

HIGH-ORDER APPROXIMATION OF STOCHASTIC DIFFERENTIAL EQUATIONS IN TOTAL VARIATION

Clément Rey ¹

ABSTRACT. We present a method for approximating solutions of Stochastic Differential Equations (SDEs) with an arbitrary weak order of convergence. This approximation is constructed for bounded and measurable test functions. Specifically, we show that, by exploiting the standard weak approximation properties of numerical schemes for smooth test functions (such as first-order weak convergence for the Euler scheme) we can achieve convergence for simply bounded and measurable test functions at any desired rate. This is done by constructing a tailored approximation for the semigroup of the SDE obtained by evaluating the scheme (e.g., Euler) on a random time grid. Convergence is established by relying on the regularization properties of the scheme, which hold under a weak uniform Hörmander condition.

Keywords : Monte Carlo methods, Numerical methods for SDE, Discrete time Markov processes, Hörmander properties, Limit theorems for bounded measurable test functions.

AMS MSC 2020: 60J05, 65C05, 65C20, 35H10, 60F17.

CONTENTS

1.	Introduction	1
2.	The distance between semigroups	3
3.	Proof of Theorem 2.1	8
	References	15

1. INTRODUCTION

In this article, we focus on the approximation in total variation of a \mathbb{R}^d -valued ($d \in \mathbb{N}^*$) random variable X_T , $T \geq 0$ where $(X_t)_{t \geq 0}$ is a solution to the inhomogeneous Stochastic Differential Equation (SDE)

$$(1.1) \quad X_t = x_0 + \int_0^t V_0(X_s, s) ds + \sum_{i=1}^N \int_0^t V_i(X_s, s) dW_s^i, \quad t \geq 0, x_0 \in \mathbb{R}^d,$$

where $((W_t^i)_{t \geq 0}, i \in \{1, \dots, N\})$ are $N \in \mathbb{N}^*$ independent one-dimensional standard Brownian motions and $V_i \in \mathcal{C}_b^\infty(\mathbb{R}^d; \mathbb{R}^d)$, $i \in \{0, \dots, N\}$. When for every $i \in \{0, \dots, N\}$, V_i does not depend on the time, *i.e.* its second variable, the SDE is said to be homogeneous.

Let $\nu > 0$ denote the desired weak rate of convergence. For any $T > 0$ and sufficiently large $n \in \mathbb{N}^*$, we demonstrate that for a specific approximation functional operator $\hat{Q}_{0,T}^n$, there exists $C > 0$ such that for every measurable and bounded function f ,

$$(1.2) \quad \sup_{x \in \mathbb{R}^d} |\mathbb{E}[f(X_T) | X_0 = x] - \hat{Q}_{0,T}^n f(x)| \leq \frac{C}{n^\nu} \|f\|_\infty,$$

where $\|f\|_\infty = \sup_{x \in \mathbb{R}^d} |f(x)|$ is the supremum norm of f . The operator $\hat{Q}_{0,T}^n$, whose construction will be made explicit below, provides a weak approximation of order ν for the semigroup associated with $(X_t)_{t \geq 0}$. Moreover, $\hat{Q}_{0,T}^n f(x)$ can be evaluated using a Monte Carlo estimator, such that the computational cost of simulating each sample is of order n , comparable to that of an Euler-Maruyama scheme with time step T/n .

The algorithm employed to construct $\hat{Q}_{0,T}^n$ is derived from the one that was presented in [3] and is based on a combination of discrete time approximations $X_T^{\Pi_h}$ of (1.1) which are built using random time grids Π_h , $h \in \{1, \dots, r\}$, for $r \in \mathbb{N}^*$ deterministic (refer to [3], Section 3.1.3 for the construction of the

¹address: CMAP, École Polytechnique, Institut Polytechnique de Paris, Route de Saclay, 91120 Palaiseau, France
e-mail: clement.rey@polytechnique.edu

random time grids). More specifically, given h , we consider a sequence of independent random variables $Z_j^h \in \mathbb{R}^N$, $j \in \mathbb{N}^*$, and we assume that Z_j^h are centered with covariance one and independent of Π_h . We construct the \mathbb{R}^d -valued process $(X_t^{\Pi_h})_{t \in \Pi_h}$ using the generic scheme

$$(1.3) \quad X_{t_{j+1}^h}^{\Pi_h} = \psi(X_{t_j^h}^{\Pi_h}, t_j^h, (t_{j+1}^h - t_j^h)^{\frac{1}{2}} Z_{j+1}^h, t_{j+1}^h - t_j^h), \quad X_0^\delta \in \mathbb{R}^d$$

for $\Pi_h = \{t_j^h, j \in \mathbb{N}\}$ with

$$\psi \in \mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}^N \times [0, 1]; \mathbb{R}^d) \quad \text{and} \quad \forall (x, t) \in \mathbb{R}^d \times \mathbb{R}_+, \psi(x, t, 0, 0) = x.$$

Considering for instance the Euler scheme of (1.1), we have

$$\psi(x, t, z, y) = x + V_0(x, t)y + \sum_{i=1}^N V_i(x, t)z^i$$

The Euler example is standard but many other scheme such as the one developed *e.g.* in [16] or [20]. In this construction Π_h is supposed independent of $((Z_j^h)_{j \in \mathbb{N}^*})_{h \in \{1, \dots, r\}}$. The time $t_{j+1}^h - t_j^h$ between two successive discrete time values is chosen randomly in $\{\frac{T}{n^l}, l \in \mathbb{N}\}$. It is shown in [3] that we can use the following representation formula: For every $x \in \mathbb{R}^d$,

$$(1.4) \quad \hat{Q}_{0,T}^n f(x) = \sum_{h=1}^r c_h \mathbb{E}[f(X_T^{\Pi_h}) | X_0^{\Pi_h} = x]$$

where the coefficients $c_h \in \mathbb{R}$, $r \in \mathbb{N}$ and the law of Π_h are given explicitly depending on the value of ν in (1.2). This representation formula enables the Monte Carlo estimation of $\hat{Q}_{0,T}^n f(x)$ by sampling $M \in \mathbb{N}^*$ independent realizations of $X_T^{\Pi_h}$, $h \in \{1, \dots, r\}$. The simulation complexity for computing $\hat{Q}_{0,T}^n$ using M Monte Carlo samples is of order $C_{rand} = M \times n \times r \times C(\nu)$ where $C(\nu)$ depends on ν . Note that a standard approach on deterministic time grid of size $\frac{T}{n}$ is $C_{det} = M \times n$ and then $C_{det} = \mathcal{O}_{n \rightarrow \infty}(C_{rand})$ underscoring the great numerical interest of this random grid approach as ν can be arbitrarily high.

In [3], the authors also proved a restrictive version of (1.2). In particular, they showed that (1.2) holds with $\|f\|_\infty$ replaced by $\|f\|_{\infty, K}$ that involves the supremum norm of f but also of its derivatives up to order $K \in \mathbb{N}^*$. More specifically,

$$(1.5) \quad \sup_{x \in \mathbb{R}^d} |\mathbb{E}[f(X_T) | X_0 = x] - \hat{Q}_{0,T}^n f(x)| \leq \frac{C}{n^\nu} \|f\|_{\infty, K}.$$

The main interest of this method arises for large values of ν . Actually, for small ν , (1.5) can usually be obtained by a Lindeberg type argument (also referred to as Duhamel's principle), which relies on a smooth short time approximation of the form

$$(1.6) \quad |\mathbb{E}[f(X_{t+\delta}) - f(\psi(x, t, \sqrt{\delta}Z, \delta)) | X_t = x]| \leq C \|f\|_{\infty, \beta} \delta^{\alpha+1}$$

where $\alpha > 0$ is referred in this paper as the weak smooth order of the scheme ψ . If $\alpha \geq \nu$, one can simply chose $\hat{Q}_{0,T}^n = Q_{0,T}^{\frac{T}{n}}$ where for every measurable f and every $x \in \mathbb{R}^d$, $Q_{s,t}^{\frac{T}{n}} f(x) = \mathbb{E}[f(X_t^{\frac{T}{n}}) | X_s^{\frac{T}{n}} = x]$, $s, t \in \pi^{\frac{T}{n}} := \{k \frac{T}{n}, k \in \mathbb{N}\}$ is the semigroup arising from the Markov process $X^{\frac{T}{n}}$ defined on the homogeneous deterministic time grid with time step $\frac{T}{n}$ (we use $X^{\frac{T}{n}}$ as short notation for $X^{\pi^{\frac{T}{n}}}$ defined as in (1.3) with $\Pi_h = \pi^{\frac{T}{n}}$). For the Euler scheme we have $\alpha = 1$ (see [21]), but various higher weak smooth order methods exist (see *e.g.* [20], [16], [2],[18]). However, the attainable value of α remains limited or achieving higher α generally requires substantially increased computational complexity

The algorithm introduced in [3] exploits relation (1.6) on random time intervals and combines numerical schemes defined on random time grids to construct \hat{Q}^n , which depends on α and ν and ensures that (1.5) holds. The idea of combining approximation methods to enhance convergence rates also appears in a broad range of techniques for computing expectations of the form $\mathbb{E}[f(X_T) | X_0 = x]$. The Multi-Level Monte Carlo (MLMC) method (see [11, 9]), which extends the statistical Romberg approach of [13], combines schemes with different and carefully chosen time steps to achieve higher-order approximations. Similarly, the Richardson-Romberg method constructs high-order approximations by exploiting asymptotic expansions of the weak error (see, *e.g.*, [17]). Moreover, MLMC and Richardson-Romberg techniques can be combined to further improve efficiency, as demonstrated in [14]. Additional developments related to these approaches are discussed in the non-exhaustive list of works [15, 7, 10, 1, 22]. We also refer the reader to [4, 12, 8] for further developments in the field of unbiased simulation method.

The main goal of this article is to show that the result from [3], namely (1.5), is true with $K = 0$ meaning that (1.2) holds. We emphasize that $\hat{Q}_{0,T}^n$ is not necessarily a probability measure. When considering total variation convergence for schemes on deterministic homogeneous time grids, several results already exist. Essentially, those results prove the convergence in total variation with rate $1/n^\alpha$ assuming that (1.6) is satisfied. In contrast with the case $K \in \mathbb{N}^*$, one way to prove total variation convergence is to use regularization properties on the semigroups arising from X or $X^{\frac{T}{n}}$, which are themselves obtained under elliptic or Hörmander type assumptions. For instance, [6] addresses the scenario where ψ is the Euler scheme of a homogeneous SDE satisfying weak uniform Hörmander property. That work also provides an expansion of the error for bounded measurable test functions, enabling the use of Richardson–Romberg extrapolation. Concerning generic schemes (which encompass the Euler scheme but also many more buildings) of inhomogeneous SDE which are simply specified by a transition function, such as $X^{\frac{T}{n}}$, some results were also already established. In [5], the uniform elliptic case is studied, while in [19], the weak local Hörmander case is addressed. For generic schemes, an additional hypothesis with form (1.6) is required to achieve order α for the total variation convergence. It is worth noting that it is shown in [19] that, even without assuming (1.6), the convergence in total variation error of such generic schemes to X_T happens with at least order $\frac{1}{n^{\frac{1}{2}-\epsilon}}$, $\epsilon > 0$, for any $\epsilon > 0$, under the same assumptions that guarantee regularization properties.

