

Approximation of Markov semigroups in total variation distance under an irregular setting: An application to the CIR process

Clément Rey

▶ **To cite this version:**

Clément Rey. Approximation of Markov semigroups in total variation distance under an irregular setting: An application to the CIR process. Stochastic Processes and their Applications, 2019, 129 (2), pp.539-571. hal-01412024v3

HAL Id: hal-01412024
https://hal.science/hal-01412024v3

Submitted on 22 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Approximation of Markov semigroups in total variation distance under an irregular setting: An application to the CIR process

Clément Rey [1]

Abstract

In this paper, we propose a method to prove the total variation convergence of approximation of Markov semigroups with singularities. In particular our approach is adapted to the study of numerical schemes for Stochastic Differential Equation (SDE) with simply locally smooth coefficients. First we present this method and then, we apply it to the CIR process. In particular, we consider the weak second order scheme introduced in [2] and we prove that it also converges towards the CIR diffusion process for the total variation distance. This convergence occurs with almost order two.

1 Introduction

In this paper we study the total variation distance between a Markov process presenting singularities and its approximation based on a discrete Markov chain. In this purpose, we follow two steps. First, we extend a result from [5]. Then, we apply this result to a second weak order scheme for CIR diffusion process based on a cubature method and introduced in [2]. While, this scheme has second weak order for smooth test functions, in this paper we are able to prove that the convergence for a class of bounded and measurable functions occurs with almost order 2. Let us be more specific. For $N \in \mathbb{N}^*$, we consider the $\mathbb{R}^d$-valued diffusion process

$$dX_t = V_0(X_t)dt + \sum_{i=1}^{N} V_i(X_t) \circ dW_t^i \tag{1}$$

with $V_i \in \mathcal{C}^\infty(\mathbb{R}^d;\mathbb{R}^d) \cap \mathcal{C}_b^\infty(\mathcal{D};\mathbb{R}^d)$, where $\mathcal{D}$ is a subset of $\mathbb{R}^d$, $(W_t)_{t\geqslant 0}$ a standard Brownian motion and $\circ dW_t^i$ denotes the Stratonovich integral with respect to W_t^i. We fix $T > 0$ and $n \in \mathbb{N}^*$ and we introduce the time grid $\pi_{T,n} = \{t_k^n = kT/n, k \in \mathbb{N}\}$. We consider the d-dimensional approximation Markov chain

$$X^n_{t^n_{k+1}} = \psi_k(X^n_{t^n_k}, \frac{Z_{k+1}}{\sqrt{n}}, \delta^n_{k+1}), \qquad k \in \mathbb{N}, \tag{2}$$

where $\psi_k : \mathbb{R}^d \times \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}^d$ is a function such that $\psi_k(x,0,0) = x$, and $Z_k \in \mathbb{R}^N$, $k \in \mathbb{N}^*$, is a sequence of independent and centered random variables and $\sup_{k\in\mathbb{N}^*} \delta_k^n \leqslant C/n$. So far, we do not

[1] clement.rey@upmc.fr, Laboratoire de Probabilités et Modèles Aléatoires-Université Paris VI, 4 place Jussieu - 75005 Paris - France.
This research benefited from the support of the "Chaire Risques Financiers", Fondation du Risque.

discuss the regularity of ψ_k. We aim to study the convergence of the law of X^n to the law of the Markov process $(X_t)_{t\geqslant 0}$ defined in (1). More precisely, we aim to estimate the weak error distance

$$\varepsilon_n(f) = \left|\mathbb{E}[f(X_t^n)] - \mathbb{E}[f(X_t)]\right|.$$

In order to obtain total variation convergence for $(\varepsilon_n)_{n\in\mathbb{N}^*}$, we have to show that $\varepsilon_n(f) \to 0$ for every bounded and measurable function f. The method we adopt in this paper is inspired from [5] and is based on the semigroup approach.

First, we introduce some notations. The semigroup of the Markov chain $(X_t^n)_{t\in\pi_{T,n}}$ is denoted by $(Q_t^n)_{t\in\pi_{T,n}}$ and its transition probabilities are given by $\nu_{k+1}^n(x,dy) = \mathbb{P}(X_{t_{k+1}^n}^n \in dy | X_{t_k^n}^n = x)$, $k \in \mathbb{N}$. We recall that for $t \in \pi_{T,n}$, $Q_t^n f(x) = \mathbb{E}[f(X_t^n)|X_0^n = x]$. We also consider a Markov process in continuous time $(X_t)_{t\geqslant 0}$ with semigroup $(P_t)_{t\geqslant 0}$ (see [31]) and we define $\mu_{k+1}^n(x,dy) = \mathbb{P}(X_{t_{k+1}^n} \in dy | X_{t_k^n}^n = x)$.

Moreover, for $f \in \mathcal{C}^\infty(\mathbb{R}^d)$ and for a multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, we denote $|\alpha| = \alpha_1 + ... + \alpha_d$ and $\partial_\alpha f = (\partial_1)^{\alpha_1} \ldots (\partial_d)^{\alpha_d} f = \partial_x^\alpha f(x) = \partial_{x_1}^{\alpha_1} \ldots \partial_{x_d}^{\alpha_d} f(x)$. We include the multi-index $\alpha = (0, ..., 0)$ and in this case $\partial_\alpha f = f$. Let us introduce the norms

$$\|f\|_{q,\infty} = \sup_{x\in\mathbb{R}^d} \sum_{0\leqslant|\alpha|\leqslant q} |\partial_\alpha f(x)|, \qquad \|f\|_{q,1} = \sum_{0\leqslant|\alpha|\leqslant q} \int_{\mathbb{R}^d} |\partial_\alpha f(x)| dx.$$

In particular $\|f\|_{0,\infty} = \|f\|_\infty$ is the usual supremum norm and we denote by $\mathcal{C}_b^q(\mathbb{R}^d) = \{f \in \mathcal{C}^q(\mathbb{R}^d), \|f\|_{q,\infty} < +\infty\}$ and by $C_c^q(\mathbb{R}^d) \subset C^q(\mathbb{R}^d)$ the set of functions with compact support. Moreover, we say that a function $f \in \mathcal{C}^q(\mathbb{R}^d)$ has polynomial growth of order $q \in \mathbb{N}$ with degree $\beta_q \in \mathbb{N}$ if there exists $C_q \geqslant 1$ such that

$$\forall x \in \mathbb{R}^d, \quad \sum_{0\leqslant|\alpha|\leqslant q} |\partial_\alpha f(x)| \leqslant C_q(1 + |x|^{\beta_q}). \tag{3}$$

Finally, we denote by $\mathcal{C}_{\text{pol}}^q(\mathbb{R}^d)$ the set of functions satisfying (3).

A first standard result is the following. Let us assume that there exists $h > 0$ and $q \in \mathbb{N}$ such that for every $f \in \mathcal{C}^q(\mathbb{R}^d)$ and every $k \in \mathbb{N}^*$,

$$\sup_{x\in\mathbb{R}^d} \left|\mu_k^n f(x) - \nu_k^n f(x)\right| = \sup_{x\in\mathbb{R}^d} \left|\int f(y)\mu_k^n(x,dy) - \int f(y)\nu_k^n(x,dy)\right| \leqslant C\|f\|_{q,\infty}/n^{1+h}.$$

Then, for every $t \in \pi_{T,n}$, we have

$$\|P_t f - Q_t^n f\|_\infty = \sup_{x\in\mathbb{R}^d} |\mathbb{E}[f(X_t^n)|X_0^n = x] - \mathbb{E}[f(X_t)|X_0 = x]| \leqslant C\|f\|_{q,\infty}/n^h. \tag{4}$$

It means that $(X_t^n)_{t\in\pi_{T,n}}$ is an approximation scheme of weak order h for the Markov process $(X_t)_{t\geqslant 0}$. In the case of the Euler scheme for diffusion processes, this result, with $h = 1$, was initially proved in the seminal papers of Milstein [27] and of Talay and Tubaro [34] (see also [19]). Later, similar results were obtained in various situations: Diffusion processes with jumps (see [32], [17]) or diffusion processes with boundary conditions (see [14], [11], [15]). An overview of this subject is given in [18]. More recently, approximation schemes of higher orders (*e.g.*, $h = 2$), based on cubature methods, have been introduced and studied by Kusuoka [23], Lyons [26], Ninomiya, Victoir [28] or Alfonsi [2].

The reader may also refer to the work of Kohatsu-Higa and Tankov [20] for a higher weak order scheme for jump processes. Despite the fact that most of these results concern diffusions with regular coefficients, some papers treat more exotic cases. For instance, in [2], Alfonsi studied the weak error for diffusion processes with coefficients that belong to $\mathcal{C}^q_{\text{pol}}(\mathbb{R}^d)$ as well as for the test functions.

Another result concerns convergence in total variation distance: We want to obtain (4) with $\|f\|_\infty$ instead of $\|f\|_{q,\infty}$, when f is a bounded and measurable function. In the case of the Euler scheme for diffusion processes, a first similar result was obtained by Bally and Talay [6], [7] using the Malliavin calculus (see also Guyon [16]). Afterwards Konakov, Menozzi and Molchanov [21], [22] obtained similar results using a parametrix method. Later, Kusuoka [24] obtained estimates of the error in total variation distance for the Victoir Ninomiya scheme (which corresponds to the case $h = 2$). More recently, in [5], a generic result ensures the total variation distance convergence for smooth schemes and for random variables $Z_k, k \in \mathbb{N}^*$, which satisfy the Doeblin condition. In [33], this method is used to prove total variation convergence with order 3 for a numerical scheme for one dimensional SDE.

However, none of the results mentioned above concerns the case of diffusion processes with irregular coefficients. In particular, in [5], the regularity in space for the functions ψ_k, which often rely on the regularity of the functions V_i, is essential to prove the total variation convergence result. The main idea of our approach is to consider that these functions V_i, $i \in \{1, \dots, N\}$, are smooth only on a subset $\mathcal{D}$ of $\mathbb{R}^d$. Simply using this local regularity, we prove the convergence for every bounded and measurable test functions with support strictly contained in $\mathcal{D}$. In order to do it, we combine some results of convergence for smooth test functions for singular diffusions on $\mathbb{R}^d$ with the total variation convergence proven in [5] for a modification of $(X_t)_{t\geqslant 0}$ with coefficients localized on $\mathcal{D}$. In a first step, we propose an abstract approach based on semigroups that leads to Theorem 2.1. Then we apply this result to the case of Markov chain approximations of diffusion processes with locally smooth coefficients and obtain Theorem 3.1. It provides an estimation of the weak error for simply bounded and measurable test functions with support contained in $\mathcal{D}$. In the rest of this paper we sometimes take the liberty to say that this is the convergence for the total variation distance.[2] It is important to notice that Theorem 3.1 is not restricted to specific diffusions or schemes.

Using this approach, we then study the CIR process. In this case, the diffusion coefficient is given by $V_1(x) = \sigma\sqrt{x}$, $\sigma \in \mathbb{R}_+$, and is singular in zero. That is why, standard estimation methods can not be applied. However, some papers manage to develop numerical analysis of this scheme using among other the close link that exists with Bessel processes. The reader may refer to [12], [3], [1], [2] or [10] for a non exhaustive list of studies concerning numerical approximation for the CIR process. In particular, in [2], the author proves the weak convergence with order 2 for smooth test functions (under polynomial growth assumptions for the test function and its derivatives), of a scheme based on cubature method (and also inspired by [3]). In [10], the authors propose an expansion of the weak error for Lipschitz test functions. However, so far, there is no study concerning the total variation convergence of a numerical scheme toward the CIR diffusion. Since the diffusion process has a singularity in zero, the result from [5] can not apply directly. Despite this singularity, we use and extend the result from [2] and [5], and prove that (4) is also satisfied for the scheme introduced in [2] under some hypotheses on the support of the test functions. More specifically we obtain the

[2] The total variation distance concerns every bounded and measurable test function with support in $\mathbb{R}^d$ (and not simply in $\mathcal{D}$). However, the total variation convergence can be deduced straightly from our result if we suppose that $\mathbb{P}(X_T \notin \mathcal{D})$ and $\mathbb{P}(X_T^n \notin \mathcal{D})$ are small enough.

following results:

First (see (62)), there exists $C, \beta \geqslant 0$ such that for every bounded and measurable test function f with $\text{supp}(f) \subset [d_1, d_2]$, $0 < d_1 \leqslant d_2 < +\infty$ and every n large enough, we have

$$|\mathbb{E}[f(X_T(x)) - f(X_T^n(x))]| \leqslant C(1 + |x|^\beta)\|f\|_\infty \ln(n)^\zeta / n^2,$$

with $(X^n_{kT/n}(x))_{k\in\mathbb{N}}$ the scheme introduced in [2].

Moreover, using the finiteness of some exponential moments of the CIR process, we obtain the following result under some appropriate assumptions (see (63)): There exists $\epsilon \in (0, 2)$ such that for every bounded and measurable test function f with $\text{supp}(f) \subset [d_1, \infty)$, $d_1 > 0$, and for every n large enough, we have

$$|\mathbb{E}[f(X_T(x)) - Y_T^n(x)]| \leqslant C \exp(\beta|x|)\|f\|_\infty / n^{2-\epsilon},$$

with $Y_T^n(x)$ a $\sigma(X_t^n(x), t \leqslant T)$-measurable random variable introduced in (63).

We begin by presenting the abstract semigroup framework of this paper in Section 2. In the same Section, we obtain Theorem 2.1 that is the abstract total variation convergence result under localized regularization properties of the semigroups. Then, in Section 3 we apply the abstract framework to the case of generic numerical schemes for diffusion processes with locally smooth coefficients and obtain Theorem 3.1. The paper ends with a theoretical application in order to obtain total variation convergence results for the scheme presented in [2] for the CIR process.

2 The distance between two semigroups

Throughout this section the following notations prevail. We fix $T > 0$ and we denote by $n \in \mathbb{N}^*$, the number of time steps between 0 and T. Then, for $k \in \mathbb{N}$ we define $t_k^n = kT/n$ and we introduce the homogeneous time grid $\pi_{T,n} = \{t_k^n = kT/n, k \in \mathbb{N}\}$ and its bounded version $\pi_{T,n}^{\tilde{T}} = \{t \in \pi_{T,n}, t \leqslant \tilde{T}\}$ for $\tilde{T} \geqslant 0$. Finally, for $S \in [0, \tilde{T})$ we denote $\pi_{T,n}^{S,\tilde{T}} = \{t \in \pi_{T,n}^{\tilde{T}}, t > S\}$. Notice that all the results from this paper remain true with non homogeneous time steps but, for sake of clarity, we do not consider this case.

2.1 Convergence of semigroups

In this section we show how regularity properties and estimation properties of the semigroups interact to lead to weak convergence results. First, we establish estimation results for smooth test functions. Then, under regularization properties of the semigroups, we provide total variation distance estimation results, *i.e.* for simply bounded and measurable test functions. We conclude by establishing Theorem 2.1 which is the main abstract result of this paper and provide an approach for total variation convergence when regularization properties are only valid in the neighborhood of the terminal date T.

2.1.1 Smooth test functions

We consider a sequence of finite transition measures $\mu_k^n(x,dy),\ k\in\mathbb{N}^*$ from $\mathbb{R}^d$ to $\mathbb{R}^d$. This means that for each fixed x and k, $\mu_k^n(x,dy)$ is a finite measure on $\mathbb{R}^d$ with the borelian σ field and for each bounded and measurable function $f:\mathbb{R}^d\to\mathbb{R}$, the application

$$x\mapsto \mu_k^n f(x):=\int_{\mathbb{R}^d} f(y)\mu_k^n(x,dy)$$

is measurable. We also denote

$$\forall x\in\mathbb{R}^d,\quad |\mu_k^n(x)|:=\sup_{\|f\|_\infty\leqslant 1}\Big|\int_{\mathbb{R}^d} f(y)\mu_k^n(x,dy)\Big|\quad\text{and}\quad |\mu_k^n|:=\sup_{x\in\mathbb{R}^d}\sup_{\|f\|_\infty\leqslant 1}\Big|\int_{\mathbb{R}^d} f(y)\mu_k^n(x,dy)\Big|,$$

and we assume that all the sequences of measures we consider in this paper satisfy

$$\sup_{k\in\mathbb{N}^*}|\mu_k^n|<\infty. \tag{5}$$

Although the main application concerns the case where $\mu_k^n(x,dy)$ is a probability measure, we do not make such assumption. We allow $\mu_k^n(x,dy)$ to be a signed measure of finite (but arbitrary) total mass. This is because one may use the results from this section not only in order to estimate the distance between two semigroups but also in order to obtain an expansion of the error. Now we associate the sequence of measures to the time grid $\pi_{T,n}$ and we define the following discrete semigroup,

$$P_0^n f(x)=f(x),\qquad P_{t_{k+1}^n}^n f(x)=P_{t_k^n}^n\mu_{k+1}^n f(x)=P_{t_k^n}^n\int_{\mathbb{R}^d} f(y)\mu_{k+1}^n(x,dy).$$

More generally, we define $(P_{t,s})_{t,s\in\pi_{T,n};t\leqslant s}$ by

$$P_{t_k^n,t_k^n}^n f(x)=f(x),\qquad \forall k,r\in\mathbb{N}^*,k\leqslant r,\quad P_{t_k^n,t_{r+1}^n}^n f(x)=P_{t_k^n,t_r^n}^n\mu_{r+1}^n f(x).$$

We notice that for $t,s,u\in\pi_{T,n}$, $t\leqslant s\leqslant u$, we have the semigroup property: $P_{t,u}^n f=P_{t,s}^n P_{s,u}^n f$. We consider the following hypothesis: Let $q\in\mathbb{N}$ and $t,s\in\pi_{T,n}$, $t\leqslant s$. If $f\in\mathcal{C}^q(\mathbb{R}^d)$ then $P_{t,s}f\in\mathcal{C}^q(\mathbb{R}^d)$ and when $f\in\mathcal{C}_b^q(\mathbb{R}^d)$,

$$\sup_{t,s\in\pi_{T,n};t\leqslant s}\|P_{t,s}^n f\|_{q,\infty}\leqslant C\|f\|_{q,\infty}. \tag{6}$$

Notice that (5) implies that (6) holds for $q=0$. We also consider the following hypothesis: Let $q\in\mathbb{N}$ and $t,s\in\pi_{T,n}$, $t\leqslant s$. If $f\in\mathcal{C}_{\rm pol}^q(\mathbb{R}^d)$ then $P_{t,s}f\in\mathcal{C}_{\rm pol}^q(\mathbb{R}^d)$ and there exists $C>1,\beta\in\mathbb{N}$ such that

$$\mathfrak{P}_q\quad\equiv\quad\forall x\in\mathbb{R}^d,\quad\sup_{t,s\in\pi_{T,n};t\leqslant s}\sum_{0\leqslant|\alpha|\leqslant q}|\partial_\alpha P_{t,s}^n f(x)|\leqslant C(1+|x|^\beta). \tag{7}$$

