

Since $|\mathrm{e}^{-\lambda/n} - 1| \leqslant K/n$, the previous equality and (4.5), (4.6) and (4.7) yield

$$\left|\Delta_{n,t}g(x) - \frac{1}{2n^2}\sum_{i=1}^{[nt}P_{(i-1)/n}H_{([nt]-i)/n}g(x)\right| \leqslant \frac{K}{n^2}\,\|g\|_\infty.$$

Furthermore it is obvious that for $s \leqslant r \leqslant t$ we have $|P_sg(x) - P_rg(x)| \leqslant K\|g\|_\infty(r-s)$, hence we also have

$$\frac{i-1}{n} \leqslant s \leqslant \frac{i}{n} \quad\Longrightarrow\quad \big|P_{(i-1)/n}H_{([nt]-i)/n}g(x) - P_sH_{[nt]/n-s}g(x)\big| \leqslant \frac{K}{n}.$$

Then the sum showing up in (4.7) is in fact equal, up to a term smaller than $K\|g\|_\infty/n^2$, to the integral $\frac{1}{2n}\int_0^{[nt]/n}P_sH_{[nt]/n-s}g(x)\,\mathrm{d}s$, and the result follows. □

As said before, there is no assumption here on the size of jumps, nor on f. On the other hand, as soon as Y is not a compound Poisson, and even if it is a "compound Poisson process *with drift*", the previous result becomes wrong, and one needs at least $\mathbf{F}(1)$, because $\int(y - \tau(y))F(\mathrm{d}y)$ comes in the explicit form of the operator $\Gamma^{(1)}_t$ (see Remark 9.1), and also $\mathbf{H}(1,1)$ of course in order to have a solution to the equation. Evidently, the operator $\Gamma^{(1)}_t$ of Theorem 2.3 formally takes the expression (4.1) when $c = 0$ and $F(\mathbb{R}^{d'}) < \infty$ and $b = \int\tau(y)F(\mathrm{d}y)$.

On the other hand if g is unbounded then the two terms on the right of (2.7) might be infinite or not defined: so if we want the previous result to hold for, say, $g \in C^0_p(\mathbb{R}^d)$ (or $g \in C^k_p(\mathbb{R}^d)$ for some k; the smoothness of g makes no difference here), that is if we want (2.8) or (2.11) for $N = 1$ to hold in the situation of Theorem 4.1, then $\mathbf{F}(p)$ is required.

5. Lévy driven stochastic differential equations

In this section we gather some results on Eq. (1.3), whose solution is denoted by X^x. These results are part of the folklore of the subject, and closely related versions of them can be found in [12]. However we need somewhat more precise estimates than in that paper and we thus give quick proofs.

Below, K_α (or $K(\alpha)$) denotes a constant which may change from line to line, and depends only on the parameter α and on the dimensions d and/or d'.

1) First we need estimates on stochastic integrals w.r.t. Y. The forthcoming result is taken from [12], but we give here a simpler proof.

Lemma 5.1. *For any predictable* (*matrix-valued*) *process* H *and any* $p \geqslant 2$, *and if* $\beta_p = \int |y|^p F(\mathrm{d}y)$, *we have* (*recall that time belongs to* $[0,1]$):

$$\mathbb{E}\left(\sup_{s\leqslant t}\left|\int_0^s H_s\,\mathrm{d}Y_s\right|^p\right) \leqslant K_p\big(|b|^p + |c|^{p/2} + (\beta_2)^{p/2} + \beta_p\big)\int_0^t \mathbb{E}\big(|H_s|^p\big)\,\mathrm{d}s.$$

Proof. It is enough to prove the result when $\beta_p < \infty$. In this case, $b' = b + \int (y - \tau(y))F(\mathrm{d}y)$ exists and satisfies $|b'| \leqslant |b| + K\sqrt{\beta_2}$ for a constant K depending on the function h only, and $Y_t = b't + Y_t^c + M_t$ where M is a purely discontinuous martingale. Then it is enough to prove our inequality separately when $Y_t = b't$, and $Y = Y^c$, and $Y = M$. In the first two cases the result is well known (and easy), so we assume that $Y = M$. It is also clearly enough to consider the case where Y and H are 1-dimensional. By a Burkholder–Davis–Gundy inequality the left side of the inequality is smaller than $K_p\mathbb{E}(Z_t^{p/2})$, where $Z_t = \sum_{s\leqslant t} H_s^2\Delta Y_s^2$. So it remains to prove that $\mathbb{E}(Z_t^{p/2}) \leqslant K_p(\beta_2^{p/2} + \beta_p)a_t$, where $a_t = \int_0^t \mathbb{E}(|H_s|^p)\,\mathrm{d}s$.

Set $q = p/2 \geqslant 1$. For all $x, z \geqslant 0$ we have first $(x+z)^q - x^q \leqslant 2^{q-1}(x^{q-1}z + z^q)$, and second $x^{q-1}z \leqslant \varepsilon x^q + z^q/\varepsilon^{q-1}$ for all $\varepsilon > 0$. Hence for all $\varepsilon > 0$ and $x, u, v \geqslant 0$ we have

$$(x+uv)^q - x^q \leqslant 2^{q-1}\left(\varepsilon x^q u + \frac{1}{\varepsilon^{q-1}}uv^q + u^q v^q\right).$$

Then if $T_m = \inf(t\colon\ Z_t \geqslant m)$, and since Z is non-decreasing and purely discontinuous,

$$Z_{t\wedge T_m}^q = \sum_{s\leqslant t\wedge T_m}\big((Z_{s-} + \Delta Z_s)^q - Z_{s-}^q\big) = \int_0^{t\wedge T_m}\big((Z_{s-} + H_s^2 y^2)^q - Z_{s-}^q\big)\mu(\mathrm{d}s,\mathrm{d}y),$$

where μ is the jump measure of Y. The predictable compensator of μ is $\nu(\mathrm{d}s,\mathrm{d}y) = \mathrm{d}s\otimes F(\mathrm{d}y)$, so we get

$$\begin{aligned}\mathbb{E}(Z_{t\wedge T_m}^q) &\leqslant 2^{q-1}\mathbb{E}\left(\int_0^{t\wedge T_m}\int\left(\varepsilon Z_{s-}^q y^2 + \frac{1}{\varepsilon^{q-1}}|H_s|^{2q}y^2 + |H_s|^{2q}|y|^{2q}\right)F(\mathrm{d}y)\,\mathrm{d}s\right)\\ &\leqslant 2^{q-1}\left(\varepsilon\beta_2\mathbb{E}(m^q \wedge Z_{t\wedge T_m}^q) + \left(\frac{1}{\varepsilon^{q-1}}\beta_2 + \beta_p\right)a_t\right),\end{aligned}$$

because Z is increasing and $Z_{s-} \leqslant m$ if $s \leqslant T_m$. The right side above is finite, hence the left side as well. Then it remains to take $\varepsilon = 1/(2^q\beta_2)$ and let $m \to \infty$ and apply the monotone convergence theorem: we get the result with $K_p = 2^{q^2}$. □

For further reference, we set

$$\eta_p = |b| + |c| + \int\big(|y|^2 1_{\{|y|\leqslant 1\}} + |y|^p 1_{\{|y|>1\}}\big)F(\mathrm{d}y) \tag{5.1}$$

so $\mathbf{F}(p)$ amounts to saying that $\eta_p < \infty$. With this notation, it follows from the previous lemma that for any $p \geqslant 2$ and any predictable process H we have

$$2 \leqslant p' \leqslant p \quad\Rightarrow\quad \mathbb{E}\left(\sup_{s\leqslant t}\left|\int_0^s H_s\,\mathrm{d}Y_s\right|^{p'}\right) \leqslant K(p,\eta_p)\int_0^t \mathbb{E}\big(|H_s|^{p'}\big)\,\mathrm{d}s, \tag{5.2}$$

where $K(p,\eta_p)$ denotes a constant which depend only on p and η_p, and on the dimensions of Y and H.

2) Now we turn to estimates on the solution X^x of (1.3). We know that it is a Markov process, whose semigroup is denoted by (P_t). The following estimates on P_t are crucial (when we write $\|g\|_{p,k} < \infty$ for a function g on $\mathbb{R}^d$, this automatically implies that $g \in C^k_p(\mathbb{R}^d)$):

Proposition 5.2. (a) *Under* $\mathbf{H}(1,1)$ *and* $\mathbf{F}(2 \vee p)$ *for some* $p \geqslant 0$, *we have for some constant* $K = K(p, f, \eta_{2\vee p})$ (*recall* (1.7)):

$$\left.\begin{array}{l} \mathbb{E}\big(\sup_s |X^x_s|^p\big) \leqslant K\alpha_p(x), \\ g \in C^0_p(\mathbb{R}^d) \quad \Rightarrow \quad \|P_t g\|_{p,0} \leqslant K\|g\|_{p,0}. \end{array}\right\} \tag{5.3}$$

(b) *Under* $\mathbf{H}(1,N)$ *and* $\mathbf{F}(N + N \vee p)$ *for some* $p \geqslant 0$ *and* $N \geqslant 1$, *we have for some constant* $K = K(p, f, \eta_{N+N\vee p})$

$$g \in C^N_p(\mathbb{R}^d) \quad \Rightarrow \quad \|P_t g\|_{p,N} \leqslant K\|g\|_{p,N}. \tag{5.4}$$

The first property in (5.3) is then a consequence of (5.2) and of Gronwall's inequality (recall that $\mathbf{F}(p) \Leftrightarrow \eta_p < \infty$). The second property in (5.3) is a trivial consequence of the first one.

For (b) above we need first some facts about the differentiability of $x \mapsto X^x$. We say that it is *continuously differentiable in* $\mathbb{L}^p$ if there are $d \times d$-dimensional processes $X^{x,(1)}$ which satisfy $\mathbb{E}(\sup_s |X^{x,(1)}_s|^p) < \infty$ and also $\mathbb{E}(\sup_s |X^y_s - X^x_s - X^{x,(1)}_s \cdot (y-x)|^p) = o(|y-x|^p)$, and $\mathbb{E}(\sup_s |X^{y,(1)}_s - X^{x,(1)}_s|^p) \to 0$ as $y \to x$. By induction, it is N times continuously differentiable in $\mathbb{L}^p$ if the $(N-1)$ derivative process $X^{x,(N-1)}$ exists and is continuously differentiable in $\mathbb{L}^p$. Observe that $X^{x,N}$ is d^N-dimensional.

It is well known, using Gronwall's Lemma and (5.2), that under $\mathbf{H}(1,1)$ and $\mathbf{F}(p)$ for some $p \geqslant 2$, then X^x is once continuously differentiable in $\mathbb{L}^p$ and $X^{x,(1)}$ is the unique solution of the following linear equation (with I_d being the $d \times d$ identity matrix):

$$X^{x,(1)}_t = I_d + \int_0^t \nabla f(X^x_{s-}) X^{x,(1)}_{s-}\, \mathrm{d}Y_s \tag{5.5}$$

and further $x \mapsto \mathbb{E}(\sup_s |X^{x,(1)}_s|^p)$ is bounded. More generally, we have:

Lemma 5.3. *Under* $\mathbf{H}(1,N)$ *for some* $N \geqslant 1$ *and* $\mathbf{F}(Np)$ *for some* $p \geqslant 2$, *then* $x \mapsto X^x$ *is* N *times continuously differentiable in* $\mathbb{L}^p$, *and we have for some constant* $K = K(N, p, f, \eta_{Np})$:

$$\mathbb{E}\Big(\sup_s |X^{x,(N)}_s|^p\Big) \leqslant K. \tag{5.6}$$

Proof. Not only do we get (5.6), but we also have that the Nth derivative is the unique solution of the following linear equation (when the $X^{x,(j)}$ for $j = 1, \ldots, N-1$ are supposed to be known):

$$X^{x,(N)}_t = \int_0^t \nabla f(X^x_{s-}) X^{x,(N)}_{s-}\, \mathrm{d}Y_s + \sum_{i=2}^N \int_0^t \nabla^i f(X^x_{s-}) F_{N,i}\big(X^{x,(1)}_{s-}, \ldots, X^{x,(N-i+1)}_{s-}\big)\, \mathrm{d}Y_s, \tag{5.7}$$

if $N \geqslant 2$ (and (5.5) if $N = 1$). Here, the components of $F_{N,i}(x^{(1)}, \ldots, x^{(N-i+1)})$ are sums of terms of the form

$$\prod_{j=1}^{N-i+1} \prod_{r=1}^{\alpha_j} x^{(j),l_r}, \quad \text{where } \sum_j j\alpha_j = N, \tag{5.8}$$

and where $x^{(j),l}$ is the lth component of $x^{(j)} \in \mathbb{R}^{d^j}$, and an "empty" product equals 1.

The proof is by induction on N, using Gronwall's Lemma and Lemma 5.1, and as it is well known it boils down to proving first that by formal differentiation of (5.7) for $N-1$ we get Eq. (5.7) for N, and to proving secondly that the solution of (5.7) satisfies the estimate (5.6).