However, the total variation order of convergence cannot go beyond the standard weak smooth order α of the scheme given by (1.6). For instance, the Euler scheme cannot converge faster than $\frac{1}{n}$ for the total variation distance and the only way to improve the convergence remains to increase the value of n . The interest of the article is to show how, without increasing n , in the weak uniform Hörmander setting, the method proposed in [3] can be applied to improve the total variation order of convergence from $\frac{1}{n^\alpha}$ to $\frac{1}{n^\nu}$ for any ν even when considering only schemes of weak smooth order α with $\alpha < \nu$.

In this paper, we do not discuss further the representation formula (1.4). We refer to an alternative equivalent representation, which only involves the discrete time semigroups

$$Q_{s,t}^{\frac{T}{n^l}} f(x) = \mathbb{E}[f(X_t^{\frac{T}{n^l}}) | X_s^{\frac{T}{n^l}} = x], \quad \forall s, t \in \pi_{\frac{T}{n^l}},$$

where $X_{\frac{T}{n^l}}$ is a short notation for $X_t^{\pi_{\frac{T}{n^l}}}$ where $\pi_{\frac{T}{n^l}} := \{k \frac{T}{n^l}, k \in \mathbb{N}\}$. The aforementioned representation formula is given in (2.2). Our approach is to show that (1.2) can be obtained using an abstract Lindeberg inspired decomposition combined with regularization properties satisfied by $Q_{s,t}^{\frac{T}{n^l}}$ (see Theorem 3.1 and Theorem 3.2). Then, we propose sufficient assumption on the function ψ and $((Z_j^h)_{j \in \mathbb{N}^*})_{h \in \{1, \dots, r\}}$ such that the required regularization properties of $Q_{s,t}^{\frac{T}{n^l}}$ hold under a weak uniform type Hörmander assumption. Those properties were demonstrated in [19] and are restated in our context in Proposition 3.3. Combining those regularization properties with the result from Theorem 3.2 we establish (1.2) in the main result Theorem 2.1.

The article is organized as follows. Section 2 introduces the framework under which Theorem 2.1 is established. Section 3 is devoted to the proof of Theorem 2.1. In particular, we develop an abstract framework to establish the convergence and derive the key intermediate results Theorem 3.1 and Theorem 3.2.

2. THE DISTANCE BETWEEN SEMIGROUPS

In this section we present our main result concerning the approximation of P_T by establishing (1.2) in Theorem 2.1. We first introduce our framework, then the construction of \hat{Q}^ν and then state this Theorem.

2.1. Framework.

Notations. Throughout this paper the following notations will prevail. We work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We fix $T > 0$ and $n \in \mathbb{N}^*$. For $l \in \mathbb{N}$ we will denote $\delta_n^l = T/n^l$ and for $\delta > 0$ we consider the time grid $\pi^\delta := \{k\delta, k \in \mathbb{N}\}$, with the convention $\pi^0 = \mathbb{R}_+$. We also fix the parameter $\alpha > 0$, $\beta \in \mathbb{N}$ and $\nu \in \mathbb{N}^*$. Along the article, we will also introduce the quantities \mathbf{m} , \mathbf{p} , ℓ , κ and \mathbf{q} that depend on α , β and ν .

For $d \in \mathbb{N}^*$, we denote by $\mathcal{M}_b(\mathbb{R}^d)$, the set of measurable and bounded functions defined on \mathbb{R}^d with values in \mathbb{R} . Moreover, $\mathcal{C}^q(\mathbb{R}^d)$, $q \in \mathbb{N} \cup \{+\infty\}$, is the set of \mathbb{R} -valued functions admitting derivatives

up to order q and such that all those derivatives (including order 0) are continuous and $\mathcal{C}_b^q(\mathbb{R}^d)$ (resp. $\mathcal{C}_c^q(\mathbb{R}^d)$, $\mathcal{C}_1^q(\mathbb{R}^d)$), $q \in \mathbb{N} \cup \{+\infty\}$, is the set of functions belonging to $\mathcal{C}^q(\mathbb{R}^d)$ such that all the derivatives (of order 0 to q) are bounded (resp. have compact support, are bounded in $L_1(\mathbb{R}^d)$).

For a multi-index $\gamma = (\gamma^1, \dots, \gamma^d) \in \mathbb{N}^d$ we denote $|\gamma| = \gamma^1 + \dots + \gamma^d$ and if $f \in \mathcal{C}^{|\gamma|}(\mathbb{R}^d)$, we define $\partial_x^\gamma f = (\partial_1)^{\gamma^1} \dots (\partial_d)^{\gamma^d} f = \partial_{x^1}^{\gamma^1} \dots \partial_{x^d}^{\gamma^d} f(x)$. Finally, we introduce the following norms for functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$, r -times differentiable with $r \in \mathbb{N}$: $\|f\|_\infty := \sup_{x \in \mathbb{R}^d} |f(x)|$, $\|f\|_{\infty, r} := \sum_{|\gamma| \leq r} \|\partial^\gamma f\|_\infty$, $\|f\|_1 := \int_{\mathbb{R}^d} |f(x)| dx$, $\|f\|_{1, r} := \sum_{|\gamma| \leq r} \|\partial^\gamma f\|_1$.

The approximated process. In this article, we aim to build a weak approximation for a solution of (1.1). In other words, we aim to construct an approximation for the semigroup associated to (1.1) *i.e.* $(P_{t,s})_{0 \leq s \leq t}$ defined for $t, s \in \mathbb{R}_+$ with $s \leq t$, $x \in \mathbb{R}^d$, by

$$P_{s,t}f(x) = \mathbb{E}[f(X_t)|X_s = x],$$

for every measurable function f from \mathbb{R}^d to \mathbb{R} such that the expectation in the *r.h.s.* above is well defined. Recall that our focus particularly lies on the case where f is a bounded and measurable function but not necessarily smooth. We recall that since $(X_t)_{t \geq 0}$ is a Markov process, P satisfies a crucial property called the semigroup property and which writes : For every $s \leq u \leq t$, we have $P_{s,u}P_{u,t} = P_{s,t}$.

The approximation procedure. We first introduce the basic processes that we use to build an approximation for P . For $\delta \in (0, 1]$ we consider a sequence of independent random variables $Z_t \in \mathbb{R}^N$, $t \in \pi^\delta \setminus \{0\}$, and we assume that Z_t^δ , are centered with $\mathbb{E}[Z_t^{\delta,i} Z_t^{\delta,j}] = \mathbf{1}_{i,j}$ for every $i, j \in \{1, \dots, N\}$. We construct the \mathbb{R}^d -valued Markov process $(X_t^\delta)_{t \in \pi^\delta}$ in the following way:

$$(2.1) \quad X_{t+\delta}^\delta = \psi(X_t^\delta, t, \delta^{\frac{1}{2}} Z_{t+\delta}^\delta, \delta), \quad t \in \pi^\delta, \quad X_0^\delta \in \mathbb{R}^d$$

where

$$\psi \in \mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}^N \times [0, 1]; \mathbb{R}^d) \quad \text{and} \quad \forall (x, t) \in \mathbb{R}^d \times \pi^\delta, \psi(x, t, 0, 0) = x.$$

Let us now define the discrete time semigroup associated to $(X_t^\delta)_{t \in \pi^\delta}$. For every measurable function f from \mathbb{R}^d to \mathbb{R} , and every $x \in \mathbb{R}^d$,

$$\forall s, t \in \pi^\delta, s \leq t, \quad Q_{s,t}^\delta f(x) = \int_{\mathbb{R}^d} f(y) Q_{s,t}^\delta(x, dy) := \mathbb{E}[f(X_t^\delta)|X_s^\delta = x].$$

We notice that $(X_t^\delta)_{t \in \pi^\delta}$ is a Markov process and then Q^δ satisfies the semigroup property : For every $s, u, t \in \pi^\delta$ with $s \leq u \leq t$, we have $Q_{s,u}^\delta Q_{u,t}^\delta = Q_{s,t}^\delta$.

We are now in position to build \hat{Q}^ν , the approximation of P with arbitrary weak order $\nu > 0$ by using the family $\left((Q_{s,t}^{\delta_n^l})_{s,t \in \pi^{\delta_n^l}; s \leq t} \right)_{l \in \mathbb{N}}$. Notice that this procedure is introduced in ([3]) (where (1.5) is demonstrated under suitable properties on Q^δ) and this section simply presents this method in our context. We consider a non-homogeneous approach but this implies no further difficulties.

We define \hat{Q}^ν in the following recursive way: For $l \in \mathbb{N}$ and $t \in \pi^{\delta_n^l}$,

$$(2.2) \quad \hat{Q}_{t,t+\delta_n^l}^\nu = Q_{t,t+\delta_n^l}^{\delta_n^{l+1}} + \sum_{i=1}^{\mathbf{m}(l)-1} \hat{I}_{t,t+\delta_n^l,i}^{\delta_n^{l+1}}$$

with, for $l \in \mathbb{N}$,

$$\hat{I}_{t,t+\delta_n^l,i}^{\delta_n^{l+1}} = \sum_{t=t_0 < t_1 < \dots < t_i \leq t+\delta_n^l} \prod_{j=1}^i Q_{t_{j-1}, t_j - \delta_n^{l+1}}^{\delta_n^{l+1}} \left(\hat{Q}_{t_j - \delta_n^{l+1}, t_j}^{\mathbf{p}(i,l)} - Q_{t_j - \delta_n^{l+1}, t_j}^{\delta_n^{l+1}} \right) Q_{t_i, t+\delta_n^l}^{\delta_n^{l+1}},$$

with

$$\mathbf{m}(l) = \left\lceil \frac{\nu}{(1+\alpha)l + \alpha} \right\rceil$$

$$\mathbf{p}(i, l) = \nu + \lceil i - (1+\alpha)(l+1)(i-1) \rceil, \quad \forall i \in \{1, \dots, \mathbf{m}(l) - 1\}.$$

