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

Clément Rey

# Total variation convergence for numerical schemes for diffusions with irregular coefficients: An application to the CIR process

Clément Rey [1]

## Abstract

In this paper, we propose a method to prove the total variation convergence for numerical schemes for Stochastic Differential Equation (SDE) with irregular coefficient. In particular, we will consider SDE with locally smooth coefficients. In a first part, we present this method and in a second time, we apply it to the CIR process. We will consider the weak second order scheme introduced in [2] and we will prove that this scheme also converges towards the diffusion for the total variation distance. This convergence will take place with almost order two.

## 1 Introduction

In this paper we study the total variation distance between a diffusion process with irregular coefficients and some numerical schemes. In order to do it, we will use a result from [4]. Then, we apply this result to a second weak order scheme for CIR process based on a cubature method and introduced in [2]. This scheme has second weak order for smooth test function and we will prove that the convergence for a class of bounded measurable functions takes place 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 $V_i \in \mathcal{C}^\infty(\mathbb{R}^d;\mathbb{R}^d) \cap \mathcal{C}_b^\infty(\mathcal{D};\mathbb{R}^d)$, with $\mathcal{D}$ a subset of $\mathbb{R}^d$, $(W_t)_{t\geqslant 0}$ a standard Brownian motion and $\circ dW_t$ the Stratonovich integral with respect to $W_t$. 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 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 $\sup_{k\in\mathbb{N}^*} \delta_k^n \leqslant C/n$. For now,

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

we do not discuss the regularity of $\psi_k$. Our aim is to study the convergence of the law of $X^n$ to the law of a Markov process $X$. More precisely, we will give estimates of the weak error

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

In order to obtain total variation convergence for $(\varepsilon_n)_{n\in\mathbb{N}^*}$, one has 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 [4] 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 will also consider a Markov process in continuous time $(X_t)_{t\geqslant 0}$ with semigroup $(P_t)_{t\geqslant 0}$ and we define $\mu_{k+1}^n(x,dy) = \mathbb{P}(X_{t_{k+1}^n} \in dy | X_{t_k^n} = x)$.

Moreover, for $f \in \mathcal{C}^\infty(\mathbb{R}^d)$ and for a multi-index $\alpha = (\alpha_1, \cdots, \alpha_d) \in \mathbb{N}^d$ we denote $|\alpha| = \alpha_1 + ... + \alpha_d$ and $\partial_\alpha f = (\partial_1)^{\alpha_1} \dots (\partial_d)^{\alpha_d} f = \partial_x^\alpha f(x) = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} f(x)$. We include the multi-index $\alpha = (0, ..., 0)$ and in this case $\partial_\alpha f = f$. We will use 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 will denote $\mathcal{C}_b^q(\mathbb{R}^d) = \{f \in \mathcal{C}^q(\mathbb{R}^d), \|f\|_{q,\infty} < \infty\}$ and $C_c^q(\mathbb{R}^d) \subset C^q(\mathbb{R}^d)$ 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 $e_q \in \mathbb{N}$ if there exists $C \geqslant 1$ such that

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

and we denote by $\mathcal{C}_{\text{pol}}^q(\mathbb{R}^d)$ the set of function satisfying (3).
A first standard result is the following: let us assume that there exists $h > 0$, $q \in \mathbb{N}$ such that for every $f \in \mathcal{C}^q(\mathbb{R}^d)$, $k \in \mathbb{N}^*$ and $x \in \mathbb{R}^d$,

$$\left|\mu_k^n f(x) - \nu_k^n f(x)\right| = \left|\int f(y)\mu_k^n(x,dy) - \int f(y)\nu_k^n(x,dy)\right| \leqslant C\|f\|_{q,\infty}/n^{1+h}. \tag{4}$$

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

$$\|P_t f - Q_t^n f\|_\infty = \sup_{x\in\mathbb{R}^d} |\mathbb{E}[f(X_t^n)|X_0^n = x] - \mathbb{E}[f(X_t)|X_0 = x]| \leqslant C\|f\|_{q,\infty}/n^h. \tag{5}$$

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 $h = 1$, has initially been proved in the seminal papers of Milstein [25] and of Talay and Tubaro [31] (see also [17]). Similar results were obtained in various situations: diffusion processes with jumps (see [29], [15]) or diffusion processes with boundary conditions (see [12], [10], [13]). An overview of this subject is given in [16]. More recently, approximation schemes of higher orders