Hence we assume the result for all $N' < N$. By formally differentiating $x \mapsto X^{n,x,(N-1)}$ in Eq. (5.7) written for $N-1$ we readily get (5.7) for N, with (using matrix notation, and $F_{N-1,1}(x^{(1)}, \dots, x^{(N-1)}) = x^{(N-1)}$):

$$F_{N,i}(x^{(1)}, \dots, x^{(N-i+1)}) = x^{(1)} F_{N-1,i-1}\big(x^{(1)}, \dots, x^{(N-i+1)}\big) + \sum_{j=1}^{N-i} \frac{\partial}{\partial x^{(j)}} F_{N-1,i}\big(x^{(1)}, \dots, x^{(N-i)}\big) x^{(j+1)}.$$

Then if all $F_{N',i}$ for $N' < N$ are sums of terms like in (5.8), the same is true of $F_{N,i}$.

Next, we prove that the solution of (5.7) satisfies (5.6) (assuming again this is true for all $N' < N$). By Gronwall's Lemma and the fact that ∇f is bounded, the only thing to prove is that

$$\mathbb{E}\left(\sup_t \left| \int_0^t \big(\nabla^i f(X^x_{s-})\big) F_{N,i}\big(X^{x,(1)}_{s-}, \dots, X^{x,(N-i+1)}_{s-}\big)\, \mathrm{d}Y_s \right|^p \right) \leqslant K$$

for all $i = 2, \dots, N$ and for some constant $K = K(N, p, f, \eta_{Np})$ (in the remainder of the proof $K = K(N, p, f, \eta_{Np})$ varies from line to line). And of course, it is enough to prove that if G is any monomial like in (5.8), then

$$\mathbb{E}\left(\sup_t \left| \int_0^t \big(\nabla^i f(X^x_{s-})\big) G\big(X^{x,(1)}_{s-}, \dots, X^{x,(N-i+1)}_{s-}\big)\, \mathrm{d}Y_s \right|^p \right) \leqslant K. \tag{5.9}$$

For this we use Lemma 5.1 and the fact that $\nabla^i f$ is bounded. By (5.2) the left side of (5.9) is smaller than

$$K\, \mathbb{E}\left(\prod_{j=1}^{N+1-i} \sup_s \big|X^{x,(j)}_s\big|^{p\alpha_j} \right) \leqslant K \prod_{j=1}^{N+1-i} \Big(\mathbb{E}\Big(\sup_s \big|X^{x,(j)}_s\big|^{Np/j} \Big) \Big)^{j\alpha_j/N}$$

by Hölder inequality, since $\sum_j j\alpha_j = N$. The recurrence assumption yields that each expectation above is smaller than some constant $K(p, N, f, \eta_{Np})$, so we obtain (5.9). □

Proof of Proposition 5.2(b). Let $g \in C^N_p(\mathbb{R}^d)$. By Lemma 5.3, for any $k = 1, \dots, N$ then $x \mapsto X^x$ is k times continuously differentiable in $\mathbb{L}^{r_k}$, where $r_k = \frac{N+N\vee p}{k}$; further if $X^{x,(0)} = X^{n,x}$ and $\widetilde{X}^{x,(j)} = \sup_t |X^{x,(j)}_t|$, then with $K = K(N, p, f, \eta_{Np})$:

$$\mathbb{E}\big(\big|\widetilde{X}^{x,(j)}\big|^r\big) \leqslant \begin{cases} K\big(1 + |x|^r\big) & \text{if } j = 0 \text{ and } r \in [0, r_1], \\ K & \text{if } j = 1, \dots, N \text{ and } r \in [0, r_j]. \end{cases} \tag{5.10}$$

Then any kth partial derivative (for $k = 1, \dots, N$) of $x \mapsto g(X^x_t)$ exists (in probability), and is continuous in probability and is smaller than a sum of terms of the form

$$Z_{x,p,k,\{\alpha_j\}} = a\big(1 + \big|\widetilde{X}^{x,(0)}\big|^p\big) \prod_{j=1}^{k} \big|\widetilde{X}^{x,(j)}\big|^{\alpha_j}, \quad \text{where } \sum_{j=1}^{k} j\alpha_j = k,\ \alpha_j \in \mathbb{N},$$

with an empty product equal to 1. Then it is enough to prove that under our assumptions, each $Z_{x,p,k,\{\alpha_j\}}$ as above has $\mathbb{E}(Z_{x,p,k,\{\alpha_j\}}) \leqslant aK(1 + |x|^p)$. But Hölder's inequality and $\sum_{j=1}^k j\alpha_j = k \leqslant r_1 - p$ yield

$$\mathbb{E}(Z_{x,p,k,\{\alpha_j\}}) \leqslant a\left[\prod_{j=1}^{k} \big(\mathbb{E}\big(\big|\widetilde{X}^{x,(j)}\big|^{k/j}\big)\big)^{j\alpha_j/k} + \big(\mathbb{E}\big(\big|\widetilde{X}^{x,(0)}\big|^{r_1 p/(r_1-k)}\big)\big)^{(r_1-k)/r_1} \prod_{j=1}^{k} \big(\mathbb{E}\big(\big|\widetilde{X}^{x,(j)}\big|^{r_j}\big)\big)^{j\alpha_j/r_1} \right].$$

Then the result readily follows from (5.10). □

3) *The generator of X^x.* As is well known, the "extended generator" of the Markov process (X_t^x) is the operator A acting on C^2 functions g on $\mathbb{R}^d$ as follows (where ∇g is a row vector; τ is the truncation function):

$$Ag(x) = \nabla g(x) f(x) b + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 g}{\partial x^i \partial x^j}(x) \big(f(x) c f(x)^\star\big)^{ij} + \int F(\mathrm{d}y) \big(g\big(x + f(x)y\big) - g(x) - \nabla g(x) f(x) \tau(y)\big). \tag{5.11}$$

In the next lemma, we denote by $C_p^{N,1}(\mathbb{R}^d \times [0,1])$ the set of all families $(g_t)_{t\in[0,1]}$ of functions on $\mathbb{R}^d$ such that $g_t'(x) = \frac{\partial}{\partial t} g_t(x)$ exists and is continuous for all x, and that the functions g_t and g_t' all belong to $C_p^N(\mathbb{R}^d)$ with $\sup_t(\|g_t\|_{p,N} + \|g_t'\|_{p,N}) < \infty$.

Lemma 5.4. *Let $l = 0$ or $l = 1$ and $p \geqslant 0$ and $N \in \mathbb{N}$.*

(a) *Under $\mathbf{H}(l, 1 \vee N)$ and $\mathbf{F}(p+N)$ there is a constant $K = K(p, N, f, \eta_{p+N})$ such that*

$$g \in C_p^{N+2}(\mathbb{R}^d) \quad \Rightarrow \quad \|Ag\|_{p+2l,N} \leqslant K \|g\|_{p,N+2}.$$

(b) *Under $\mathbf{H}(l,1)$ and $\mathbf{F}(p)$, for any $(g_t) \in C_p^{2,1}(\mathbb{R}^d \times [0,1])$, the function $t \mapsto Ag_t(x)$ is continuously differentiable and its derivative is $Ag_t'(x)$.*

Proof. We prove (b) first. Observe that under our assumptions on (g_t) the partial derivatives of order 1 and 2 w.r.t. x commute with the partial derivative w.r.t. t, hence the claim readily follows from (5.11) and the dominated convergence theorem. It is even simpler to check that $\|Ag\|_{p+2l,0} \leqslant K\|g\|_{p,2}$ for some $K = K(p, f, \eta_p)$ when $g \in C_p^2(\mathbb{R}^d)$.

It remains to prove (a) when $N \geqslant 1$, and this is proved by induction on N. For example if $N = 1$ and if we denote by ∂_k the derivative w.r.t. the kth coordinate of x, we have $\partial_k Ag = A\partial_k g + A_k' g$ (by applying (b) and again the dominated convergence theorem, and using $\mathbf{H}(l,1)$), where

$$A_k' g(x) = \nabla g(x) \partial_k f(x) b + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 g}{\partial x^i \partial x^j}(x) \partial_k \big(f(x) c f(x)^\star\big)^{ij} + \int F(\mathrm{d}y) \big(\nabla g\big(x + f(x)y\big) \partial_k f(x) y - \nabla g(x) \partial_k f(x) \tau(y)\big). \tag{5.12}$$

We have seen already that $\|A\partial_k g\|_{p+2l,0} \leqslant K\|g\|_{p,N+2}$, and the same argument shows that $\|A_k' g\|_{p+2l,0} \leqslant K\|g\|_{p,N+2}$ as well: hence the result for $N = 1$. We can obviously iterate the procedure and get the result for N arbitrary; details are left to the reader. □

We denote by A^k the kth iterate of A (and A^0 is the identity). A straightforward iteration of the above result yields the

Lemma 5.5. *Let $l = 0$ or $l = 1$ and $p \geqslant 0$ and $k \geqslant 1$, and assume $\mathbf{H}(l, 1 \vee (N + 2k - 2))$ and $\mathbf{F}(p + N + 2k - 2)$ for some $N \in \mathbb{N}$.*

(a) *There is a constant $K = K(p, N, k, \eta_{p+N+2k-2})$ such that*

$$g \in C_p^{N+2k}(\mathbb{R}^d) \quad \Rightarrow \quad \|A^k g\|_{p+2lk,N} \leqslant K \|g\|_{p,N+2k}.$$

(b) *If* $(g_t) \in C_p^{2k,1}(\mathbb{R}^d \times [0,1])$, *then the function* $t \mapsto A^k g_t(x)$ *is continuously differentiable and its derivative is* $A^k g_t'(x)$.

Another very important property for us is the next one, well known in general but perhaps not under these hypotheses:

Lemma 5.6. *Let* $l = 0$ *or* $l = 1$ *and* $p \geqslant 0$.

(a) *Assume* $\mathbf{H}(l,1)$ *and* $\mathbf{F}(2 \vee (p+2l))$. *For any* $g \in C_p^2(\mathbb{R}^d)$ *we have*

$$P_t g(x) = g(x) + \int_0^t P_s A g(x)\,\mathrm{d}s. \tag{5.13}$$

In particular, the map $t \mapsto P_t g(x)$ *is differentiable and*

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t g(x) = P_t A g(x) = A P_t g(x). \tag{5.14}$$

(b) *Let* $N, N' \geqslant 0$ *and assume* $\mathbf{H}(l, 1 \vee (2N+N'))$ *and* $\mathbf{F}(2 \vee (p+2l) \vee (p+2N+N'))$. *If* $g \in C_p^{2N+2+N'}(\mathbb{R}^d)$ *we have*

$$P_t g(x) = \sum_{k=0}^{N} \frac{t^k}{k!} A^k g(x) + \frac{1}{(N)!} \int_0^t (t-s)^N P_s A^{N+1} g(x)\,\mathrm{d}s. \tag{5.15}$$

Proof. (a) An application of Itô's formula yields that the process

$$M_t = g(X_t^x) - g(x) - \int_0^t A g(X_{s-}^x)\,\mathrm{d}s$$

is a local martingale. Further Lemma 5.4 yields that $Ag \in C_{p+2l}^0(\mathbb{R}^d)$, hence $\sup_t |M_t| \leqslant K(1 + |x|^p + \sup_t |X_t^x|^{p+2l})$ for some constant K, and this quantity is integrable by (5.3). Hence M is a martingale, and taking expectations above yields (5.13). This gives that the map $t \mapsto P_t g(x)$ is first continuous, and second differentiable with derivative $P_t A g(x)$. For any given s the function $g' = P_s g$ is also in $C_p^2(\mathbb{R}^d)$ and the derivative of $t \mapsto P_{t+s} g(x) = P_t g'(x)$ at $t = 0$ is $P_s A g(x)$ and also $Ag'(x) = A P_s g(x)$, so that (5.14) holds.