It is worth noting that if $\mathbf{m}(l) = 1$, we have $\hat{Q}_{t,t+\delta_n^l}^\nu = Q_{t,t+\delta_n^l}^{\delta_n^{l+1}}$. When $\mathbf{m}(l) > 1$, the recursion also ends and $\hat{Q}_{t,t+\delta_n^l}^{\delta_n^l}$ is well-defined. This is a direct consequence of Proposition 3.1 (see also Lemma 3.8. in [3]). In particular, $\hat{Q}_{0,T}^\nu$ is well defined and can be built from the family $\left((Q_t^{\delta_n^l})_{t \in \pi^{\delta_n^l}} \right)_{l \in \{1, \dots, \ell\}}$ with

$$\ell := \mathbf{m}(0) = \lceil \frac{\nu}{\alpha} \rceil.$$

The functional operator $\hat{Q}_{0,T}^n := \hat{Q}_{0,T}^\nu$ is then used as an approximation of P_T . On the numerical viewpoint, a strong interest of this approach is that, in [3], the authors provide a random grid approach to compute $\hat{Q}_{0,T}^\nu$ (see the representation formula (1.4)). Furthermore, $\hat{Q}_{0,T}^\nu$ can be computed with complexity of order n in terms of the number of simulation of random variables with the transition laws with form $Q_{t,t+\delta_n^l}^{\delta_n^l}$, $t \in \pi^{\delta_n^l}$, $l \in \mathbb{N}$ that can be performed using (2.1). This allows to approximate $\hat{Q}_{0,T}^\nu$ using a Monte Carlo estimator where each sample requires an order n of simulations, similar to standard discretization methods such as the Euler scheme with n time steps. The choice of $t \in \pi^{\delta_n^l}$ and $l \in \mathbb{N}$ for the simulation is given by the random grid Π_h in the representation formula (1.4). We do not discuss further the simulation and complexity of $\hat{Q}_{0,T}^\nu$ in this paper. It would be exactly similar to the study provided in [3] since it does not depend on the test function f or its regularity and we refer to this article for more details. We merely note that the estimator of $\hat{Q}_{0,T}^\nu$ based on random grids, as proposed in [3] for smooth f , remains applicable in our framework when f is only assumed to be bounded and measurable. In [3], the authors established that (1.5) holds for smooth f . The novelty of our approach lies in extending this result to show that (1.2) holds under the weaker assumption that f is simply bounded and measurable.

2.2. Main result. In this section, we present Theorem 2.1 where we derive an estimate of the form (1.2), which holds for simply bounded and measurable test functions f .

We recall that for two probability measures μ_1 and μ_2 on \mathbb{R}^d , the total variation distance between μ_1 and μ_2 is given by

$$\begin{aligned} d_{TV}(\mu_1, \mu_2) &= \sup_{A \in \mathcal{B}(\mathbb{R}^d)} |\mu_1(A) - \mu_2(A)| = \sup_{f \in \mathcal{M}_b(\mathbb{R}^d; \mathbb{R}), \|f\|_\infty \leq 1} \frac{1}{2} |\mu_1(f) - \mu_2(f)| \\ &= \sup_{f \in \mathcal{C}^\infty(\mathbb{R}^d; \mathbb{R}), \|f\|_\infty \leq 1} \frac{1}{2} |\mu_1(f) - \mu_2(f)| \end{aligned}$$

where $\mu_1(f) = \int_{\mathbb{R}^d} f(y) \mu_1(dy)$ and similarly for $\mu_2(f)$, $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ -algebra on \mathbb{R}^d . The last equality above is a direct consequence of the Lusin's Theorem. In our setting, for a given $x \in \mathbb{R}^d$, $\hat{Q}_{0,T}^n(x, dy)$ is not necessarily a probability measure. Nevertheless, by a slight abuse of terminology, we will refer to (1.2) as a total variation convergence result. We explicitly mean that the total variation norm of the finite signed measure $\mu^n(dy) = \hat{Q}_{0,T}^n(x, dy) - P_T(x, dy)$, $x \in \mathbb{R}^d$, tends to zero when n goes to infinity.

2.2.1. Assumption. In this section, we introduce the assumptions that will be employed to establish the total variation convergence result. We first assume that we have $V_0 = \partial_y \psi(\cdot, \cdot, 0, 0) + \frac{1}{2} \sum_{i=1}^N \partial_{z^i}^2 \psi(\cdot, \cdot, 0, 0)$, $V_i = \partial_{z^i} \psi(\cdot, \cdot, 0, 0)$, $i \in \{1, \dots, N\}$. This assumption is actually not restrictive. Within our framework, the convergence result we establish cannot occur under any other configuration (see Remark 3.1)). We now present the precise set of hypotheses required to establish Theorem 2.1. In the following, $n \in \mathbb{N}^*$ and $\delta \in (0, 1]$.

Hypotheses on ψ and $V_i, i \in \{0, \dots, N\}$. Approximation.

In this first part, we introduce the following short-time estimate assumptions. For $n, l \in \mathbb{N}$:

$\mathbf{E}_n(l)$. For every $r \in \mathbb{N}$, and every $f \in \mathcal{C}_b^{\beta+r}(\mathbb{R}^d)$

$$(2.3) \quad \forall t \in \pi^{\delta_n^l}, \quad \|P_{t,t+\delta_n^l} - Q_{t,t+\delta_n^l}^{\delta_n^l} f\|_{\infty,r} \leq C \|f\|_{\infty,\beta+r} (\delta_n^l)^{\alpha+1}.$$

We also introduce a dual version of (2.3). We specify that for a functional operator F on $\mathcal{M}_b(\mathbb{R}^d)$, we denote by F^* its dual operator satisfying, for every $g \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ and $f \in \mathcal{M}_b(\mathbb{R}^d)$, $\int_{\mathbb{R}^d} g(x) F f(x) dx = \int_{\mathbb{R}^d} F^* g(x) f(x) dx$. For $n, l \in \mathbb{N}$:

$\mathbf{E}_n^*(l)$. For every $r \in \mathbb{N}$, and every $f \in \mathcal{C}_1^{\beta+r}(\mathbb{R}^d)$, then

$$(2.4) \quad \forall t \in \pi^{\delta_n^l}, \quad \|P_{t,t+\delta_n^l}^* - Q_{t,t+\delta_n^l}^{\delta_n^l,*} f\|_{1,r} \leq C \|f\|_{1,\beta+r} (\delta_n^l)^{\alpha+1}.$$

Hypotheses on ψ and $V_i, i \in \{0, \dots, N\}$. Regularity and Hörmander property.

We first consider a boundary assumption concerning the derivatives of ψ : For $r \in \mathbb{N}^*$,

A₁^δ(r). There exists $\mathfrak{D}_r \geq 1, \mathbf{p}_r \in \mathbb{N}$ such that for every $(x, t, z, y) \in \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}^N \times [0, 1]$,

$$(2.5) \quad \sum_{|\gamma^x|+|\gamma^t|=0}^r \sum_{|\gamma^z|+|\gamma^y|=1}^{r-|\gamma^x|-|\gamma^t|} |\partial_x^{\gamma^x} \partial_t^{\gamma^t} \partial_z^{\gamma^z} \partial_y^{\gamma^y} \psi|_{\mathbb{R}^d}(x, t, z, y) \leq \mathfrak{D}_r (1 + \delta^{-\frac{\mathbf{p}_r}{2}} |z|_{\mathbb{R}^N}^{\mathbf{p}_r}),$$

Without loss of generality, we assume that the sequences $(\mathfrak{D}_r)_{r \in \mathbb{N}^*}$ and $(\mathbf{p}_r)_{r \in \mathbb{N}^*}$ are non decreasing.

The second hypothesis we need on ψ is a uniform weak Hörmander property on some vector fields we now introduce. We denote the Lie bracket of two \mathcal{C}^1 vector fields in \mathbb{R}^d , $[\cdot, \cdot] : (\mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d))^2 \rightarrow \mathcal{C}^0(\mathbb{R}^d, \mathbb{R}^d)$, $f_1, f_2 \mapsto [f_1, f_2] := \nabla_x f_2 f_1 - \nabla_x f_1 f_2$.

We denote $\bar{V}_0 = V_0 - \frac{1}{2} \sum_{i=1}^N \nabla_x V_i V_i$. For $\alpha \in \{0, \dots, N\}^m$ with $m =: \|\alpha\| \in \mathbb{N}$ and $V : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}^d$, we define also $V^{[\alpha]}$ using the recurrence relation $V^{[(\alpha, 0)]} = [\bar{V}_0, V^{[\alpha]}] + \partial_t V^{[\alpha]} + \frac{1}{2} \sum_{i=1}^N [V_i, [V_i, V^{[\alpha]}]]$ and $V^{[(\alpha, j)]} := [V_j, V^{[\alpha]}]$ if $j \in \{1, \dots, N\}$ with the convention $V^{[0]} = V$. We are now in a position to introduce our Hörmander hypothesis on ψ : For $L \in \mathbb{N}$, the order of our Hörmander condition, let us define for every $(x, t) \in \mathbb{R}^d \times \mathbb{R}_+$,

$$(2.6) \quad \mathcal{V}_L(x, t) := 1 \wedge \inf_{\mathbf{b} \in \mathbb{R}^d, |\mathbf{b}|_{\mathbb{R}^d} = 1} \sum_{\substack{\alpha \in \{0, \dots, N\}^{\|\alpha\|}; \\ \|\alpha\| \leq L}} \sum_{i=1}^N \langle V_i^{[\alpha]}(x, t), \mathbf{b} \rangle_{\mathbb{R}^d}^2.$$

A₂(L). We introduce the uniform weak Hörmander property of order L ,

$$(2.7) \quad \mathcal{V}_L := \inf_{t \in \mathbb{R}_+} \inf_{x \in \mathbb{R}^d} \mathcal{V}_L(x, t) > 0.$$

When $L = 0$ this hypothesis is also called uniform ellipticity.

Remark 2.1. *The reason we use the term Hörmander resides in the fact (2.7) exactly corresponds to the commonly known uniform weak Hörmander property the solution to the SDE (1.1), written in Stratonovich form as*

$$(2.8) \quad X_t = X_0 + \int_0^t \bar{V}_0(X_s, s) ds + \sum_{i=1}^N \int_0^t V_i(X_s, s) \circ dW_s^i, \quad t \geq 0, X_0 \in \mathbb{R}^d$$

where $\circ dW_s^i$ stands for the Stratonovich integral w.r.t. $(W_t^i)_{t \geq 0}$. When $L = 0$, (2.7) is also called uniform elliptic property. It is worth noting that (2.7) is equivalent to assume that \mathbb{R}^d is spanned by $\{V_i^{[\alpha]}(x, t), i \in \{1, \dots, N\}, \alpha \in \{0, \dots, N\}^{\|\alpha\|}, \|\alpha\| \leq L\}$ uniformly in $(x, t) \in \mathbb{R}^d \times \mathbb{R}_+$.