(*e.g.*, $h = 2$), based on cubature methods, have been introduced and studied by Kusuoka [21], Lyons [24], Ninomiya, Victoir [26] or Alfonsi [2]. The reader may also refer to the work of Kohatsu-Higa and Tankov [18] for a higher weak order scheme for jump processes. Despite the fact that most of these results concern diffusions with regular coefficients, some authors tackle some more exotic cases. For instance, in [2], Alfonsi has studied the weak error for diffusions 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 (5) with $\|f\|_{q,\infty}$ replaced by $\|f\|_\infty$ when $f$ is a bounded and measurable function. In the case of the Euler scheme for diffusion processes, a first result of this type has been obtained by Bally and Talay [5], [6] using the Malliavin calculus (see also Guyon [14]). Afterwards Konakov, Menozzi and Molchanov [19], [20] obtained similar results using a parametrix method. Later,Kusuoka [22] obtained estimates of the error in total variation distance for the Victoir Ninomiya scheme (which corresponds to the case $h = 2$). More recently, in [4], a generic result ensures the total variation distance convergence as soon as we consider smooth schemes and the random variables $Z_k, k \in \mathbb{N}^*$ satisfy the Doeblin condition. In [30], 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, have treated the case of diffusion processes with irregular coefficients. In particular, in [4], the regularity in space for the functions $\psi_k$, which often rely on the regularity of the functions $V_i$, is essential to prove the total variation convergence result. The main difference here, is that we will consider that these functions are smooth only on a subset $\mathcal{D}$ of $\mathbb{R}^d$. However, we will be able to exploit this local regularity in order to prove the convergence for every bounded and measurable test function with support strictly contained in $\mathcal{D}$. In order to do it, we will mix some results of convergence for smooth test functions for irregular diffusions on $\mathbb{R}^d$ with the total variation convergence proven in [4] for a modification of $X$ with coefficients localized on $\mathcal{D}$. The main result concerning this approach is given in Theorem 3.1 and provides a estimation of the weak error for simply bounded measurable test function with support contained in $\mathcal{D}$. In this paper we will sometimes take the liberty to say that this is the convergence for the total variantion distance.[2] It is important to notice that this result do not concern specific diffusions or scheme but treats general case.

Using this approach, we will focus on the CIR processes. In this case, the diffusion coefficient is $V_1(x) = \sigma\sqrt{x}$, $\sigma \in \mathbb{R}_+$, and is singular in zero. That is why, standard estimation methods do not apply straightly. However, some authors manage to develop numerical analysis of this scheme using among other the close link that exists with Bessel processes. The reader may refer to [11], [3], [1], [2] or [9] for a non exhaustive list of study concerning numerical approximation for the CIR processes. In particular, in [2], the author proves the weak convergence with order 2 for smooth test function (polynomial growth for the test function and its derivaties), of a scheme based on cubature method (and also inspired by [3]). In [9], the authors propose an expansion of the weak error for Lipschitz test functions. However, until now, there is no study

[2]The total variation distance oncerns 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 straightfully from our result if we suppose that $\mathbb{P}(X_T \notin \mathcal{D})$ and $\mathbb{P}(X^n_T \notin \mathcal{D})$ are small enough

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 [4] can not apply directly. Despite this singularity, we will use and extend the result from [2] and [4] , and prove that (5) is also satisfied for the scheme introduced in [2] under some hypothesi on the support of the test function. More specifically we will obtain the following results:
First (see (104)), there exist $C, \beta \geqslant 0$ such that for every bounded and measurable test function $f$ with $\text{supp}(f) \subset [d_1, d_2]$, $0 < d_1 \leqslant d_2 < \infty$ we obtain, for $n$ large enough,

$$\mathbb{E}[f(X_T(x)) - f(X_T^n(x))] \leqslant C(1 + |x|^\beta)\|f\|_\infty \ln(n)^\zeta / n^2, \tag{6}$$

with $(X^n_{kT/n}(x))_{k\in\mathbb{N}}$ the scheme introduced in [2].
Moreover (see (106)), using the finitness of some exponential moments for the CIR processes, we also obtain the following result: There exists $\epsilon \in (0, 2)$ such that for every bounded and measurable test function $f$ with $\text{supp}(f) \subset [d_1, \infty)$, $d_1 > 0$, then, for $n$ large enough, we have

$$\mathbb{E}[f(X_T(x)) - Y_T^n(x))] \leqslant C \exp(\beta|x|)\|f\|_\infty / n^{2-\epsilon}, \tag{7}$$

with $Y_T^n(x)$ a $\sigma(X_T^n(x))$-measurable function defined in (106).