Moreover we assume that there exists $C>1$ and $\beta\in\mathbb{N}$ such that for every $f\in\mathcal{C}_b^q(\mathbb{R}^d)$, we have

$$\mathfrak{P}'_q\quad\equiv\quad\forall x\in\mathbb{R}^d,\quad\sup_{t,s\in\pi_{T,n};t\leqslant s}\sum_{0\leqslant|\alpha|\leqslant q}|\partial_\alpha P_{t,s}^n f(x)|\leqslant C(1+|x|^\beta)\|f\|_{q,\infty}. \tag{8}$$

We consider now a second sequence of finite transition measures $\nu_k^n(x,dy), k \in \mathbb{N}^*$ and the corresponding semigroup $(Q_t^n)_{t\in\pi_{T,n}}$ defined as above. We aim to estimate the distance between $P^n f$ and $Q^n f$ in terms of the distance between the transition measures $\mu_k^n(x,dy)$ and $\nu_k^n(x,dy)$, so we denote

$$\Delta_k^n = \mu_k^n - \nu_k^n.$$

$(P_t^n)_{t\in\pi_{T,n}}$ can be seen as a semigroup in continuous time $(P_t)_{t\geqslant 0}$ considered on the time grid $\pi_{T,n}$, while $(Q_t)_{t\in\pi_{T,n}}$ would be its approximation discrete semigroup. Let $q \in \mathbb{N}$ and $h \geqslant 0$ be fixed. We introduce a short time error approximation assumption: There exists a constant $C > 0$ (depending on q only) such that for every $f \in \mathcal{C}_b^q(\mathbb{R}^d)$ and every $k \in \mathbb{N}^*$, we have

$$E_n(h,q) \quad \equiv \quad \|\Delta_k^n f\|_\infty \leqslant C\|f\|_{q,\infty}/n^{h+1}. \tag{9}$$

We also introduce a short time error approximation assumption for test functions with polynomial growth: If $f \in \mathcal{C}_{\mathrm{pol}}^q(\mathbb{R}^d)$, then there exists $C > 1$ and $\beta \in \mathbb{N}$ such that for every $k \in \mathbb{N}^*$,

$$E_{n,\mathrm{pol}}(h,q) \quad \equiv \quad \forall x \in \mathbb{R}^d, \quad |\Delta_k^n f(x)| \leqslant C(1+|x|^\beta)/n^{h+1}. \tag{10}$$

Moreover, we assume that there exists $C > 1$ and $\beta \in \mathbb{N}$ such that for every $f \in \mathcal{C}_b^q(\mathbb{R}^d)$ and every $k \in \mathbb{N}^*$,

$$E'_{n,\mathrm{pol}}(h,q) \quad \equiv \quad \forall x \in \mathbb{R}^d, \quad |\Delta_k^n f(x)| \leqslant C(1+|x|^\beta)\|f\|_{q,\infty}/n^{h+1}. \tag{11}$$

At this point, we establish weak convergence results for smooth test functions.

Proposition 2.1. *Let $q, h \in \mathbb{N}$ be fixed.*

A. *Assume that ν^n satisfies (6) for this q, μ^n satisfy (5) and that we have $E_n(h,q)$ (see (9)). Then, there exists $C > 1$ such that for every $f \in \mathcal{C}_b^q(\mathbb{R}^d)$, we have*

$$\sup_{t\in\pi_{T,n}^T} \|P_t^n f - Q_t^n f\|_\infty \leqslant C\|f\|_{q,\infty}/n^h. \tag{12}$$

B. *Assume that μ^n and ν^n satisfy respectively $\mathfrak{P}_0$ (see (7)) and $\mathfrak{P}_q$, and that $E_{n,\mathrm{pol}}(h,q)$ (see (10)) holds. Then, for every $f \in \mathcal{C}_{\mathrm{pol}}^q(\mathbb{R}^d)$, there exists $C > 1$ and $\beta \in \mathbb{N}$ such that*

$$\forall x \in \mathbb{R}^d, \quad \sup_{t\in\pi_{T,n}^T} |P_t^n f(x) - Q_t^n f(x)| \leqslant C(1+|x|^\beta)/n^h. \tag{13}$$

C. *Assume that μ^n and ν^n satisfy respectively $\mathfrak{P}_0$ (see (7)) and $\mathfrak{P}'_q$ (see (8)) and that $E'_{n,\mathrm{pol}}(h,q)$ (see (11)) holds. Then, there exists $C > 1$ and $\beta \in \mathbb{N}$ such that for every $f \in \mathcal{C}_b^q(\mathbb{R}^d)$, we have*

$$\forall x \in \mathbb{R}^d, \quad \sup_{t\in\pi_{T,n}^T} |P_t^n f(x) - Q_t^n f(x)| \leqslant C(1+|x|^\beta)\|f\|_{q,\infty}/n^h. \tag{14}$$

Proof. We simply prove (13) and (14).The proof of (12) is similar but simpler so we leave it to the reader. Let $m \in \mathbb{N}^*$, $m \leqslant n$. Using the Lindeberg decomposition, we have from the semigroup property

$$\begin{aligned} |P_{t_m^n}^n f(x) - Q_{t_m^n}^n f(x)| &\leqslant \sum_{k=0}^{m-1} |P_{t_k^n}^n P_{t_k^n,t_{k+1}^n}^n Q_{t_{k+1}^n,t_m^n}^n f(x) - P_{t_k^n}^n Q_{t_k^n,t_{k+1}^n}^n Q_{t_{k+1}^n,t_m^n}^n f(x)| \\ &= \sum_{k=0}^{m-1} |P_{t_k^n}^n \Delta_{k+1}^n Q_{t_{k+1}^n,t_m^n}^n f(x)|. \end{aligned} \tag{15}$$

Now, we prove (13). Since $f \in \mathcal{C}^q_{\rm pol}(\mathbb{R}^d)$ then, using $\mathfrak{P}_q$ (see (7)) for Q^n, we have $Q^n_{t^n_{k+1},t^n_m} f \in \mathcal{C}^q_{\rm pol}(\mathbb{R}^d)$ and then using (10) and the linearity of the semigroup P^n together with $\mathfrak{P}_0$ (see (7)), we obtain

$$|P^n_{t^n_k} \Delta^n_{k+1} Q^n_{t^n_{k+1},t^n_m} f(x)| \leqslant C(1+|x|^\beta)/n^{h+1}.$$

Summing over $k = 0, ..., m-1$, yields (13). In order to prove (14), we use (10) to obtain $|\Delta_{k+1} Q^n_{t^n_{k+1},t^n_m} f(x)| \leqslant C(1+|x|^\beta)\|Q^n_{t^n_{k+1},t^n_m} f\|_{q,\infty}/n^{h+1}$ where C and β do not depend on f. Using once again the linearity of the semigroup P^n and $\mathfrak{P}_0$ (see (7)), it follows that $|P^n_{t^n_k} \Delta_{k+1} Q^n_{t^n_{k+1},t^n_m} f(x)| \leqslant C(1+|x|^\beta)\|Q^n_{t^n_{k+1},t^n_m} f\|_{q,\infty}/n^{h+1}$. Finally, the assumption $\mathfrak{P}'_q$ (see (8)) for Q^n gives (14). □

2.1.2 Measurable test functions (convergence in total variation distance)

The estimates (12), (13) and (14) require quite strong regularity properties for the test function f. We aim to show that, if the underlying semigroups have a regularization property, then we may obtain estimates of the error for simply bounded and measurable test functions. In this section, we describe how the regularization properties of a semigroup lead to total variation convergence based on results from [5]. Our purpose is then to obtain similar results when regularization properties only holds closely to the terminal date T.

A first hypothesis concerns the adjoint semigroup approximation. Let $q \in \mathbb{N}$. We assume that there exists a constant $C \geqslant 1$ such that for every bounded and measurable function f and every $g \in \mathcal{C}^q(\mathbb{R}^d)$

$$E^*_n(h,q) \quad \equiv \quad |\langle g, \Delta^n_k f\rangle| \leqslant C\|g\|_{q,1}\|f\|_\infty/n^{1+h}. \tag{16}$$

where $\langle g, f\rangle = \int g(x)f(x)dx$ is the scalar product in $\mathrm{L}^2(\mathbb{R}^d)$.

Our regularization hypothesis is the following. Let $q \in \mathbb{N}$, $S > 0$ and let $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing function. We assume that there exists a constant $C \geqslant 1$ such that for every bounded and measurable function f,

$$R_{q,\eta}(S) \quad \equiv \quad \forall t, s \in \pi_{T,n},\ \text{with } S \leqslant s-t, \quad \|P^n_{t,s} f\|_{q,\infty} \leqslant \frac{C}{S^{\eta(q)}}\|f\|_\infty. \tag{17}$$

We also consider the "adjoint regularization hypothesis". We assume that there exists an adjoint semigroup $(P^{n,*}_{t,s})_{t,s\in\pi_{T,n};t\leqslant s}$, defined by

$$\forall t, s \in \pi_{T,n},\ t \leqslant s, \quad \left\langle P^{n,*}_{t,s} g, f\right\rangle = \left\langle g, P^n_{t,s} f\right\rangle$$

for every bounded and measurable function f and every function $g \in \mathcal{C}^\infty_c(\mathbb{R}^d)$. In addition we suppose that for every $g \in \mathrm{L}^1(\mathbb{R}^d)$, we have

$$R^*_{q,\eta}(S) \quad \equiv \quad \forall t, s \in \pi_{T,n},\ \text{with } S \leqslant s-t, \quad \|P^{n,*}_{t,s} g\|_{q,1} \leqslant \frac{C}{S^{\eta(q)}}\|g\|_1. \tag{18}$$

Notice that a sufficient condition for $R^*_{q,\eta}(S)$ to hold is the following: For every bounded and measurable function $f \in \mathcal{C}^{|\alpha|}(\mathbb{R}^d)$ and for every multi index α with $|\alpha| \leqslant q$, assume that

$$\forall t, s \in \pi_{T,n},\ \text{with } S \leqslant s-t, \quad \|P^n_{t,s}\partial_\alpha f\|_\infty \leqslant \frac{C}{S^{\eta(q)}}\|f\|_\infty.$$

Indeed, for every $g \in \mathrm{L}^1(\mathbb{R}^d)$, it follows that

$$\begin{aligned}\|\partial_\alpha P^{n,*}_{t,s} g\|_1 &\leqslant \sup_{\|f\|_\infty \leqslant 1} |\left\langle \partial_\alpha P^{n,*}_{t,s} g, f\right\rangle| = \sup_{\|f\|_\infty \leqslant 1} |\left\langle g, P^n_{t,s}(\partial_\alpha f)\right\rangle| \\ &\leqslant \|g\|_1 \sup_{\|f\|_\infty \leqslant 1} \|P^n_{t,s}(\partial_\alpha f)\|_\infty \leqslant \frac{C}{S^{\eta(q)}}\|g\|_1.\end{aligned}$$

Finally, we consider the following stronger regularization property: For every multi-index α, β with $|\alpha| + |\beta| = q$ and every bounded and measurable function $f \in \mathcal{C}^q(\mathbb{R}^d)$,

$$\overline{R}_{q,\eta}(S) \quad \equiv \quad \forall t, s \in \pi_{T,n}, \text{ with } S \leqslant s - t, \quad \|\partial_\alpha P^n_{t,s} \partial_\beta f\|_\infty \leqslant \frac{C}{S^{\eta(q)}}\|f\|_\infty. \tag{19}$$

We notice that $\overline{R}_{q,\eta}(S)$ implies both $R_{q,\eta}(S)$ and $R^*_{q,\eta}(S)$ and that a semigroup satisfying $\overline{R}_{q,\eta}$ is absolutely continuous with respect to the Lebesgue measure.

Now we can state our first result concerning total variation convergence between P^n and Q^n.

Proposition 2.2. *Let $q \in \mathbb{N}$, $h \geqslant 0$, $S \in [T/n, T/2)$ and let $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing function. Assume that $E_n(h, q)$ (see (9)) and $E^*_n(h, q)$ (see (16)) hold for P^n and Q^n. Also assume that P^n satisfies $R_{q,\eta}(S)$ (see (17)) and Q^n satisfies $R^*_{q,\eta}(S)$ (see (18)) and that (5) and (6) with $q = 0$ hold for both of them. Then, for every bounded and measurable function f, we have*

$$\sup_{t \in \pi^{2S,T}_{T,n}} \|P^n_t f - Q^n_t f\|_\infty \leqslant \frac{C}{S^{\eta(q)}}\|f\|_\infty / n^h.$$

In concrete applications the following slightly more general variant of the above proposition will be useful.

Proposition 2.3. *Let $q \in \mathbb{N}$, $h \geqslant 0$, $S \in [T/n, T/2)$ and let $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing function. Assume that $E_n(h, q)$ (see (9)) and $E^*_n(h, q)$ (see (16)) hold for P^n and Q^n. Moreover, assume that there exists $(P^{reg,n}_{t,s})_{t,s \in \pi_{T,n}; t \leqslant s}$ which satisfies $R_{q,\eta}(S)$ (see(17)) and $(Q^{reg,n}_{t,s})_{t,s \in \pi_{T,n}; t \leqslant s}$ which satisfies $R^*_{q,\eta}(S)$ (see (18)) and such that (5) and (6) with $q = 0$ hold for both of them. Also assume that for every bounded and measurable function f and every $t, s \in \pi_{T,n}$ with $s - t \geqslant S$, we have*

$$\|Q^n_{t,s} f - Q^{reg,n}_{t,s} f\|_\infty + \|P^n_{t,s} f - P^{reg,n}_{t,s} f\|_\infty \leqslant C S^{-\eta(q)} \|f\|_\infty / n^{h+1}. \tag{20}$$

Then, for every bounded and measurable function f,

$$\sup_{t \in \pi^{2S,T}_{T,n}} \|P^n_t f - Q^n_t f\|_\infty \leqslant C \sup_{k \leqslant n}(|\mu^n_k| + |\nu^n_k|) S^{-\eta(q)} \|f\|_\infty / n^h.$$

Remark 2.1. *Notice that $P^{reg,n}$ and $Q^{reg,n}$ are not supposed to satisfy the semigroup property and are not directly related to μ^n and ν^n.*

The proof of those results can be found in [5] (see Proposition 2.3 and Proposition 2.4) and follows similar ideas from the one of Proposition 2.1 combined with regularizations properties.

2.1.3 Total variation convergence under mixing regularization properties

In this section, we consider semigroups with mixing regularization properties. In particular, we study two semigroups P^n and Q^n which satisfy the regularization properties only closely to the date T and we show that the convergence for bounded and measurable test functions still holds.

Theorem 2.1. *Let $(P^n_{t,s})_{t,s\in\pi_{T,n};t\leqslant s}$ and $(Q^n_{t,s})_{t,s\in\pi_{T,n};t\leqslant s}$ be two semigroups with transition measures μ^n and ν^n. Let $q\in\mathbb{N}$, h, $\delta\in[2T/n,T]$ and let $\eta:\mathbb{R}_+\to\mathbb{R}_+$ be an increasing function. Also define*

$$\tilde{\delta}_n=\inf\{t;t>\delta,T-t\in\pi_{T,n}\},\qquad t_{\delta,n}=T-\tilde{\delta}_n,\qquad k_{\delta,n}=t_{\delta,n}n/T. \tag{21}$$

On the interval $[0,t_{\delta,n}]$, assume that $E'_{n,\mathrm{pol}}(h,q)$ (see (11)) holds between $(P^n_{t,s})_{t,s\in\pi^{t_{\delta,n}}_{T,n};t\leqslant s}$ and $(Q^n_{t,s})_{t,s\in\pi^{t_{\delta,n}}_{T,n};t\leqslant s}$ and that one of those semigroups satisfies $\mathfrak{P}_0$ (see (7)) while the other satisfies $\mathfrak{P}'_q$ (see (8)). Then we have the following properties:

A. *On the interval $[t_{\delta,n},T]$, assume that both $(\mu^n_k)_{k\in\{1,\dots,k_{\delta,n}\}}$ and $(\nu^n_k)_{k\in\{1,\dots,k_{\delta,n}\}}$ satisfy (5) and that $E_n(h,q)$ (see (9)) and $E^*_n(h,q)$ (see (16)) hold between $(P^n_{t,s})_{t,s\in\pi^T_{T,n};t_{\delta,n}\leqslant t\leqslant s}$ and $(Q^n_{t,s})_{t,s\in\pi^T_{T,n};t_{\delta,n}\leqslant t\leqslant s}$. Also assume that there exists some $(Q^{reg,n}_{t,s})_{t,s\in\pi^T_{T,n};t_{\delta,n}\leqslant t\leqslant s}$ which satisfies $R_{q,\eta}(\delta/2)$ (see (17)) and $(P^{reg,n}_{t,s})_{t,s\in\pi^T_{T,n};t_{\delta,n}\leqslant t\leqslant s}$ which satisfies $R^*_{q,\eta}(\delta/2)$ (see (18)) and such that (20) holds. Then, there exists $C\geqslant 1$ and $\beta\in\mathbb{N}$, such that for every bounded and measurable function f on $\mathbb{R}^d$, we have*

$$\forall x\in\mathbb{R}^d,\quad |P^n_Tf(x)-Q^n_Tf(x)|\leqslant C\frac{1+|x|^\beta}{\delta^{\eta(q)}}\|f\|_\infty/n^h.$$

B. *On the interval $[t_{\delta,n},T]$, assume that for every $m\in n\mathbb{N}^*=\{nk,k\in\mathbb{N}^*\}$, both $(\mu^m_k)_{k\in\{k_{\delta,n}m/n+1,\dots,m\}}$ and $(\nu^m_k)_{k\in\{k_{\delta,n}m/n+1,\dots,m\}}$ satisfy (5), and that $E_m(h,q)$ (see (9)) and $E^*_m(h,q)$ (see (16)) hold between $(P^m_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s}=(P_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s}$ and $(Q^m_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s}$. Also suppose that there exists a family $((Q^{reg,m}_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s})_{m\in n\mathbb{N}^*}$ which satisfies $R_{q,\eta}(\delta/2)$ (see (17)) and $R^*_{q,\eta}(\delta/2)$ (see (18)) (with n replaced by m for every $m\in n\mathbb{N}^*$) and such that (20) holds between $((Q^{reg,m}_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s})_{m\in n\mathbb{N}^*}$ and $((Q^m_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s})_{m\in n\mathbb{N}^*}$ (with n replaced by m for every $m\in n\mathbb{N}^*$).*

Then, there exists $C\geqslant 1$ and $\beta\in\mathbb{N}$, such that for every bounded and measurable function f on $\mathbb{R}^d$, we have

$$\forall x\in\mathbb{R}^d,\quad |P^n_Tf(x)-Q^n_Tf(x)|\leqslant C\frac{1+|x|^\beta}{\delta^{\eta(q)}}\|f\|_\infty/n^h.$$

Proof. We prove **A.** For sake of clarity, assume that $P^{\mathrm{reg}}=P$ and that $Q^{\mathrm{reg}}=Q$. The proof is very similar otherwise. Using the semigroup property, it follows that

$$\begin{aligned}|P^n_{0,T}f(x)-Q^n_{0,T}f(x)|=&|P^n_{0,t_{\delta,n}}P^n_{t_{\delta,n},T}f(x)-Q^n_{0,t_{\delta,n}}Q^n_{t_{\delta,n},T}f(x)|\\ =&|P^n_{0,t_{\delta,n}}P^n_{t_{\delta,n},T}f(x)-P^n_{0,t_{\delta,n}}Q^n_{t_{\delta,n},T}f(x)+P^n_{0,t_{\delta,n}}Q^n_{t_{\delta,n},T}f(x)-Q^n_{0,t_{\delta,n}}Q^n_{t_{\delta,n},n}f(x)|\\ =&|P^n_{0,t_{\delta,n}}(P^n_{t_{\delta,n},n}-Q^n_{t_{\delta,n},T})f(x)+(P^n_{0,t_{\delta,n}}-Q^n_{0,t_{\delta,n}})Q^n_{t_{\delta,n},T}f(x)|\end{aligned}$$

Since P^n and Q^n satisfy (5), (9), (16), and $R_{q,\eta}(\delta/2)$ (see (17)) holds for Q^n while $R^*_{q,\eta}(\delta/2)$ (see (18)) holds for P^n (with $t,s\in[t_{\delta,n},T]$), using Proposition 2.2, we derive: $|P^n_{0,t_{\delta,n}}(P^n_{t_{\delta,n},T}-P^n_{t_{\delta,n},T})f(x)|\leqslant\|(P^n_{t_{\delta,n},T}-Q^n_{t_{\delta,n},T})f\|_\infty\leqslant C\delta^{-\eta(q)}\|f\|_\infty/n^h$. In order to bound the second term of the *r.h.s* of the above equation, we use Proposition 2.1 together with $\|Q^n_{t_{\delta,n},T}f\|_{q,\infty}\leqslant C\delta^{-\eta(q)}\|f\|_\infty$ which follows from $R_{q,\eta}(\delta/2)$.