(b) Observe that (5.15) for $N = 0$ is indeed (5.13), and the proof for N arbitrary is by induction. In fact, it is clearly enough to prove

$$\int_0^t (t-s)^{N-1} P_s A^N g(x)\,\mathrm{d}s = \frac{1}{N} A^N g(x) + \frac{1}{N} \int_0^t (t-s)^N P_s A^{N+1} g(x)\,\mathrm{d}s.$$

But this follows from (5.14) applied to $A^N g$, which is in $C_{p+2lN}^2(\mathbb{R}^d)$ by Lemma 5.5. □

4) *The generator of* Y. All the previous results hold of course when $d' = d$ and $f(x)$ is equal to the identity matrix for all x: we then get $X^0 = Y$: so (P_t) is the semigroup of Y, and A is replaced by the generator B of Y which acts on C^2 functions h on $\mathbb{R}^{d'}$ as follows:

$$Bh(x) = \nabla h(x) b + \frac{1}{2} \sum_{i,j=1}^{d'} \frac{\partial^2 h}{\partial y^i \partial y^j}(x) c^{ij} + \int F(\mathrm{d}y)\big(h(x+y) - h(x) - \nabla h(x)\tau(y)\big). \tag{5.16}$$

In this case, observe that in Lemma 5.4(a) we need only $\mathbf{F}(p)$: indeed $\mathbf{H}(0,N)$ is trivially fulfilled for all N; and we have $A'_k g = 0$ in (5.12) so $\partial_k Ag = A\partial_k g$, hence in order to have $Ag = Bg \in C^1_p(\mathbb{R}^{d'})$ we need only $\mathbf{F}(p)$ and $g \in C^3_p(\mathbb{R}^{d'})$, and our claim follows by a trivial induction. Therefore, with B^k denoting the kth iterate of B, Lemma 5.5 reads as follows:

Lemma 5.7. *Let $p \geqslant 0$ and $k, N \in \mathbb{N}$ with $k \geqslant 1$, and assume $\mathbf{F}(p)$.*

(a) *There is a constant $K = K(p, N, k, \eta_p)$ such that*

$$h \in C^{N+2k}_p(\mathbb{R}^{d'}) \quad \Rightarrow \quad \|B^k h\|_{p,N} \leqslant \|h\|_{p,N+2k}.$$

(b) *If $(h_t) \in C^{2k,1}_p(\mathbb{R}^{d'} \times [0,1])$, then the function $t \mapsto B^k h_t(y)$ is continuously differentiable and its derivative is $B^k h'_t(y)$.*

Similarly, Lemma 5.6 is true with $\mathbf{F}(2 \vee p)$ as the only assumption, and for all N. Therefore, using the previous lemma, we readily obtain:

Lemma 5.8. *Under $\mathbf{F}(2 \vee p)$, for any $k \in \mathbb{N}$ there is a constant $K = K(p, k, \eta_p)$ such that if $h \in C^{2k+4}_p(\mathbb{R}^{d'})$, then*

$$\left| n\big(\mathbb{E}\big(h(Y_{1/n})\big) - h(0)\big) - \sum_{i=0}^{k} \frac{1}{(i+1)!\, n^i} B^{i+1} h(0) \right| \leqslant \frac{K}{n^{k+1}} \|h\|_{p,2k+4}. \tag{5.17}$$

6. Some technical lemmas

6.1. Some consequences of the assumptions on ζ^n_1

Let us associate with ζ^n_1 its "normalized" distribution F_n, and also the vector b_n and the matrix $\tilde{c}_n$, as follows:

$$F_n(A) = n\mathbb{P}(\zeta^n_1 \in A), \quad b_n = F_n(\tau), \quad \tilde{c}_n = F_n(\tau\tau^\star). \tag{6.1}$$

By results in [7] (see Theorem VII-3-4), the convergence $Y^n \xrightarrow{\mathcal{L}} Y$ is equivalent to having

$$b_n \to b, \quad \tilde{c}_n \to \tilde{c}, \quad h \text{ bounded continuous null around } 0 \quad \Rightarrow \quad F_n(h) \to F(h). \tag{6.2}$$

We also introduce an operator B_n acting on C^1 functions h on $\mathbb{R}^{d'}$ as follows:

$$B_n(h) = \nabla h(0) b_n + \int F_n(\mathrm{d}y)\big(h(y) - h(0) - \nabla h(0)\tau(y)\big) = n\mathbb{E}\big(h(\zeta^n_1) - h(0)\big). \tag{6.3}$$

Let us also recall that $C'^k_p(\mathbb{R}^{d'})$ is the set of all functions in $C^k_p(\mathbb{R}^{d'})$ which vanish at 0, as well as their first and second derivatives. The next lemma shows in particular that $\mathbf{G}(\{u_n\}, p)$ for any $p \geqslant 2$ and any sequence $u_n \to 0$ implies (6.2).

Lemma 6.1. *If u_n is a sequence satisfying $u_n \geqslant \frac{1}{n}$ and if $p \geqslant 2$, then Assumption $\mathbf{G}(\{u_n\}, p)$ is equivalent to each one of the following two properties*:

(a) *We have $\mathbf{F}(p)$ and there is a constant K such that*

$$h \in C^4_p(\mathbb{R}^{d'}) \quad \Rightarrow \quad \big|B_n(h) - Bh(0)\big| \leqslant K u_n \|h\|_{p,4}. \tag{6.4}$$

(b) *We have* $\mathbf{F}(p)$ *and there is a constant* K *such that* (*recall* (2.14) *for* $\tilde{c}$):

$$\left.\begin{aligned} &|b_n - b| \leqslant K u_n, \quad |\tilde{c}_n - \tilde{c}| \leqslant K u_n, \\ &h \in C_p'^4(\mathbb{R}^{d'}) \quad \Rightarrow \quad \bigl|F_n(h) - F(h)\bigr| \leqslant K u_n \|h\|_{p,4}. \end{aligned}\right\} \tag{6.5}$$

Proof. First, we have

$$\begin{aligned} &B\tau(0) = b, \qquad B_n(\tau) = b_n, \\ &B(\tau\tau^\star)(0) = \tilde{c}, \qquad B_n(\tau\tau^\star) = \tilde{c}_n, \\ &h \in C_p'^4(\mathbb{R}^{d'}) \quad \Rightarrow \quad Bh(0) = F(h), \quad B_n(h) = F_n(h). \end{aligned}$$

Since the components of τ and $\tau\tau^\star$ belong to $C_p^4(\mathbb{R}^{d'})$ for all $p \geqslant 0$, we get (a) $\Rightarrow$ (b).

Next, we can rewrite $Bh(0)$ and $B_n(h)$ as follows:

$$\begin{aligned} Bh(0) &= \nabla h(0) b + \frac{1}{2} \sum_{i,j=1}^{d'} \frac{\partial^2 h}{\partial y^i \partial y^j}(0) \tilde{c}^{ij} + F(\tilde{h}), \\ B_n(h) &= \nabla h(0) b_n + \frac{1}{2} \sum_{i,j=1}^{d'} \frac{\partial^2 h}{\partial y^i \partial y^j}(0) \tilde{c}_n^{ij} + F_n(\tilde{h}), \end{aligned} \tag{6.6}$$

where

$$\tilde{h}(y) = h(y) - h(0) - \nabla h(0)\tau(y) - \frac{1}{2} \sum_{i,j=1}^{d'} \frac{\partial^2 h}{\partial y^i \partial y^j}(0) \tau^i(y) \tau^j(y). \tag{6.7}$$

Observe that there is a constant C such that $\|\tilde{h}\|_{p,4} \leqslant C \|h\|_{p,4}$ and $\nabla^i \tilde{h}(0) = 0$ for $i = 0, 1, 2$ (recall τ is C^∞ with compact support and $\tau(y) = y$ for $|y|$ small). Thus (b) $\Rightarrow$ (a).

Third, (1.6) and (6.3) yield

$$B_n(h) = n\bigl(\mathbb{E}\bigl(h(\zeta_1^n)\bigr) - h(0)\bigr) = n\delta_n(h) + n\bigl(\mathbb{E}\bigl(h(Y_{1/n})\bigr) - h(0)\bigr). \tag{6.8}$$

Combining this with (5.17) for $k = 0$ immediately yields the equivalence of $\mathbf{G}(\{u_n\}, p)$ with (a), since $u_n \geqslant 1/n$. □

In the next corollary we use the notation (to be compared with (5.1)):

$$\eta_p' = \sup_n \left(|b_n| + \int F_n(\mathrm{d}y) \bigl(|y|^2 1_{\{|y| \leqslant 1\}} + |y|^p 1_{\{|y|>1\}} \bigr) \right). \tag{6.9}$$

Corollary 6.2. *Suppose that* $\mathbf{G}(\{u_n\}, p)$ *holds for some sequence* $u_n \to 0$ *and some* $p \geqslant 2$. *Then* $\eta_p' < \infty$.

Proof. It is of course no restriction here to assume that $u_n \geqslant 1/n$. Hence we have (6.5) by the previous lemma. Since we can find a function $h \in C_p'^4(\mathbb{R}^{d'})$ such that

$$|y|^2 1_{\{|y| \leqslant 1\}} + |y|^p 1_{\{|y|>1\}} \leqslant \bigl|\tau(y)\bigr|^2 + h(y),$$

the result readily follows from (6.5) and $\mathbf{F}(p)$. □

We have seen that $\mathbf{G}(\{u_n\}, p)$ gives us an estimate on the difference $B_n(h) - Bh(0)$. Our other assumptions will give us expansions of $B_n(h)$ around $Bh(0)$:

Lemma 6.3. (a) *Under* $\mathbf{G}'(\{u_n\}, \{u'_n\}, p \vee 2)$ *for some* $p \geqslant 0$ *we have a constant* K *such that for all* $h \in C_p^6(\mathbb{R}^{d'})$ (*recall* (2.5) *for* ϕ):

$$\left.\begin{aligned} &\bigl|\phi(g)\bigr| + \bigl|B^2h(0)\bigr| \leqslant K\|h\|_{p,6}, \\ &\left|B_n(h) - Bh(0) - u_n\phi(h) - \frac{1}{2n}B^2h(0)\right| \leqslant K\left(u'_n \vee \frac{1}{n^2}\right)\|h\|_{p,6}. \end{aligned}\right\} \tag{6.10}$$

(b) *Under* $\mathbf{G}''(N, p\vee 2)$ *for some* $p \geqslant 0$ *and some* $N \geqslant 1$ *we have a constant* K *such that for all* $k = 0, \ldots, N+1$ *and* $g \in C_p^{2k+2}(\mathbb{R}^{d'})$:

$$\left.\begin{aligned} &\bigl|B^{(k)}(h)\bigr| \leqslant K\|h\|_{p,2k+2}, \\ &\left|B_n(h) - \sum_{i=1}^{k} \frac{1}{n^{i-1}i!}B^{(i)}(h)\right| \leqslant \frac{K}{n^k}\|h\|_{p,2k+2}, \end{aligned}\right\} \tag{6.11}$$

where $B^{(1)}(h) = Bh(0)$ *and* $B^{(k)}(h) = B^k h(0) + k!\,\phi_{k-1}(h)$ *for* $k \geqslant 2$.

Proof. (a) The first inequality follows from combining (2.4) and (2.5) plus the fact that $u'_n/u_n \to 0$, and from Lemma 5.7. The second inequality follows from combining (6.8) with (2.5) and (5.17) for $k = 1$.

(b) The second inequality follows from combining (6.8) with (2.6) for $k-1$ and (5.17) for k. For the first inequality, in view of Lemma 5.7 it suffices to prove that $|\phi_{k-1}(h)| \leqslant K\|h\|_{p,2k+2}$ for $k \geqslant 2$. We set $\Phi_{n,k} = \delta_n - \sum_{i=1}^{k-1} \frac{1}{n^{i+1}}\phi_i$. We know that $|\Phi_{n,k}(h)| \leqslant K\|h\|_{p,2k+2}/n^{k+1}$ and also $|\Phi_{n,k-1}(h)| \leqslant K\|h\|_{p,2k+2}/n^k$. Since $\phi_{k-1} = n^k(\Phi_{n,k-1} - \Phi_{n,k})$, the result is then obvious. □

The operators $B^{(k)}$ and ϕ above are linear, and we need to check that they commute with differentiation. This is obvious for $B^{(1)}$ by Lemma 5.7, but otherwise it needs a proof.

Lemma 6.4. *Let* $p \geqslant 0$ *and* $k \geqslant 2$ *and* $(h_t) \in C_p^{2k+2,1}(\mathbb{R}^{d'} \times [0,1])$.

(a) *Under* $\mathbf{G}'(\{u_n\}, \{u'_n\}, p \vee 2)$ *and if* $k = 2$, *the function* $t \mapsto \phi(h'_t)$ *is continuous and is the derivative of* $t \mapsto \phi(h_t)$.
(b) *Under* $\mathbf{G}''(k-1, p \vee 2)$ *the function* $t \mapsto B^{(k)}(h'_t)$ *is continuous and is the derivative of* $t \mapsto B^{(k)}(h_t)$.

Proof. We prove only (b), since for (a) the proof is similar (simpler in fact because we do not need the induction step).

In view of Lemma 5.7, it is enough to prove the result with ϕ_{k-1} instead of $B^{(k)}$, and for this we use an induction: we suppose that the result holds for all $k' \leqslant k-1$. We consider the operators $\Phi_{n,k}$ of the previous proof, with $\Phi_{n,1} = 0$.