We will begin presenting the framework of this paper. In Section 3., we will give some convergence results for smooth test functions and for bounded measurable test functions. In the end of this Section, we give the main result of this paper concerning total variation convergence for schemes with singular coefficients (see Theorem 3.1). The paper ends with a theoretical application in order to obtain a total variation convergence result for the scheme presented in [2] for the CIR process.

# 2 The distance between two Markov semigroups

Throughout this section the following notations will prevail. We fix $T > 0$ and we denote $n \in \mathbb{N}^*$, the number of time step 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 will 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 step but, for sake of simplicity, we will not consider this case. First, we state some results for smooth test functions.

## 2.1 Convergence of semigroups

### 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 $\sigma$ field and for each bounded 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

$$\forall x \in \mathbb{R}^d, \quad |\mu_k^n(x)| := \sup_{\|f\|_\infty \leqslant 1} \Big| \int_{\mathbb{R}^d} f(y)\mu_k^n(x,dy)\Big|, \tag{8}$$

and

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

and, we assume that all the sequences of measures we consider in this paper satisfy the following property:

$$\sup_{k\in\mathbb{N}^*} |\mu_k^n| < \infty. \tag{9}$$

Although the main application concerns the case where $\mu_k^n(x,dy)$ is a probability measure, we do not assume this here: 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

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

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 will consider the following hypothesis: let $q\in\mathbb{N}$ and $t\leqslant s\in\pi_{T,n}$. If $f\in\mathcal{C}^q(\mathbb{R}^d)$ then $P_{t,s}f\in\mathcal{C}^q(\mathbb{R}^d)$ and

$$\sup_{t,s\in\pi_{T,n};t\leqslant s} \|P_{t,s}^n f\|_{q,\infty} \leqslant C\|f\|_{q,\infty}. \tag{10}$$

Notice that (9) implies that (10) holds for $q=0$. We will also consider the following hypothesis: let $q\in\mathbb{N}$ and $t\leqslant s\in\pi_{T,n}$. If $f\in\mathcal{C}^q_{\mathrm{pol}}(\mathbb{R}^d)$ then $P_{t,s}f\in\mathcal{C}^q_{\mathrm{pol}}(\mathbb{R}^d)$ and there exists $C>1,\beta\in\mathbb{N}$ such that

$$\forall x\in\mathbb{R}^d, \quad \sup_{t,s\in\pi_{T,n};t\leqslant s} \sum_{0\leqslant|\alpha|\leqslant q} |\partial_\alpha P_{t,s}^n f(x)| \leqslant C(1+|x|^\beta). \tag{11}$$

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

$$\forall x\in\mathbb{R}^d, \quad \sup_{t,s\in\pi_{T,n};t\leqslant s} \sum_{0\leqslant|\alpha|\leqslant q} |\partial_\alpha P_{t,s}^n f(x)| \leqslant C(1+|x|^\beta)\|f\|_{q,\infty}. \tag{12}$$

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. Our aim is 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 $q \in \mathbb{N}$, $h \geqslant 0$ be fixed. We introduce a short time error approximation assumption: there exists a constant $C > 0$ (depending on $q$ only) such that for every $k \in \mathbb{N}^*$, we have

$$E_n(h, q) \qquad \|\Delta_k^n f\|_\infty \leqslant C\|f\|_{q,\infty}/n^{h+1}. \tag{13}$$

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

$$E_{n,\mathrm{pol}}(h, q), \qquad \forall x \in \mathbb{R}^d, \quad |\Delta_k^n f(x)| \leqslant C(1 + |x|^\beta)/n^{h+1}. \tag{14}$$

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

$$E'_{n,\mathrm{pol}}(h, q), \qquad \forall x \in \mathbb{R}^d, \quad |\Delta_k^n f(x)| \leqslant C(1 + |x|^\beta)\|f\|_{q,\infty}/n^{h+1}. \tag{15}$$

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

***A.*** *Suppose that $\nu^n$ satisfies (10) for this $q$, $\mu^n$ satisfy (9) and that we have $E_n(h, q)$ (see (13)). Then for every $f \in \mathcal{C}^q(\mathbb{R}^d)$,*

$$\sup_{t\in\pi^T_{T,n}} \|P_t^n f - Q_t^n f\|_\infty \leqslant C\|f\|_{q,\infty}/n^h. \tag{16}$$