Now, we prove **B.** Using the same decomposition as for the proof of **A.** we observe that the only change in the proof concerns the study of the term $\|(Q^n_{t_{\delta,n},T} - P^n_{t_{\delta,n},T})f\|_\infty$. First, we notice that $t \in \pi_{T,m}$ for every $t \in \pi_{T,n}$ and every $m \in n\mathbb{N}^*$. We introduce the sequence of discrete semigroups $((Q^{n,m}_t)_{t\in\pi^T_{T,nm};t\geqslant t_{\delta,n}})_{m\in\mathbb{N}^*}$ defined in the following way: For every $t \in \pi^T_{T,nm}, t \geqslant t_{\delta,n}$, let $Q^{n,m}_t f(x) := Q^{nm}_t f(x)$. In the same way, let $((P^{n,m}_t)_{t\in\pi^T_{T,nm};t\geqslant t_{\delta,n}})_{m\in\mathbb{N}^*}$ the sequence of discrete semigroups such that for every $t \in \pi_{T,nm}, t \geqslant t_{\delta,n}$ we have $P^{n,m}_t f(x) := P^{nm}_t f(x) = P_t f(x)$. Let $m' \geqslant m$. Since $t^{nm}_{mk} = t^n_k$ for every $m \in \mathbb{N}^*$ and every $k \in \mathbb{N}$, it follows that

$$\begin{aligned}\|Q^{n,m}_{t^n_k,t^n_{k+1}} f - Q^{n,m'}_{t^n_k,t^n_{k+1}} f\|_\infty &= \|Q^{n,m}_{t^{nm}_{mk},t^{nm}_{m(k+1)}} f - Q^{n,m'}_{t^{nm'}_{m'k},t^{nm'}_{m'(k+1)}} f\|_\infty \\ &\leqslant \|Q^{nm}_{t^{nm}_{mk},t^{nm}_{m(k+1)}} f - P^{nm}_{t^{nm}_{mk},t^{nm}_{m(k+1)}} f\|_\infty + \|P^{nm'}_{t^{nm'}_{m'k},t^{nm'}_{m'(k+1)}} f - Q^{nm'}_{t^{nm'}_{m'k},t^{nm'}_{m'(k+1)}} f\|_\infty,\end{aligned}$$

Since Q^{nm} and $Q^{nm'}$ verify respectively $E_{nm}(h,q)$ and $E_{nm'}(h,q)$ and both Q^{nm} and $Q^{nm'}$ satisfy (6), we use the Lindeberg decomposition (15) in order to derive: $\|Q^{n,m}_{t^n_k,t^n_{k+1}} f - Q^{n,m'}_{t^n_k,t^n_{k+1}} f\|_\infty \leqslant C\|f\|_{q,\infty}/(n^{h+1}m^h)$. In the same way we obtain $|\langle g, Q^{n,m}_{t^n_k,t^n_{k+1}} f - Q^{n,m'}_{t^n_k,t^n_{k+1}} f\rangle| \leqslant C\|g\|_{1,q}\|f\|_\infty/(n^{h+1}m^h)$. Now, since both Q^{nm} and $Q^{nm'}$ have modifications, $Q^{\mathrm{reg},nm}$ and $Q^{\mathrm{reg},nm'}$, which satisfy both $R_{q,\eta}(\delta/2)$ (see (17)) and $R^*_{q,\eta}(\delta/2)$ (see (18)) (with n replaced by mn and by $m'n$), we can show that: $\|Q^{n,m}_{t_{\delta,n},T} f - Q^{n,m'}_{t_{\delta,n},T} f\|_\infty \leqslant C\delta^{-\eta(q)}\|f\|_\infty/(n^h m^h)$. The sequence $(Q^{n,m}_{t_{\delta,n},T})_{m\in\mathbb{N}^*}$ is thus Cauchy and it converges toward $P^n_{t_{\delta,n},T}$ for smooth test functions using Proposition 2.1. In particular, taking $m = 1$ and letting m' tend to infinity in the previous inequality we have

$$\|Q^{n,1}_{t_\delta,T} f - P^n_{t_\delta,T} f\|_\infty \leqslant C\delta^{-\eta(q)}\|f\|_\infty/n^h,$$

where the *l.h.s* of the above inequality is exactly the term that we study and then the proof is completed. □

This result is crucial to prove the weak convergence for bounded and measurable test functions for diffusion processes with simply locally smooth coefficients. The idea consists in introducing a regularized version of the underlying process in the neighborhood of T with smooth coefficients in $\mathcal{C}^q_b(\mathbb{R}^d)$ and then to consider scheme of this regularized process. This scheme (or its modification) has mixing regularization properties and we can use Theorem 2.1. It then remains to control the error committed between the real process and its modification.

Notice also that from point **B.** we can focus exclusively on proving the regularization properties for the modification of the approximation semigroup Q^n. In particular there is no regularization property to prove on P^n (or its modification) which is very useful. This is the method we employ in the following. Based on Theorem 2.1, we now present our method to prove total variation convergence of numerical schemes of diffusion processes with singularities.

3 Total variation convergence for locally smooth diffusions

A main application of Theorem 2.1 concerns the use of Markov chains with form (2) for the approximation, in total variation distance, of d-dimensional diffusion processes with form

$$dX_t = V_0(X_t)dt + \sum_{i=1}^{N} V_i(X_t) \circ dW_t^i, \tag{22}$$

with $(W_t)_{t\geqslant 0}$ a standard Brownian motion and $\circ dW_t^i$ the Stratonovich integral with respect to W_t^i. Moreover, we assume that the coefficients $V_i \in \mathcal{C}_b^\infty(\mathcal{D};\mathbb{R}^d)$ for a subset $\mathcal{D}$ of $\mathbb{R}^d$. The innovation of our approach relies on the regularity of the coefficients V_i. Existing results concerning total variation convergence require that $V_i \in \mathcal{C}_b^q(\mathbb{R}^d;\mathbb{R}^d)$ for some q large enough whereas in our approach we simply assume that $V_i \in \mathcal{C}_b^q(\mathcal{D};\mathbb{R}^d)$. In the same way we will denote $V_{0,\mathrm{Ito}}$ such that

$$dX_t = V_{0,\mathrm{Ito}}(X_t)dt + \sum_{i=1}^{N} V_i(X_t)dW_t^i. \tag{23}$$

The infinitesimal generator of this Markov process reads

$$A = V_0 + \frac{1}{2}\sum_{i=1}^{N} V_i^2,$$

with the notation $Vf(x) = \langle V(x), \nabla f(x)\rangle$. In the latter, when it is relevant, we denote by $X_t(x)$ the process starting from x. Moreover we denote by $(P_t)_{t\geqslant 0}$ (see [31]) the semigroup of this process.

We now introduce a discrete time approximation process for $(X_t)_{t\geqslant 0}$. Let $n \in \mathbb{N}^*$ be the number of time steps between 0 and T equipped with the time grid $\pi_{T,n} = \{t_k^n = kT/n, k \in \mathbb{N}\}$. We consider a sequence of independent random variables $Z_k^n = (Z_k^{n,1}, \ldots, Z_k^{n,N}) \in \mathbb{R}^N$, $k \in \{1, \ldots, n\}$ and we denote $Z^n = (Z_1^n, ..., Z_n^n)$. We assume that Z^n is centered and that there exists two deterministic sequences $(b_k^l(Z^n))_{k\in\mathbb{N}^*} \leqslant (b_k^u(Z^n))_{k\in\mathbb{N}^*}$ such that for every $k \in \{k_{\delta,n}, \ldots, n\}$, $\mathbb{P}(Z_k^n \in [b_k^l(Z^n), b_k^u(Z^n)]) = 1$, with $k_{\delta,n}$ defined in (21). We also assume that n and Z satisfy

$$n \geqslant \sup_{k\in\{k_{\delta,n},\ldots,n\}} T^2/(b_k^u(Z^n) - b_k^l(Z^n))^2. \tag{24}$$

Finally, we consider the sequence of independent random variables $\kappa_k \in \mathbb{R}$, $k \in \mathbb{N}^*$, (and independent from Z^n) and we define the $\mathbb{R}^d$ valued Markov chain

$$X_{t_{k+1}}^n = \psi\big(\kappa_{k+1}, X_{t_k}^n, \frac{Z_{k+1}^n}{\sqrt{n}}, t_{k+1}^n - t_k^n\big), \quad k \in \mathbb{N}, \quad \text{with} \quad \forall(\kappa, x) \in \mathbb{R}\times\mathbb{R}^d, \quad \psi(\kappa, x, 0, 0) = x. \tag{25}$$

In this section we provide an approach to show the convergence (as n tends to infinity), for the total variation distance, of $(X_T^n)_{n\in\mathbb{N}^*}$ to X_T defined in (22).

Remark 3.1. *The reason to consider the random variables κ_k is the following. In the Victoir Ninomiya scheme, at each time step k, one throws a coin $\kappa_k \in \{1, -1\}$ and employs different form of the function ψ according to the fact that κ_k is equal to 1 or to -1.*

Now let $Q^n_{s,t}f(x) = \mathbb{E}[f(X^n_t)|X^n_s = x]$, $s,t \in \pi_{T,n}$, $s \leqslant t$. This semigroup is not supposed to have mixing regularization properties and to satisfy every required approximation assumptions with $(P^n_t)_{t\in\pi_{T,n}}$, (with notation $(P^n_t)_{t\in\pi_{T,n}} = (P_t)_{t\in\pi_{T,n}}$) to use Theorem 2.1. Let $\delta \in [2T/n, T]$ and consider $t_{\delta,n}$ defined in (21). In our approach, we simply suppose that Q^n and P^n satisfy the assumptions from Theorem 2.1 on the time interval $[0, t_{\delta,n}]$. The next step thus consists in introducing regularized modifications for Q^n and P^n which satisfy the hypotheses of Theorem 2.1 on $[0, t_{\delta,n}]$ but also on the interval $(t_{\delta,n}, T]$.

First we introduce a regularized version of $(X_t)_{t\geqslant 0}$ with smooth coefficients such that we can simulate an approximation Markov chain with suitable regularization and approximation properties. Let $v \geqslant 0$ and $\mathcal{D} \subset \mathbb{R}^d$. We begin by defining $\mathcal{D}^v := \{x \in \mathcal{D}, \inf_{y\in\mathbb{R}^d\setminus\mathcal{D}} |x-y| \geqslant v\}$ with the convention $\inf_{y\in\emptyset} |x-y| = +\infty$. Now we introduce $\phi^{\mathcal{D}}_v \in \mathcal{C}^\infty(\mathbb{R}^d;\mathcal{D})$ such that $\phi^{\mathcal{D}}_v(x) = x$ if $x \in \mathcal{D}^v$ and has null derivatives when $x \in \mathbb{R}^d \setminus \mathcal{D}$. Moreover we assume that for every multi-index α with $|\alpha| \geqslant 1$, $\partial_\alpha \phi^{\mathcal{D}}_v \in \mathcal{C}^\infty_b(\mathbb{R}^d \setminus \mathcal{D}^v;\mathcal{D})$ and that for every $x \in \mathbb{R}^d \setminus \mathcal{D}^v$, we have

$$|\partial_\alpha \phi^{\mathcal{D}}_v(x)| \leqslant \frac{C}{v^{|\alpha|}},$$

where C does not depend on v. Now, we can introduce a regularized version of $(X_t)_{t\geqslant 0}$ (see (22)). For $i = 0,\dots,N$, we denote $V_i^{\mathcal{D},v} = V_i \circ \phi^{\mathcal{D}}_v \in \mathcal{C}^\infty_b(\mathbb{R}^d;\mathbb{R}^d)$ and consider

$$dX^{\mathcal{D},v}_t = V_0^{\mathcal{D},v}(X^{\mathcal{D},v}_t)dt + \sum_{i=1}^N V_i^{\mathcal{D},v}(X^{\mathcal{D},v}_t) \circ dW^i_t. \tag{26}$$

Using this regularized process with smooth coefficients, we build the following locally regularized version of $(X_t)_{t\geqslant 0}$ (see (22)):

$$\overline{X}^{\mathcal{D},v}_t(x) = \begin{cases} X_t(x), & \text{if } t \leqslant t_{\delta,n}, \\ X^{\mathcal{D},v}_{t-t_{\delta,n}}(X_{t_{\delta,n}}(x)), & \text{if } \quad t \geqslant t_{\delta,n}, \end{cases} \tag{27}$$

with $t_{\delta,n}$ defined in (21). We denote by $(\overline{P}^{\mathcal{D},v}_t)_{t\geqslant 0}$, the semigroup of this process such that $\overline{P}^{\mathcal{D},v}_{s,t} f(x) = \mathbb{E}[f(X^n_t)|X^n_s = x]$, $s \leqslant t$ and also denote $(\overline{P}^{\mathcal{D},v,n}_t)_{t\in\pi_{T,n}} = (\overline{P}^{\mathcal{D},v}_t)_{t\in\pi_{T,n}}$. At this point, it simply remains to introduce a Markov chain approximation for $(\overline{X}^{\mathcal{D},v}_t(x))_{t\geqslant 0}$ which is well suited to the use of Theorem 2.1. Since our approach requires regularization properties, we introduce a modification $(\tilde{Z}^n_k)_{k\in\mathbb{N}^*}$ of $(Z^n_k)_{k\in\mathbb{N}^*}$ as a sequence of centered and independent random variables with $\tilde{Z}^n_k = (\tilde{Z}^{n,1}_k,\dots,\tilde{Z}^{n,N}_k) \in \mathbb{R}^N$ such that for every $k \in \{k_{\delta,n},\dots,n\}$, $\mathbb{P}(\tilde{Z}^n_k \in [b^l_k(\tilde{Z}^n), b^u_k(\tilde{Z}^n)]) = 1$. This sequence varies from $(Z^n_k)_{k\in\mathbb{N}^*}$ in the following way. Let $\varepsilon_*, r_* > 0$. We assume that for every $m \in n\mathbb{N}^*$, the sequence $(\tilde{Z}^m_k)_{k\in\mathbb{N}^*}$ satisfies the following Doeblin property: There exists $(z^m_{*,k})_{k\in\mathbb{N}^*} \in (\mathbb{R}^N)^{\otimes\mathbb{N}^*}$ such that for every Borel set $A \subset \mathbb{R}^N$ and every $k \in \mathbb{N}^*$

$$L_{z^m_*}(\varepsilon_*, r_*) \quad \equiv \quad \mathbb{P}(\tilde{Z}^m_k \in A) \geqslant \varepsilon_* \lambda(A \cap B_{r_*}(z^m_{*,k})), \tag{28}$$

and we denote $L_{z_*}(\varepsilon_*, r_*) \equiv \cap_{m\in n\mathbb{N}} L_{z^m_*}(\varepsilon_*, r_*)$. In the sequel, we will denote

$$m_* = \varepsilon_* \int_{\mathbb{R}^N} \Big(\mathbb{1}_{|z|\leqslant r_*/2} + \exp\Big(1 - \frac{r_*^2}{r_*^2 - (2|z| - r_*)^2}\Big) \mathbb{1}_{r_*/2<|z|<r_*} \Big) dz \tag{29}$$

Moreover, we assume that the following moment finiteness assumption holds:

$$\forall p \geqslant 1, \qquad M_p(\tilde{Z}) := 1 \vee \sup_{n\in\mathbb{N}^*} \sup_{k\leqslant n} \mathbb{E}[|\tilde{Z}_k^n|^p] < +\infty. \tag{30}$$