We have $\mathbf{F}(p)$ and (2.4), thus $|Y_{1/n}|^p$ and $|\zeta_1^n|^p$ are integrable. It follows from Lebesgue's theorem that $t \mapsto \delta_n(h'_t)$ is continuous and is the derivative of $t \mapsto \delta_n(h_t)$. Then the induction hypothesis yields that

$$\left.\begin{aligned} &t \mapsto \Phi_{n,k-1}(h'_t) \quad \text{is continuous, and} \\ &\Phi_{n,k-1}(h_{t+s}) - \Phi_{n,k-1}(h_t) - \textstyle\int_0^s \Phi_{n,k-1}(h'_{t+u})\,\mathrm{d}u = 0. \end{aligned}\right\} \tag{6.12}$$

Next, (2.6) for $k-1$ yields for all t, and for some constant K:

$$\bigl|\Phi_{n,k}(h_t)\bigr| + \bigl|\Phi_{n,k}(h'_t)\bigr| \leqslant \frac{K}{n^{k+1}}. \tag{6.13}$$

Moreover, $\phi_{k-1} = n^k(\Phi_{n,k-1} - \Phi_{n,k})$. Hence we first deduce from (6.13) and $t \in [0,1]$ that

$$\bigl|\phi_{k-1}(h'_t) - \phi_{k-1}(h'_s)\bigr| \leqslant n^k\bigl|\Phi_{n,k-1}(h'_t) - \Phi_{n,k-1}(h'_s)\bigr| + \frac{2K}{n}.$$

If we use the first part of (6.12) and let first $s \to t$ and next $n \to \infty$, we deduce that $\phi_{k-1}(h'_s) \to \phi_{k-1}(h'_t)$ as $s \to t$. Second, taking account of the second part of (6.12), we see that

$$\phi_{k-1}(h_{t+s}) - \phi_{k-1}(h_t) - \int_0^s \phi_{k-1}(h'_{t+u})\,\mathrm{d}u = -n^k\biggl(\Phi_{n,k}(h_{t+s}) - \Phi_{n,k}(h_t) - \int_0^s \Phi_{n,k}(h'_{t+u})\,\mathrm{d}u\biggr).$$

By (6.13), the right side above is smaller than a constant times $1/n$. This being true for all n, we get $\phi_{k-1}(h_{t+s}) - \phi_{k-1}(h_t) - \int_0^s \phi_{k-1}(h'_{t+u})\,\mathrm{d}u = 0$: this finishes the proof. □

6.2. Proof of Proposition 2.9

In Case (a) we assume $\widehat{\mathbf{G}}(\{u_n\}, p)$ and $u_n \to 0$; in Case (b) we assume $\widehat{\mathbf{G}}'(\{u_n\}, \{u'_n\}, p)$ with $u_n \geqslant 1/n$ and $u_n \to 0$ and $u'_n/u_n \to 0$; and in both cases we suppose $p \geqslant 2$.

1) Set

$$\eta''_p = \sup_n\biggl(|b'_n| + |c'_n| + \int F'_n(\mathrm{d}y)\bigl(|y|^2 1_{\{|y|\leqslant 1\}} + |y|^p 1_{\{|y|>1\}}\bigr)\biggr).$$

Exactly as in Corollary 6.2 we see that in both cases we have $\eta''_p < \infty$. In the remainder of the proof K denotes a constant which changes from line to line and depends on p and η_p and η''_p only.

We denote by B'_n the generator of the Lévy process Y'^n. Lemmas 5.7 and 5.8 and the fact that $Y'_{1/n} = \zeta_1^n$, hence $B_n(h) = n(\mathbb{E}(h(Y'^n_{1/n})) - h(0))$, yield for $k = 0$ and $k = 1$:

$$h \in C_p^6(\mathbb{R}^{d'}) \quad \Rightarrow \quad \|B'_n h\|_{p,4} \leqslant K\|h\|_{p,6}, \tag{6.14}$$

$$h \in C_p^{4+2k}(\mathbb{R}^{d'}) \quad \Rightarrow \quad \biggl|B_n(h) - \sum_{i=0}^{k} \frac{1}{(i+1)!\,n^i}\, B_n'^{i+1} h(0)\biggr| \leqslant \frac{K}{n^{k+1}}\,\|h\|_{p,4+2k}. \tag{6.15}$$

On the other hand, similar to (6.6), we have

$$(B'_n - B)h(x) = \nabla h_x(0)(b'_n - b) + \frac{1}{2}\sum_{i,j=1}^{d'} \frac{\partial^2 h_x}{\partial y^i\,\partial y^j}(0)(\tilde{c}'^{ij}_n - \tilde{c}^{ij}) + (F'_n - F)(\widetilde{h_x}), \tag{6.16}$$

where $h_x(y) = h(x+y)$ and $\widetilde{h_x}$ is the transform of h_x given by (6.7). Then, comparing this with (6.6), and since $h \in C_p^k(\mathbb{R}^{d'})$ yields $\|\tilde{h}\|_{p,k} \leqslant C\|h\|_{p,k}$ for some constant C, we immediately deduce from (6.16) with $x = 0$ that

$$h \in C_p^4(\mathbb{R}^{d'}) \quad \Rightarrow \quad \bigl|B'_n h(0) - Bh(0)\bigr| \leqslant K u_n \|h\|_{p,4} \tag{6.17}$$

in Case (a). In Case (b) we have the same, and also

$$h \in C_p^6(\mathbb{R}^{d'}) \quad \Rightarrow \quad \bigl|B'_n h(0) - Bh(0) - u_n\phi(h)\bigr| \leqslant K u'_n \|h\|_{p,6}, \tag{6.18}$$

provided we have set (recall (2.16) for β, σ and Φ)

$$\phi(h) = \nabla h(0)\beta + \frac{1}{2}\sum_{i,j=1}^{d'} \frac{\partial^2 h}{\partial y^i\,\partial y^j}(0)\sigma^{ij} + \Phi(\tilde{h}).$$

2) In Case (a), the result is then a trivial consequence of (6.17) and (5.17) with $k=0$ and (6.15) with $k=0$ as well, applied to the equality (6.8).

3) Now we assume that we are in Case (b). The assumptions of this case imply those of Case (a), so (2) above yields $\mathbf{G}(\{u_n\}, p)$ (recall that now $u_n \geqslant 1/n$), so it remains to prove (2.5) with $u'_n \vee (u_n/n)$ instead of u'_n.

Let us consider (6.16) with some $h \in C^6_p(\mathbb{R}^{d'})$. Exactly as in Lemma 5.7, we can differentiate up to 2 times in x, and any partial derivative of the left side is given by the right side applied to the same partial derivatives of $x \mapsto h_x$ or $x \mapsto \widetilde{h}_x$, and $\widetilde{h}_x$ belongs to $C'^4_p(\mathbb{R}^{d'})$ and satisfies $\|\widetilde{h}_x\|_{p,4} \leqslant (1+|x|^p)\|h\|_{p,6}$. Then we deduce from (6.17) that

$$\big\|(B'_n - B)h\big\|_{p,2} \leqslant K u_n \|h\|_{p,6}.$$

Next, Lemma 5.7 yields

$$\big|B(B'_n - B)h(0)\big| \leqslant K u_n \|h\|_{p,6}.$$

On the other hand combining (6.14) and (6.17) gives us

$$\big|(B'_n - B)B'_n h(0)\big| \leqslant K u_n \|h\|_{p,6}$$

as well, and since $B'^2_n - B^2 = (B'_n - B)B'_n + B(B'_n - B)$ we finally get:

$$h \in C^6_p(\mathbb{R}^{d'}) \quad \Rightarrow \quad \big|B'^2_n h(0) - B^2 h(0)\big| \leqslant K u_n \|h\|_{p,6}. \tag{6.19}$$

At this point we can inject (5.17) for $k=1$ and (6.15) for $k=1$ as well into (6.8); in view of (6.18) and (6.19) we obtain:

$$\left|\delta_n(h) - \frac{u_n}{n}\phi(h)\right| \leqslant K\left(\frac{u'_n}{n} + \frac{u_n}{n^2} + \frac{1}{n^3}\right)\|h\|_{p,6},$$

and since $u_n \geqslant 1/n$ the result readily follows.

6.3. Estimates for $X^{n,x}$

Next we turn to studying the solution $X^{n,x}$ of Eq. (2.2), with Y^n given by (2.1). We first give a result similar to Lemma 5.1:

Lemma 6.5. *For any adapted (matrix-valued) process H and any $p \geqslant 2$, and if $\beta^n_p = \int |y|^p F_n(\mathrm{d}x)$, we have*

$$\mathbb{E}\left(\sup_{s\leqslant t}\left|\int_0^s H_{\varphi_n(s)}\,\mathrm{d}Y^n_s\right|^p\right) \leqslant K_p\big(|b_n|^p + (\beta^n_2)^{p/2} + \beta^n_p\big)\int_0^t \mathbb{E}\big(|H_{\varphi_n(s)}|^p\big)\,\mathrm{d}s.$$

Proof. It is enough to prove the result when $\beta^n_p < \infty$, and in the 1-dimensional case. Then $b'_n = b_n + \int (y - \tau(y))F_n(\mathrm{d}y)$ exists and satisfies $|b'_n| \leqslant |b_n| + K\sqrt{\beta^n_2}$ for a constant K depending on the function h only, and $Y^n_t = b'_n \frac{[nt]}{n} + M^n_t$, where $M^n_t = \sum_{i=1}^{[nt]} \xi^n_i$ and $\xi^n_i = \zeta^n_i - \mathbb{E}(\zeta^n_i)$. As in Lemma 5.1, the result is obvious when $Y^n_t = b'_n \frac{[nt]}{n}$, hence we can assume $Y^n = M^n$. Note that M^n is a martingale w.r.t. the filtration $(\mathcal{F}_{[nt]/n})_{t\geqslant 0}$. So we reproduce the proof of Lemma 5.1 with $Z_t = \sum_{i=1}^{[nt]} H^2_{(i-1)/n}\xi^n_i$ and $a_t = \frac{1}{n}\sum_{i=1}^{[nt]} \mathbb{E}(|H_{(i-1)/n}|^p) \leqslant \int_0^t \mathbb{E}(|H_{\varphi_n(s)}|^p)\,\mathrm{d}s$, and we have to prove again that $\mathbb{E}(Z^{p/2}_t) \leqslant K_p((\beta^n_2)^{p/2} + \beta^n_p)a_t$. With T_m as in Lemma 5.1, we get

$$\begin{aligned}\mathbb{E}(Z^q_{t\wedge T_m}) &= \mathbb{E}\left(\sum_{i=1}^{[nt]}\big((Z_{(i-1)/n} + H^2_{(i-1)/n}(\xi^n_i)^2)^q - (Z_{(i-1)/n})^q\big)1_{\{nT_m \geqslant i\}}\right)\\ &= \frac{1}{n}\int F_n(\mathrm{d}y)\mathbb{E}\left(\sum_{i=1}^{[nt]}\big((Z_{(i-1)/n} + H^2_{(i-1)/n}y^2)^q - (Z_{(i-1)/n})^q\big)1_{\{nT_m \geqslant i\}}\right),\end{aligned}$$

because the set $\{nT_m \geqslant i\} = \{nT_m > i-1\}$ is $\mathcal{F}_{(i-1)/n}$-measurable. Then we finish as in Lemma 5.1 again. □

As a consequence we get, using the notation (6.9), and similarly to (5.2):

$$2 \leqslant p' \leqslant p \quad \Rightarrow \quad \mathbb{E}\left(\sup_{s \leqslant t} \left| \int_0^s H_{\varphi_n(s)} \, \mathrm{d}Y_s^n \right|^{p'} \right) \leqslant K(p, \eta'_p) \int_0^t \mathbb{E}\big(|H_{\varphi_n(s)}|'^p \big) \, \mathrm{d}s. \tag{6.20}$$

At this point we can do for $X^{n,x}$ exactly what we have done for X^x in the previous section. First, although $X^{n,x}$ is not a Markov process, we introduce the analogue of its semigroup by putting

$$P_t^n g(x) = \mathbb{E}\big[g(X_t^{n,x}) \big]. \tag{6.21}$$

Observe that $P_t^n g(x) = P_{i/n}^n g(x)$ whenever $i \leqslant nt < i+1$. Then the analogue of Proposition 5.2 reads as:

Proposition 6.6. (a) *Under* $\mathbf{H}(1,1)$ *and* $\eta'_{2\vee p} < \infty$ *for some* $p \geqslant 0$, *we have for some constant* $K = K(p, f, \eta'_{2\vee p})$:

$$\left.\begin{array}{l} \mathbb{E}\big(\sup_s |X_s^{n,x}|^p\big) \leqslant K \alpha_p(x), \\ g \in C_p^0(\mathbb{R}^d) \quad \Rightarrow \quad \|P_t^n g\|_{p,0} \leqslant K \|g\|_{p,0}. \end{array}\right\} \tag{6.22}$$

(b) *Under* $\mathbf{H}(1,N)$ *and* $\eta'_{N+N\vee p} < \infty$ *for some* $p \geqslant 0$ *and* $N \geqslant 1$, *we have for some constant* $K = K(p, f, \eta'_{N+N\vee p})$:

$$g \in C_p^N(\mathbb{R}^d) \quad \Rightarrow \quad \|P_t^n g\|_{p,N} \leqslant K \|g\|_{p,N}. \tag{6.23}$$

Let us define the following operators A_n acting on C^1 functions:

$$\begin{aligned} A_n g(x) &= n\mathbb{E}\big(g(x + f(x)\zeta_1^n) - g(x)\big) \\ &= \nabla g(x) f(x) b_n + \int F_n(\mathrm{d}y)\big(g\big(x + f(x)y\big) - g(x) - \nabla g(x) f(x) \tau(y)\big). \end{aligned} \tag{6.24}$$

This operator obviously satisfies (by (2.2) and (6.24)):

$$P_{(i+1)/n}^n g(x) = P_{i/n}^n g(x) + \frac{1}{n} P_{i/n}^n A_n g(x). \tag{6.25}$$

So it plays the role of the generator for the process $X^{n,x}$. The proof of Lemma 5.4 holds ("uniformly" in n) in that case as well, and we can state the

Lemma 6.7. *Let* $l = 0$ *or* $l = 1$ *and* $p \geqslant 0$ *and* $N \in \mathbb{N}$.