Hypotheses on Z^δ . Lebesgue lower bounded distributions. A first assumption concerns the finiteness of the moment of Z^δ : For $p \geq 0$:

A₃^δ. For every $p \geq 0$,

$$(2.9) \quad \sup_{t \in \pi^\delta \setminus \{0\}} \mathbb{E}[|Z_t^\delta|_{\mathbb{R}^N}^p] < \infty.$$

A second assumption is made on the distribution of Z^δ . We suppose that the distribution of Z^δ is Lebesgue lower bounded:

A₄^δ. There exists $z_* = (z_*, t)_{t \in \pi^\delta, *}$ taking its values in \mathbb{R}^N and $\varepsilon_*, r_* > 0$ such that for every Borel set $A \subset \mathbb{R}^N$ and every $t \in \pi^\delta \setminus \{0\}$,

$$(2.10) \quad L_{z_*}^\delta(\varepsilon_*, r_*) \quad \mathbb{P}(Z_t^\delta \in A) \geq \varepsilon_* \lambda_{\mathcal{L}}(A \cap B_{r_*}(z_*, t))$$

where $\lambda_{\mathcal{L}}$ is the Lebesgue measure on \mathbb{R}^N .

2.2.2. Total variation convergence result. Under suitable assumptions on the approximation schemes $(X^\delta)_{\delta > 0}$ the next result shows the total variation convergence of $\hat{Q}_{0,T}^n$, built by using (2.2), towards $P_T f(x) := \mathbb{E}[f(X_T) | X_0 = x]$ where X is the solution to (2.8). Before we state this result we define recursively, for $l \in \mathbb{N}$,

$$\kappa(l, \nu) = \max \left\{ \beta \mathbf{m}(l), \max_{i=1}^{\mathbf{m}(l)-1} \{i \kappa(l+1, \mathbf{p}(i, l))\} \right\},$$

with the convention $\max \emptyset = -\infty$.

Theorem 2.1. *Let $T > 0$, $n \in \mathbb{N}^*$, $L \in \mathbb{N}$, and $\mathbf{q} = \kappa(0, \nu)$. We assume that $\mathbf{A}_1^{\delta_1^n}(\max(\mathbf{q} + 3, 2L + 5))$, $\mathbf{A}_2(L)$, $\mathbf{A}_3^{\delta_1^n}$ and $\mathbf{A}_4^{\delta_1^n}$ hold and that $\mathbf{E}_n(l)$ and $\mathbf{E}_n^*(l)$ are satisfied for every $l \in \mathbb{N}$. If n is large enough, then for every $f \in \mathcal{M}_b(\mathbb{R}^d)$,*

$$(2.11) \quad \|P_T f - \hat{Q}_{0,T}^\nu f\|_\infty \leq \frac{1}{n^\nu} \|f\|_\infty \frac{C \exp(CT)}{(\mathcal{V}_L T(\nu))^\eta}.$$

where $T(\nu) := \inf \left\{ t \in \pi^{\delta_1^n}, t \geq T \frac{n-\ell}{n(\ell+1)} \right\}$ and $\eta, C \geq 0$ do not depend on n , T or f .

Remark 2.2. *We recall that, for bounded and measurable f , we can build a Monte Carlo estimator for $\hat{Q}_{0,T}^\nu f$ based on random grids similarly as in [3] when f is smooth. As expected, the complexity of the algorithm is not affected but one may wonder about its variance. In [3], Theorem 4.4, the variance is shown to be bounded when ψ is the Euler scheme and f is smooth. The proof of this result exploits the regularity of f and the bound derived on the variance depends on the supremum norm of f but also of its derivatives. Such a strategy is not appropriate when f is simply bounded and measurable and extension to such results is not discussed in this article.*

Example 2.1. *Let us consider $X = (X^1, X^2)$, the solution of the 2-dimensional system of \mathbb{R} valued SDE, starting at point $x_0 = (x_0^1, x_0^2) \in \mathbb{R}^2$ and given by*

$$\begin{aligned} dX_t^1 &= b(X_t^1, t) dt + \sigma(X_t^1, t) dW_t \\ dX_t^2 &= X_t^1 dt \end{aligned}$$

where $(W_t)_{t \geq 0}$ is a one dimensional standard Brownian motion, b and σ continuous and bounded and their derivatives of any order are also continuous and bounded. Moreover, let us assume that $\underline{\sigma} := \inf_{(x^1, t) \in \mathbb{R} \times \mathbb{R}_+} |\sigma(x^1, t)| \neq 0$. This two-dimensional system naturally models position and velocity in physical dynamics. It also arises in finance for the valuation of Asian options whose price depends on the time average value, modeled by X^2 , of a financial object (e.g. a stock) modeled by X^1 .

In the setting from (1.1), we have $V_0 : (x, t) \mapsto (b(x^1, t), x^1)$ and $V_1 : (x, t) \mapsto (\sigma(x^1, t), 0)$. In this example uniform ellipticity holds for X^1 since $\underline{\sigma} \neq 0$. However ellipticity does not hold for X since $\dim(\text{span}((\sigma, 0)))(x, t) \leq 1 < 2$ for any $(x, t) \in \mathbb{R}^2 \times \mathbb{R}_+$. Nevertheless, let us compute the Lie brackets. In particular

$$[V_0, V_1] : (x, t) \mapsto (\partial_{x^1} \sigma(x^1, t) b(x^1, t) - \partial_{x^1} b(x^1, t) \sigma(x^1, t), -\sigma(x^1, t)),$$

and, for $\sigma(x^1, t) \neq 0$, $\text{span}((\sigma, 0), (\partial_{x^1} \sigma b - \partial_{x^1} b \sigma + \partial_t \sigma, -\sigma)(x, t)) = \mathbb{R}^2$ so that weak Hörmander condition holds.

Now, let us consider the Euler scheme of X , given by $(X_0^{\delta, 1}, X_0^{\delta, 2}) = x_0$ and by (2.1) when we consider the transition function

$$\begin{aligned} \psi(x, t, z, y)^1 &= x^1 + b(x, t)y + \sigma(x, t)z^i \\ \psi(x, t, z, y)^2 &= x^2 + x^1 y \end{aligned}$$

for $(x, t, z, y) \in \mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$. Moreover, we choose $Z_{t+\delta}^\delta \in \mathbb{R}$, $t \in \pi^\delta$, independent and centered with variance one, Lebesgue lower bounded distribution and moment of order three equal to zero. In this setting, all the assumption from Theorem 2.1 (with $L = 1$, $\alpha = 1$ and $\beta = 4$, ν being chosen arbitrarily) are verified. We give further details concerning the uniform weak Hörmander hypothesis $\mathbf{A}_2(1)$. With notations introduced in (2.6), for

$$\begin{aligned} \mathcal{V}_1(x, t) &= 1 \wedge \inf_{\mathbf{b} \in \mathbb{R}^2, |\mathbf{b}|_{\mathbb{R}^d} = 1} \langle V_1(x, t), \mathbf{b} \rangle_{\mathbb{R}^2}^2 + \langle [V_0 - \frac{1}{2} \nabla_x V_1 V_1, V_1](x, t) + \partial_t V_1(x, t), \mathbf{b} \rangle_{\mathbb{R}^2}^2 \\ &= 1 \wedge \inf_{\mathbf{b} \in \mathbb{R}^2, |\mathbf{b}|_{\mathbb{R}^2} = 1} \langle (\sigma, 0), \mathbf{b} \rangle_{\mathbb{R}^2}^2 + \langle (\partial_{x^1} \sigma b - \partial_{x^1} b \sigma + \frac{1}{2} \sigma^2 \partial_{x^1}^2 \sigma + \partial_t \sigma, -\sigma), \mathbf{b} \rangle_{\mathbb{R}^2}^2(x^1, t) \\ &> 1 \wedge C \frac{\underline{\sigma}^4}{1 + \|b\|_{\infty, 1}^4 + \|\sigma\|_{\infty, 2}^6}. \end{aligned}$$

where $C > 0$ is a universal constant and then assumption $\mathbf{A}_2(1)$ holds. We are now in a position to state our convergence result deduced from Theorem 2.1. For a fixed time step $\delta > 0$, we introduce

$$\forall s, t \in \pi^\delta, s \leq t, \quad Q_{s,t}^\delta f(x) := \mathbb{E}[f(X_t^\delta) | X_s^\delta = x].$$

Given $T > 0$, $n \in \mathbb{N}^*$, $\nu > 0$, we build $\hat{Q}_{0,T}^\nu$ using (2.2) and for n large enough, for every $f \in \mathcal{M}_b(\mathbb{R}^d)$, we have

$$\sup_{x \in \mathbb{R}^d} |\mathbb{E}[f(X_T)|X_0 = x] - \hat{Q}_{0,T}^\nu f(x)| \leq \frac{1}{n^\nu} \|f\|_\infty \frac{C \exp(CT)}{(\underline{\sigma}T(\nu))^\eta},$$

where C and η do not depend on n , T or f .

3. PROOF OF THEOREM 2.1

In this section, we demonstrate our main theorem. The original approach we develop consists in two steps. First we consider abstract properties on the semigroups involved in our approximation result and demonstrate a total variation convergence abstract result under these hypotheses. Then, we show that in our framework the abstract properties and more specifically the regularization properties required on the semigroups are satisfied (this result is actually established in [19]). Then, we use the abstract total variation convergence result to deduce (2.11).

We emphasize that our semigroup approach is generic and can apply to any family of probability measures used to build P or $(Q_n^{\delta_n^l})_{l \in \mathbb{N}}$ as long as they satisfy the semigroup property.

3.1. Total variation convergence from regularization properties of semigroup. In this section we present the methodology to derive weak estimates based on short time estimate assumptions and regularity assumptions on the semigroups P and Q^δ . When the semigroup preserves the regularity, the weak estimate for smooth test function is established in Proposition 3.1 as a consequence of results from [3]. The novelty of our approach is to demonstrate that if we assume that our semigroups satisfy regularization properties, then we can derive total variation convergence results. We begin by introducing abstract regularization properties on our semigroup. First, we suppose that the semigroups we study are such that for every $r \in \mathbb{N}$ with $r \leq \mathbf{q} = \kappa(0, \nu)$, if $f \in C_b^r(\mathbb{R}^d)$ then $P_{s,t}f \in C_b^r(\mathbb{R}^d)$ and

$$(3.1) \quad \sup_{t \geq s \geq 0} \|P_{s,t}f\|_{r,\infty} \leq C \|f\|_{r,\infty}.$$

Moreover, the approximation of P we consider in this paper is built from a family of discrete semigroups $Q := (Q_n^{\delta_n^l})_{l \in \mathbb{N}} = ((Q_{s,t}^{\delta_n^l})_{s,t \in \pi_n^{\delta_n^l}; s \leq t})_{l \in \mathbb{N}}$ such that for every $f \in C_b^r(\mathbb{R}^d)$ then $Q_{s,t}^{\delta_n^l} f \in C_b^r(\mathbb{R}^d)$ and

$$(3.2) \quad \forall s, t \in \pi_n^{\delta_n^l}, s \leq t, \quad \|Q_{s,t}^{\delta_n^l} f\|_{r,\infty} \leq C \|f\|_{\infty,r}.$$

We also consider a dual version of (3.1) and (3.2). We assume that for every $f \in C_1^r(\mathbb{R}^d)$, then $P_{s,t}^* f \in C_1^r(\mathbb{R}^d)$ and

$$(3.3) \quad \sup_{t \geq s \geq 0} \|P_{s,t}^* f\|_{r,1} \leq C \|f\|_{r,1},$$

and for the family of semigroups $Q = ((Q_{s,t}^{\delta_n^l})_{s,t \in \pi_n^{\delta_n^l}; s \leq t})_{l \in \mathbb{N}}$, $Q_{s,t}^{\delta_n^l,*} f \in C_1^r(\mathbb{R}^d)$ and

$$(3.4) \quad \forall s, t \in \pi_n^{\delta_n^l}, s \leq t, \quad \|Q_{s,t}^{\delta_n^l,*} f\|_{r,1} \leq C \|f\|_{r,1}.$$

Using those properties combined with short time estimates (2.3) and (2.4), we can establish weak estimates in the case where f is a smooth test function.