We construct the $\mathbb{R}^d$ valued Markov chain approximation of $(X_t^{\mathcal{D},v})_{t\geqslant 0}$ (see (26)) as follows

$$\tilde{X}^n_{t^n_{k+1}} = \tilde{\psi}\big(\kappa_{k+1}, \tilde{X}^n_{t^n_k}, \frac{\tilde{Z}^n_{k+1}}{\sqrt{n}}, t^n_{k+1} - t^n_k\big), \qquad k \in \mathbb{N}, \tag{31}$$

where

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

The semigroup of this discrete process reads: $\tilde{Q}^n_{s,t}f(x) = \mathbb{E}[f(\tilde{X}^n_t)|\tilde{X}^n_s = x]$, $s,t\in\pi_{T,n}$, $s\leqslant t$. Using this regular approximation we construct the locally regularized version of $(X^n_t)_{t\in\pi_{T,n}}$ (see (25)) which we will use to approximate $(\overline{X}^{\mathcal{D},v}_t)_{t\geqslant 0}$ (see (27)) for the total variation distance with Theorem 2.1. Let

$$\overline{X}^n_{t^n_{k+1}}(x) = \begin{cases} \psi(\kappa_k, \overline{X}^n_{t^n_k}(x), Z^n_{k+1}/\sqrt{n}, t^n_{k+1} - t^n_k), & \text{if} \quad t^n_k \leqslant t_{\delta,n}, \\ \tilde{\psi}(\kappa_k, \overline{X}^n_{t^n_k}(x), \tilde{Z}^n_{k+1}/\sqrt{n}, t^n_{k+1} - t^n_k) & \text{if} \quad t_{\delta,n} \leqslant t^n_k \leqslant T, \end{cases} \tag{32}$$

We define its semigroup: $\overline{Q}^n_{s,t}f(x) = \mathbb{E}[f(\overline{X}^n_t)|\overline{X}^n_s = x]$,$s,t\in\pi_{T,n}$, $s\leqslant t$. Notice that for every $n\in\mathbb{N}^*$ we have $(\overline{Q}^n_{s,t})_{s,t\in\pi^{t_{\delta,n}}_{T,n};s\leqslant t} = (Q^n_{s,t})_{s,t\in\pi^{t_{\delta,n}}_{T,n};s\leqslant t}$ and for every $m\in n\mathbb{N}^*$, we have $(\overline{Q}^m_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s} = (\tilde{Q}^m_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s}$ and similar identifications holds when Q is replaced by P. Notice also that from this definition, the Doeblin assumption (28) is only necessary for every $m\in n\mathbb{N}^*$ and every $k\in\{k_{\delta,n}m/n,\dots,m\}$. At this point, we introduce the following quantities: Let $r\in\mathbb{N}^*$ and let $\tilde{\psi}\in\mathcal{C}^r(\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^N\times\mathbb{R}_+;\mathbb{R}^d)$ and define

$$\|\tilde{\psi}\|_{1,r,\infty} = 1 \vee \sum_{|\alpha|=0}^{r} \sum_{|\beta|+|\gamma|=1}^{r-|\alpha|} \|\partial_x^\alpha\partial_z^\beta\partial_t^\gamma\tilde{\psi}\|_\infty.$$

For $r\in\mathbb{N}^*$, we also denote

$$\mathfrak{K}_r(\tilde{\psi}) = (1+\|\tilde{\psi}\|_{1,r,\infty})\exp(\|\tilde{\psi}\|^2_{1,3,\infty}). \tag{33}$$

Before we establish our main result concerning numerical schemes, some assumptions are still to be introduced. Assume that

$$\frac{3\|\tilde{\psi}\|_{1,3,\infty}}{n^{1/4}} + \frac{M_8(\tilde{Z})}{n} + \exp(-m_*^2 n\delta/(4T)) \leqslant \frac{1}{2}. \tag{34}$$

Moreover we assume that there exists $\lambda_* > 0$ such that

$$\inf_{\kappa\in\mathbb{R}}\inf_{x\in\mathbb{R}^d}\inf_{|\xi|=1} \sum_{i=1}^N \Big\langle \partial_{z_i}\tilde{\psi}(\kappa,x,0,0),\xi\Big\rangle^2 \geqslant \lambda_*. \tag{35}$$

and

$$n^{1/2} \geqslant 3\frac{2^{N+2}}{\lambda_*}\|\tilde{\psi}\|^2_{1,3,\infty}. \tag{36}$$

We now introduce our main hypothesis concerning the fact that $(\overline{X}^n_t)_{t\in\pi_{T,n}}$ (see (32)) is a well adapted modification for $(X^n_t)_{t\in\pi_{T,n}}$ (see (31)):

$$(\mathbf{SG_{Reg}}(\overline{P}^{\mathcal{D},v,n}, \overline{Q}^n, h, q, n, \delta, T, \mathcal{D}, v))$$
$$\equiv$$

Assume that $L_{z_*}(\varepsilon_*, r_*)$ (see (28)), (30), (34), (35) and (36) hold and that:

I) Localized identity of the regularized approximation. Let us denote $\Omega_{X^n} = \{\sup_{t\in\pi^{\tilde{\delta}_n}_{T,n}} |X^n_{t_{\delta,n}+t} - X^n_T| < v\}$ and $\Omega_{\overline{X}^n} = \{\sup_{t\in\pi^{\tilde{\delta}_n}_{T,n}} |\overline{X}^n_{t_{\delta,n}+t} - \overline{X}^n_T| < v\}$ and assume that:

$$(X^n_{t_{\delta,n}+t}\mathbb{1}_{\Omega_{X^n}})_{t\in\pi^{\tilde{\delta}_n}_{T,n}} \overset{Law}{=} (\overline{X}^n_{t_{\delta,n}+t}\mathbb{1}_{\Omega_{\overline{X}^n}})_{t\in\pi^{\tilde{\delta}_n}_{T,n}},$$

on the event $\{X^n_{t_{\delta,n}} = \overline{X}^n_{t_{\delta,n}}\} \cap \{X^n_T, \overline{X}^n_T \in \mathcal{D}^{2v}\}$, with $\tilde{\delta}_n$ defined in (21).

II) Approximation. On the interval $[0, t_{\delta_n}]$, assume that $(\overline{P}^{\mathcal{D},v,n}_{s,t})_{s,t\in\pi^{t_{\delta,n}}_{T,n};s\leqslant t}$ and $(\overline{Q}^n_{s,t})_{s,t\in\pi^{t_{\delta,n}}_{T,n};s\leqslant t}$ satisfy respectively $\mathfrak{P}_0$ (see (7)) and $\mathfrak{P}'_q$ (see (8)) and $E'_{n,\mathrm{pol}}(h,q)$ (see (11)) holds between them.

On the interval $[t_{\delta,n}, T]$, assume that for every $m \in n\mathbb{N}^*$, the assumptions $E_m(h,q)$ (see (9)) and $E^*_m(h,q)$ (see (16)) hold between $(\overline{P}^{\mathcal{D},v,m}_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s}$ and $(\overline{Q}^m_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi^{\tilde{\delta}_n}_{T,m};t\leqslant s}$.

We now establish the total variation convergence result that we derive from this approach and from Theorem 2.1.

Theorem 3.1. *Let $T > 0$ and $n \in \mathbb{N}^*$. Let $q \in \mathbb{N}$, $h > 0$ and $\delta \in [2T/n, T)$. Assume that (24) and $(\mathbf{SG_{Reg}}(\overline{P}^{\mathcal{D},v,n}, \overline{Q}^n, h, q, n, \delta, T, \mathcal{D}, v))$ hold. Then, there exists $l \in \mathbb{N}^*$, $C, \beta \geqslant 1$ such that we have for every bounded and measurable function $f : \mathbb{R}^d \to \mathbb{R}$, with $\mathrm{supp}(f) \subset \mathcal{D}^{2v}$,*

$$|\mathbb{E}[f(X_T(x)) - f(X^n_T(x))]| \leqslant 8\Big(\exp\Big(-\frac{v^2/2 - \delta^2\|V_{0,\mathrm{Ito}}\mathbb{1}_{\mathcal{D}^v}\|^2_\infty}{\delta\sum_{i=1}^N \|V_i\mathbb{1}_{\mathcal{D}^v}\|^2_\infty}\Big) + \exp\Big(-\frac{C_{Z,n}v^2}{\delta\|\psi\mathbb{1}_{\mathcal{D}^v}\|^2_{1,1,\infty}}\Big)\Big)\|f\|_\infty$$
$$+ C(1+|x|^\beta)\frac{\mathfrak{K}_{q+3}(\tilde{\psi})^l}{(\lambda_*\delta)^{\eta(q)}}\|f\|_\infty/n^h \tag{37}$$

with $\eta(q) = q(q+1)$, $C_{Z,n} = \inf_{k\in\{k_{\delta,n},\ldots,n\}} T/(12(b^u_k(Z^n) - b^l_k(Z^n))^2)$ and $\mathfrak{K}_r(\psi)$, $r \in \mathbb{N}^*$, defined in (33).

Remark 3.2. *Notice that from Lemma 3.1, then (37) remains true if we replace V by $V^{\mathcal{D},v}$ or if we replace $C_{Z,n}/\|\psi\mathbb{1}_{\mathcal{D}^v}\|^2_{1,1,\infty}$ by $C_{Z,n}/\|\psi\|^2_{1,1,\infty}$ or by $C_{\tilde{Z},n}/\|\tilde{\psi}\mathbb{1}_{\mathcal{D}^v}\|^2_{1,1,\infty}$, $C_{\tilde{Z},n}/\|\tilde{\psi}\|^2_{1,1,\infty}$ as soon as (24) holds with Z^n replaced by $\tilde{Z}^n$ (when those quantities are finite).*

Remark 3.3. *We can obtain the same result as (37) for bounded and measurable test functions f with* $\mathrm{supp}(f) \subset \mathbb{R}^d$. *In order to do it, we simply have to assume that $\mathbb{P}(X_T \notin \mathcal{D}^{2v})$ and $\mathbb{P}(X_T^n \notin \mathcal{D}^{2v})$ are small enough. In this case we can rigorously speak of total variation distance estimation. However, there is no particular interest in doing it if we cannot explicit $\mathbb{P}(X_T \notin \mathcal{D}^{2v})$ or $\mathbb{P}(X_T^n \notin \mathcal{D}^{2v})$. That is why, we simply provide (37) and consider this result as a total variation distance estimation between the process and its scheme.*

In this section we present the tools from which we derive Theorem 3.1. We first present a generic class of Markov chains which have regularization properties and can be used to simulate $(\tilde{X}_t^n)_{t\in\pi_{T,n}^T;t\geqslant t_{\delta,n}}$. This provides appropriate mixing regularization properties for $(\overline{X}_t^n)_{t\in\pi_{T,n}^T}$ necessary to use Theorem 2.1 point **B.** with P^n replaced by $\overline{P}^{\mathcal{D},v,n}$ and Q^n replaced by $\overline{Q}^n$. Finally we give concentration inequalities for Markov diffusion processes and for Markov chains to control the total variation distance between P^n and $\overline{P}^{\mathcal{D},v,n}$ and between Q^n and $\overline{Q}^n$. We conclude by establishing Theorem 3.1.

3.1 Regularization semigroups

In this section, we present a class of approximation Markov chains that can be used to build a semigroup and its modification with adapted regularization properties. In particular this class is large enough to contain classical numerical schemes (Euler, Ninomiya-Victoir...) for diffusion processes with smooth coefficients such as $(X_t^{\mathcal{D},v}(x))_{t\geqslant 0}$. In this section we fix $n \in \mathbb{N}^*$. We are going to build $(\tilde{Q}_{t,s}^n)_{t,s\in\pi_{T,n};t\leqslant s}$ and prove that there exists a family of kernels $(\tilde{Q}_{t,s}^{\mathrm{reg},n})_{t,s\in\pi_{T,n};t\leqslant s}$ which satisfies $R_{q,\eta}$ (see (17)) and $R_{q,\eta}^*$ (see (18)) and such that (20) holds between $\tilde{Q}^n$ and $\tilde{Q}^{\mathrm{reg},n}$.

3.1.1 A class of random tools

First, we present the random tools that appear in the construction of our class of Markov chains and which provide regularization properties through Doeblin properties. Let $n \in \mathbb{N}^*$ and $N \in \mathbb{N}^*$ be fixed. We consider a sequence of independent random variables $\tilde{Z}_k^n = (\tilde{Z}_k^{n,1}, \ldots, \tilde{Z}_k^{n,N}) \in \mathbb{R}^N$, $k \in \{1, \ldots, n\}$ and we denote $\tilde{Z}^n = (\tilde{Z}_1^n, ..., \tilde{Z}_n^n)$.
Our aim is to settle an integration by parts formula based on the law of $\tilde{Z}^n$. Let $\varepsilon_*, r_* > 0$. The basic assumption is the Doeblin condition $L_{z_*^n}(\varepsilon_*, r_*)$ (see (28) with this fixed n) that reads: There exists $z_{*,k}^n \in \mathbb{R}^N$, $k \in \{1, \ldots, n\}$, such that for every Borel set $A \subset \mathbb{R}^N$ and every $k \in \{1, \ldots, n\}$: $\mathbb{P}(\tilde{Z}_k^n \in A) \geqslant \varepsilon_* \lambda(A \cap B_{r_*}(z_{*,k}^n))$. We also say that the random variables $\tilde{Z}_k^n, k \in \{1, \ldots, n\}$ are lower bounded by the Lebesgue measure. We also suppose that (30) holds. It is easy to check that $L_{z_*^n}(\varepsilon_*, r_*)$ holds if and only if there exists some non negative measures π_k^n with total mass $\pi_k^n(\mathbb{R}^N) < 1$ and a lower semi-continuous function $\varphi^n \geqslant 0$ such that $\mathbb{P}(\tilde{Z}_k^n \in dz) = \pi_k^n(dz) + \varphi^n(z - z_{*,k}^n)dz$. Notice that the random variables $(\tilde{Z}_1^n, \ldots, \tilde{Z}_n^n)$ are not assumed to be identically distributed. However, the fact that $r_* > 0$ and $\varepsilon_* > 0$ are the same for every k (and also for every n) represents a mild substitute of this property. From $L_{z_*^n}(\varepsilon_*, r_*)$, we introduce a variant of φ^n (which does not depend on n) for the representation of the law of Z^n. Let $\varphi_{r_*/2} : \mathbb{R}^N \to \mathbb{R}$ be the function defined by

$$\varphi_{r_*/2}(z) = \mathbb{1}_{|z|\leqslant r_*/2} + \exp\Big(1 - \frac{r_*^2}{r_*^2 - (2|z| - r_*)^2}\Big)\mathbb{1}_{r_*/2<|z|<r_*}. \tag{38}$$

Then $\varphi_{r_*/2} \in \mathcal{C}_b^\infty(\mathbb{R}^N)$, $0 \leqslant \varphi_{r_*/2} \leqslant 1$ and we have the following crucial property: For every $p, q \in \mathbb{N}$ there exists a universal constant $C_{q,p}$ such that for every $z \in \mathbb{R}^N$, and every $i_1, \ldots, i_q \in \{1, \ldots, N\}$,

we have

$$\varphi_{r_*/2}(z)\Big|\frac{\partial^q}{\partial z^{i_1}...\partial z^{i_q}}(\ln\varphi_{r_*/2})(z)\Big|^p \leqslant \frac{C_{q,p}}{r_*^{pq}},$$

with the convention $\ln\varphi_{r_*/2}(z)=0$ for $|z|\geqslant r_*$. As an immediate consequence of $L_{z_*^n}(\varepsilon_*,r_*)$ (see (28)), for every non negative function $f:\mathbb{R}^N\to\mathbb{R}_+$, we have

$$\mathbb{E}[f(\tilde{Z}_k^n)]\geqslant \varepsilon_*\int_{\mathbb{R}^N}\varphi_{r_*/2}(\ z-z_{*,k}^n\ )f(z)dz.$$

By a change of variables, it follows that

$$\mathbb{E}\Big[f\Big(\frac{\tilde{Z}_k^n}{\sqrt{n}}\Big)\Big]\geqslant \varepsilon_*\int_{\mathbb{R}^N} n^{N/2}\varphi_{r_*/2}\big(\sqrt{n}(z-\frac{z_{*,k}}{\sqrt{n}})\big)f(z)dz=\varepsilon_*\int_{\mathbb{R}^N}\phi_n(z)f(z)dz, \tag{39}$$

with $\phi_n(z)=n^{N/2}\varphi_{r_*/2}(\sqrt{n}z)$. We notice that $\int\phi_n(z)dz=m_*\varepsilon_*^{-1}$, with m_* defined in (29).

We consider a sequence of independent random variables $\chi_k\in\{0,1\}$, $U_k,V_k\in\mathbb{R}^N$, $k\in\{1,\dots,n\}$ with laws given by

$$\begin{aligned}
&\mathbb{P}(\chi_k=1)=m_*,\qquad \mathbb{P}(\chi_k=0)=1-m_*,\\
&\mathbb{P}(U_k\in dz)=\frac{\varepsilon_*}{m_*}\phi_n\Big(z-\frac{z_{*,k}^n}{\sqrt{n}}\Big)dz,\\
&\mathbb{P}(V_k\in dz)=\frac{1}{1-m_*}\Big(\mathbb{P}\Big(\frac{\tilde{Z}_k^n}{\sqrt{n}}\in dz\Big)-\varepsilon_*\phi_n\Big(z-\frac{z_{*,k}^n}{\sqrt{n}}\Big)dz\Big).
\end{aligned}$$

Notice that (39) guarantees that $\mathbb{P}(V_k\in dz)\geqslant 0$. Then a direct computation shows that

$$\mathbb{P}(\chi_kU_k+(1-\chi_k)V_k\in dz)=\mathbb{P}\Big(\frac{\tilde{Z}_k^n}{\sqrt{n}}\in dz\Big).$$

This is the splitting procedure for $\frac{\tilde{Z}_k^n}{\sqrt{n}}$ and the regularization properties follow from the following representation,

$$\frac{\tilde{Z}_k^n}{\sqrt{n}}=\chi_kU_k+(1-\chi_k)V_k,$$

through the random variables U_k, $k\in\mathbb{N}^*$.