(a) *Under* $\mathbf{H}(l, 1\vee N)$ *and* $\eta'_{(p+N)\vee 2} < \infty$ *there is a constant* $K = K(p, N, f, \eta'_{(p+N)\vee 2})$ *such that*

$$g \in C_p^{N+2}(\mathbb{R}^d) \quad \Rightarrow \quad \|A_n g\|_{p+2l,N} \leqslant K \|g\|_{p,N+2}.$$

(b) *Under* $\mathbf{H}(l,1)$ *and* $\eta'_{p\vee 2} < \infty$, *for any* $(g_t) \in C_p^{2,1}(\mathbb{R}^d \times [0,1])$ *the function* $t \mapsto A_n g_t(x)$ *is continuously differentiable and its derivative is* $A_n g'_t(x)$.

6.4. Expansion of the generators

Observe that we can write A_n in a different form. For any C^2 function g we put

$$L_x g(y) = g\big(x + f(x)y\big). \tag{6.26}$$

Then we have for $n \geqslant 1$ (recall (6.3)):

$$A_n g(x) = B_n(L_x g). \tag{6.27}$$

Note that we also have $Ag(x) = BL_x g(0)$ (see (5.11) and (5.16)). Then under $\mathbf{G}''(N, p)$, and similarly to (6.27) it is natural to set for $k = 1, \ldots, N+1$:

$$A^{(k)} g(x) = B^{(k)}(L_x g), \tag{6.28}$$

while under $\mathbf{G}'(\{u_n\}, \{u'_n\}, p)$ we set

$$\mathcal{U} g(x) = \phi(L_x g), \qquad \mathcal{V} g(x) = B^2 L_x g(0). \tag{6.29}$$

Since $B^{(1)}(g) = Bg(0)$, we see that

$$A^{(1)} = A. \tag{6.30}$$

Lemma 6.8. *Let $l = 0$ or $l = 1$ and $p \geqslant 0$ and $N \in \mathbb{N}$ and $k \geqslant 2$. Assume $\mathbf{H}(l, N \vee 1)$ and $\mathbf{G}''(k-1, (p+N) \vee 2)$.*

(a) *There is a constant K such that*

$$g \in C_p^{N+2k+2}(\mathbb{R}^d) \quad \Rightarrow \quad \|A^{(k)} g\|_{p+2(k+1)l, N} \leqslant K \|g\|_{p, N+2k+2}.$$

(b) *If $(g_t) \in C_p^{2k+2,1}(\mathbb{R}^d \times [0,1])$, then $t \mapsto A^{(k)} g_t(x)$ is continuously differentiable and its derivative is $A^{(k)} g'_t(x)$.*

Proof. We denote by ∇_x^r (resp. ∇_y^r) the rth iterate of the gradient w.r.t. x (resp. y). Let $g \in C_p^{N+2k+2}(\mathbb{R}^d)$. We clearly have for $0 \leqslant i + r \leqslant N + 2k + 2$ and $r \leqslant N$ and some constant K (which varies from line to line in this proof):

$$\left|\nabla_x^r \nabla_y^i L_x g(y)\right| \leqslant K \alpha_{p+il}(x)\left(1 + |y|^{p+r}\right) \|g\|_{p, N+2k+2}.$$

Therefore

$$r = 0, \ldots, N \quad \Rightarrow \quad \|\nabla_x^r L_x g\|_{p+N, 2k+2} \leqslant K \alpha_{p+2l(k+1)}(x) \|g\|_{p, N+2k+2}. \tag{6.31}$$

Then applying Lemma 6.4 r times, with t replaced by the component of x w.r.t. which we differentiate, we obtain that

$$\nabla_x^r B^{(k)}(L_x g) = B^{(k)}(\nabla_x^r L_x g) \tag{6.32}$$

as soon as $r \leqslant N$ and $\mathbf{G}''(k-1, p+N)$ holds. In view of (6.28), the properties (6.31) and (6.32) and (6.11) imply (a).

If further $g = g_t$ depends on $t \in [0,1]$ in a continuously differentiable way, we can add a derivation w.r.t. t above, and this derivation again commutes with $B^{(k)}$, hence with $A^{(k)}$: so we have (b). □

Remark 6.9. For the genuine Euler scheme $A^{(k)}(g) = B^k L_x g(0)$. Hence by Lemma 5.7 the above result holds with $g \in C_p^{N+2k}(\mathbb{R}^d)$ instead of $g \in C_p^{N+2k+2}(\mathbb{R}^d)$, and then $\|A^{(k)} g\|_{p+2kl, N} \leqslant K \|g\|_{p, N+2k}$.

The same proof, based on (a) of Lemmas 5.7 and 6.4, yields also the following:

Lemma 6.10. *Let $l = 0$ or $l = 1$ and $p \geqslant 0$ and $N \in \mathbb{N}$. Assume $\mathbf{H}(l, N \vee 1)$ and $\mathbf{G}'(\{u_n\}, \{u'_n\}, (p+N) \vee 2)$.*

(a) *There is a constant K such that*

$$g \in C_p^{N+6}(\mathbb{R}^d) \quad \Rightarrow \quad \|\mathcal{U} g\|_{p+6l, N} \leqslant K \|g\|_{p, N+6}, \quad \|\mathcal{V} g\|_{p+6l, N} \leqslant K \|g\|_{p, N+6}.$$

(b) *If* $(g_t) \in C_p^{6,1}(\mathbb{R}^d \times [0,1])$, *then* $t \mapsto \mathcal{U}g_t(x)$ *and* $t \mapsto \mathcal{V}g_t(x)$ *are continuously differentiable and their derivatives are* $\mathcal{U}g_t'(x)$ *and* $\mathcal{V}g_t'(x)$.

Lemma 6.11. *Let* $l = 0$ *or* $l = 1$, *and* $p \geqslant 0$, *and* $N \in \mathbb{N}$, *and assume* $\mathbf{H}(l, N \vee 1)$.

(a) *Under* $\mathbf{G}(\{u_n\}, (p+N) \vee 2)$ *there is a constant* K *such that*

$$g \in C_p^{N+4}(\mathbb{R}^d) \quad \Rightarrow \quad \|A_n g - Ag\|_{p+4l,N} \leqslant K u_n \|g\|_{p,N+4}.$$

(b) *Under* $\mathbf{G}'(\{u_n\}, \{u_n'\}, (p+N) \vee 2)$ *there is a constant* K *such that*

$$g \in C_p^{N+6}(\mathbb{R}^d) \quad \Rightarrow \quad \left\| A_n g - Ag - u_n \mathcal{U}g - \frac{1}{2n}\mathcal{V}g \right\|_{p+6l,N} \leqslant K\left(u_n' \vee \frac{1}{n^2}\right)\|g\|_{p,N+6}.$$

(c) *Under* $\mathbf{G}''(k-1, (p+N) \vee 2)$ *for some* $k \geqslant 2$ *there is a constant* K *such that*

$$g \in C_p^{N+2k+2}(\mathbb{R}^d) \quad \Rightarrow \quad \left\| A_n g - \sum_{i=1}^{k} \frac{1}{n^{i-1} i!} A^{(i)} g \right\|_{p+2l(k+1),N} \leqslant \frac{K}{n^k}\|g\|_{p,N+2k+2}.$$

Proof. (a) Let $g \in C_p^{N+4}(\mathbb{R}^d)$. Since B_n and B commute with derivations we have (as in (6.32)) for $r = 0, \ldots, N$:

$$\nabla_x^r\big(B_n(L_x g) - BL_x g(0)\big) = B_n(\nabla_x^r L_x g) - B\nabla_x^r L_x g(0).$$

We also have (6.31) with $k = 1$, hence $\|\nabla_x^r L_x g\|_{p+N,4} \leqslant K\alpha_{p+4l}(x)\|g\|_{p,N+4}$ for $r = 0, \ldots, N$. Then the result readily follows from (6.4).

(b) Let $g \in C_p^{N+6}(\mathbb{R}^d)$. Exactly as above (and using Lemma 6.10) we get

$$\begin{aligned}&\nabla_x^r\left(B_n(L_x g) - BL_x g(0) - u_n\phi(L_x g) - \frac{1}{2n}B^2 L_x g(0)\right)\\ &\quad = B_n(\nabla_x^r L_x g) - B\nabla_x^r L_x g(0) - u_n\phi(\nabla_x^r L_x g) - \frac{1}{2n}\,B^2\nabla_x^r L_x g(0).\end{aligned}$$

By (6.31) for $k = 2$, we have $\|\nabla_x^r L_x g\|_{p+N,6} \leqslant K\alpha_{p+6l}(x)\|g\|_{p,N+6}$ for $r = 0, \ldots, N$. Hence the result readily follows from (6.10).

(c) Let $g \in C_p^{N+2k+2}(\mathbb{R}^d)$. Using now Lemma 6.8, we get

$$\nabla_x^r\left(B_n - \sum_{i=1}^{k} \frac{1}{n^{i-1} i'!} B^{(i)}\right)(L_x g) = \left(B_n - \sum_{i=1}^{k} \frac{1}{n^{i-1} i!} B^{(i)}\right)(\nabla_x^r L_x g)$$

and also $\|\nabla_x^r L_x g\|_{p+N,2k+2} \leqslant K\alpha_{p+2l(k+1)}(x)\|g\|_{p,N+2k+2}$ for $r = 0, \ldots, N$. Then the result follows from (6.11). □

Now we define the operators which come in the definition of $\Gamma_t^{(k)}$ in the expansion (2.11). We set, as soon as $A^{(k)}$ is well defined:

$$D_k = \frac{1}{k!}(A^{(k)} - A^k). \tag{6.33}$$

Observe that $D_1 = 0$. By combining Lemmas 5.5 and 6.8, we readily get:

Lemma 6.12. *Let* $l = 0$ *or* $l = 1$, *and* $p \geqslant 0$, *and* $N \in \mathbb{N}$, *and assume* $\mathbf{H}(l, N + 2k - 2)$ *and* $\mathbf{G}''(k, p + N + 2k - 2)$ *for some* $k \geqslant 2$.

(a) *There is a constant K such that*

$$g \in C_p^{N+2k+2}(\mathbb{R}^d) \quad \Rightarrow \quad \|D_k g\|_{p+2(k+1)l,N} \leqslant K\|g\|_{p,N+2k+2}.$$

(b) *If $(g_t) \in C_p^{2k+2,1}(\mathbb{R}^d \times [0,1])$, then $t \mapsto D_k g_t$ is continuously differentiable and its derivative is $D_k g_t'$.*

6.5. The operators U_t, V_t and $\Gamma_t^{(k)}$

At this point we can define the operators U_t, V_t and $\Gamma_t^{(k)}$ coming in (2.10) and (2.11). First, U_t and V_t are defined as follows:

$$U_t g(x) = \int_0^t P_s \mathcal{U} P_{t-s} g(x)\,\mathrm{d}s, \qquad V_t g(x) = \int_0^t P_s(\mathcal{V} - A^2) P_{t-s} g(x)\,\mathrm{d}s. \tag{6.34}$$

For $\Gamma_t^{(k)}$ we start by defining a sequence of numbers by induction on n:

$$d_0 = 1, \qquad d_{n+1} = -\sum_{k=1}^{n+1} \frac{d_{n+1-k}}{(k+1)!}. \tag{6.35}$$

Then $\Gamma_t^{(q)}$ is defined by induction on q, starting with $\Gamma_t^{(0)} = P_t$, and setting for $q \geqslant 1$:

$$\Gamma_t^{(q)} g(x) = \sum_{k\geqslant 1,\, u,r\geqslant 0:\ k+u+r\leqslant q} (-1)^r \frac{d_{q-k-r-u}}{r!} \int_0^t \frac{\partial^{q-k-r-u}}{\partial s^{q-k-r-u}} \big(\Gamma_s^{(u)} D_{k+1} P_{t-s} A^r g(x)\big)\,\mathrm{d}s. \tag{6.36}$$

Of course one has to prove that this makes sense. For $q = 1$, the previous equation takes a simpler form:

$$\Gamma_t^{(1)} g(x) = \int_0^t P_s D_2 P_{t-s} g(x)\,\mathrm{d}s, \tag{6.37}$$

More generally, the right side of (6.36) involves the operators $\Gamma_u^{(i)}$ for $i = 0, \ldots, q-1$, so this formula is indeed an induction formula.