Proposition 3.1. *We have the following properties:*

- A.** *Assume that (3.1), (3.2) and $\mathbf{E}_n(l)$ hold for every $l \in \{0, \dots, \ell\}$. Then, for every $l \in \{0, \dots, \ell-1\}$ and $f \in C_b^{\kappa(0,\nu)}(\mathbb{R}^d)$, we have*

$$(3.5) \quad \hat{\mathbf{E}}_n(l, \nu) \equiv \|P_{t,t+\delta_n^l} f - \hat{Q}_{t,t+\delta_n^l}^\nu f\|_{\infty,r} \leq C \|f\|_{\infty, \kappa(l,\nu)+r} \frac{1 + T^p}{n^\nu}.$$

for every $t \in \pi_n^{\delta_n^l}$, every $v \in \{1, \dots, l + \nu\}$ and every $r \in \mathbb{N}$ such that $\kappa(l, \nu) + r \leq \mathbf{q}$.

- B.** *Assume that (3.3), (3.4) and $\mathbf{E}_n^*(l)$ hold for every $l \in \{0, \dots, \ell\}$. Then, for every $l \in \{0, \dots, \ell-1\}$, and $f \in C_1^{\kappa(0,\nu)}(\mathbb{R}^d)$,*

$$(3.6) \quad \hat{\mathbf{E}}_n^*(l, \nu) \equiv \forall t \in \pi_n^{\delta_n^l}, \|P_{t,t+\delta_n^l}^* f - \hat{Q}_{t,t+\delta_n^l}^{\nu,*} f\|_{1,r} \leq C \|f\|_{1, \kappa(l,\nu)+r} \frac{1 + T^p}{n^\nu}.$$

for every $t \in \pi_n^{\delta_n^l}$, every $v \in \{1, \dots, l + \nu\}$ and every $r \in \mathbb{N}$ such that $\kappa(l, \nu) + r \leq \mathbf{q}$.

Notice that in both (3.5) and (3.6), $p \geq 0$ is a constant that may depend on ν and α .

This result shows among others, that (1.5) holds with $K = \kappa(0, \nu)$. In addition, relying on its proof, below, we observe that $\hat{Q}_{0,T}^\nu$ is built using the family of semigroups $Q^{\delta_n^l}$ with l limited to $\{1, \dots, \ell\}$.

Proof of Proposition 3.1. The proof of this result relies on Lemma 2.4 in [3]. More precisely, assuming that $\hat{\mathbf{E}}_n(\ell - 1, \nu)$ and $\hat{\mathbf{E}}_n^*(\ell - 1, \nu)$ hold, the same argument as in Lemma 2.4 of [3] yields, for $\ell \geq 2$, that $\hat{\mathbf{E}}_n(\ell - 2, \nu)$ also holds. By a similar reasoning, one deduces $\hat{\mathbf{E}}_n^*(\ell - 2, \nu)$. Iterating this step shows that $\hat{\mathbf{E}}_n(l, \nu)$ and $\hat{\mathbf{E}}_n^*(l, \nu)$ are valid for all $l \in 0, \dots, \ell - 1$.

Therefore, it remains to establish $\hat{\mathbf{E}}_n(\ell - 1, \nu)$ and $\hat{\mathbf{E}}_n^*(\ell - 1, \nu)$, to which we now turn. We observe that $\mathbf{m}(\ell - 1) = 1$ so that $\hat{Q}_{t,t+\delta_n^{\ell-1}}^\nu = Q_{t,t+\delta_n^{\ell-1}}^{\delta_n^\ell}$ (see (2.2)) for every $t \in \pi^{\delta_n^{\ell-1}}$ and every $v \in \{1, \dots, \ell + \nu - 1\}$. Moreover, following an argument similar to that used in the proof of Proposition 2.2 in [5], if one combines the Lindeberg-type decomposition arising from the semigroup property,

$$\forall t \in \pi^{\delta_n^{\ell-1}}, P_{t,t+\delta_n^{\ell-1}} f - Q_{t,t+\delta_n^{\ell-1}}^{\delta_n^\ell} f = \sum_{j=1}^n Q_{t,t_j-\delta_n^\ell}^{\delta_n^\ell} (P_{t_j-\delta_n^{\ell+1},t_j} - Q_{t_j-\delta_n^{\ell+1},t_j}^{\delta_n^\ell}) P_{t_j,t+\delta_n^{\ell-1}},$$

where $t_0 = t$ and $t_j = t_0 + j\delta_n^\ell$, and the assumptions (3.1), (3.2), (3.3), (3.4), $\mathbf{E}_n(\ell)$ and $\mathbf{E}_n^*(\ell)$, one obtains that

$$\|P_{t,t+\delta_n^{\ell-1}} f - Q_{t,t+\delta_n^{\ell-1}}^{\delta_n^\ell} f\|_{\infty,r} \leq C \|f\|_{\infty,\beta+r} \frac{1 + T^p}{n^{\ell(\alpha+1)-1}},$$

and

$$\|P_{t,t+\delta_n^{\ell-1}}^* f - Q_{t,t+\delta_n^{\ell-1}}^{\delta_n^{\ell,*}} f\|_{1,r} \leq C \|f\|_{1,\beta+r} \frac{1 + T^p}{n^{\ell(\alpha+1)-1}}.$$

Moreover, $\ell(\alpha + 1) - 1 \geq \ell + \nu - 1$, and $\beta = \kappa(\ell - 1, \ell + \nu - 1) = \dots = \kappa(\ell - 1, 1)$. Therefore, $\hat{\mathbf{E}}_n(\ell - 1, \nu)$ and $\hat{\mathbf{E}}_n^*(\ell - 1, \nu)$ hold and the proof is completed. \square

The estimates derived in Proposition 3.1 still involve the derivatives of f in the upper bound. To address this issue and establish the total variation convergence, our approach consists in considering stronger assumptions than (3.1), (3.2), (3.3) and (3.4) which we refer to as regularization properties. In concrete applications, those properties are not necessarily satisfied by the discrete semigroup Q^δ but rather by a family of functional operators that are sufficiently close to Q^δ in total variation norm. We refer to this family, as a modification of Q^δ . It is not necessarily a semigroup. Consequently, this hypothesis is formulated not only for discrete semigroups but for discrete families of functional operators.

Let $q \in \mathbb{N}$, $\eta \geq 0$ and $S > 0$ be fixed. For $\delta > 0$, let $(Q_{s,t}^\delta)_{s,t \in \pi^\delta, t > s}$, be a family of functional operators. We consider the following regularization property :

For every $r \in \mathbb{N}$ and every multi-index γ with $|\gamma| + r \leq q$, and $f \in \mathcal{C}_b^\infty(\mathbb{R}^d)$, then

$$(3.7) \quad R_{q,\eta,S}(Q^\delta) \equiv \quad \forall s, t \in \pi^\delta, t - s \geq S, \quad \|Q_{s,t}^\delta \partial_x^\gamma f\|_{r,\infty} \leq \frac{C}{(t-s)^\eta} \|f\|_\infty,$$

where C does not depend on δ , f , η or S . In our approach we will not necessarily use directly (3.7) but an estimate that it implies on the adjoint semigroup. Actually, remembering that $\|f\|_1 \leq \sup_{g \in \mathcal{M}_b(\mathbb{R}^d), \|g\|_\infty = 1} \langle g, f \rangle$, we notice that $R_{q,\eta,S}(Q^\delta)$ implies that for every $r \in \mathbb{N}$ and every multi-index γ with $|\gamma| + r \leq q$ and $f \in \mathcal{C}^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$

$$(3.8) \quad \forall s, t \in \pi^{\delta_n^\ell}, t - s \geq S, \quad \|Q_{s,t}^{\delta_n^{\ell,*}} \partial_x^\gamma f\|_{r,1} \leq \frac{C}{(t-s)^\eta} \|f\|_1.$$

where C does not depend on δ , f , η or S .

Assumptions (3.8) and (3.7) are crucial to derive total variation convergence results. In essence, our strategy is to apply the short-time estimates not directly to the test function f , but to $Q_{s,t}^\delta \partial^\gamma f$. This requires controlling the supremum norm of $Q_{s,t}^\delta \partial^\gamma f$ and its derivatives. Fortunately, this quantity is itself controlled by the supremum norm of f by exploiting (3.8). As a result, only the supremum norm of f appears in the final estimate, yielding a total variation convergence result. Nevertheless, we notice that the control provided by (3.8) is singular when s and t are close. To handle this, we adopt a dual approach, which necessitates both a dual short-time estimate and the dual regularization property (3.8). Moreover, we introduce a threshold parameter $S > 0$ for our regularization properties. This threshold is required because, in practical applications, the regularization properties can be established for any fixed

$S > 0$, provided that δ is sufficiently small.

Using these hypotheses, we can derive the following crucial total variation convergence towards the semigroup P with rate $1/n^\nu$.

Theorem 3.1. *Let $T > 0$, $n \in \mathbb{N}^*$ and $\eta \geq 0$. Assume that (3.1), (3.2), (3.3) and (3.4) hold and that $\mathbf{E}_n(l)$ and $\mathbf{E}_n^*(l)$ hold for every $l \in \{1, \dots, \ell\}$. Moreover, assume that $R_{\mathbf{q}, \eta, T(\nu)}(Q_n^{\delta_n^1})$ and $R_{\mathbf{q}, \eta, T(\nu)}(P)$ hold. Then, for every $f \in \mathcal{M}_b(\mathbb{R}^d)$,*

$$\|P_T f - \hat{Q}_{0,T}^\nu f\|_\infty \leq \frac{1}{n^\nu} \|f\|_\infty \frac{C(1+T^p)}{T(\nu)^\eta}.$$

where $T(\nu) = \inf \left\{ t \in \pi^{\delta_n^1}, t \geq T \frac{n-\ell}{n(\ell+1)} \right\}$ and $p, C \geq 0$ do not depend on n, T or f .