Remark 3.4. *The above splitting procedure has already been widely used in the litterature and is called the Nummelin splitting: In [30] and [25], it is used to prove convergence to equilibrium of Markov processes. In [8], [9] and [36], it is used to study the Central Limit Theorem. Besides, in [29], the above splitting method (with $\mathbb{1}_{B_{r_*}(z_{*,k}^n)}$ instead of $\phi_n(z-\frac{z_{*,k}^n}{\sqrt{n}})$) is used in a framework which is similar to the one in this paper. Last in [5], this exact framework is used to prove total variation convergence under a regular setting.*

3.1.2 A class of Markov chains

We now present the class of discrete time random processes that is used to build the semigroup $(\tilde{Q}_{t,s}^n)_{t,s\in\pi_{T,n};t\leqslant s}$ and its modification $(\tilde{Q}_{t,s}^{\mathrm{reg},n})_{t,s\in\pi_{T,n};t\leqslant s}$ which satisfies regularization properties. We

recall that $n \in \mathbb{N}^*$ is the number of time steps between 0 and T equipped with the time grid $\pi_{T,n}$. We consider two sequences of independent random variables $\tilde{Z}_k^n \in \mathbb{R}^N, \kappa_k \in \mathbb{R}$, $k \in \mathbb{N}^*$ and we assume that $\tilde{Z}_k^n$ are centered and verifies $L_{z_*}(\varepsilon_*, r_*)$ (see (28)) and (34). As in (31), we construct the $\mathbb{R}^d$ valued Markov chain

$$\tilde{X}_{t_{k+1}}^n = \psi\big(\kappa_{k+1}, X_{t_k}^n, \frac{\tilde{Z}_{k+1}^n}{\sqrt{n}}, t_{k+1}^n - t_k^n\big), \qquad k \in \mathbb{N},$$

where

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

3.1.3 The regularization property

In the following, we will not work under $\mathbb{P}$, but under a localized probability measure defined as follows. For $t, s \in \pi_{T,n}$, $t < s$, we consider the set

$$\Lambda_{t,s} = \Big\{ \frac{1}{(s-t)n/T} \sum_{t \leqslant t_k^n \leqslant s} \chi_k \geqslant \frac{m_*}{2} \Big\}.$$

Using the Hoeffding inequality and the fact that $\mathbb{E}[\chi_k] = m_*$, it can be checked that

$$\mathbb{P}(\Lambda_{t,s}^c) \leqslant \exp(-m_*^2(s-t)n/(2T)).$$

We consider also the localization function $\varphi_{n^{1/4}/2}$, defined as in (38) with r_* replaced by $n^{1/4}$, and we construct the random variable

$$\Theta_{t,s,n} = \mathbb{1}_{\Lambda_{t,s}} \times \prod_{k=1}^n \varphi_{n^{1/4}/2}(\tilde{Z}_k^n).$$

Since Z_k^n has finite moments of any order, one can show the following inequality: For every $l \in \mathbb{N}$, we have

$$\mathbb{P}(\Theta_{t,s,n} = 0) \leqslant \mathbb{P}(\Lambda_{t,s}^c) + \sum_{k=1}^n \mathbb{P}(|\tilde{Z}_k^n| \geqslant n^{1/4}) \leqslant \exp(-m_*^2(s-t)n/(2T)) + \frac{M_{4(l+1)}(\tilde{Z})}{n^l},$$

with the notation $M_{4(l+1)}(\tilde{Z})$ introduced in (30). We define the probability measure

$$d\mathbb{P}_{\Theta_{t,s,n}} = \frac{1}{\mathbb{E}[\Theta_{t,s,n}]} \Theta_{t,s,n} d\mathbb{P}.$$

We consider the Markov chain $(\tilde{X}_t^n)_{t \in \pi_{T,n}}$, defined in (31) and we introduce $(\tilde{Q}_{t,s}^{\Theta,n})_{t,s \in \pi_{T,n}; t \leqslant s}$ (that replace notation $\tilde{Q}^{\mathrm{reg},n}$) such that,

$$\forall t, s \in \pi_T^n; t \leqslant s, \quad \tilde{Q}_{t,s}^{\Theta,n} f(x) := \mathbb{E}_{\Theta_{t,s,n}}[f(\tilde{X}_s^n) | \tilde{X}_t^n = x] = \frac{1}{\mathbb{E}[\Theta_{t,s,n}]} \mathbb{E}[\Theta_{t,s,n} f(\tilde{X}_s^n) | \tilde{X}_t^n = x].$$

Notice that $(\tilde{Q}_{t,s}^{\Theta,n})_{t,s \in \pi_{T,n}; t \leqslant s}$ is not a semigroup, but this is not necessary. We are not able to prove the regularization property for $\tilde{Q}^n$ but for its modification $\tilde{Q}^{\Theta,n}$ that we now establish. Notice that considering the hypotheses of Theorem 2.1, this is sufficient to obtain total variation convergence.

Proposition 3.1. ***A.*** *Let $T > 0$ and $n \in \mathbb{N}^*$. We assume that $n \in \mathbb{N}^*$ and $t, s \in \pi_{T,n}^T$, $t < s$, satisfy:*

$$\frac{3\|\tilde{\psi}\|_{1,3,\infty}}{n^{1/4}} + \frac{M_8(\tilde{Z}^n)}{n} + \exp(-m_*^2 n(s-t)/(2T)) \leqslant \frac{1}{2}$$

and

$$n^{1/2} \geqslant 3\frac{2^{N+2}}{\lambda_*}\|\tilde{\psi}\|_{1,3,\infty}^2.$$

Moreover we assume that

$$\inf_{\kappa\in\mathbb{R}} \inf_{x\in\mathbb{R}^d} \inf_{|\xi|=1} \sum_{i=1}^N \Big\langle \partial_{z_i}\tilde{\psi}(\kappa, x, 0, 0), \xi \Big\rangle^2 \geqslant \lambda_*.$$

Then for every $q \in \mathbb{N}$ and multi index α, β with $|\alpha| + |\beta| \leqslant q$, there exists $l \in \mathbb{N}^*$ and $C \geqslant 1$ which depend on m_*, r_* and on the moments of $\tilde{Z}^n$ such that, for every bounded and measurable function $f \in \mathcal{C}^q(\mathbb{R}^d)$, we have

$$\|\partial_\alpha \tilde{Q}_{t,s}^{n,\Theta} \partial_\beta f\|_\infty \leqslant C\frac{\mathfrak{K}_{q+3}(\tilde{\psi})^l}{(\lambda_*(s-t))^{q(q+1)}}\|f\|_\infty \tag{40}$$

with $\mathfrak{K}_r(\tilde{\psi})$, $r \in \mathbb{N}^*$, defined in (33). In particular, $\tilde{Q}_t^{n,\Theta}(x, dy) = p_t^{n,\Theta}(x, y)dy$ and $(x, y) \mapsto p_t^{n,\Theta}(x, y)$ belongs to $\mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{R}^d)$.

B. *For every $l \in \mathbb{N}$ and every $t, s \in \pi_{T,n}^T; t \leqslant s$, we have, for every bounded and measurable function f,*

$$\|\tilde{Q}_{t,s}^n f - \tilde{Q}_{t,s}^{n,\Theta} f\|_\infty \leqslant 4\Big(\exp(-m_*^2 n(s-t)/(2T)) + \frac{M_{4(l+1)}(\tilde{Z})}{n^l}\Big)\|f\|_\infty.$$

Remark 3.5. *The estimation (40) means that the strong regularization property $\overline{R}_{q,\eta}$ (see 19), with $\eta(q) = q(q+1)$, holds for $(\tilde{Q}_{t,s}^{\Theta,n})_{t,s\in\pi_{T,n};t\leqslant s}$. In particular $R_{q,\eta}(\delta/2)$ (see (17)) and $R_{q,\eta}^*(\delta/2)$ (see (18)) hold for $(\tilde{Q}^{\Theta,n}_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi_{T,n}^{\tilde{\delta}_n};t\leqslant s}$ which is a modification of $(\overline{Q}^n_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi_{T,n}^{\tilde{\delta}_n};t\leqslant s}$. (see (32) for definition). Moreover, using this result we deduce from $L_{z_*}(\varepsilon_*, r_*)$ (see (28)), (30), (34), (35) and (36 that for every $m \in n\mathbb{N}^*$ then $R_{q,\eta}(\delta/2)$ and $R_{q,\eta}^*(\delta/2)$ hold, with n replaced by m, for $(\tilde{Q}^{\Theta,m}_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi_{T,m}^{\tilde{\delta}_n};t\leqslant s}$ which is a modification of $(\overline{Q}^m_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi_{T,m}^{\tilde{\delta}_n};t\leqslant s}$. Finally using point **B.**; it follows that that (20) holds between $(\tilde{Q}^{\Theta,m}_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi_{T,m}^{\tilde{\delta}_n};t\leqslant s}$ and $(\overline{Q}^m_{t_{\delta,n}+t,t_{\delta,n}+s})_{t,s\in\pi_{T,m}^{\tilde{\delta}_n};t\leqslant s}$, with n replaced by m.*

3.2 Concentration inequalities

The concentration inequalities are crucial tools in the proof of Theorem 3.1. In particular, they provide a solution to estimate the total variation distance between P^n and $\overline{P}^{\mathcal{D},v,n}$ and between Q^n and $\overline{Q}^n$. We begin with a first practical lemma.

Lemma 3.1. *Let $\mathcal{T} \subset \mathbb{R}_+$ and let $(H_t)_{t\in\mathcal{T}}$ and $(K_t)_{t\in\mathcal{T}}$ two processes taking values in $\mathbb{R}^d$. Let $y \in \mathbb{R}^d$ and $v > 0$. We define the processes $(H(y, v)_t)_{t\in\mathcal{T}}$ and $(K(y, v)_t)_{t\in\mathcal{T}}$ with $H(y, v)_t = H_t \mathbb{1}_{\sup_{t\in\mathcal{T}}|H_t-y|<v}$ and $K(y, v)_t = K_t \mathbb{1}_{\sup_{t\in\mathbb{T}}|K_t-y|<v}$. We assume that $(H(y, v)_t)_{t\in\mathcal{T}}$ and $(K(y, v)_t)_{t\in\mathcal{T}}$ follow the same law. Then*

$$\mathbb{P}(\sup_{t\in\mathcal{T}}|H_t - y| < v) = \mathbb{P}(\sup_{t\in\mathcal{T}}|K_t - y| < v).$$

Proof. We consider that $(H_t)_{t\in\mathcal{T}}$ and $(K_t)_{t\in\mathcal{T}}$ are non null processes. Otherwise the proof is straightforward. We have $\{\sup_{t\in\mathcal{T}}|H_t-y|<v\}=\{\sup_{t\in\mathcal{T}}|H_t-y|<v\}\cap(\{H=H(y,v)\}\cup\{H=H\mathbb{1}_{\sup_{t\in\mathcal{T}}|H_t-y|\geqslant v}\})$. Moreover, since the process H is not zero, we have $\{H=H\mathbb{1}_{\sup_{t\in\mathcal{T}}|H_t-y|\geqslant v}\}=\{\sup_{t\in\mathcal{T}}|H_t-y|\geqslant v\}$ and $\{\sup_{t\in\mathcal{T}}|H_t-y|<v\}\cap\{H=H\mathbb{1}_{\sup_{t\in\mathcal{T}}|H_t-y|\geqslant v}\}=\emptyset$. We thus derive

$$\mathbb{P}(\sup_{t\in\mathcal{T}}|H_t-y|<v)=\mathbb{P}(\sup_{t\in\mathcal{T}}|H(y,v)_t-y|<v)=\mathbb{P}(\sup_{t\in\mathcal{T}}|K(y,v)_t-y|<v)=\mathbb{P}(\sup_{t\in\mathcal{T}}|K_t-y|<v),$$

where we use the fact that $(H(y,v)_t)_{t\in\mathcal{T}}$ and $(K(y,v)_t)_{t\in\mathcal{T}}$ follow the same law.

□

We now establish concentration inequalities for both continuous and discrete time processes.

3.2.1 The continuous case - The Bernstein's inequality

Proposition 3.2. *Let $(M_t)_{t\geqslant 0}$ be a continuous local martingale such that $M_0=0$ and $\langle M\rangle_\infty=+\infty$ a.s. Then, for every $c>0$ and every $v\geqslant 0$,*

$$\mathbb{P}\Big(\sup_{0\leqslant s\leqslant t}|M_s|>v|\langle M\rangle_t\leqslant c\Big)\leqslant 2\exp\Big(-\frac{v^2}{2c}\Big). \tag{41}$$

Proof. In order to prove (41), we will use the following result which concerns the specific case of the Brownian motion.

Lemma 3.2. *Let $(W_t)_{t\geqslant>0}$ a standard Brownian motion. Then, for every $v\geqslant 0$,*

$$\forall t>0,\quad \mathbb{P}(\sup_{0\leqslant s\leqslant t}|W_s|\geqslant v)\leqslant 2\exp\Big(-\frac{v^2}{2t}\Big). \tag{42}$$

Proof of Lemma 3.2. We recall that for every $\alpha>0$, $\xi^\alpha(W)_t=\exp(\alpha W_t-\frac{\alpha^2 t}{2})$ is a $\sigma(W_s,s\leqslant t)$-martingale. Using the symmetry of the Brownian motion and the Doob maximum inequality for non negative martingales, we derive

$$\begin{aligned}\mathbb{P}(\sup_{0\leqslant s\leqslant t}|W_s|\geqslant v)=&2\mathbb{P}(\sup_{0\leqslant s\leqslant t}W_s\geqslant v)=2\mathbb{P}\Big(\exp\Big(\alpha\sup_{0\leqslant s\leqslant t}W_s-\frac{\alpha^2t}{2}\Big)\geqslant\exp\Big(\alpha v-\frac{\alpha^2t}{2}\Big)\Big)\\ \leqslant&2\mathbb{P}\Big(\sup_{0\leqslant s\leqslant t}\xi^\alpha(W)_s\geqslant\exp\Big(\alpha v-\frac{\alpha^2t}{2}\Big)\Big)\leqslant 2\frac{\mathbb{E}[\xi^\alpha(W)_t]}{\exp(\alpha v-\frac{\alpha^2t}{2})}=2\exp\Big(-\alpha v+\frac{\alpha^2t}{2}\Big).\end{aligned}$$

The function $\alpha\mapsto\exp(-\alpha v+\frac{\alpha^2t}{2})$ being convex, we obtain $\inf_{\alpha>0}\exp(-\alpha v+\frac{\alpha^2t}{2})=\exp(-\frac{v^2}{2t})$ and (42) follows. □

Now, since $M_0=0$ and $\langle M\rangle_\infty=+\infty$, we can use the Dambis-Dubins-Schwarz Theorem. Let us define $T_t=\inf\{s:\langle M\rangle_s>t\}$. Then $W_t=M_{T_t}$ is a $\mathcal{F}_{T_t}$- Brownian motion and $M_t=W_{\langle M\rangle_t}$. It follows from Lemma 3.2 that

$$\begin{aligned}\mathbb{P}(\sup_{0\leqslant s\leqslant t}|M_s|\geqslant v|\langle M\rangle_t\leqslant c)=&\mathbb{P}(\sup_{0\leqslant s\leqslant\langle M\rangle_t}|W_s|\geqslant v|\langle M\rangle_t\leqslant c)=\mathbb{E}[\mathbb{P}(\sup_{0\leqslant s\leqslant\langle M\rangle_t}|W_s|\geqslant v|\langle M\rangle_t)|\langle M\rangle_t\leqslant c]\\ \leqslant&\mathbb{E}\Big[2\exp\Big(-\frac{v^2}{2\langle M\rangle_t}\Big)\Big|\langle M\rangle_t\leqslant c\Big]\leqslant 2\exp\Big(-\frac{v^2}{2c}\Big).\end{aligned}$$

□

Applying this result to Markov diffusions we derive the following result.

Corollary 3.1. *Let $\delta \in (0,T]$. Let $(X_t)_{t\geqslant 0}$ be a diffusion process with form (23) with $V_{0,\mathrm{Ito}}, V_i \in \mathcal{C}_b^0(\mathbb{R}^d;\mathbb{R}^d)$, $i \in \{1,\dots,N\}$. Then, for every $v \geqslant 0$, we have*

$$\mathbb{P}(\sup_{T-\delta\leqslant t\leqslant T}|X_T - X_t| \geqslant v|X_T) \leqslant 2\exp\Big(-\frac{v^2/2 - \delta^2\|V_{0,\mathrm{Ito}}\mathbb{1}_{B_v(X_T)}\|_\infty^2}{\delta\sum_{i=1}^N \|V_i\mathbb{1}_{B_v(X_T)}\|_\infty^2}\Big). \tag{43}$$

with $B_r(x) = \{y\in\mathbb{R}^d, |y-x| < r\}$ for every $x\in\mathbb{R}^d$ and every $r\geqslant 0$.

Proof. Applying Lemma 3.1, we derive

$$\begin{aligned}
\mathbb{P}(\sup_{T-\delta\leqslant t\leqslant T}|X_T - X_t| \geqslant v|X_T) &= \mathbb{P}\Big(\sup_{T-\delta\leqslant t\leqslant T}\Big|\int_t^T V_{0,\mathrm{Ito}}(X_s)ds + \sum_{i=1}^N\int_t^T V_i(X_s)dW_s^i\Big| \geqslant v\Big|X_T\Big)\\
&=\mathbb{P}\Big(\sup_{T-\delta\leqslant t\leqslant T}\Big|\int_t^T V_{0,\mathrm{Ito}}(X_s)\mathbb{1}_{B_v(X_T)}(X_s)ds + \sum_{i=1}^N\int_t^T V_i(X_s)\mathbb{1}_{B_v(X_T)}(X_s)dW_s^i\Big| \geqslant v\Big|X_T\Big)\\
&\leqslant\mathbb{P}\Big(\sup_{T-\delta\leqslant t\leqslant T}\Big|\sum_{i=1}^N\int_t^T V_i(X_s)\mathbb{1}_{B_v(X_T)}(X_s)dW_s^i\Big| \geqslant v - \delta\|V_{0,\mathrm{Ito}}\mathbb{1}_{B_v(X_T)}\|_\infty\Big).
\end{aligned}$$

Now, (43) follows from Proposition 3.2 with $(M_t)_{0\leqslant t\leqslant\delta} = \sum_{i=1}^N\int_{T-t}^T V_i(X_s)\mathbb{1}_{B_v(X_T)}(X_s)dW_s^i$ using that $\langle M\rangle_t \leqslant \delta\sum_{i=1}^N\|V_i\mathbb{1}_{B_v(X_T)}\|_\infty^2$.

□

3.2.2 The discrete case - The Hoeffding inequality

Before we establish a well adapted concentration inequality for our discrete time approximation, we recall the Hoeffding inequality.

Proposition 3.3. ***(The Hoeffding inequality).*** *Let $(M_n)_{n\in\mathbb{N}}$ be a discrete centered Markov process such that there exists two sequences $(b_n^l)_{n\in\mathbb{N}} \leqslant (b_n^u)_{n\in\mathbb{N}}$ such that for every $n\in\mathbb{N}^*$, $\mathbb{P}(M_n - M_{n-1} \in [b_n^l, b_n^u]) = 1$. Then, for every $v\geqslant 0$,*

$$\mathbb{P}(|M_n| \geqslant v) \leqslant 2\exp\Big(-\frac{2v^2}{\sum_{k=1}^N (b_l^u - b_k^l)^2}\Big).$$

Corollary 3.2. *We fix $T>0$ and $n\in\mathbb{N}^*$. Let $\delta\in[2T/n, T]$. Assume that $(X_t^n)_{t\in\pi_{T,n}}$ is defined as (25) with $\psi\in\mathcal{C}^1(\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^N\times\mathbb{R}_+;\mathbb{R}^d)$ and $\mathbb{P}(Z_k^n \in [b_k^l(Z^n), b_k^u(Z^n)]) = 1$ and such that (24) holds. Then, for every $v\geqslant 0$,*

$$\mathbb{P}(\sup_{t\in\pi_{T,n}^T; t\geqslant T-\tilde\delta_n}|X_T^n - X_t^n| \geqslant v|X_T^n) \leqslant 2\exp\Big(-\frac{C_{Z,n}v^2}{\delta\|\psi\mathbb{1}_{B_v(X_T)}\|_{1,1,\infty}^2}\Big), \tag{44}$$

with $\tilde\delta_n$ defined in (21) and $C_{Z,n} = \inf_{k\in\{k_{\delta,n},\dots,n\}} T/(12(b_k^u(Z^n) - b_k^l(Z^n))^2)$.