***B.*** *Suppose that $\mu^n$ and $\nu^n$ satisfy respectively (11) for $q = 0$ and for this $q$, and that $E_{n,\mathrm{pol}}(h, q)$ (see (14)) holds. Then for every $f \in \mathcal{C}^q_{\mathrm{pol}}(\mathbb{R}^d)$, there exists $C > 1, \beta \in \mathbb{N}$ such that*

$$\sup_{t\in\pi^T_{T,n}} \|P_t^n f - Q_t^n f\|_\infty \leqslant C(1 + |x|^\beta)/n^h. \tag{17}$$

***C.*** *Suppose that $\mu_k^n$ and $\nu_k^n$ satisfy respectively (11) for $q = 0$ and (12) for this $q$ and if $E'_{n,\mathrm{pol}}(h, q)$ (see (15)) holds, there exists $C > 1, \beta \in \mathbb{N}$ such that for all $f \in \mathcal{C}^q_b(\mathbb{R}^d)$, then*

$$\sup_{t\in\pi^T_{T,n}} \|P_t^n f - Q_t^n f\|_\infty \leqslant C(1 + |x|^\beta)\|f\|_{q,\infty}/n^h. \tag{18}$$

*Proof.* Let $m \in \mathbb{N}^*$, $m \leqslant n$. We have

$$\begin{aligned} |P^n_{t^n_m} f - Q^n_{t^n_m} f|_\infty &\leqslant \sum_{k=0}^{m-1} |P^n_{t^n_k} P^n_{t^n_k, t^n_{k+1}} Q^n_{t^n_{k+1}, t^n_m} f(x) - P^n_{t^n_k} Q^n_{t^n_k, t^n_{k+1}} Q^n_{t^n_{k+1}, t^n_m} f(x)|_\infty \\ &= \sum_{k=0}^{m-1} |P^n_{t^n_k} \Delta^n_{k+1} Q^n_{t^n_{k+1}, t^n_m} f(x)|. \end{aligned} \tag{19}$$

Now, since $f \in \mathcal{C}^q_{\rm pol}(\mathbb{R}^d)$ then, using (11) 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 (14) and the linearity of the semigroup $P^n$ together with (11) for $q = 0$, 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 $k = 0, ..., m-1$, (17) follows. In order to prove (18), we use (14) 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 $\beta$ do not depend on $f$. Using once again the linearity of the semigroup $P^n$ and (11) for $q = 0$, it follows $|P^n_{t^n_k} \Delta_{k+1} Q^n_{t^n_{k+1},t^n_m} f(x)| \leqslant C(1 + |x|^\beta)\|Q^n_{t^n_{k+1},t^n_m} f\|_{q,\infty}/n^{h+1}$. Then, property (12) for $Q^n$ gives (18). The proof of (16) is similar but simplier so we leave it out.

□

### 2.1.2 Measurable test functions (convergence in total variation distance)

The estimates (16), (17) and (18) requires a lot of regularity for the test function $f$. We aim to show that, if the semigroups at work have a regularization property, then we may obtain estimates of the error for measurable and bounded test functions. In order to state this result we have to give some hypothesis on the adjoint semigroup. Let $q \in \mathbb{N}$. We assume that there exists a constant $C \geqslant 1$ such that for every measurable and bounded function $f$ and any $g \in \mathcal{C}^q(\mathbb{R}^d)$

$$E^*_n(h, q) \qquad |\langle g, \Delta^n_k f\rangle| \leqslant C\|g\|_{q,1}\|f\|_\infty/n^{1+h}. \tag{20}$$

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 $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ an increasing function be given. We assume that there exists a constant $C \geqslant 1$ such that

$$R_{q,\eta}(S) \qquad \forall t, s \in \pi_{T,n}, \text{ with } S \leqslant s - t, \quad \|P^n_{t,s} f\|_{q,\infty} \leqslant \frac{C}{S^{\eta(q)}}\|f\|_\infty. \tag{21}$$