Proof. First we note that if $\alpha \geq \nu$, $\mathbf{m}(0) = \ell = 1$ and we have $\hat{Q}_{0,T}^\nu = Q_{0,T}^{\delta_n^1}$ (see (2.2)) and the proof is similar to the one of Proposition 2.3 in [5]. In the following proof we thus consider the case $\mathbf{m}(0) > 1$.

Step 1. The first step is to provide a representation formula for the semigroup $(P_t)_{t \geq 0}$ which relies on the family of semigroup $((Q_t^{\delta_n^1})_{t \in \pi^{\delta_n^1}})_{l \in \mathbb{N}}$. In particular, we deduce from the semigroup property that, for every $t \in \pi^{\delta_n^1}$,

$$P_T = Q_{0,T}^{\delta_n^1} + \sum_{i=1}^{\mathbf{m}(0)-1} I_{T,i}^{\delta_n^1} + R_{T,\mathbf{m}(0)}^{\delta_n^1}$$

with

$$I_{T,i}^{\delta_n^1} = \sum_{0=t_0 < \dots < t_i \leq T \in \pi^{\delta_n^1}} \left(\prod_{j=1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \left(P_{t_j - \delta_n^1, t_j} - Q_{t_j - \delta_n^1, t_j}^{\delta_n^1} \right) \right) Q_{t_i, T}^{\delta_n^1}.$$

and

$$R_{T,m}^{\delta_n^1} = \sum_{0=t_0 < \dots < t_m \leq T \in \pi^{\delta_n^1}} \left(\prod_{j=1}^m Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \left(P_{t_j - \delta_n^1, t_j} - Q_{t_j - \delta_n^1, t_j}^{\delta_n^1} \right) \right) P_{t_m, T},$$

where $t_j = j\delta_n^1$. Moreover, it follows from (2.2) that

$$\begin{aligned} I_{T,i}^{\delta_n^1} - \hat{I}_{0,T,i}^{\delta_n^1} &= \sum_{0=t_0 < \dots < t_i \leq T \in \pi^{\delta_n^1}} \left(\prod_{j=1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \left(P_{t_j - \delta_n^1, t_j} - Q_{t_j - \delta_n^1, t_j}^{\delta_n^1} \right) \right) Q_{t_i, T}^{\delta_n^1} \\ &- \sum_{0=t_0 < \dots < t_i \leq T \in \pi^{\delta_n^1}} \left(\prod_{j=1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \left(P_{t_j - \delta_n^1, t_j} - Q_{t_j - \delta_n^1, t_j}^{\delta_n^1} + \hat{Q}_{t_j - \delta_n^1, t_j}^{\mathbf{P}(i,0)} - P_{t_j - \delta_n^1, t_j} \right) \right) Q_{t_i, T}^{\delta_n^1}, \end{aligned}$$

which can be written in compact formulation

$$I_{T,i}^{\delta_n^1} - \hat{I}_{0,T,i}^{\delta_n^1} = - \sum_{0=t_0 < \dots < t_i \leq T \in \pi^{\delta_n^1}} \sum_{h=1}^{2^i-1} \left(\prod_{j=1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \right) Q_{t_i, T}^{\delta_n^1},$$

with, for every $j \in \{1, \dots, i\}$, $\Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \in \left\{ P_{t_j - \delta_n^1, t_j} - Q_{t_j - \delta_n^1, t_j}^{\delta_n^1}, \hat{Q}_{t_j - \delta_n^1, t_j}^{\mathbf{P}(i,0)} - P_{t_j - \delta_n^1, t_j} \right\}$. Importantly, the case $\Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} = P_{t_j - \delta_n^1, t_j} - Q_{t_j - \delta_n^1, t_j}^{\delta_n^1}$ for every $j \in \{1, \dots, i\}$ is excluded. Then, it follows from (2.2) that, we have the representation formula

$$P_T f - \hat{Q}_{0,T}^\nu f = - \sum_{i=1}^{\mathbf{m}(0)-1} \sum_{0=t_0 < \dots < t_i \leq T \in \pi^{\delta_n^1}} \sum_{h=1}^{2^i-1} \left(\prod_{j=1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \right) Q_{t_i, T}^{\delta_n^1} + R_{T,\mathbf{m}(0)}^{\delta_n^1}$$

Step 2. In this second step, we aim to estimate each term appearing in the representation formula established in **Step 1**. Using this decomposition, it is sufficient to prove that

$$\left\| \left(\prod_{j=1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \right) Q_{t_i, T}^{\delta_n^1} f \right\|_\infty \leq \frac{C(1+T^p)}{T(\nu)^\eta} \|f\|_\infty / n^{\nu+i},$$

and, for the remainder,

$$\left\| \left(\prod_{j=1}^{\mathbf{m}(0)} Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \left(P_{t_j - \delta_n^1, t_j} - Q_{t_j - \delta_n^1, t_j}^{\delta_n^1} \right) \right) P_{t_{\mathbf{m}(0)}, T} f \right\|_{\infty} \leq \frac{C(1+T^p)}{T(\nu)^\eta} \|f\|_{\infty} / n^{\nu+i}.$$

We focus on the study of $\sum_{i=1}^{\mathbf{m}(0)-1} I_{T,i}^{\delta_n^1} - \hat{I}_{0,T,i}^{\delta_n^1}$. The study of $R_{T,\mathbf{m}(0)}^{\delta_n^1}$ is similar so we leave it out. First we notice that, using the convention $t_{i+1} = T + \delta_n^1$, for $j_i = \operatorname{argsup}_{j \in \{1, \dots, i+1\}} \{t_j - \delta_n^1 - t_{j-1}\}$, we have $t_{j_i} - \delta_n^1 - t_{j_i-1} \geq T(\nu) = \inf \left\{ t \in \pi^{\delta_n^1}, t \geq T \frac{n-\mathbf{m}(0)}{n(\mathbf{m}(0)+1)} \right\}$ (recall that $\mathbf{m}(0) = \ell$).

Let $f \in C_c^\infty(\mathbb{R}^d)$. We note that $\mathbf{p}(i, 0) \in \{(1+\alpha)(1+l), \dots, \nu+1\}$ since $\mathbf{m}(0) > 1$. Moreover, using Proposition 3.1, the property $\hat{\mathbf{E}}_n(1, \nu)$ holds and then (3.5) holds with $l = 1$ and $\nu = \mathbf{p}(i, 0)$. Using successively this property and $\mathbf{E}_n(1)$ with (3.2), it follows that

$$\begin{aligned} & \left\| \left(\prod_{j=1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \right) Q_{t_i, T}^{\delta_n^1} f \right\|_{\infty} \\ &= \left\| \left(\prod_{j=1}^{j_i-1} Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \right) \left(Q_{t_{j_i-1}, t_{j_i} - \delta_n^1}^{\delta_n^1} \Lambda_{t_{j_i} - \delta_n^1, t_{j_i}}^{\delta_n^1, h} \right) \left(\prod_{j=j_i+1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \right) Q_{t_i, T}^{\delta_n^1} f \right\|_{\infty} \\ &\leq C \left\| \left(Q_{t_{j_i-1}, t_{j_i} - \delta_n^1}^{\delta_n^1} \Lambda_{t_{j_i} - \delta_n^1, t_{j_i}}^{\delta_n^1, h} \right) \left(\prod_{j=j_i+1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \right) Q_{t_i, T}^{\delta_n^1} f \right\|_{(j_i-1) \max(\beta, \kappa(1, \mathbf{p}(i, 0))), \infty} \\ &\quad \times \frac{1+T^p}{n \sum_{j=1}^{j_i-1} p_i^h(j)}, \end{aligned}$$

with $p_i^h(j) = \mathbf{p}(i, 0)$ if $\Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} = \hat{Q}_{t_j - \delta_n^1, t_j}^{\mathbf{p}(i, 0)} - P_{t_j - \delta_n^1, t_j}$ and $p_i^h(j) = \alpha + 1$ if $\Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} = P_{t_j - \delta_n^1, t_j} - Q_{t_j - \delta_n^1, t_j}^{\delta_n^1}$ and p is a constant depending on ν and α (and which may change value in the following lines or according to i). Notice that it is not possible to have $p_i^h(j) = \alpha + 1$ for every $j \in \{1, \dots, i\}$ and that there exists $\hat{j}_i \in \{1, \dots, i\}$, such that

$$\sum_{j=1}^i p_i^h(j) = \hat{j}_i \mathbf{p}(i, 0) + (i - \hat{j}_i)(\alpha + 1) \geq \nu + i,$$

Now, for $\varepsilon > 0$, $y \in \mathbb{R}^d$, we consider $\phi_\varepsilon(y) = \varepsilon^{-d} \phi(\varepsilon^{-1}y)$ with $\phi \in C_c^\infty(\mathbb{R}^d)$, $\phi \geq 0$, $\|\phi\|_1 = 1$ and for a fixed $x \in \mathbb{R}^d$, we define $\phi_{\varepsilon, x}(y) = \phi_\varepsilon(x - y)$. Moreover, we denote

$$\Gamma_i = \Lambda_{t_{j_i-1}, t_{j_i} - \delta_n^1}^{\delta_n^1, h} \left(\prod_{j=j_i+1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \right) Q_{t_i, T}^{\delta_n^1}$$

Since we have (3.1) and (3.2), $Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Gamma_i f$ belongs to $\mathcal{C}_b^\infty(\mathbb{R}^d)$.

Similarly as earlier in the proof, since $\mathbf{p}(i, 0) \leq \nu + 1$, it follows from Proposition 3.1 that, $\hat{\mathbf{E}}_n^*(1, \nu)$ holds and then (3.6) holds with $l = 1$ and $\nu = \mathbf{p}(i, 0)$. Using successively this property and $\hat{\mathbf{E}}_n^*(1, \mathbf{p}(i, 0))$ with (3.4), it follows that for a multi-index $\gamma \in \mathbb{N}^d$, $x \in \mathbb{R}^d$,

$$\begin{aligned} |\partial_x^\gamma Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Gamma_i f(x)| &= \lim_{\varepsilon \rightarrow 0} |\langle \Gamma_i^* Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1, *} \partial^\gamma \phi_{\varepsilon, x}, f \rangle| \\ &\leq C \sup_{\varepsilon > 0} \left\| Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1, *} \partial^\gamma \phi_{\varepsilon, x} \right\|_{(i-j_i+1) \max(\beta, \kappa(1, \mathbf{p}(i, 0))), 1} \|f\|_{\infty} \frac{1+T^p}{n \sum_{i=1}^{j_i-1} p_i^h(j)} \end{aligned}$$

Our concern is the case $|\gamma| \leq (j_i - 1) \max(\beta, \kappa(1, \mathbf{p}(i, 0)))$. Using $R_{\mathbf{q}, \eta, T(\nu)}(Q_{\delta_n^1}^{\delta_n^1})$ (see (3.7)) and more particularly the implication (3.8) on $Q_{\delta_n^1, *}^{\delta_n^1}$, it follows that,

$$\left\| \left(\prod_{j=1}^i Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} \right) Q_{t_i, T}^{\delta_n^1} f \right\|_{\infty} \leq C \sup_{\varepsilon > 0} \frac{1}{T(\nu)^\eta} \sup_{x \in \mathbb{R}^d} \|\phi_{\varepsilon, x}\|_1 \|f\|_{\infty} \frac{1+T^p}{n \sum_{j=1}^i p_i^h(j)}$$

and we conclude the estimate noticing that $\sum_{j=1}^i p_i^h(j) \geq \nu + i$ and $\|\phi_{\varepsilon,x}\|_1 = \|\phi\|_1 = 1$ for every $x \in \mathbb{R}^d$.