Proof. Let $N=1$ for sake of clarity in the writing. Using twice the Taylor expansion at order one of ψ with respect to the third variable and then to the fourth variable, we derive

$$X_{t_{k+1}^n}^n - X_{t_k^n}^n = w_{k+1}^0\int_0^1(1-\lambda)\partial_t\psi(\kappa_k, X_{t_k}^n, \lambda w_{k+1}^0, w_{k+1}^1)d\lambda + w_{k+1}^1\int_0^1(1-\lambda)\partial_z\psi(\kappa_k, X_{t_k}^n, 0, \lambda w_{k+1}^1)d\lambda,$$

with $w_{k+1}^0 = T/n$ and $w_{k+1}^1 \in [b_{k+1}^l(Z^n)/\sqrt{n}, b_{k+1}^u(Z^n)/\sqrt{n}]$. We apply Lemma 3.1 and Proposition 3.3 in order to obtain

$$\mathbb{P}\Big(\sup_{t\in\pi_{T,n}^T; t\geqslant T-\tilde{\delta}_n} |X_T^n - X_t^n| \geqslant v | X_T^n\Big) \leqslant 2\exp\Big(\frac{-v^2/4}{\sum_{k=k_{\delta,n}}^n (\frac{T}{n}\|\partial_t\psi \mathbb{1}_{B_v(X_T^n)}\|_\infty)^2 + (\frac{b_k^u(Z^n)-b_k^l(Z^n)}{\sqrt{n}}\|\partial_z\psi\mathbb{1}_{B_v(X_T^n)}\|_\infty)^2}\Big),$$

with $k_{\delta,n} = n(T-\tilde{\delta}_n)/T$ defined in (21). Since $|\delta - \tilde{\delta}_n| \leqslant T/n$ and $2T/n \leqslant \delta$, it follows that $n - k_\delta \leqslant 3n\delta/(2T)$. Therefore, using (24), we rearrange the terms and the proof of (44) is completed. □

3.3 Proof of Theorem 3.1

Using those different presented tools we are now able to prove our main result concerning total variation approximation for diffusion processes with locally smooth coefficients.

Proof of Theorem 3.1. In a first step we decompose the total variation distance $|\mathbb{E}[(f(X_T) - f(X_T^n))]|$. Then, using Theorem 2.1 and concentration inequalities (see Corollary 3.1 and Corollary 3.2), we study each terms appearing in this decomposition.

Step 1. Decomposition of the total variation distance. Let us define $\Omega_X = \{\sup_{T-\tilde{\delta}_n\leqslant t\leqslant T} |X_t - X_T| < v\} \cap \{\sup_{T-\tilde{\delta}_n\leqslant t_k^n\leqslant T} |X_t^n - X_T^n| < v\}$ and $\overline{\Omega}_X = \{\sup_{T-\tilde{\delta}_n\leqslant t\leqslant T} |\overline{X}_t^{\mathcal{D},v} - \overline{X}_T^{\mathcal{D},v}| < v\} \cap \{\sup_{T-\tilde{\delta}_n\leqslant t_k^n\leqslant T} |\overline{X}_t^n - \overline{X}_T^n| < v\}$. From $(\mathbf{SG_{Reg}}(\overline{P}^{\mathcal{D},v,n}, \overline{Q}^n, h, q, n, \delta, T, \mathcal{D}, v))$ and since $\mathrm{supp}(f) \subset \mathcal{D}^{2v}$, it follows that

$$\mathbb{E}[(f(X_T) - f(X_T^n))\mathbb{1}_{\Omega_X}] = \mathbb{E}[(f(\overline{X}_T^{\mathcal{D},v}) - f(\overline{X}_T^n))\mathbb{1}_{\overline{\Omega}_X}],$$

which yields

$$|\mathbb{E}[(f(X_T) - f(X_T^n))]| \leqslant |\mathbb{E}[f(\overline{X}_T^{\mathcal{D},v}) - f(\overline{X}_T^n)]| + 2\|f\|_\infty\big(\mathbb{P}(\Omega_X^c | X_T, X_T^n \in \mathcal{D}^{2v}) + \mathbb{P}(\overline{\Omega}_X^c | \overline{X}_T^{\mathcal{D},v}, \overline{X}_T^n \in \mathcal{D}^{2v})\big).$$

Step 2. Analysis of the terms composing the total variation distance estimation. To complete the proof, it remains to estimate each term of the *r.h.s.* of this inequality. We focus on the first one. As a direct consequence of Proposition 3.1 (see Remark 3.5) and Theorem 2.1 (which follow from $L_{z_*}(\varepsilon_*, r_*)$ (see (28)), (34), (30), (36) and (35) and $(\mathbf{SG_{Reg}}(\overline{P}^{\mathcal{D},v,n}, \overline{Q}^n, h, q, n, \delta, T, \mathcal{D}, v))$ **II)**), we deduce from the mixing regularization properties of $(\overline{X}_t^n)_{t\in\pi_{T,n}}$ that

$$|\mathbb{E}[f(\overline{X}_T^{\mathcal{D},v}) - f(\overline{X}_T^n)]| \leqslant C(1+|x|^\beta)\frac{\mathfrak{K}_{q+3}(\tilde{\psi})^l}{(\lambda_*\delta)^{\eta(q)}}\|f\|_\infty/n^h$$

Now, we study the second term of the *r.h.s* of the decomposition of the error. First, let us observe that

$$\begin{aligned}\mathbb{P}(\Omega_X^c|X_T, X_T^n \in \mathcal{D}^{2v}) \leqslant &\mathbb{P}(\sup_{T-\tilde{\delta}_n\leqslant t\leqslant T}|X_t - X_T| \geqslant v|X_T \in \mathcal{D}^{2v}) \\ &+ \mathbb{P}(\sup_{T-\tilde{\delta}_n\leqslant t_k^n\leqslant T}|X_{t_k^n}^n - X_T^n| \geqslant v|X_T^n \in \mathcal{D}^{2v}),\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}\big(\overline{\Omega}_X^c|\overline{X}_T^{\mathcal{D},v}, \overline{X}_T^n \in \mathcal{D}^{2v}\big) \leqslant &\mathbb{P}(\sup_{T-\tilde{\delta}_n\leqslant t\leqslant T}|\overline{X}_t^{\mathcal{D},v} - \overline{X}_T^{\mathcal{D},v}| \geqslant v|\overline{X}_T \in \mathcal{D}^{2v}) \\ &+ \mathbb{P}(\sup_{T-\tilde{\delta}_n\leqslant t_k^n\leqslant T}|\overline{X}_{t_k^n}^n - \overline{X}_T^n| \geqslant v|\overline{X}_T^n \in \mathcal{D}^{2v}).\end{aligned}$$

Now, we notice that for every $x \in \mathcal{D}^{2v}$, then $B_v(x) \subset \mathcal{D}^v$. Using Lemma 3.1 and Corollary 3.1, we derive

$$\begin{aligned}\mathbb{P}(\sup_{T-\tilde{\delta}_n\leqslant t\leqslant T}|\overline{X}_t^{\mathcal{D},v} - \overline{X}_T^{\mathcal{D},v}| \geqslant v|\overline{X}_T \in \mathcal{D}^{2v}) =&\mathbb{P}(\sup_{T-\tilde{\delta}_n\leqslant t\leqslant T}|X_t - X_T| \geqslant v|X_T \in \mathcal{D}^{2v}) \\ \leqslant &2\exp\Big(-\frac{v^2/2 - \delta^2\|V_{0,\mathrm{Ito}}\mathbb{1}_{\mathcal{D}^v}\|_\infty^2}{\delta\sum_{i=1}^N\|V_i\mathbb{1}_{\mathcal{D}^v}\|_\infty^2}\Big),\end{aligned}$$

In the same way, using Lemma 3.1 with the hypothesis $(\mathbf{SG_{Reg}}(\overline{P}^{\mathcal{D},v,n}, \overline{Q}^n, h, q, n, \delta, T, \mathcal{D}, v))$ **I)**, it follows from Corollary 3.2, that

$$\begin{aligned}\mathbb{P}(\sup_{T-\tilde{\delta}_n\leqslant t_k^n\leqslant T}|\overline{X}_{t_k^n}^n - \overline{X}_T^n| \geqslant v|\overline{X}_T^n \in \mathcal{D}^{2v}) =&\mathbb{P}(\sup_{T-\tilde{\delta}_n\leqslant t_k^n\leqslant T}|X_{t_k^n}^n - X_T^n| \geqslant v|X_T^n \in \mathcal{D}^{2v}) \\ \leqslant &2\exp\Big(-\frac{C_{Z,n}v^2}{\delta\|\psi\mathbb{1}_{\mathcal{D}^v}\|_{1,1,\infty}^2}\Big).\end{aligned}$$

Gathering all the terms together yields (37). □

4 Second order total variation convergence towards CIR processes

In this section, we apply Theorem 3.1 obtained in a general setting to the case of a second weak order scheme for the CIR process. This scheme was first introduced in [2] and is built using cubature methods. It is inspired from the approach proposed in [28] to build the so called Ninomiya Victoir schemes for SDE with smooth coefficients. Finally, in [23] or [5], the authors showed that the total variation convergence occurs for those cubature schemes as soon as the coefficients of the SDE are smooth. In this section, our purpose is to exploit and extend those results in order to obtain total variation convergence results for the CIR diffusion process which is singular in the neighborhood of zero.

4.1 The Ninomiya Victoir scheme

We begin by presenting the Ninomiya Victoir scheme (see [28]). Let us define $\exp(V)(x) := \Phi_V(x,1)$ where Φ_V solves the deterministic equation

$$\Phi_V(x,t) = x + \int_0^t V(\Phi_V(x,s))ds. \tag{45}$$

By a change of variables, we obtain $\Phi_{\varepsilon V}(x,t) = \Phi_V(x,\varepsilon t)$ so we have

$$\exp(\varepsilon V)(x) := \Phi_{\varepsilon V}(x,1) = \Phi_V(x,\varepsilon).$$

We also notice that the semigroup of the above Markov process is given by $P_t^V f(x) = f(\Phi_V(x,t))$ and has the infinitesimal generator $A_V f(x) = Vf(x)$. In particular the relation $P_t^V A_V = A_V P_t^V$ reads

$$Vf(\Phi_V(x,t)) = A_V P_t^V f = P_t^V A_V f = V(x)\partial_x \left(f \circ \Phi_V\right)(x,t).$$

Using m times Dynkin's formula $P_t^V f(x) = f(x) + \int_0^t P_s^V A_V f(x) ds$ we obtain

$$f(\Phi_V(x,t)) = f(x) + \sum_{r=1}^{m} \frac{t^r}{r!} V^r f(x) + \frac{1}{m!}\int_0^t (t-s)^m V^{m+1} P_s^V f(x) ds. \tag{46}$$

We present now a second order scheme introduced in [28] and also used in [2]. We consider a sequence $\kappa_k,\ k \in \mathbb{N}^*$, of independent Bernoulli random variables. Let us define $\psi : \mathbb{Z} \times \mathbb{R}^d \times \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}^d$ by:

$$\psi(\kappa, x, w^1, w^0) = \begin{cases} \exp(w^0 V_0) \circ \exp(w^{1,1} V_1) \circ \cdots \circ \exp(w^{1,N} V_N) \circ \exp(w^0 V_0)(x), & \text{if } \kappa = 1, \\ \exp(w^0 V_0) \circ \exp(w^{1,N} V_N) \circ \cdots \circ \exp(w^{1,1} V_1) \circ \exp(w^0 V_0)(x), & \text{if } \kappa = -1. \end{cases} \tag{47}$$

Moreover, we denote $w_k^0 = T/2n$ and $w_k^1 = (w_k^{1,i})_{i=1,..,N}$ with $w_k^i = \sqrt{T} Z_k^i/\sqrt{n}$, $i = 1,\ldots,N$ and we assume that $Z_k,\ k \in \mathbb{N}$ are independent random variables which are lower bounded by the Lebesgue measure *i.e.* $L_{z_*}(\varepsilon_*, r_*)$ (see (28)) holds: There exists $z_{*,k} \in \mathbb{R}^N$ and $\varepsilon_*, r_* > 0$ such that for every Borel set $A \subset \mathbb{R}^N$ and every $k \in \mathbb{N}^*$: $\mathbb{P}(Z_k \in A) \geqslant \varepsilon_* \lambda(A \cap B_{r_*}(z_{*,k}))$. Finally, we assume that the sequence Z_k satisfies the following moment condition:

$$\begin{aligned} \mathbb{E}[Z_k^i] = \mathbb{E}[(Z_k^i)^3] = \mathbb{E}[(Z_k^i)^5] = 0, \qquad \mathbb{E}[(Z_k^i)^2] = 1, \qquad \mathbb{E}[(Z_k^i)^4] = 3, \\ \forall p \geqslant 1, \qquad \mathbb{E}[|Z_k|^p] < \infty. \end{aligned} \tag{48}$$

We recall that $T > 0$, $n \in \mathbb{N}^*$, and that $t_k^n = Tk/n$. One step of the scheme for diffusion with regular coefficients (between times t_k and t_{k+1}) is thus given by

$$X_{t_{k+1}^n}^n = \psi(\kappa_k, X_{t_k^n}^n, w_{k+1}^1, w_{k+1}^0). \tag{49}$$

4.1.1 Convergence results for diffusion with smooth coefficients

We begin by recalling some convergence results concerning this numerical scheme.

Smooth test functions

Here, we assume that the test functions are smooth. We state a first result, which is the starting point in order to prove the convergence for the total variation distance.

Theorem 4.1. *Let $(X_t)_{t\geqslant 0}$ be the process defined by (22) and $(X_t^n)_{t\in\pi_{T,n}}$ given by (49). Assume that (48) holds. We have the following properties:*

A. *Assume that for every $l \leqslant 3$, we have $V_0^l : \mathcal{C}^\infty_{\rm pol}(\mathbb{R}^d;\mathbb{R}^d) \to \mathcal{C}^\infty_{\rm pol}(\mathbb{R}^d;\mathbb{R}^d)$ and for every $i = 1,\dots N$, we have $V_i^{2l}, A^l : \mathcal{C}^\infty_{\rm pol}(\mathbb{R}^d;\mathbb{R}^d) \to \mathcal{C}^\infty_{\rm pol}(\mathbb{R}^d;\mathbb{R}^d)$. We also assume that $\sum_{i=0}^N V_i(x) \leqslant C(1+|x|)$ and that $\sup_{t\in[0,T]} |\Phi_{V_i}(x,t)| \leqslant C_i(1+|x|^{\beta_i})$, $C_i, \beta_i \geqslant 1$. Then, $E_{n,\rm pol}(2,6)$ (see (10)) and $E'_{n,\rm pol}(2,6)$ (see (11)) are satisfied between $(X_t)_{t\in\pi_{T,n}}$ and $(X_t^n)_{t\in\pi_{T,n}}$.*

B. *Suppose that $V_i \in \mathcal{C}_b^\infty(\mathbb{R}^d;\mathbb{R}^d)$. Then, there exists some universal constants $C, l \geqslant 1$ such that for every $f \in \mathcal{C}_b^6(\mathbb{R}^d)$, we have*

$$\sup_{t\in\pi_{T,n}^T} |\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_t^n)]| \leqslant CC_6(V)^l \|f\|_{6,\infty}/n^2, \tag{50}$$

with $C_k(V) := \sup_{i=0,\dots,N} \|V_i\|_{k,\infty}$.

Proof. We prove only point **A.** The proof of **B.** can be found in [4] (see Theorem 5.1.). We focus on the proof of $E'_{n,\rm pol}(2,6)$ (see (10)). We assume that $N = 1$ and $T = 1$ for sake of clarity, the proof being similar otherwise. It is sufficient to prove that the schemes $X^{0,n}_{t_{k+1}} = \Phi_{V_0}(X^{0,n}_{t_k}, t_{k+1} - t_k)$ and $X^{1,n}_{t_{k+1}} = \Phi_{V_1}(X^{1,n}_{t_k}, \frac{Z^1_k}{\sqrt{n}})$, $k \in \mathbb{N}$, are weak second order schemes. We prove that they are in fact h-weak order schemes, for every integer $h \in \mathbb{N}^*$ as soon as $Z_k = Z_k^1$ matches the $2h+1$-th moments of the centered normal distribution and has finite moments of any order. First, we notice that the sublinear growth of the coefficients implies that for every $t \in \mathbb{R}_+$, the moments functions $x \mapsto \mathbb{E}[|X_t(x)|^q]$, $q \in \mathbb{N}$, belong to $\mathcal{C}^0_{\rm pol}(\mathbb{R}^d)$ (see [1]). Let us consider $f \in \mathcal{C}_b^\infty(\mathbb{R}^d)$. According to the definition of V_0, we notice that for every $l \in \mathbb{N}, V_0^l f \in \mathcal{C}^\infty_{\rm pol}(\mathbb{R}^d)$. In this case, expansion (46) reads

$$\forall h \in \mathbb{N}, \quad f(\Phi_{V_0}(x,t)) = f(x) + \sum_{l=1}^h \frac{t^l}{l!} V_0^l f(x) + R_{0,t}^{h+1} f(x)$$

with

$$R_{0,t}^{h+1} f(x) = \int_0^t \frac{(t-s)^h}{h!} V_0^{h+1} f(\Phi_{V_0}(x,s)) ds$$

As already mentioned, $V_0^{h+1} f \in \mathcal{C}^\infty_{\rm pol}(\mathbb{R}^d)$. Moreover, for every $x \in \mathbb{R}^d$, $\sup_{t\in[0,T]} |\Phi_{V_0}(x,t)| \leqslant C_0(1 + |x|^{\beta_0})$ and then for every $t \in [0,1]$, there exists $\beta \in \mathbb{N}$ such that

$$|R_{0,t}^{h+1} f(x)| \leqslant C(1+|x|^\beta) t^{h+1} \|f\|_{h+1,\infty}.$$

The scheme with transition function Φ_{V_0} is thus a h-weak order scheme for the operator V_0. Now, let $H_k = \sqrt{t} Z_k$. Still using (46), we derive

$$\mathbb{E}[f(\Phi_{V_1}(H_k, x))] = \sum_{2l\leqslant 2h+1} \frac{t^l \mathbb{E}[|Z_k|^{2l}]}{(2l)!} V_1^{2l} f(x) + \mathbb{E}[R_{1,H_k}^{h+1} f(x_1)]$$

with

$$R^{h+1}_{1,H_k} f(x) = \frac{H_k^{2h+2}}{(2h+1)!} \int_0^1 (1-s)^{2h+1} V_1^{2h+2} f(\Phi_{V_1}(x, sH_k)) ds.$$

Moreover, $V_1^{2h+2} f \in \mathcal{C}^\infty_{\rm pol}(\mathbb{R}^d;\mathbb{R})$ and for every $x \in \mathbb{R}^d$, $\sup_{t\in[0,T]} |\Phi_{V_1}(x,t)| \leqslant C_1(1+|x|^{\beta_1})$. Once again, there exists $\beta_1 \in \mathbb{N}^*$, $C_1 > 0$, such that

$$\mathbb{E}[|R^{h+1}_{1,H_k} f(x)|] \leqslant \frac{\mathbb{E}[|H_k|^{2h+2}]}{(2h+1)!} C_1(1+|x|^{\beta_1}) \|f\|_{2h+2,\infty} \leqslant C(1+|x|^{\beta_1}) t^{h+1} \|f\|_{2h+2,\infty},$$

and the scheme with transition function Φ_{V_1} is a h-order scheme for the operator V_1.