We also consider the "adjoint regularization hypothesis". We assume that there exists an adjoint semigroup $P^{n,*}_{t,s}$, that is

$$\left\langle P^{n,*}_{t,s} g, f\right\rangle = \left\langle g, P^n_{t,s} f\right\rangle$$

for every measurable and bounded function $f$ and every function $g \in \mathcal{C}^\infty_c(\mathbb{R}^d)$. We assume that $P^{n,*}_{t,s}$ satisfies

$$R^*_{q,\eta}(S) \qquad \forall t, s \in \pi_{T,n}, \text{ with } S \leqslant s - t, \quad \|P^{n,*}_{t,s} f\|_{q,1} \leqslant \frac{C}{S^{\eta(q)}}\|f\|_1. \tag{22}$$

Notice that a sufficient condition in order that $R^*_{q,\eta}(S)$ holds is the following: for every multi index $\alpha$ with $|\alpha| \leqslant q$

$$\forall t, s \in \pi_{T,n}, \text{ with } S \leqslant s - t, \quad \|P^n_{t,s} \partial_\alpha f\|_\infty \leqslant \frac{C}{S^{\eta(q)}}\|f\|_\infty. \tag{23}$$

Indeed:

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

Now we can state our first result for total variation convergence between $P^n$ and $Q^n$. Those results will be valid as soon as both $P^n$ and $Q^n$ satisfy (9).

**Proposition 2.2.** *Let $q \in \mathbb{N}$, $h \geqslant 0$, $S \in [T/n, T/2)$ and $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ an increasing function be fixed. We assume that $E_n(h,q)$ (see (13)) and $E^*_n(h,q)$ (see (20)) hold for $P^n$ and $Q^n$. We also suppose that $P^n$ satisfies $R_{q,\eta}(S)$ (see (21)) and $Q^n$ satisfies $R^*_{q,\eta}(S)$ (see (22)) and that (10) hold with $q = 0$ for both of them. Then,*

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

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 $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ an increasing function be fixed. We assume that $E_n(h,q)$ (see (13)) and $E^*_n(h,q)$ (see (20)) hold for $P^n$ and $Q^n$. Moreover, we assume that there exists some kernels $(\overline{P}^n_{t,s})_{t,s \in \pi_{T,n}; t \leqslant s}$ which satisfies $R_{q,\eta}(S)$ (see(21)) and $(\overline{Q}^n_{t,s})_{t,s \in \pi_{T,n}; t \leqslant s}$ which satisfies $R^*_{q,\eta}(S)$ (see (22)) and that (10) hold with $q = 0$ for both of them. We also assume that for every $t, s \in \pi_{T,n}$ with $s - t \geqslant S$,*

$$\|Q^n_{t,s} f - \overline{Q}^n_{t,s} f\|_\infty + \|P^n_{t,s} f - \overline{P}^n_{t,s} f\|_\infty \leqslant C S^{-\eta(q)} \|f\|_\infty / n^{h+1}. \tag{24}$$

*Then,*

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

**Remark 2.1.** *Notice that $\overline{P}^n$ and $\overline{Q}^n$ are not supposed to satisfy the semigroup property and are not directly related to $\mu^n$ and $\nu^n$.*

The proof of those results can be found in [4] and follows similar ideas from the one of Proposition 2.1.

### 2.1.3 Mixing regularity properties

In this section, we will consider semigroups with mixing regularity properties. We will study two semigroups $P^n$ and $Q^n$ which satisfy the hypothesis for total variation convergence only closely to the date $T$ and we will show that we the convergence for bounded test functions holds.

**Theorem 2.1.** *Let $P^n$ and $Q^n$ two semigroups with transition measures $\mu^n$ and $\nu^n$. Let $q \in \mathbb{N}$, $h, \eta > 0$, $2T/n \leqslant \delta \leqslant T$ and define*

$$\tilde{\delta}_n = \inf\{t; t > \delta, T - t \in \pi_{T,n}\}. \tag{25}$$

*We suppose that, on the interval $[0, T - \tilde{\delta}_n]$, $\mu^n$ and $\nu^n$ satisfy respectively (11) for $q = 0$ and (12) for this $q$ and that $E'_{n,\mathrm{pol}}(h, q)$ (see (15)) holds. Then we have the following properties;*