Step 3. We gather all the estimates together and the proof is completed as long as $f \in \mathcal{C}_c(\mathbb{R}^d)$. The extension to $f \in \mathcal{M}_b(\mathbb{R}^d)$ follows from the Lusin's theorem. \square

We now provide an extension of Theorem 3.1 in which the regularization hypothesis is not required for P or $Q^{\delta_n^1}$ themselves but rather for some modifications of those semigroups close enough in total variation norm. This result is usefull in concrete applications. For instance, it may happens that the measures $Q^\delta(x, dy)$, $x \in \mathbb{R}^d$, are not absolutely continuous *w.r.t.* the Lebesgue measure and then do not satisfy the regularization property. This is also usefull in our approach to prove Theorem 2.1. In particular we will use a modification with form $Q^{\delta,\theta}$ of which regularization properties are stated in Proposition 3.3.

Proposition 3.2. *Let $T > 0$, $n \in \mathbb{N}^*$ and $\eta > 0$. Assume that (3.1) and (3.2) and (3.3), (3.4) hold and that the short time estimates $\mathbf{E}_n(l)$ and $\mathbf{E}_n^*(l)$ hold for every $l \in \{1, \dots, \ell\}$. Also, assume that there exists a modification $\bar{Q}^{\delta_n^1}$ (respectively \bar{P}) of $Q^{\delta_n^1}$ (resp. P) which satisfy $R_{\mathbf{q},\eta,T(\nu)}(\bar{Q}^{\delta_n^1})$ (resp. $R_{\mathbf{q},\eta,T(\nu)}(\bar{P})$) and such that for every $f \in \mathcal{M}_b(\mathbb{R}^d)$ and every $s, t \in \pi^{\delta_n^1}$, $t - s \geq T(\nu)$,*

$$(3.9) \quad \|Q_{s,t}^{\delta_n^1} f - \bar{Q}_{s,t}^{\delta_n^1} f\|_\infty + \|P_{s,t} f - \bar{P}_{s,t} f\|_\infty \leq \frac{C(1+T^{\bar{p}})}{(t-s)^{\eta n^{\nu+\ell}}} \|f\|_\infty$$

where \bar{p} can depend on ν and α . Then, for every $f \in \mathcal{M}_b(\mathbb{R}^d)$,

$$\|P_T f - \hat{Q}_{0,T}^\nu f\|_\infty \leq \frac{1}{n^\nu} \|f\|_\infty \frac{C(1+T^p)}{T(\nu)^\eta}.$$

where $p, C \geq 0$ do not depend on n , T or f .

Proof. The proof follows the same line as the one of the previous Theorem 3.1. Consequently, we only focus on the specificity of this proof, avoiding arguments which are similar to the previous proof. In particular we study, for every $i \in \{1, \dots, \mathbf{m}(0)\}$,

$$\begin{aligned} & \|Q_{t_i, T}^{\delta_n^1} \prod_{j=1}^i \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} f\|_\infty \\ & \leq \|Q_{t_i, T}^{\delta_n^1} \prod_{j=j_i+1}^i \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_{j_i} - \delta_n^1, t_{j_i}}^{\delta_n^1, h} \bar{Q}_{t_{j_i-1}, t_{j_i} - \delta_n^1}^{\delta_n^1} \prod_{j=1}^{j_i-1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} f\|_\infty \\ & \quad + \|Q_{t_i, T}^{\delta_n^1} \prod_{j=j_i+1}^i \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} \Lambda_{t_{j_i} - \delta_n^1, t_{j_i}}^{\delta_n^1, h} \left(Q_{t_{j_i-1}, t_{j_i} - \delta_n^1}^{\delta_n^1} - \bar{Q}_{t_{j_i-1}, t_{j_i} - \delta_n^1}^{\delta_n^1} \right) \\ & \quad \quad \quad \prod_{j=1}^{j_i-1} \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} f\|_\infty. \end{aligned}$$

The first term is studied similarly as in Theorem 3.1 We use the same notations as introduced in this proof. For the second term we use (3.9) together with successive application of (3.2) and it follows that

$$\|Q_{t_i, T}^{\delta_n^1} \prod_{j=1}^i \Lambda_{t_j - \delta_n^1, t_j}^{\delta_n^1, h} Q_{t_{j-1}, t_j - \delta_n^1}^{\delta_n^1} f\|_\infty \leq \frac{C(1+T^{\nu, \alpha})}{T(\nu)^\eta} \|f\|_\infty \frac{1}{n^{\nu+i}} + \frac{C(1+T^{\bar{p}})}{T(\nu)^\eta} \|f\|_\infty \frac{1}{n^{\nu+\ell}}.$$

Notice that the study of the remainder $R_{0,T,\mathbf{m}(0)}^{\delta_n^1}$ which appears in the proof of Proposition 3.1 is similar. Rearranging the terms completes the proof. \square

At this stage, we establish a total variation convergence result that requires the regularization property only for the family of semigroups $(Q^\delta)_{\delta>0}$, rather than for P itself. In the context of the proof of Theorem 2.1, this approach allows us to formulate a set of assumptions that impose conditions exclusively on the function ψ and the random variables Z^δ , as in Theorem 2.1. More specifically, we consider the following hypothesis:

$$\begin{aligned} & \bar{R}_{\mathbf{q},\eta,n}(Q) \\ & \equiv \\ & \text{For every } l \in \mathbb{N}^*, \end{aligned}$$

$\bar{R}_{\mathbf{q},\eta,n}(Q)$.i. (3.2) and (3.4) hold.

$\bar{R}_{\mathbf{q},\eta,n}(Q)$.ii. There exists a modification $\bar{Q}^{\delta_n^l}$ of $Q^{\delta_n^l}$ which satisfies $R_{\mathbf{q},\eta,T(\nu)}(\bar{Q}^{\delta_n^l})$ (see (3.7)) and such that: $\forall s, t \in \pi^{\delta_n^l}, t - s \geq T(\nu)$,

$$(3.10) \quad \|Q_{s,t}^{\delta_n^l} f - \bar{Q}_{s,t}^{\delta_n^l} f\|_{\infty} \leq C(t-s)^{-\eta} \frac{\|f\|_{\infty}}{n^{\nu+\ell}}.$$

Theorem 3.2. *Let $T > 0$, $n \in \mathbb{N}^*$ and $\eta \geq 0$. Assume that $\bar{R}_{\mathbf{q},\eta,n}(Q)$ holds and that for every $l \in \mathbb{N}^*$, $\mathbf{E}_n(l)$ and $\mathbf{E}_n^*(l)$ hold. Then, for every $f \in \mathcal{M}_b(\mathbb{R}^d)$,*

$$(3.11) \quad \|P_T f - \hat{Q}_{0,T}^{\nu} f\|_{\infty} \leq \frac{1}{n^{\nu}} \|f\|_{\infty} \frac{C(1+T^p)}{T(\nu)^{\eta}},$$

where $p, C \geq 0$ do not depend on n, T or f .

Remark 3.1. *The proof of the inequality (3.11) is inspired from Theorem 3.1. However, we cannot use directly this result, since we assume the regularization properties only on Q^{δ} and not for P . Note that our method does not allow one to establish the same result under regularization hypotheses on P instead of Q^{δ} . The reason is that in our proof, P is built as the limit of Q^{δ} as δ tends to zero, and such an argument cannot be reversed because P itself does not depend on δ .*

Finally, we emphasize that this approach highlights that the limit semigroup P is uniquely determined. In particular, in the setting of Theorem 2.1, $V_0 = \partial_y \psi(\cdot, \cdot, 0, 0) + \frac{1}{2} \sum_{i=1}^N \partial_{z_i}^2 \psi(\cdot, \cdot, 0, 0)$, $V_i = \partial_{z_i} \psi(\cdot, \cdot, 0, 0)$, $i \in \{1, \dots, N\}$ as a consequence of Theorem 2.2 in [19].

Proof of Theorem 3.2. We fix $n \in \mathbb{N}^*$ and we study the sequence of discrete semigroups $\left(\left(Q_{s,t}^{\delta_n^l} \right)_{s,t \in \pi^{\delta_n^l}; s \leq t} \right)_{l \in \mathbb{N}^*}$.

Step 1. We show that for every bounded and measurable test function f , $\left(Q_{0,T}^{\delta_n^l} f \right)_{l \in \mathbb{N}^*}$ is Cauchy in $(\mathcal{M}_b(\mathbb{R}^d), \|\cdot\|_{\infty})$ and that, for $l \geq \ell$,

$$(3.12) \quad \left\| P_T f - Q_{0,T}^{\delta_n^l} f \right\|_{\infty} \leq \frac{1}{n^{\nu}} \|f\|_{\infty} \frac{CT^{\alpha+1}}{T(\nu)^{\eta}},$$

where C is a constant depending on ν and α (and which may change from line to line in the following calculus).