In order to give a precise meaning to the previous formulas, we need some prerequisites. Let $j \geqslant 1$ and $n, q \geqslant 2$. We say that an operator $Q_{u_1,\ldots,u_m}$ acting on functions over $\mathbb{R}^d$, where $u_i \in [0,1]$ and $m \geqslant 1$, is "of type $\mathrm{A}_q(n, j)$" if it is the composition (in an arbitrary order) of the operators $P_{u_i - u_{i'}}$, and j operators D_{k_i} with $2 \leqslant k_i \leqslant q$, and j' times the operator A, with $j' + k_1 + \cdots + k_j = n$.

Lemma 6.13. *Let $l = 0$ or $l = 1$, and $p \geqslant 0$, and $j \geqslant 1$ and $n, q \geqslant 2$. Let $Q_{u_1,\ldots,u_m}$ be an operator of type $A_q(n, j)$. Assume $\mathbf{H}(l, N)$ and $\mathbf{G}''(q, N + N \vee p)$ for some N.*

(a) *If $N \geqslant 2n + 2j$ there is a constant K such that*

$$g \in C_p^N(\mathbb{R}^d) \quad \Rightarrow \quad \|Q_{u_1,\ldots,u_m} g\|_{p+2l(n+j),\,N-2n-2j} \leqslant K\|g\|_{p,N}.$$

(b) *If $N \geqslant 2n + 2j + 2$, then $(u_1, \ldots, u_m) \mapsto Q_{u_1,\ldots,u_m} g(x)$ is continuously differentiable, and any one of the partial derivatives is the action over g and at point x of a linear combination of operators of type $A_q(n+1, j)$, containing exactly the same D_k's as $Q_{u_1,\ldots,u_m}$ does.*

Proof. $Q_{u_1,\ldots,u_m}$ is a product $R_s B_s R_{s-1} B_{s-1} \ldots R_1 B_1 R_0$, where each R_k is either P_{u_1} or $P_{u_i - u_{i-1}}$ or the identity, and each B_i is either some D_{k_i} (we then set $k_i' = k_i + 1$), or A (we then set $k_i = k_i' = 1$), with $k_1 + \cdots + k_s = n$

and $2 \leqslant k_i \leqslant q$ if $B_i = D_{k_i}$. We write $T_k = B_k R_{k-1} B_{k-1} \dots R_1 B_1 R_0$, and also $r_i = k'_1 + \dots + k'_i$. Observe that $n + j = r_s$. Below, the constant K will change from line to line.

Then apply repeatedly Lemmas 5.2(b), 5.4(a) and 6.12(a): first, R_0 sends $C_p^N(\mathbb{R}^d)$ continuously into $C_p^N(\mathbb{R}^d)$; then $T_1 = B_1 R_0$ sends $C_p^N(\mathbb{R}^d)$ continuously into $C_{p+2lr_1}^{N-2r_1}(\mathbb{R}^d)$ because $N + N \vee p \geqslant p + (N - 2r_1)$.

Suppose that T_i sends $C_p^N(\mathbb{R}^d)$ continuously into $C_{p+2lr_i}^{N-2r_i}(\mathbb{R}^d)$. If $i \leqslant s$, then $R_s T_s$ sends $C_p^N(\mathbb{R}^d)$ continuously into $C_{p+2lr_i}^{N-2r_i}(\mathbb{R}^d)$ as well by (5.4), because $2 \vee (N - 2r_i + (N - 2r_i)) \vee (p + 2lr_i)) \leqslant N + N \vee p$; and if $i < s$, then $T_{i+1} = B_{i+1} R_i T_i$ sends $C_p^N(\mathbb{R}^d)$ continuously into $C_{p+2lr_{i+1}}^{N-2r_{i+1}}(\mathbb{R}^d)$ by Lemmas 6.12 or 5.4. Since $Q_{u_1,\dots,u_n} = R_s T_s$ and $n + j = r'_s$, we finally get $\|Q_{u_1,\dots,u_m} g\|_{p+2l(n+j),N-2n-2j} \leqslant K \|g\|_{p,N}$.

For (b) we can apply repeatedly (5.14): with the notation above, only the operators R_i have to be differentiated w.r.t. some given u_j. Assume that $R_i = P_{u_j - u_{j-1}}$ (resp. $= P_{u_{j+1} - u_j}$). Then we differentiate R_i applied to $T_i g$, which is in $C_{p+2lr_i}^{N-2r_i}(\mathbb{R}^d)$, so we need $\mathbf{F}(p + 2lr_i + 2l)$, which holds because $2r_s + 2 = 2n + 2j \leqslant N$, and the differential is $A R_i T_i g$ (resp. $-A R_i T_i g$), which belongs to $C_{p+2lr_i+2l}^{N-2r_i-2}(\mathbb{R}^d)$; then we have to check that this differentiation commutes with the action of $R_s B_s \dots B_{i+1}$: for this we use Lemmas 5.4(b) or 6.12(b) and we need $N \geqslant 2n + 2j + 2$. Hence the partial derivative of $Q_{u_1,\dots,u_m} g(x)$ w.r.t. u_j is the sum, over all i such that R_i is as above, of the same operator except that we introduce an additional operator A or $-A$ at the i'th place. □

In a similar way, but with the help of Lemma 6.10, we get:

Lemma 6.14. *Let $l = 0$ or $l = 1$, and $p \geqslant 0$, and $N \geqslant 6$. Let $Q_{s,t} = P_s \mathcal{U} P_{t-s}$ or $Q_{s,t} = P_s(\mathcal{V} - A^2) P_{t-s}$. Assume $\mathbf{H}(l, N)$ and $\mathbf{G}'(\{u_n\}, \{u'_n\}, N + N \vee p)$ for some N.*

(a) *We have a constant K such that*

$$g \in C_p^N(\mathbb{R}^d) \quad \Rightarrow \quad \|Q_{s,t} g\|_{p+6l,N-6} \leqslant K \|g\|_{p,N}.$$

(b) *If $N \geqslant 8$, then $s \mapsto Q_{s,t} g(x)$ is continuously differentiable, and its derivative is the action over g and at a point x of an operator which sends $C_p^N(\mathbb{R}^d)$ into $C_{p+8l}^{N-8}(\mathbb{R}^d)$.*

Lemma 6.15. *Let $N \geqslant 6$ and $l = 0$ or $l = 1$ and $p \geqslant 0$, and assume $\mathbf{H}(l, N)$.*

(a) *Under $\mathbf{G}'(\{u_n\}, \{u'_n\}, N + N \vee p)$ there is a constant K such that*

$$g \in C_p^N(\mathbb{R}^d) \quad \Rightarrow \quad \|U_t g\|_{p+6l,N-6} \leqslant K t \|g\|_{p,N}, \quad \|V_t g\|_{p+6l,N-6} \leqslant K t \|g\|_{p,N}.$$

(b) *Under $\mathbf{G}''(q, N + N \vee p)$ for some $q \geqslant 1$ and $N \geqslant 6q$ the formula* (6.36) *defines an operator $\Gamma_t^{(q)}$ on $\mathcal{C}_p^N$, and there is a constant K such that*

$$g \in C_p^N(\mathbb{R}^d) \quad \Rightarrow \quad \|\Gamma_t^{(q)} g\|_{p+6lq,N-6q} \leqslant K t \|g\|_{p,N}.$$

Proof. (a) is obvious (because of the previous lemma), so we concentrate on (b). In all the proof we assume $\mathbf{H}(l, N)$ and $\mathbf{G}''(q + 1, N + N \vee p)$ with $N \geqslant 6q$, and r ranges through $\{2, \dots, q + 1\}$. Here again K changes from line to line.

(1) An operator $R_{t,v}$ is said to be "of type $\mathrm{B}_r(n, j)$" if its action over g is a linear combination of terms of the form

$$\int_0^t \mathrm{d}u_1 \dots \int_0^{u_{m-1}} \mathrm{d}u_m \, Q_{u_1,\dots,u_m,t,v} g(x) \tag{6.38}$$

(when $m = 0$ this is just $R_t = Q_t$), where each $Q_{u_1,\dots,u_m,t,v}$ is of type $\mathrm{A}_r(n', j')$ for some $n' \leqslant n$ and $j' \leqslant j$. Of course, the second argument v may be lacking, and then we just write R_t.

If $R_{t,v}$ is of type $\mathrm{B}_q(n, j)$, the previous lemma readily gives, for any $t > 0$:

$$N \geqslant 2n + 2j, \quad g \in C_p^N(\mathbb{R}^d) \quad \Rightarrow \quad \|R_{t,v}g\|_{p+2l(n+j),N-2n-2j} \leqslant Kt\|g\|_{p,N}. \tag{6.39}$$

Next, if we formally differentiate the expression (6.38), say $\Pi_{t,v}g(x)$, we get

$$\frac{\partial}{\partial v}\Pi_{t,v}g(x) = \int_0^t \mathrm{d}u_1 \cdots \int_0^{u_{m-1}} \mathrm{d}u_m \, \frac{\partial}{\partial v} Q_{u_1,\dots,u_m,t,v}g(x),$$

$$\frac{\partial}{\partial t}\Pi_{t,v}g(x) = \int_0^t \mathrm{d}u_1 \cdots \int_0^{u_{m-1}} \mathrm{d}u_m \, \frac{\partial}{\partial t} Q_{u_1,\dots,u_m,t,v}g(x) + \int_0^t \mathrm{d}u_2 \cdots \int_0^{u_{m-1}} \mathrm{d}u_m \, Q_{t,u_2,\dots,u_m,t,v}g(x).$$

Therefore the second part of the previous lemma gives us for any operator $R_{t,v}$ of type $\mathrm{B}_r(n, j)$:

$$\left.\begin{array}{l} N \geqslant 2n + 2j + 2, \quad g \in C_p^N(\mathbb{R}^d) \quad \Rightarrow \quad (t, v) \mapsto R_{t,v}g(x) \text{ is continuously} \\ \text{differentiable, and its partial derivatives are the action over } g \\ \text{and at point } x \text{ of another operator of type } \mathrm{B}_r(n + 1, j). \end{array}\right\} \tag{6.40}$$

Two other trivial facts are as follows:

$$R_{t,v} \text{ is of type } B_r(n, j) \quad \Rightarrow \quad R'_t = \int_0^t R_{s,t}\,\mathrm{d}s \text{ is of type } B_r(n, j). \tag{6.41}$$

$$R_t \text{ is of type } B_r(n, j) \text{ and } Q_{t,v} \text{ is of type } A_r(n', j') \quad \Longrightarrow \quad R_t Q_{t,v} \text{ is of type } B_r(n + n', j + j'). \tag{6.42}$$

(2) Now we prove by induction on m that

$$\Gamma_t^{(m)} \text{ is of type } B_{m+1}(2m, m) \tag{6.43}$$

for all $m = 1, \dots, q$. Observe that this is true for $m = 1$, in an obvious way, by (6.37).

Let us assume that (6.43) holds for all $m' \leqslant m - 1$, for some m between 2 and q. In order to prove (6.43) for m, and in view of (6.36), it is enough to prove that for any $k \geqslant 1$ and $i, w, r \geqslant 0$ with $i + k + r + w = m$, then the operator

$$R_t = \int_0^t \frac{\partial^i}{\partial s^i}(\Gamma_s^{(w)} D_{k+1} P_{t-s} A^r)\,\mathrm{d}s \quad \text{is of type } B_{m+1}(2m, m) \tag{6.44}$$

(recall that $\Gamma_s^{(0)} = P_s$). For $w \geqslant 1$ our induction hypothesis yields that $\Gamma_s^{(w)}$ is of type $\mathrm{B}_{w+1}(2w, w)$, hence $\Gamma_s^{(w)} D_{k+1} P_{t-s} A^r$ is of type $\mathrm{B}_{1+k\vee w}(2w + k + r + 1, w + 1)$ by (6.42); and the same is obviously true when $w = 0$. Therefore (6.40) applied repeatedly and (6.41) imply that, provided $N \geqslant 6w + 2k + 2r + 2i + 2$, then R_t is of type $\mathrm{B}_{1+k\vee w}(2w + k + r + i + 1, w + 1)$. Since the maxima of w (resp. $2w + k + r + i + 1 = m + w + 1$, resp. $6w + 2k + 2r + 2i + 2$) over our possible choices of (w, k, i, r) are achieved simultaneously and are equal to $m - 1$ (resp. $2m$, resp. $6m - 2$), and since $k \vee w \leqslant m$, we deduce from $N \geqslant 6m - 2$ that indeed (6.44) holds: hence we get (6.43) whenever $m \leqslant q$.