***A.*** *On the interval $[T - \tilde{\delta}_n, T]$, we assume that both $\mu^n$ and $\nu^n$ satisfy (9), we assume that $E_n(h, q)$ (see (13)) and $E^*_n(h, q)$ (see (20)) hold for $P^n$ and $Q^n$. We also suppose that there exists some kernels $\overline{Q}^n$ which satisfies $R_{q,\eta}(\delta/2)$ (see (21)) and $\overline{P}^n$ which satisfies $R^*_{q,\eta}(\delta/2)$ (see (22)) and such that (24) holds on the interval $[T - \tilde{\delta}_n, T]$.*
*Then there exists $C \geqslant 1$, $\beta \in \mathbb{N}$, such that for all measurable test function $f$ on $\mathbb{R}^d$, we have*

$$\forall x \in \mathbb{R}^d, \quad |P^n_T f(x) - Q^n_T f(x)| \leqslant C \frac{1 + |x|^\beta}{\delta^{\eta(q)}} \|f\|_\infty / n^h. \tag{26}$$

***B.*** *On the interval $[T - \tilde{\delta}_n, T]$, we assume that for every $m \in \mathbb{N}$, $m \geqslant n$, both $\mu^m$ and $\nu^m$ satisfy (9), and that $E_m(h, q)$ (see (13)) and $E^*_m(h, q)$ (see (20)) hold between $(P^m_t)_{t \in \pi_{T,m}} = (P_t)_{t \in \pi_{T,m}}$ and $(Q^m_t)_{t \in \pi_{T,m}}$. We also suppose that there exists a family of kernels $(\overline{Q}^m)_{m \in \mathbb{N}; m \geqslant n}$ which satisfies $R_{q,\eta}(\delta/2)$ (see (21)) and $R^*_{q,\eta}(\delta/2)$ (see (22)) and such that (24) holds on the interval $[T - \tilde{\delta}_n, T]$ between $Q^m$ and $\overline{Q}^m$ for every $m \geqslant n$.*

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

$$\forall x \in \mathbb{R}^d, \quad |P^n_T f(x) - Q^n_T f(x)| \leqslant C \frac{1 + |x|^\beta}{\delta^{\eta(q)}} \|f\|_\infty / n^h. \tag{27}$$

*Proof.* We prove **A.**. For sake of clarity, we suppose that $\overline{P} = P$ and that $\overline{Q} = Q$. The proof is very similar otherwise. We denote $t_\delta = T - \tilde{\delta}_n \in \overline{\pi}^T_{T,n}$.

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

Since $P^n$ and $Q^n$ satisfy (9), (13), (20), with also $R_{q,\eta}(\delta/2)$ (see (21)) for $Q^n$ and $R^*_{q,\eta}(\delta/2)$ (see (22)) for $P^n$ when $T - \tilde{\delta} \leqslant t^n_k \leqslant t^n_m \leqslant T$, using Property 2.2, we have: $|P^n_{0,t_\delta}(P^n_{t_\delta,T} - P^n_{t_\delta,T}) f(x)| \leqslant \|(P^n_{t_\delta,T} - Q^n_{t_\delta,T}) f\|_\infty \leqslant C\delta^{-\eta(q)} \|f\|_\infty / n^h$. In order to bound the second term we use Proposition 2.1 together with $\|Q^n_{t_\delta,T} f\|_{q,\infty} \leqslant C\delta^{-\eta(q)} \|f\|_\infty$ 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 $\|(Q^n_{t_\delta,T} - P^n_{t_\delta,T}) f\|_\infty$. we introduce

the sequence of discrete semigroups $((Q_t^{n,m})_{t\in\pi_{T,n};t\geqslant t_\delta})_{m\in\mathbb{N}^*}$ defined in the following way: For all $t\in\pi_{T,n}, t\geqslant t_\delta$ we have $Q_t^{n,m}f(x)=Q_t^{nm}f(x)$. Let $m'\geqslant m$, then

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

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 (10), we use the Lindeberg decomposition (19) in order to obtain: $\|Q^{n,m}_{t_k^n,t_{k+1}^n}f-Q^{n,m'}_{t_k^n,t_{k+1}^n}f\|_\infty\leqslant C\|f\|_{q,\infty}/(n^{h+1}m^h)$. In the same way we obtain $|\langle g,Q^{n,m}_{t_k^n,t_{k+1}^n}f-Q^{n,m'}_{t_k^n,t_{k+1}^n}f\rangle|\leqslant C\|g\|_{1,q}\|f\|_\infty/(n^{h+1}m^h)$. Now, since both $Q^{nm}$ and $Q^{nm'}$ have modifications which satisfy both $R_{q,\eta}(\delta/2)$ (see (21)) and $R^*_{q,\eta}(\delta/2)$ (see (22), we can show that: $\forall t\in\pi_{T,n}^{t_\delta,T}$, $\|Q_t^{n,m}f-Q_t^{n,m'}f\|_\infty\leqslant C\delta^{-\eta(q)}\|f\|_\infty/(n^hm^h)$. The sequence $((Q_t^{n,m})_{t\in\pi_{T,n};t\geqslant t_\delta})_{m\in\mathbb{N}^*}$ is thus Cauchy and it converges toward $(P_t^n)_{t\in\pi_{T,n}}$ 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