Finally, for every $f \in \mathcal{C}^\infty_{\rm pol}$ then $Af \in \mathcal{C}^\infty_{\rm pol}$ and we obtain $E'_{n,\rm pol}(2,6)$ (see (11)) using the Ninomiya Victoir composition (47) and the polynomial control of the moments of the diffusion and of Φ_{V_i}, $i \in \{0,\dots,N\}$. The proof of $E_{n,\rm pol}(2,6)$ (see (10)) is very similar and left to the reader. □

Remark 4.1. *Notice that property $\mathfrak{P}'_q$ (see (8)) has already been studied in [1] for the CIR and since $\sup_{t\in[0,T]} \Phi_{V_i}(.,t) \in \mathcal{C}^0_{\rm pol}(\mathbb{R}^d)$, we can use Property 2.1 in order to obtain the weak convergence for smooth test functions with polynomial growth.*

Bounded measurable test functions

Under an ellipticity condition, we can control the total variation distance between a diffusion process with form (22) and its second order scheme (49).

Theorem 4.2. *Assume that $V_i \in \mathcal{C}^\infty_b(\mathbb{R}^d;\mathbb{R}^d), i = 0,\dots,N$, and*

$$\inf_{|\xi|=1} \sum_{i=1}^N \langle V_i(x), \xi\rangle^2 \geqslant \lambda_* > 0 \qquad \forall x \in \mathbb{R}^d. \tag{51}$$

Let $S \in (0, T/2)$. Then there exists $n_0 \in \mathbb{N}^*$, $l \in \mathbb{N}^*$ and $C \geqslant 1$ such that for every $n \geqslant n_0$ and for every bounded and measurable function f defined on $\mathbb{R}^d$,

$$\sup_{t\in\pi^{2S,T}_{T,n}} |\mathbb{E}[f(X_t)] - \mathbb{E}[f(X^n_t)]| \leqslant C \frac{C_6(V)^l \mathfrak{K}_9(\psi)^l}{(\lambda_* S)^{42}} \|f\|_\infty / n^2. \tag{52}$$

Remark 4.2. *This result has already been obtained in [5]. The result (52) signifies the convergence in total variation distance for the weak error with order 2. We notice that, the key point of this proof does not rely on the weak order of the scheme. This is the fact that, the splitting procedure (47) in order to build the scheme, always includes a diffusion part (through $\exp(\frac{Z^i_k}{\sqrt{n}} V^i)$) together with the ellipticity condition (51). Consequently, a similar procedure could be used in order to prove the convergence in total variation for even higher order schemes as soon as we control this error for smooth test functions. Finally, it is important to notice that the generic property $L_{z_*}(\varepsilon_*, r_*)$ (see (28)) is crucial here. On the one hand, it enables to apply a Malliavin inspired calculus crucial to achieve total variation convergence. On the other hand, since the random variables $(Z_k)_{k\in\mathbb{N}^*}$ do not have a specific law but only satisfy the Doeblin condition $L_{z_*}(\varepsilon_*, r_*)$ (see (28)) and the moment condition (48), the result can be seen as an invariance principle.*

4.2 The CIR model

The CIR process is an $\mathbb{R}_+$-valued diffusion process driven by the following SDE,

$$dX_t = (a - kX_t)dt + \sigma\sqrt{X_t}dW_t. \tag{53}$$

This model was first presented in 1985 in [13] and was inspired by Vasicek (1977) models [35] by modifying the volatility term introducing a 'square root' term, among others in order to guarantee non-negativity. In this paper, we suppose that $a, k, \sigma > 0$. In this case, it is important to notice that the model does not reach 0 for $2a \geqslant \sigma^2$.

4.2.1 Second weak order scheme for the CIR process

The Ninomiya Victoir scheme for the CIR

Applying the notations from (22), we have

$$\begin{aligned} \forall x \in \mathbb{R}, \quad V_{0,cir}f(x) &= (a - kx - \frac{\sigma^2}{4})\partial_x f(x) \\ \forall x \in \mathbb{R}_+, \quad V_{1,cir}f(x) &= \sigma\sqrt{x}\partial_x f(x) \end{aligned}$$

Solving the PDE (45) yields the following flows

$$\begin{aligned} \forall x \in \mathbb{R}, \quad \Phi_{0,cir}(t,x) &= xe^{-kt} + (a - \frac{\sigma^2}{4})\frac{1-e^{-kt}}{k} \\ \forall x \in \mathbb{R}_+, \quad \Phi_{1,cir}(t,x) &= (\sqrt{x} + \frac{\sigma}{2}t)^2. \end{aligned}$$

At this point, we distinguish two cases. Indeed, we notice that if $\sigma^2 > 4a$ and $x \leqslant x^*(t) := k^{-1}(\frac{\sigma^2}{4} - a)(e^{kt}-1)$, then $\Phi_0^{CIR}(t,x)$ takes negative values and then the scheme (49) is not well defined anymore. In this case, we have to introduce another scheme in the neighborhood of zero (when $x \leqslant x^*(t)$) and to use a switching procedure in this area. Otherwise, we will prove that, as soon as the scheme (49) is well defined, then it is a second weak order scheme. As a consequence, if $4a \geqslant \sigma^2$, we define

$$\psi_{cir}(x, w^1, w^0) = \exp(w^0 V_{0,cir}) \circ \exp(w^1 V_{1,cir}) \circ \exp(w^0 V_{0,cir})(x). \tag{54}$$

Now, we introduce $w_k^0 = T/n (= t_k^n - t_{k-1}^n)$ and $w_k^1 = \sqrt{T}Z_k/\sqrt{n}$, where Z_k, $k \in \mathbb{N}^*$, are independent $\mathbb{R}$-valued random variables which are lower bounded by the Lebesgue measure *i.e.* $L_{z_*}(\varepsilon_*, r_*)$ (see (28)) holds. Finally, we assume that the sequence Z_k satisfies the moment conditions (48) and that $\mathbb{P}(Z_k \in [-\sqrt{3}, \sqrt{3}]) = 1$ for every $k \in \mathbb{N}^*$.

One step of the scheme for the CIR diffusion (between times t_k^n and t_{k+1}^n) is given by

$$X_{t_{k+1}}^n = \psi_{cir}(\kappa_k, X_{t_k}^n, w_{k+1}^1, w_{k+1}^0). \tag{55}$$

4.2.2 Second weak order scheme in the neighborhood of zero

This section is dedicated to the introduction of a specific scheme in the neighborhood of zero when $\sigma^2 > 4a$. In this case, the scheme defined in (55) is indeed not well defined. First of all, we have to

identify the threshold such that, for every time step we use the scheme (55) or we introduce another scheme. To this end, we consider some bounded random variables Z_k, in order to use the following result that determine this switching threshold.

Lemma 4.1. *Assume that $\sigma^2 > 4a$. Let $t > 0$, $A > 0$, $w^1 \in [-A, A]$ and define*

$$K(t,A) = e^{\frac{kt}{2}} \left[\left(\frac{\sigma^2}{4} - a \right) \frac{1 - e^{\frac{-kt}{2}}}{k} + \left(\sqrt{e^{\frac{kt}{2}} \left(\frac{\sigma^2}{4} - a \right) \frac{1 - e^{-\frac{kt}{2}}}{k}} + \frac{\sigma}{2} A \right)^2 \right]. \tag{56}$$

Then

$$\forall x \geqslant K(t,A), \quad \Phi_{0,cir}(\frac{t}{2}, .) \circ \Phi_{1,cir}(w^1, .) \circ \Phi_{0,cir}(\frac{t}{2}, x) \geqslant 0.$$

Proof. We first notice that $K(t,A) \geqslant x^*(\frac{t}{2})$ and then for every $t \in [0,1]$, the function $x \mapsto \Phi_{0,cir}(\frac{t}{2}, .) \circ \Phi_{1,cir}(w^1, .) \circ \Phi_{0,cir}(\frac{t}{2}, x)$, $x \in [K(t,A), +\infty)$, is well defined. Moreover, $\Phi_{0,cir}$ is increasing with respect to its space variable x and

$$\Phi_{0,cir}\Big(\frac{t}{2}, K(t,A)\Big) = \left(\sqrt{e^{\frac{kt}{2}} \left(\frac{\sigma^2}{4} - a \right) \frac{1 - e^{-\frac{kt}{2}}}{k}} + \frac{\sigma}{2} A \right)^2.$$

Since all the terms inside the parenthesis are positive and $w^1 \geqslant -A$, we deduce that

$$\begin{aligned} \Phi_{1,cir}(w^1, .) \circ \Phi_{0,cir}(\frac{t}{2}, K(t,A)) &= \left[\sqrt{e^{\frac{kt}{2}} \left(\frac{\sigma^2}{4} - a \right) \frac{1 - e^{\frac{kt}{2}}}{k}} + \frac{\sigma}{2}(A + w^1) \right]^2 \\ &\geqslant e^{\frac{kt}{2}} \left(\frac{\sigma^2}{4} - a \right) \frac{1 - e^{-\frac{kt}{2}}}{k} = x^*\Big(\frac{t}{2}\Big). \end{aligned}$$

Finally, notice that $\Phi_{1,cir}$ is increasing with respect to its space variable x and the proof is completed. □

Now, it remains to introduce the scheme that we will use in the neighborhood of zero.

Moments matching approach. Our approach consists in checking, at each step of the Ninomiya Victoir scheme, if it may take negative values. In this case, we switch with a scheme based on moment matching approach and inspired from Andersen [3] and introduced in [2]. One step of this scheme consists in simulating a discrete random process $(\zeta_t)_{t \geqslant 0} \in \{y_1, y_2\} \in \mathbb{R}_+^2$, which depends on the current position, at the selected date $t \geqslant 0$ which corresponds to the time step. We recall that the first two moments of the CIR starting from $X_0 = x \in \mathbb{R}_+$, are given by

$$\begin{aligned} \mathbb{E}^x[X_t] &= x e^{-kt} + \frac{a}{k}(1 - e^{-kt}) \\ \mathbb{E}^x[X_t^2] &= x \frac{\sigma^2}{k}(e^{-kt} - e^{-2kt}) + a \frac{\sigma^2}{2k^2}(1 - e^{-kt})^2 + \mathbb{E}^x[X_t]^2, \end{aligned}$$

and we denote $u_q(x,t) = \mathbb{E}^x[X_t^q]$. In the neighborhood of zero, one step of the scheme $(X_t^n)_{t \in \pi_{T,n}^T}$, between t_k^n and t_{k+1}^n consists in simulating a random variable distributed under the same law as

$\zeta_{t^n_{k+1}-t^n_k}$ with (y_1, y_2) given by the solution of the following equation:

$$\begin{cases} P(\zeta_t = y_1)y_1 + P(\zeta_t = y_2)y_2 & = u_1(x,t) \\ P(\zeta_t = y_1)y_1^2 + P(\zeta_t = y_2)y_2^2 & = u_2(x,t), \end{cases}$$

on the event $\{\overline{X}^n_{t_k} = x\}$. In this equation the only fixed parameters are $u_1(t,x)$ and $u_2(t,x)$, and obviously $\mathbb{P}(\zeta_t = y_2) = 1 - \mathbb{P}(\zeta_t = y_1) \in]0,1[$. Thus we can fix values for y_1 and y_2 to solve the first equation and then the second one gives a second order equation to solve to find $\mathbb{P}(Y_t = y_1)$. For instance, let $v \in]0,1[$ and let us choose

$$y_1 = v\frac{u_1(t,x)}{\mathbb{P}(\zeta_t = y_1)}, y_2 = (1-v)\frac{u_1(t,x)}{\mathbb{P}(\zeta_t = y_2)}.$$

The second equation then rewrites

$$u_2(t,x)\mathbb{P}(\zeta_t = y_1)^2 + \left[(1-2v)u_1(t,x) - u_2(t,x)\right]\mathbb{P}(\zeta_t = y_1) + v^2 u_1(t,x)^2 = 0.$$

We thus consider the second order equation which depends on the parameter v:

$$u_2(t,x)\rho^2 + \left[(1-2v)u_1(t,x) - u_2(t,x)\right]\rho + v^2 u_1(t,x)^2 = 0. \tag{57}$$

We want to find a couple $(\rho(v), v)$, where $\rho(v)$ is a solution of the equation above, such that $\rho(v) \in]0,1[$. Let us denote

$$\Delta_v(t,x) = \left[(1-2v)u_1(t,x) - u_2(t,x)\right]^2 - 4v^2 u_1(t,x)^2 u_2(t,x)$$

Now we chose $v = \frac{1}{2}$. We compute the following solution for (57):

$$\rho(1/2) = \frac{u_2(t,x) \pm \sqrt{u_2(t,x)(u_2(t,x) - u_1(t,x)^2)}}{2u_2(t,x)}$$

At this point, we thus put

$$\mathbb{P}(\zeta_t = y_1) = \frac{1}{2}\left(1 - \sqrt{1 - \frac{u_1(t,x)^2}{u_2(t,x)}}\right).$$

Since $u_1(t,x) \geqslant \max\left(a^2\left(\frac{1-e^{-kt}}{k}\right), 2x\frac{a}{k}(e^{-kt} - e^{-2kt})\right)$, we derive the following lower bound,

$$P(\zeta_t = y_1) \geqslant \frac{1}{2}\left(1 - \sqrt{1 - \frac{a}{a+\sigma^2}}\right).$$

This provides the following crucial property in order to prove the second order convergence: Assume that there exists $C > 0$ such that $0 < K(t,A) < Ct$. Then, for every $q \in \mathbb{N}$,

$$\forall t \in]0,1], \forall x \in [0, K(t,A)], \quad \exists C > 1, \mathbb{E}[\zeta_t^q] < Ct^q$$

We define the transition function in the neighborhood of zero by

$$\widehat{\psi}_{cir}(\rho, x, w^0) = \begin{cases} \dfrac{u_1(w^0,x)}{2p(x,w^0)}, & \text{if } \rho = 1, \\ \dfrac{u_1(w^0,x)}{2(1-p(x,w^0))}, & \text{if } \rho = -1, \end{cases}$$

with $p(x,w^0)=\frac{1}{2}(1-\sqrt{1-\frac{u_1(w^0,x)^2}{u_2(w^0,x)}})$. Now let $w_k^0=T/N$. We define a step of the second order scheme for the CIR in the neighborhood of zero by

$$X_{t_{k+1}}^n=\widehat{\psi}_{cir}(\rho_k,X_{t_k}^n,w_{k+1}^0),$$

with $(\rho_k)_{k\in\mathbb{N}}$ the sequence of random variables such that $\mathbb{P}(\rho_k=1|X_{t_k}^n)=p(X_{t_k}^n,w_{k+1}^0)$ and $\mathbb{P}(\rho_k=-1|X_{t_k}^n)=1-p(X_{t_k}^n,w_{k+1}^0)$. Finally in the case $\sigma^2>4a$, we use the Lemma 4.1 and we define the CIR scheme by

$$X_{t_{k+1}}^n=\begin{cases}\psi_{cir}(\kappa_k,X_{t_k}^n,w_{k+1}^1,w_{k+1}^0), & \text{if } X_{t_k}^n\geqslant K(T/n,\sqrt{3T/n}),\\ \widehat{\psi}_{cir}(\rho_k,X_{t_k}^n,w_{k+1}^0), & \text{if } X_{t_k}^n<K(T/n,\sqrt{3T/n}),\end{cases}\tag{58}$$

where $w_k^1=\sqrt{T}Z_k/\sqrt{n}$ is defined as in (54).

4.3 Convergence results

4.3.1 Smooth test functions

We focus on the convergence of the CIR schemes defined in (55) and (58). Using Property 2.1. In order to apply this result we have to establish the following straightforward property.

Lemma 4.2. *Let $l\in\mathbb{N}$. Then, we have*

$$V_{1,cir}^{2l}:\mathcal{C}_p^\infty(\mathbb{R}_+^*)\to\mathcal{C}_p^\infty(\mathbb{R}_+^*).\tag{59}$$

Proof. Notice that it is sufficient to show the result for $l=1$. In this case, the result is straightforward since

$$\forall f\in\mathcal{C}_p^\infty(\mathbb{R}_+^*),\quad V_{1,cir}^2f(x)=\sigma^2\sqrt{x}\left(\frac{1}{2\sqrt{x}}\partial_xf(x)+\sqrt{x}\partial_x^2f(x)\right)=\frac{\sigma^2}{2}\partial_xf(x)+\sigma^2x\partial_x^2f(x).$$

□

The property (59) leads the following short time estimate.