At this stage, (6.43) with $m = q$ and (6.39) gives the result. □

7. Proof of Theorem 2.1

Let us set for $n \geqslant 1$ and $j = 1, \ldots, n$ and $i = 0, \ldots, j$:

$$\beta_{n,i,j} g(x) = P^n_{i/n} P_{(j-i)/n} g(x),$$

and also for $i = 1, \ldots, j$:

$$\gamma_{n,i,j} g(x) = \beta_{n,i,j} g(x) - \beta_{n,i-1,j} g(x).$$

Observe that:

$$\Delta_{n,t} g(x) = \beta_{n,[nt],[nt]} g(x) - \beta_{n,[nt],0} g(x) = \sum_{i=1}^{[nt]} \gamma_{n,i,[nt]} g(x). \tag{7.1}$$

Below, we assume $\mathbf{H}(l, 4)$ for $l = 0$ or $l = 1$, and also $\mathbf{G}(\{u_n\}, 4 + 4 \vee p)$ for some $p \geqslant 0$ and for some sequence (u_n) decreasing to 0. By Corollary 6.2 we have $\eta'_{4+4\vee p} < \infty$. We also take $g \in C^4_p(\mathbb{R}^d)$. In view of (5.15) for $N = 0$ and of (6.25), a simple computation shows that

$$\gamma_{n,i,j} g(x) = \frac{1}{n} P^n_{(i-1)/n} A_n P_{(j-i)/n} g(x) - \int_0^{1/n} P^n_{(i-1)/n} P_s A P_{(j-i)/n} g(x)\, \mathrm{d}s. \tag{7.2}$$

Proposition 5.2 and Lemma 6.11 for $N = 0$ and $N' = 1$ and (6.30) yield that $\|(A_n - A) P_t g\|_{p+2l,0} \leqslant K \|g\|_{p,4}$ for all t and some constant K. Hence if

$$\begin{aligned} \gamma'_{n,i,j} g(x) &= \frac{1}{n} P^n_{(i-1)/n} A P_{(j-i)/n} g(x) - \int_0^{1/n} P^n_{(i-1)/n} P_s A P_{(j-i)/n} g(x)\, \mathrm{d}s \\ &= \int_0^{1/n} P^n_{(i-1)/n} (I - P_s) A P_{(j-i)/n} g(x)\, \mathrm{d}s, \end{aligned} \tag{7.3}$$

where I denotes the identity operator, then by virtue of (6.23) and (7.1), we clearly have

$$\left| \Delta_{n,t} g(x) - \sum_{i=1}^{[nt]} \gamma'_{n,i,[nt]} g(x) \right| \leqslant K t u_n \alpha_{p+2l}(x) \|g\|_{p,4}. \tag{7.4}$$

Next we apply (5.15) for $N = 0$ again and to the function $g' = A P_{(j-i)/n} g$, which satisfies $\|g'\|_{p+2l,2} \leqslant K \|g\|_{p,4}$ by Proposition 5.2 and Lemma 5.4, to get that $\|P_s g' - g'\|_{p+4l,0} \leqslant K \|g\|_{p,4}$ (uniformly in i and j). Using also (6.23), we readily deduce that

$$\left| \gamma'_{n,i,j} g(x) \right| \leqslant \frac{K}{n^2} \alpha_{p+4l}(x) \|g\|_{p,4},$$

and the estimate (2.8) thus follows from (7.4).

8. Proof of Theorem 2.2

Let us state first a (probably) well known lemma, which shows in particular how the constants d_n of (6.35) come into the picture through expansions of some integrals. This lemma is a simple variation on Taylor's formula and its proof is left to the reader.

Lemma 8.1. *Let $M \in \mathbb{N}$, and h be an $M+1$ times differentiable function over $[0,1]$, whose derivatives of order $0, 1, \ldots, M+1$ are all bounded by a constant ρ. Then we have for all $t \in [0,1]$:*

$$\left| \frac{1}{n} \sum_{i=1}^{[nt]} h\left(\frac{i-1}{n}\right) - \sum_{r=0}^{M} \frac{d_r}{n^r} \int_0^{[nt]/n} h^{(r)}(s)\, \mathrm{d}s \right| \leqslant \frac{2Mt\rho}{n^{M+1}}. \tag{8.1}$$

Now we assume $\mathbf{H}(l,10)$ and $\mathbf{G}'(\{u_n\}, \{u'_n\}, 10 + 10 \vee p)$ for some $p \geqslant 0$. Recall that u_n and u'_n/u_n go to 0. Take a function $g \in C^{10}_p(\mathbb{R}^d)$. For simplicity we also write $u''_n = u'_n \vee u_n^2 \vee \frac{1}{n^2}$.

We still have (7.1) and (7.2). By Proposition 5.2 and Lemma 6.11(b) we have for some constant K (which below will change from to line) that $\|(A_n - A - u_n\mathcal{U} - \frac{1}{2n}\mathcal{V})P_t g\|_{p+6l,0} \leqslant K u''_n \|g\|_{p,10}$. Hence, if instead of (7.3) we set

$$\gamma'_{n,i,j} g(x) = \int_0^{1/n} P^n_{(i-1)/n} \left((I - P_s)A + u_n\mathcal{U} + \frac{1}{2n}\mathcal{V} \right) P_{(j-i)/n} g(x)\, \mathrm{d}s,$$

by virtue of (6.23) and (7.1), we get

$$\left| \gamma_{n,i,[nt]} g(x) - \gamma'_{n,i,[nt]} g(x) \right| \leqslant \frac{K u''_n}{n} \alpha_{p+6l}(x) \|g\|_{p,10}. \tag{8.2}$$

Next we apply (5.15) with $N' = 0$ and $N = 1$, and Lemma 5.5(b) with $k = 2$ and $N = 0$, to the function $AP_{(j-i)/n}g$, which satisfies $\|AP_{(j-i)/n}g\|_{p+2l,4} \leqslant K\|g\|_{p,10}$ (by Proposition 5.2 and Lemma 5.4), to get that $\|(P_s - I - sA)AP_{(j-i)/n}g\|_{p+6l,0} \leqslant K\|g\|_{p,10}$. Using also (6.23), we readily deduce that if

$$\begin{aligned} \gamma''_{n,i,j} g(x) &= \int_0^{1/n} P^n_{(i-1)/n} \left(u_n\mathcal{U} + \frac{1}{2n}\mathcal{V} - sA^2 \right) P_{(j-i)/n} g(x)\, \mathrm{d}s \\ &= \frac{1}{n}\, P^n_{(i-1)/n} \left(u_n\mathcal{U} + \frac{\mathcal{V} - A^2}{2n} \right) P_{(j-i)/n} g(x), \end{aligned}$$

then

$$\left| \gamma'_{n,i,[nt]} g(x) - \gamma''_{n,i,[nt]} g(x) \right| \leqslant \frac{K}{n^3} \alpha_{p+6l}(x) \|g\|_{p,10}. \tag{8.3}$$

Next we apply (5.15) again with $N = 0$, and Lemma 5.5(b) with $k = 1$, to get that

$$\|P_{(j-i)/n} g - P_{(j-i+1)/n} g\|_{p+2l,6} \leqslant \frac{K}{n} \|g\|_{p,10}.$$

Since by Lemmas 5.5 and 6.8 the operators $\mathcal{U}$ and $\mathcal{V} - A^2$ send $C^6_{p+2l}(\mathbb{R}^d)$ continuously into $C^0_{p+8l}(\mathbb{R}^d)$, and by (6.23), we obtain

$$\left| \gamma''_{n,i,j} g(x) - \frac{1}{n} P^n_{(i-1)/n} \left(u_n\mathcal{U} + \frac{\mathcal{V} - A^2}{2n} \right) P_{(j-i+1)/n} g(x) \right| \leqslant \frac{K u''_n}{n} \alpha_{p+8l}(x) \|g\|_{p,10}. \tag{8.4}$$

Next we observe once more that the $\|\cdot\|_{p+4l,4}$ norms of the functions $\mathcal{U}P_s g$ and $(\mathcal{V} - A^2)P_s g$ are smaller than $K\|g\|_{p,10}$: we can apply Theorem 2.1 to these functions and, since u_n^2 and $1/n^2$ are smaller than u''_n, we get

$$\begin{aligned} &\left| P^n_{(i-1)/n} \left(u_n\mathcal{U} + \frac{\mathcal{V} - A^2}{2n} \right) P_{(j-i+1)/n} g(x) - P_{(i-1)/n} \left(u_n\mathcal{U} + \frac{\mathcal{V} - A^2}{2n} \right) P_{(j-i+1)/n} g(x) \right| \\ &\quad \leqslant K u''_n \alpha_{p+8l}(x) \|g\|_{p,10}. \end{aligned}$$

Then putting this together with (7.1), (8.2), (8.3) and (8.4) gives us, with the notation $G_{x;t}(s) = P_s\mathcal{U}P_{t-s}g(x)$ and $H_{x;t}(s) = P_s(\mathcal{V} - A^2)P_{t-s}g(x)$:

$$\left|\Delta_{n,t}g(x) - \frac{u_n}{n}\sum_{i=1}^{[nt]} G_{x;[nt]/n}\left(\frac{i-1}{n}\right) - \frac{1}{2n^2}\sum_{i=1}^{[nt]} H_{x;[nt]/n}\left(\frac{i-1}{n}\right)\right| \leqslant tKu_n''\alpha_{p+8l}(x)\|g\|_{p,10}. \tag{8.5}$$

It remains to observe that the two functions $G_{x;t}$ and $H_{x;t}$ (on the interval $[0,t]$) satisfy the assumptions of Lemma 8.1 with $M = 0$ and for some constant ρ, by virtue of Lemma 6.14(b). Then with the notation (6.34) we readily deduce (2.10) from the previous inequality and the fact that $u_n/n \leqslant u_n''$.

9. Proof of Theorem 2.3

The proof of this theorem is similar to the proof of the previous one, except that here we need an induction on N, after observing that the result for $N = 0$ is nothing else than Theorem 2.1.

So below we assume that $N \geqslant 1$ and that $\mathbf{H}(l, 6N+4)$ and $\mathbf{G}(N, 6N+4+(6N+4)\vee p)$ hold, and we take $g \in C_p^{6N+4}(\mathbb{R}^d)$. We also assume that the expansion (2.11) with a remainder satisfying (2.13) holds for all integers from 0 up to $N-1$. The claims concerning the operators $\Gamma_t^{(k)}$ are in Lemma 6.15, so we concentrate on the expansion.

We still have (7.1) and (7.2). By Proposition 5.2 and Lemma 6.11(b), and since $3 + N \leqslant 6N + 4$, we have

$$\left\|\left(A_n - \sum_{k=1}^{N+1}\frac{1}{n^{k-1}\,k!}A^{(k)}\right)P_t g\right\|_{p+2l(N+2),0} \leqslant \frac{K}{n^{N+1}}\|g\|_{p,6N+4}$$

for all t, for some constant K (which again changes from line to line). Hence, recalling that $A^{(1)} = A$, and if instead of (7.3) we set

$$\gamma'_{n,i,j}g(x) = \int_0^{1/n} P^n_{(i-1)/n}\left((I - P_s)A + \sum_{k=2}^{N+1}\frac{1}{n^{k-1}\,k!}A^{(k)}\right)P_{(j-i)/n}g(x)\,\mathrm{d}s,$$

by virtue of (6.23) and (7.1), we get

$$\left|\gamma_{n,i,[nt]}g(x) - \gamma'_{n,i,[nt]}g(x)\right| \leqslant \frac{K}{n^{N+2}}\,\alpha_{p+2l(N+2)}(x)\|g\|_{p,6N+4}. \tag{9.1}$$

Next we apply (5.15) with $N' = 0$, to the function $AP_{(j-i)/n}g$: taking advantage of Proposition 5.2 and Lemma 5.4, we get that

$$\left\|\left(P_s - I - \sum_{k=1}^{N}\frac{s^k}{k!}A^k\right)AP_{(j-i)/n}g\right\|_{p+2l(N+2),0} \leqslant Ks^{N+1}\|g\|_{p,6N+4}.$$

Using also (6.23) and the notation D_k of (6.33), we readily deduce that if

$$\begin{aligned}\gamma''_{n,i,j}g(x) &= \int_0^{1/n} P^n_{(i-1)/n}\left(\sum_{k=2}^{N+1}\frac{1}{n^{k-1}\,k!}A^{(k)} - \sum_{k=1}^{N}\frac{s^k}{k!}A^{k+1}\right)P_{(j-i)/n}g(x)\,\mathrm{d}s\\ &= \frac{1}{n}P^n_{(i-1)/n}\left(\sum_{k=1}^{N}\frac{1}{n^k}D_{k+1}\right)P_{(j-i)/n}g(x),\end{aligned}$$

then

$$\left|\gamma'_{n,i,[nt]}g(x) - \gamma''_{n,i,[nt]}g(x)\right| \leqslant \frac{K}{n^{N+2}}\alpha_{p+2l(N+2)}(x)\|g\|_{p,6N+4}. \tag{9.2}$$

We easily deduce from (5.15) that, under the same assumptions than in Lemma 5.6(b), we have

$$g(x) = \sum_{k=0}^{N}(-1)^k\frac{t^k}{k!}P_tA^kg(x) - \frac{1}{(N)!}\int_0^t(2t-s)^NA^{N+1}P_sg(x)\,\mathrm{d}s. \tag{9.3}$$

(For checking this, we use $P_tA^k = A^kP_t$ and we can replace P_tg in the kth summand above by the right side of (5.15) written for $N-k$ instead of N and then compute the right side of (6.21); a repeated use of the binomial formula gives us that this right side equals $g(x)$.)