For $l' \geq l \in \mathbb{N}^*$, following the Lindeberg decomposition yields

$$\left\| Q_{0,T}^{\delta_n^{l'}} f - Q_{0,T}^{\delta_n^l} f \right\|_{\infty} \leq \sum_{m=1}^n \left\| Q_{0,(m-1)\delta_n^1}^{\delta_n^{l'}} \left(Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^{l'}} - Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^l} \right) Q_{m\delta_n^1, T}^{\delta_n^{l'}} f \right\|_{\infty}.$$

Now notice that for $g \in \mathcal{C}_b^{\beta}(\mathbb{R}^d)$,

$$\begin{aligned} \left\| Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^{l'}} g - Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^l} g \right\|_{\infty} &\leq \left\| P_{(m-1)\delta_n^1, m\delta_n^1} - Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^{l'}} \right\|_{\infty} \\ &\quad + \left\| P_{(m-1)\delta_n^1, m\delta_n^1} - Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^l} \right\|_{\infty} \end{aligned}$$

with, as a consequence of $\mathbf{E}_n(1)$,

$$\left\| P_{(m-1)\delta_n^1, m\delta_n^1} - Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^k} \right\|_{\infty} \leq C \|g\|_{\infty, \beta} \frac{T^{\alpha+1}}{n^{\alpha+1}},$$

if $l = 1$. If $l > 1$, as a consequence of $\mathbf{E}_n(l)$ and (3.2),

$$\begin{aligned} &\left\| P_{(m-1)\delta_n^1, m\delta_n^1} f - Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^l} f \right\|_{\infty} \\ &\leq \sum_{u=1+n^{l-1}(m-1)}^{mn^{l-1}} \left\| P_{(m-1)\delta_n^1, (u-1)\delta_n^k} \left(Q_{(u-1)\delta_n^k, u\delta_n^k}^{\delta_n^l} - P_{(u-1)\delta_n^k, u\delta_n^k} \right) Q_{u\delta_n^k, m\delta_n^1}^{\delta_n^l} f \right\|_{\infty} \\ &\leq C \|g\|_{\infty, \beta} \frac{T^{\alpha+1}}{n^{k\alpha+1}}. \end{aligned}$$

Therefore, using (3.2), $\mathbf{E}_n(l)$ and $\mathbf{E}_n(l')$, we obtain

$$\left\| Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^{l'}} g - Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^l} g \right\|_{\infty} \leq C \|g\|_{\infty, \beta} \frac{T^{\alpha+1}}{n^{l\alpha+1}}.$$

In the same way we deduce from (3.4), $\mathbf{E}_n^*(l)$ and $\mathbf{E}_n^*(l')$ that

$$\left\| Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^{l'}, *}, g - Q_{(m-1)\delta_n^1, m\delta_n^1}^{\delta_n^l, *}, g \right\|_1 \leq C \|g\|_{1, \beta} \frac{T^{\alpha+1}}{n^{\alpha+1}}.$$

Combining those estimates with $R_{\mathbf{q}, \eta, T(\nu)}(\overline{Q}^{\delta_n^l})$ and $R_{\mathbf{q}, \eta, T(\nu)}(\overline{Q}^{\delta_n^{l'}})$ together with (3.10), the same approach as in the proof of Proposition 3.2 yields, for every $f \in \mathcal{M}_b(\mathbb{R}^d)$,

$$(3.13) \quad \left\| Q_{0, T}^{\delta_n^{l'}} f - Q_{0, T}^{\delta_n^l} f \right\|_\infty \leq \frac{1}{n^{l\alpha}} \|f\|_\infty \frac{C(1+T^p)}{T(\nu)^\eta}.$$

where p depends on l and α . The sequence $(Q_{0, T}^{\delta_n^l} f)_{l \in \mathbb{N}^*}$ is thus Cauchy in the Banach space $(\mathcal{M}_b(\mathbb{R}^d), \|\cdot\|_\infty)$ and then $\lim_{l \rightarrow \infty} Q_{0, T}^{\delta_n^l} f$ exists and belongs to $\mathcal{M}_b(\mathbb{R}^d)$. Moreover, remember that as soon as $f \in C_c^\infty(\mathbb{R}^d)$, (3.5) holds and then $Q_{0, T}^{\delta_n^l} f$ converges to $P_T f$ as l tends to infinity so that it is also the case when f is simply bounded and measurable. Taking $l \geq \ell$ in (3.13) and letting l' tends to infinity, it follows that (3.12) holds.

Step 2. We now show that for every $l \geq \ell$ and every $f \in \mathcal{M}_b(\mathbb{R}^d)$,

$$(3.14) \quad \|Q_{0, T}^{\delta_n^l} f - \hat{Q}_{0, T}^\nu f\|_\infty \leq \frac{1}{n^\nu} \|f\|_\infty \frac{C(1+T^p)}{T(\nu)^\eta},$$

where p may depend on ν or α .

To begin, since $l \geq \ell$, the triangle inequality implies that replacing P by $Q^{\delta_n^l}$ preserves the validity of the short time estimates $\mathbf{E}_n(l')$ and $\mathbf{E}_n^*(l')$ for every $l' \in \{1, \dots, \ell\}$.

In addition, from $\overline{R}_{\mathbf{q}, \eta, n}(Q)$.ii., the property $R_{\mathbf{q}, \eta, T(\nu)}(\overline{Q}^{\delta_n^1})$ (resp. $R_{\mathbf{q}, \eta, T(\nu)}(\overline{Q}^{\delta_n^l})$) (see (3.7)) hold for a modification $\overline{Q}^{\delta_n^1}$ (resp. $\overline{Q}^{\delta_n^l}$) of $Q^{\delta_n^1}$ (resp. $Q^{\delta_n^l}$) which satisfies (3.10). Therefore, all the assumptions of Proposition 3.2 are fulfilled when we replace P by $Q^{\delta_n^l}$, so that 3.14 holds.

Step 3. We combine (3.12) and (3.14) and (3.11) follows. □

3.2. Regularization properties of our the approximating semigroup. The following result is a direct application of Theorem 2.1 in [19] to our setting. It demonstrates that one can construct a modification of Q^δ that satisfies the regularization property, and more generally ensures that $\overline{R}_{\mathbf{q}, \eta, T(\nu)}(Q)$ holds. Before stating the result, we introduce this modification. For $s, t \in \pi^\delta$, $s \leq t$, $\theta > 0$, and G a d -dimensional Gaussian random variable with mean 0 and covariance matrix identity and independent of $(Z_{t+\delta}^\delta)_{t \in \pi^\delta}$, we define

$$Q_{s, t}^{\delta, \theta} f(x) = \int_{\mathbb{R}^d} f(y) Q_{s, t}^{\delta, \theta}(x, dy) := \mathbb{E}[f(X_t^\delta + \delta^\theta G) | X_s^\delta = x].$$

Proposition 3.3. *Let $L \in \mathbb{N}$, $S > 0$ and let $f \in C_b^\infty(\mathbb{R}^d; \mathbb{R})$. Then we have the following properties: Let $q \in \mathbb{N}$, let $\gamma, \zeta \in \mathbb{N}^d$ such that $|\gamma| + |\zeta| \leq q$. Assume that $\mathbf{A}_1^\delta(\max(q+3, 2L+5))$, $\mathbf{A}_2(L)$, \mathbf{A}_3^δ and \mathbf{A}_4^δ hold and that δ is small enough. Then, for every $x \in \mathbb{R}^d$ and $s, t \in \pi^\delta$ with $t-s \geq S$,*

$$(3.15) \quad |\partial_x^\zeta Q_{s, t}^{\delta, \theta} \partial^\gamma f(x)| \leq \|f\|_\infty \frac{C \exp(C(t-s))}{((t-s)\mathcal{V}_L)^\eta},$$

where $\eta \geq 0$ and $C \geq 0$ do not depend on δ, S or f . Moreover

$$|Q_{s, t}^\delta f(x) - Q_{s, t}^{\delta, \theta} f(x)| \leq \delta^\theta \|f\|_\infty \frac{C \exp(C(t-s))}{((t-s)\mathcal{V}_L)^\eta}$$

where $\eta \geq 0$ and $C \geq 0$ do not depend on δ, S or f .

Now we are able to prove our main result.

3.3. Proof of Theorem 2.1. The proof is now a direct consequence of the combination of the result obtained in Section 3.1 and Section 3.2. More specifically, we have seen in Proposition 3.3 that $Q^{\delta,\theta}$ verifies the regularization property. We also remark that (3.2) and (3.4) are satisfied as a consequence of Theorem 4.2 in [19]. We chose $S = T(\nu)$ and $\theta \geq \nu + \ell$ so that $\bar{R}_{\mathbf{q},\eta,n}(Q)$ holds if we set $\bar{Q}^\delta = Q^{\delta,\theta}$. The proof of (2.11) is then an immediate consequence of Theorem 3.2.

REFERENCES

- [1] A. Al Gerbi, B. Jourdain, and E. Clément. Ninomiya–victoir scheme: Strong convergence, antithetic version and application to multilevel estimators. *Monte Carlo Methods and Applications*, 22(3):197–228, 2016.
- [2] A. Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comp.*, 79(269):209–237, 2010.
- [3] A. Alfonsi and V. Bally. A generic construction for high order approximation schemes of semigroups using random grids. *Numerische Mathematik*, 148:743 – 793, 2019.
- [4] V. Bally and A. Kohatsu-Higa. A probabilistic interpretation of the parametrix method. *The Annals of Applied Probability*, 25(6):3095 – 3138, 2015.
- [5] V. Bally and C. Rey. Approximation of markov semigroups in total variation distance. *Electron. J. Probab.*, 21:44 pp., 2016.
- [6] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Related Fields*, 104(1):43–60, 1996.
- [7] M. Ben Alaya and A. Kebaier. Central limit theorem for the multilevel Monte Carlo Euler method. *The Annals of Applied Probability*, 25(1):211 – 234, 2015.
- [8] N. Frikha, A. Kohatsu-Higa, and L. Li. Integration by parts formula for killed processes: a point of view from approximation theory. *Electronic Journal of Probability*, 24:1 – 44, 2019.
- [9] M.B. Giles. Multilevel monte carlo path simulation. *Oper. Res.*, 56:607–617, 2008.
- [10] P.W. Glynn and C-H Rhee. Unbiased estimation with square root convergence for sde models. *Operations Research*, 63(5):1026–1043, 2015.
- [11] S. Heinrich. Multilevel monte carlo methods. In *Large-Scale Scientific Computing*, pages 58–67, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.
- [12] P. Henry-Labordère, X Tan, and N. Touzi. Unbiased simulation of stochastic differential equations. *The Annals of Applied Probability*, 27(6):3305 – 3341, 2017.
- [13] A. Kebaier. Statistical romberg extrapolation: A new variance reduction method and applications to option pricing. *The Annals of Applied Probability*, 15(4):2681–2705, 2005.
- [14] V. Lemaire and G. Pagès. Multilevel Richardson–Romberg extrapolation. *Bernoulli*, 23(4A):2643 – 2692, 2017.
- [15] D. McLeish. A general method for debiasing a monte carlo estimator. *Monte Carlo Methods and Applications*, 17(4):301–315, 2011.
- [16] S. Ninomiya and N. Victoir. Weak approximation of stochastic differential equations and application to derivative pricing. *Appl. Math. Finance*, 15(1-2):107–121, 2008.
- [17] G. Pagès. Multi-step richardson-romberg extrapolation: Remarks on variance control and complexity. *Monte Carlo Methods and Applications*, 13(1):37–70, 2007.
- [18] C. Rey. Convergence in total variation distance of a third order scheme for one-dimensional diffusion processes. *Monte Carlo Methods and Applications*, 23(1):1–12, 2017.
- [19] C. Rey. Hörmander properties of discrete time markov processes, 2024.
- [20] D. Talay. Second-order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stochastics and Stochastic Reports*, 29(1):13–36, 1990.
- [21] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.
- [22] M. Vihola. Unbiased estimators and multilevel monte carlo. *Operations Research*, 66(2):pp. 448–462, 2018.