Theorem 4.3. *Let $(X_t)_{t\geqslant0}$ be the process defined by (53) and let $(X_{t_k}^n)$ be defined by (55) if $4a\geqslant\sigma^2$ and by (58) otherwise. Then, there exists $l\in\mathbb{N}^*$, $C,\beta>1$, such for every $f\in\mathcal{C}_b^6(\mathbb{R}_+;\mathbb{R})$, $E'_{n,\mathrm{pol}}(2,6)$ (see (11)) holds and*

$$\forall x\in\mathbb{R}_+,\quad|\mathbb{E}[f(X_T(x)]-\mathbb{E}[f(X_T^n(x)]|\leqslant C(1+|x|^\beta)\|f\|_{6,\infty}/n^2.\tag{60}$$

Proof. First we recall that the proof of $\mathfrak{P}'_q$ (see (8)) for the CIR diffusion and its scheme are given in [2]. If $4a\geqslant\sigma^2$, it is sufficient to use Theorem 4.1 point **A.** to obtain (11) and then (60) follows from Property 2.1. Now, let $\sigma^2>4a$. Here, the only thing we still have to check is that (11) is satisfied for the moment matching scheme as soon as $x\in[0,K(T/n,\sqrt{3T/n})]$ with K defined in (56). This is a consequence of the two following results.

Lemma 4.3. *Let $\sigma^2>4a$. Let K be the function defined in (56) and let $\lambda\in\mathbb{R}$. Then, for every $q\in\mathbb{N}$,*

$$\exists C_q>0,t_q\in[0,1],\quad\forall t\in[0,t_q],\forall x\in[0,K(t,\lambda\sqrt{t})[,\quad\mathbb{E}[X_t^q|X_0=x]\leqslant C_qt^q$$

Proof. Let us consider the function $g(x) = x^{q+1}$, $x \in \mathbb{R}_+$. Applying Ito's formula to g for the CIR process gives:

$$X_t^{q+1} = x^{q+1} + \int_0^t \left[(q+1)X_s^q(a - kX_s) + \frac{1}{2}q(q+1)\sigma^2 X_s^q\right] ds + \int_0^t (q+1)X_s^{q+\frac{1}{2}}\sigma dW_s.$$

Using localization by considering the stopping time $\tau_m = \inf\{t \geqslant 0 : |X_t| \geqslant m\}$ and the Fubini theorem, yields

$$\begin{aligned}\mathbb{E}[|X_{t\wedge\tau_m}^{q+1}|] &\leqslant x^{q+1} + \mathbb{E}\Big[\Big|\int_0^{t\wedge\tau_m} \big[(q+1)X_s^q(a-kX_s) + \frac{1}{2}q(q+1)\sigma^2 X_s^q\big]ds\Big|\Big] \\ &\leqslant x^{q+1} + \int_0^t \left[(q+1)a + \frac{1}{2}q(q+1)\sigma^2\right]\mathbb{E}[|X_{s\wedge\tau_m}^q|] - k\mathbb{E}[X_{s\wedge\tau_m}^{q+1}]ds\end{aligned}$$

Reasoning by induction, we assume that: $\exists C_q > 0, t_q \in [0,1], \mathbb{E}[X_{s\wedge\tau_m}^q] \leqslant C_q t^q$. Since $x \in [0, K(t,\lambda\sqrt{t})[$ with $K(t,\lambda\sqrt{t}) = \underset{t\to 0}{O}(t)$, there exists $t_{q+1} \in [0,1]$ such that

$$\forall t \in [0,t_{q+1}], \mathbb{E}[|X_{t\wedge\tau_m}^{q+1}|] \leqslant \quad K(t,\lambda\sqrt{t})^{q+1} + C_q\left((q+1)a + \tfrac{1}{2}q(q+1)\sigma^2\right)t^{q+1} + \int_0^t k\mathbb{E}[|X_{s\wedge\tau_m}^{q+1}|]ds$$

Applying Gronwall's lemma and the fact that $t \in [0,t_{q+1}]$, and that there exists $C > 0$ such that $K(t,\lambda\sqrt{t}) < Ct$, we deduce that

$$\mathbb{E}[|X_{t\wedge\tau_m}^{q+1}|] \leqslant \left(C_q\left[(q+1)a + \frac{1}{2}q(q+1)\sigma^2\right] + C\right)e^k t^{q+1}$$

Finally, the continuity of the flow and Fatou's lemma give the result. □

Using this result the property $E'_{n,\mathrm{pol}}(2,6)$ (see (11)) is a consequence of the following theorem

Lemma 4.4. *Let $h \in \mathbb{N}$, $x \in [0, C^*(t)[$ with $C^*(t) = \underset{t\to 0}{O}(t)$ and let $(\zeta_t)_{t\geqslant 0}$ be a $\mathbb{R}_+$ valued random process such that $\forall q = 1,\dots,h, \mathbb{E}[\zeta_t^q] = \mathbb{E}[X_t^q]$ with $(X_t)_{t\geqslant 0}$ the CIR process defined in (53) and that for every $q \leqslant h+1$, there exists $C_q > 0$ and $t_q^\zeta \in [0,1]$, such that for every $t \in [0,t_q^\zeta], \mathbb{E}[|\zeta_t|^q] \leqslant C_q t^q$. Then the scheme with transition probability $\mathbb{P}(\zeta_{T/n} \in dx)$ satisfies $E'_{n,\mathrm{pol}}(h,h+1)$ (see (11)).*

Proof. Let us write the Taylor expansion of f at order h:

$$f(\zeta_t) = f(0) + \sum_{l=1}^h \frac{\zeta_t^l}{l!}f^{(l)}(0) + \int_0^{\zeta_t} \frac{(\zeta_t - y)^h}{h!}f^{(h+1)}(y)dy.$$

Since $(\zeta_t)_{t\geqslant 0}$ matches the first h moments of the CIR we have :

$$\mathbb{E}[f(X_t) - f(\zeta_t)] = \mathbb{E}\left[\int_0^{X_t} \frac{(X_t - y)^h}{h!}f^{(h+1)}(y)dy - \int_0^{\zeta_t} \frac{(\zeta_t - y)^h}{h!}f^{(h+1)}(y)dy\right]$$

Moreover $f \in \mathcal{C}_b^\infty(\mathbb{R}_+)$, so using the Lemma 4.3 it follows that

$$\mathbb{E}[f(X_t) - f(\zeta_t)] \leqslant C\|f\|_{h+1,\infty}\mathbb{E}\Big[\Big|\frac{X_t^{h+1}}{h!} + \frac{\zeta_t^{h+1}}{h!}\Big|\Big] \leqslant C\|f\|_{h+1,\infty}t^{h+1}$$

and the proof is completed.

□

The proof of (60) is a direct application of this theorem for $h = 2$ and ζ is replaced by X^n defined by (58). □

4.3.2 Convergence for measurable test function

Using the results from the previous sections (in particular Theorem 3.1), we are now able to study the total variation convergence of the CIR scheme with almost order 2. We prove that convergence for bounded and measurable functions but with support $\mathcal{D}_{cir}$ strictly contained in $\mathbb{R}_+$. We now introduce this support. Let $d_2 > d_1 > 0$ and define

$$\mathcal{D}_{cir} = [d_1, d_2]$$

From Theorem 3.1 we deduce the following total variation distance estimation for the CIR process.

Theorem 4.4. *Let $T > 0$. Let $\delta \in (0, T)$ and let $n \in \mathbb{N}^*$ such that $2T/n \leqslant \delta$. Let $(X_t)_{t\geqslant 0}$ be the process defined by (53) and let $(X^n_t)_{t\in\pi_{T,n}}$ be the process defined by (55) if $4a \geqslant \sigma^2$ and (58) otherwise. Let $\tilde{\psi}_{cir}$ the function defined as in (54) with $V_{0,cir}$ and $V_{1,cir}$ replaced by $V_{0,cir}^{\mathcal{D}_{cir},v}$ and $V_{1,cir}^{\mathcal{D}_{cir},v}$. Assume that $v \leqslant (d_2 - d_1)/4$, $(K(T/n, \sqrt{3T/n}) - 2v)_+ < d_1$ and that assumptions (34) and (36) hold with $s - t = \delta/2$ and $\tilde{\psi}$ replaced by $\tilde{\psi}_{cir}$ that respectively read in this case*

$$\frac{3\|\tilde{\psi}_{cir}\|_{1,3,\infty}}{n^{1/4}} + \frac{M_8(Z)}{n} + \exp(-m_*^2 n\delta/(4T)) \leqslant \frac{1}{2} \quad \text{and} \quad n^{1/2} \geqslant 24\frac{\|\tilde{\psi}_{cir}\|^2_{1,3,\infty}}{\lambda_{*,cir}}.$$

with $\lambda_{*,cir} := \inf_{\mathcal{D}} V_{1,cir}^{\mathcal{D}_{cir},v}(x) \geqslant \sigma d_1$. Then, there exists $\mathfrak{l}_* \in \mathbb{N}^*$, $C, \beta \geqslant 1$ such that we have for every bounded and measurable function $f : \mathbb{R}^d \to \mathbb{R}$, with $\mathrm{supp}(f) \subset \mathcal{D}^{2v} = [d_1 + 2v, d_2 - 2v]$,

$$|\mathbb{E}[f(X_T(x)) - f(X^n_T(x))]| \leqslant 8\Big(\exp\Big(-\frac{v^2/2 - \delta^2 a^2 \vee (kd_2)^2}{\delta\sigma^2 d_2}\Big) + \exp\Big(-\frac{v^2}{\delta 12^2 \|\psi_{cir}\mathbb{1}_{\mathcal{D}^v_{cir}}\|^2_{1,1,\infty}}\Big)\Big)\|f\|_\infty$$
$$+ C(1 + |x|^\beta)\frac{\mathfrak{K}_9(\tilde{\psi}_{cir})^{\mathfrak{l}_*}}{(\lambda_{*,cir}\delta)^{42}}\|f\|_\infty/n^2$$

with $\mathfrak{K}_r(\tilde{\psi}_{cir})$, $r \in \mathbb{N}^*$, defined in (33).

Now, we give a structural result in order to obtain convergence for the total variation distance for the CIR process with almost order 2.

Corollary 4.1. *Let $T > 0$, $n \in \mathbb{N}^*$, $v \in (0, 1]$, $0 < d_1 \leqslant d_2$ and $\delta > 0$. We assume that the hypotheses from Theorem 4.4 are fulfilled with those parameters and that there exists a sequence $(\rho_n)_{n\in\mathbb{N}}$ taking strictly positive values such that,*

$$\delta \in [\underline{\delta}(n, v, d_1, d_2), \overline{\delta}(n, v, d_1, d_2)] \tag{61}$$

with

$$\underline{\delta}(n, v, d_1, d_2) = \frac{\mathfrak{K}_9(\tilde{\psi}_{cir})^{\mathfrak{l}_*/42}}{\lambda_{*,cir}\rho_n^{1/42}}$$

and

$$\overline{\delta}(n,v,d_1,d_2)=\frac{v^2}{4\sigma^2 d_2\ln(n)}\wedge\frac{\ln(\rho_n)\sigma^2 d_2}{(a\vee kd_2)^2}\wedge\frac{v^2}{12^2\|\psi_{cir}\mathbb{1}_{\mathcal{D}_{cir}^v}\|_{1,1,\infty}^2\ln(n)}$$

Then, there exists $C,\beta\geqslant 0$*, such that for every bounded and measurable function* f *with* $\mathrm{supp}(f)\subset\mathcal{D}_{cir}^{2v}=[d_1+2v,d_2-2v]$*,*

$$|\mathbb{E}[f(X_T(x))-f(X_T^n(x))]|\leqslant C(1+|x|^\beta)\|f\|_\infty\rho_n/n^2.$$

The reader may first notice that for $\rho_n=\ln(n)^\zeta$, $\zeta>42$, we have $\underline{\delta}(n,v,d_1,d_2)\leqslant\overline{\delta}(n,v,d_1,d_2)$ for n large enough and we can find δ which satisfies the hypotheses from Theorem 4.4 and (61) and for every bounded and measurable test function f with $\mathrm{supp}(f)\subset[d_1+2v,d_2-2v]$,

$$|\mathbb{E}[f(X_T(x))-f(X_T^n(x))]|\leqslant C(1+|x|^\beta)\|f\|_\infty\ln(n)^\zeta/n^2. \tag{62}$$

Moreover, through the sequence $(\rho_n)_{n\in\mathbb{N}}$, this result shows that we can consider asymptotic cases that are $v\to 0$, $d_1\to 0$ or $d_2\to\infty$. For instance, it is possible to do it expressing those parameters as functions of n. Using the definition of $\underline{\delta}$ and $\overline{\delta}$, we can identify the rates of convergence of $v(n)\to 0$, $d_1(n)\to 0$ or $d_2(n)\to\infty$, with respect to n, such that (61) and the hypotheses of Theorem 4.4 hold and then obtain similar results as (62).

In particular, let $\epsilon>0$ and define $\rho_n=n^\epsilon$. We fix d_1 and v and choose $d_2(n)$ such that $\underline{\delta}(n,v,d_1,d_2(n))=o_{n\to\infty}\overline{\delta}(n,v,d_1,d_2(n))$ and that we can find $\delta>0$ which satisfies the hypotheses from Theorem 4.4 and (61). Moreover, since $(X_t)_{t\geqslant 0}$ is a CIR process, we have $\mathbb{P}(X_t\geqslant d_2(n)-v)\leqslant\mathbb{E}[\exp(\lambda X_t)]\exp(-\lambda(d_2(n)-v))$, with, for every $t\geqslant 0$, $\mathbb{E}[\exp(\lambda X_t(x))]\leqslant C\exp(Cx)<+\infty$ for every $\lambda\leqslant 2(1-\exp(-kt))/(k\sigma^2)$. Now assume that there exists $\epsilon\in(0,2)$ and such a $\lambda:=\lambda_{v,\epsilon}$, such that $d_2(n)$ satisfies $\exp(-\lambda_{v,\epsilon}(d_2(n)-v))\leqslant C/n^{2-\epsilon}$. Then, there exists $n_0\in\mathbb{N}$, $C,\beta\geqslant 0$, such that for every bounded and measurable function f with $\mathrm{supp}(f)\subset[d_1+2v,\infty)$, and $n\geqslant n_0$, we obtain

$$|\mathbb{E}[f(X_T(x))-Y_T^n(x)]|\leqslant C(1+|x|^\beta)\|f\|_\infty/n^{2-\epsilon}+C\exp(\beta x)\|f\|_\infty/n^{2-\epsilon} \tag{63}$$

with $Y_T^n(x)=f(X_T^n(x))\mathbb{1}_{X_T^n(x)\in[d_1+2v,d_2(n)-2v]}$. It is much more difficult to obtain this type of result for test functions with support contained in $(0,\infty)$ since we do not have such estimates in the neighborhood of zero.

References

[1] A. Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.*, 11(4):355–384, 2005.

[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] L. Andersen. Simple and efficient simulation of the heston stochastic volatility model. *Journal of Comp. Finance*, 11(3), 2008.

[4] V. Bally and C. Rey. Approximation of Markov semigroup in total variation disctance. January 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] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density. *Monte Carlo Methods Appl.*, 2(2):93–128, 1996.

[8] S. G. Bobkov, G. P. Chistyakov, and F. Götze. Berry–Esseen bounds in the entropic central limit theorem. *Probab. Theory Related Fields*, 159(3-4):435–478, 2014.

[9] S. G. Bobkov, G. P. Chistyakov, and F. Götze. Fisher information and the central limit theorem. *Probab. Theory Related Fields*, 159(1-2):1–59, 2014.

[10] R. Bompis and E. Gobet. Analytical approximations of local-Heston volatility model and error analysis. working paper or preprint, March 2015.

[11] M. Bossy, E. Gobet, and D. Talay. A symmetrized Euler scheme for an efficient approximation of reflected diffusions. *J. Appl. Probab.*, 41(3):877–889, 2004.

[12] M. Broadie and Ö. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54(2):217–231, 2006.

[13] J. C. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[14] E. Gobet. Weak approximation of killed diffusion using Euler schemes. *Stochastic Process. Appl.*, 87(2):167–197, 2000.

[15] E. Gobet and St. Menozzi. Stopped diffusion processes: boundary corrections and overshoot. *Stochastic Process. Appl.*, 120(2):130–162, 2010.

[16] J. Guyon. Euler scheme and tempered distributions. *Stochastic Process. Appl.*, 116(6):877–904, 2006.

[17] J. Jacod, T. G. Kurtz, S. Méléard, and P Protter. The approximate Euler method for Lévy driven stochastic differential equations. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(3):523–558, 2005.

[18] B. Jourdain and A. Kohatsu-Higa. *A Review of Recent Results on Approximation of Solutions of Stochastic Differential Equations*, volume 65 of *Progress in Probability*. Springer Basel, 2011.

[19] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.

[20] A. Kohatsu-Higa and P. Tankov. Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Process. Appl.*, 120(11):2258–2285, 2010.

[21] V. Konakov and S. Menozzi. Weak error for stable driven stochastic differential equations: expansion of the densities. *J. Theoret. Probab.*, 24(2):454–478, 2011.

[22] V. Konakov, S. Menozzi, and S. Molchanov. Explicit parametrix and local limit theorems for some degenerate diffusion processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(4):908—923, 2010.

[23] S. Kusuoka. Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus. In *Advances in mathematical economics. Vol. 6*, volume 6 of *Adv. Math. Econ.*, pages 69–83. Springer, Tokyo, 2004.

[24] S. Kusuoka. Gaussian K-scheme: justification for KLNV method. In *Advances in mathematical economics. Vol. 17*, volume 17 of *Adv. Math. Econ.*, pages 71–120. Springer, Tokyo, 2013.

[25] E. Löcherbach and D. Loukianova. On Nummelin splitting for continuous time Harris recurrent Markov processes and application to kernel estimation for multi-dimensional diffusions. *Stochastic Process. Appl.*, 118(8):1301–1321, 2008.

[26] T. Lyons and N. Victoir. Cubature on Wiener space. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):169–198, 2004. Stochastic analysis with applications to mathematical finance.

[27] G.N. Milstein. Weak approximation of solutions of systems of stochastic differential equations. In *Numerical Integration of Stochastic Differential Equations*, volume 313 of *Mathematics and Its Applications*, pages 101–134. Springer Netherlands, 1995.

[28] 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.

[29] I. Nourdin and G. Poly. An invariance principle under the total variation distance. 15 pages, October 2013.

[30] E. Nummelin. A splitting technique for Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete*, 43(4):309–318, 1978.

[31] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Applied Mathematical Sciences. Springer New York, 1992.

[32] P. Protter and D. Talay. The Euler scheme for Lévy driven stochastic differential equations. *Ann. Probab.*, 25(1):393–423, 1997.

[33] C. Rey. Convergence in total variation distance for a third order scheme for one dimensional diffusion process. working paper or preprint, February 2016.

[34] 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.

[35] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177 – 188, 1977.

[36] A. Yu. Zaĭtsev. Approximation of convolutions of probability distributions by infinitely divisible laws under weakened moment constraints. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 194(Problemy Teorii Veroyatnost. Raspred. 12):79–90, 177–178, 1992.