Coming back to our problem, we deduce from (9.3) that the function $\psi_{n,N-k} = g - \sum_{r=0}^{N-k}(-1)^k\frac{1}{n^r r!}P_{1/n}A^rg$, for any $k \leqslant N$, satisfies

$$\|\psi_{n,N-k}\|_{p+2l(N-k+1),4N+2k+2} \leqslant \frac{K}{n^{N-k+1}}\|g\|_{p,6N+4},$$

and so does the function $\phi_{i,j,n,N-k} = P_{(j-i)/n}\psi_{n,N-k}$. Hence

$$\|P^n_{(i-1)/n}D_{k+1}\phi_{i,j,n,N-k}\|_{p+2l(N+2),4N-2} \leqslant \frac{K}{n^{N-k+1}}\ \|g\|_{p,6N+4}$$

by Lemmas 6.6 and 6.12, and if

$$\gamma'''_{n,i,j}\,g(x) = \frac{1}{n}\sum_{k=1}^{N}\sum_{r=0}^{N-k}\frac{(-1)^r}{n^{k+r}r!}P^n_{(i-1)/n}D_{k+1}P_{(j-i+1)/n}A^rg(x)$$

we get

$$\left|\gamma''_{n,i,[nt]}g(x) - \gamma'''_{n,i,[nt]}g(x)\right| \leqslant \frac{K}{n^{N+2}}\alpha_{p+2l(N+2)}(x)\|g\|_{p,6N+4}. \tag{9.4}$$

Now we apply the induction hypothesis. Observe that if $k+r \leqslant N$ the function $g' = D_{k+1}P_tA^rg$ satisfies $\|g'\|_{p+2l(k+r+1),6N-2k-2r} \leqslant K\|g\|_{p,6N+4}$ by Lemma 6.12. So our assumptions and the fact that $6(N-k-r)+4 \leqslant 6N-2k-2r$ (remember that $k+r \geqslant 1$) allow us to apply the expansion (2.11) to this function at the order $N-k-r$, which gives

$$\left|\left(P^n_{(i-1)/n} - P_{(i-1)/n} - \sum_{u=1}^{N-k-r}\frac{1}{n^u}\Gamma^{(u)}_{(i-1)/n}\right)D_{k+1}P_{(j-i+1)/n}A^rg(x)\right|$$
$$\leqslant \frac{K}{n^{N-k-r+1}}\alpha_{p+2l(2N+3-k-r)}(x)\|g\|_{p,6N+4}.$$

Henceforth, if we set $\Gamma^{(0)}_t = P_t$ and

$$\xi^{k,r,u}_{n,i,j}\,g(x) = \Gamma^{(u)}_{(i-1)/n}D_{k+1}P_{(j-i+1)/n}A^rg(x)$$

for $1 \leqslant k \leqslant N$ and $0 \leqslant r \leqslant N-k$ and $0 \leqslant u \leqslant N-k-r$, then

$$\left|\gamma'''_{n,i,[nt]}g(x) - \frac{1}{n}\sum_{k=1}^{N}\sum_{r=0}^{N-k}\sum_{u=0}^{N-r-k}\frac{(-1)^r}{n^{k+r+u}r!}\xi^{k,r,u}_{n,i,[nt]}g(x)\right| \leqslant \frac{K}{n^{N+2}}\,\alpha_{p+4l(N+1)}(x)\|g\|_{p,6N+4}. \tag{9.5}$$

In other words, if we fix t and introduce the functions

$$g_{x;k,r,u,t}(s) = \Gamma^{(u)}_s D_{k+1}P_{t-s}A^rg(x),$$

by putting together (7.1), (9.1), (9.2), (9.4) and (9.5), we obtain

$$\left|\Delta_{n,t}g(x)-\sum_{k=1}^{N}\sum_{r=0}^{N-k}\sum_{u=0}^{N-r-k}\frac{(-1)^r}{n^{k+r+u}r!}\frac{1}{n}\sum_{i=1}^{[nt]}g_{x;k,r,u,t}\left(\frac{i-1}{n}\right)\right|\leqslant\frac{Kt}{n^{N+1}}\alpha_{p+4l(N+1)}(x)\|g\|_{p,6N+4}. \tag{9.6}$$

Now, by (6.43) $g_{x;k,r,u,t}(s)$ is the action over g and at point x of an operator of type $\mathrm{B}_{1+u\vee k}(2u+k+r+1, u+1)$. Hence (6.40) applied repeatedly and the fact that $6N+4\geqslant 2(2u+k+r+1)+2(u+1)+2(N-k-r-u)$ for all u,k,r with $k\geqslant 1$ and $k+r+u\leqslant N$ show that the function $g_{x;k,r,u,t}$ is differentiable up to order $N-k-r-u+1$, with all partial derivatives up to that order being bounded by $K\alpha_{p+4l(N+1)}(x)\|g\|_{p,6N+4}$. Then Lemma 8.1 applied to $M=N-k-r-u$ gives:

$$\left|\frac{1}{n}\sum_{i=1}^{[nt]}g_{x;k,r,u,t}\left(\frac{i-1}{n}\right)-\sum_{v=0}^{N-k-r-u}\frac{d_v}{n^v}\int_0^{[nt]/n}g^{(v)}_{x;k,r,u,t}(s)\,\mathrm{d}s\right|\leqslant\frac{K\alpha_{p+4l(N+1)}(x)}{n^{N+1-k-r-u}}\|g\|_{p,6N+4}.$$

Injecting this into (9.6) gives

$$\left|\Delta_{n,t}g(x)-\sum_{k=1}^{N}\sum_{r=0}^{N-k}\sum_{u=0}^{N-r-k}\sum_{v=0}^{N-k-r-u}\frac{(-1)^r d_v}{n^{k+r+u+v}r!}\frac{1}{n}\int_0^{[nt]/n}g^{(v)}_{x;k,r,u,t}(s)\,\mathrm{d}s\right|$$
$$\leqslant\frac{Kt}{n^{N+1}}\alpha_{p+4l(N+1)}(x)\|g\|_{p,6N+4}.$$

At this point, it suffices to use the definition (6.36) of $\Gamma_t^{(q)}$ and to reorder the sums above: we readily get (2.11) and (2.13), and we are done.

Remark 9.1. We can compute "explicitly" the operators $\Gamma_t^{(k)}$, although this becomes incredibly tedious when k grows. For example, in the 1-dimensional case ($r=d=1$), and using (6.37) and the fact that

$$\Gamma_t^{(1)}g(x)=\frac{1}{2}\int_0^t P_s D_2 P_{t-s}g(x)\,\mathrm{d}s,$$

where Π is the following operator:

$$\begin{aligned}D_2g(x)=&-b^2(g'ff')(x)-\frac{bc}{2}(4g''f^2f'+g'f^2f'')(x)-\frac{c^2}{2}(2g'''f^3f'+g''f^3f''+g''f^2f'^2)(x)\\
&-bf(x)f'(x)\int F(\mathrm{d}y)\big(g'\big(x+f(x)y\big)y-g'(x)\tau(y)\big)\\
&-b\int F(\mathrm{d}y)\big(g'\big(x+f(x)y\big)\big(f\big(x+f(x)y\big)-f(x)\big)-g'(x)f(x)f'(x)\tau(y)\big)\\
&-cf^2(x)f'(x)\int F(\mathrm{d}y)\big(g''\big(x+f(x)y\big)y-g''(x)\tau(y)\big)\\
&-\frac{c}{2}f^2(x)f''(x)\int F(\mathrm{d}y)\big(g'\big(x+f(x)y\big)y-g'(x)\tau(y)\big)\\
&-\frac{c}{2}f^2(x)f'(x)^2\int F(\mathrm{d}y)g''\big(x+f(x)y\big)y^2\\
&-\frac{c}{2}\int F(\mathrm{d}y)\big(g''\big(x+f(x)y\big)\big(f\big(x+f(x)y\big)^2-f(x)^2\big)\big)-2g''(x)f(x)^2f'(x)\tau(y)\big)\end{aligned}$$

$$
\begin{aligned}
&- \int F(\mathrm{d}y) \int F(\mathrm{d}y') \big(g\big(x + f(x)y' + f\big(x + f(x)y'\big)y\big) - g\big(x + f(x)y' + f(x)y\big) \\
&- g'\big(x + f(x)y'\big)\big(f\big(x + f(x)y'\big) - f(x)\big)\tau(y) \\
&- \big(g'\big(x + f(x)y\big)y - g'(x)\tau(y)\big) f(x) f'(x)\tau(y')\big).
\end{aligned}
$$

If we are in the compound Poisson case (i.e. $c = 0$ and $F(\mathbb{R}) < \infty$ and $b = \int \tau(y) F(\mathrm{d}y)$), we see that $D_2 P_s g = H_s g$ where H_s is defined by (4.2), as it should be.

On the other hand, as soon as $b \neq \int \tau(y) F(\mathrm{d}y)$, and even if $F(\mathbb{R}) < \infty$, then $D_2 g$ is well defined only under $\mathbf{F}(1)$, and we even need $\mathbf{F}(2)$ if further $c \neq 0$. So, although this is no true proof, it seems quite unlikely that Theorems 2.2 and 2.3 stay true when we drop all integrability assumption on the jumps of Y, even when g is bounded.

References

[1] S. Asmussen, J. Rosiński, Approximations of small jumps of Lévy processes with a view towards simulation, J. Appl. Probab. 38 (2001) 482–493.
[2] V. Bally, D. Talay, The law of the Euler scheme for stochastic differential equations (I): convergence rate of the distribution function, Probab. Theory Related Fields 104 (1996) 43–60.
[3] S.N. Ethier, M.F. Norman, Error estimate for the diffusion approximation of the Wright–Fisher model, Proc. Nat. Acad. Sci. USA 74 (11) (1977) 5096–5098.
[4] S.N. Ethier, T.G. Kurtz, Characterization and Convergence, Wiley, New York, 1986.
[5] J. Jacod, P. Protter, Asymptotic error distributions for the Euler method for stochastic differential equations, Ann. Probab. 26 (1998) 267–307.
[6] J. Jacod, The Euler scheme for Lévy driven stochastic differential equations, Prépublication LPMS 711, 2002.
[7] J. Jacod, A.N. Shiryaev, Limit Theorems for Stochastic Processes, second ed., Springer-Verlag, Heidelberg, 2003.
[8] A. Kohatsu-Higa, N. Yoshida, On the simulation of some functionals of solutions of Lévy driven sde's, Preprint, 2001.
[9] T.G. Kurtz, P. Protter, Wong–Zakai corrections, random evolutions, and simulation schemes for SDEs, in: E. Mayer-Wolf, E. Merzbach, A. Shwartz (Eds.), Stochastic Analysis, Academic Press, Boston, MA, 1991, pp. 331–346.
[10] T.G. Kurtz, P. Protter, Weak error estimates for simulation schemes for SDEs, 1999.
[11] G.C. Papanicolaou, W. Kohler, Asymptotic theory of mixing stochastic ordinary differential equations, Comm. Pure Appl. Math. 27 (1974) 641–668.
[12] P. Protter, D. Talay, The Euler scheme for Lévy driven stochastic differential equations, Ann. Probab. 25 (1997) 393–423.
[13] J. Rosiński, Series representations of Lévy processes from the perspective of point processes, in: O.E. Barndorff-Nielsen, T. Mikosch, S.I. Resnick (Eds.), Lévy Processes – Theory and Applications, Birkhäuser, Boston, 2001, pp. 401–415.
[14] S. Rubenthaler, Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process, Stochastic Process. Appl. 103 (2003) 311–349.
[15] L. Słominski, Stability of strong solutions of stochastic differential equations, Stochastic Process. Appl. 31 (1989) 173–202.
[16] D. Talay, L. Tubaro, Expansion of the global error for numerical schemes solving stochastic differential equations, Stochastic Anal. Appl. 8 (1990) 94–120.