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

where the left hand side of the above inequality is exactly the term that we study and then the proof is completed. □

This result will be very useful in order to prove the convergence for bounded measurable test functions for diffusions with simpy locally smooth coefficients. The method consists in introducing a modification of the underlying process in the neighborhood of $T$ with smooth coefficient in $\mathcal{C}_b^q(\mathbb{R}^d)$. Then, we can use Theorem 2.1 and it remains to control the error committed between the real process and its modification. Moreover, using **B.**, we see that we can focus exclusively on proving regularization properties for the modification of the approximation semigroup $Q$. In particular there is no regularization property to prove on $P$ (or its modification) which is quite useful. Using this observation, we now focus on the regularization property for a modification of $Q$ for a class of Markov chain.

## 2.2 A class of random tools

In this section we consider a sequence of independent random variables $Z_k=(Z_k^1,\cdots,Z_k^N)\in\mathbb{R}^N$, $k\in\{1,\cdots,n\}$ and we denote $Z=(Z_1,...,Z_n)$. The number $n$ is fixed throughout this section (so there is no asymptotic procedure going on; but morally $n$ is large because we are interested in estimating the error as $n\to\infty$). Our aim is to settle an integration by parts formula based on the law of $Z$. The basic assumption is the following: 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 $k\in\{1,\cdots,n\}$

$$L_{z_*}(\varepsilon_*,r_*)\qquad \mathbb{P}(Z_k\in A)\geqslant\varepsilon_*\lambda(A\cap B_{r_*}(z_{*,k})) \tag{28}$$

where $\lambda$ is the Lebesgue measure on $\mathbb{R}^N$. This condition is sometimes called the Doeblin condition. One can also say that the random variables $Z_k, k\in\{1,\ldots,n\}$ are lower bounded

by the Lebesgue measure. We also define

$$M_p(Z) := 1 \vee \sup_{k \leqslant n} \mathbb{E}[|Z_k|^p] \tag{29}$$

and assume that $M_p(Z) < \infty$ for every $p \geqslant 1$.
It is easy to check that (28) holds if and only if there exists some non negative measures $\mu_k$ with total mass $\mu_k(\mathbb{R}^N) < 1$ and a lower semi-continuous function $\varphi \geqslant 0$ such that $\mathbb{P}(Z_k \in dz) = \mu_k(dz) + \varphi(z - z_{*,k})dz$. Notice that the random variables $Z_1, \cdots, Z_n$ are not assumed to be identically distributed. However, the fact that $r_* > 0$ and $\varepsilon_* > 0$ are the same for all $k$ represents a mild substitute of this property. In order to construct $\varphi$ we have to introduce the following function: For $v > 0$, set $\varphi_v : \mathbb{R}^N \to \mathbb{R}$ defined by

$$\varphi_v(z) = \mathbb{1}_{|z| \leqslant v} + \exp\Big(1 - \frac{v^2}{v^2 - (|z| - v)^2}\Big)\mathbb{1}_{v<|z|<2v}. \tag{30}$$

Then $\varphi_v \in \mathcal{C}_b^\infty(\mathbb{R}^N)$, $0 \leqslant \varphi_v \leqslant 1$ and we have the following crucial property: for every $p, k \in \mathbb{N}$ there exists a universal constant $C_{q,p}$ such that for every $x \in \mathbb{R}^N$, $q \in \mathbb{N}$ and $i_1, \cdots, i_q \in \{1, \cdots, N\}$, we have

$$\varphi_v(z)|\frac{\partial^q}{\partial z^{i_1} \cdot \partial z^{i_q}}(\ln \varphi_v)(z)|^p \leqslant \frac{C_{q,p}}{v^{pq}}, \tag{31}$$

with the convention $\ln \varphi_v(z) = 0$ for $|z| \geqslant 2v$. As an immediate consequence of (71), for every non negative function $f : \mathbb{R}^N \to \mathbb{R}_+$

$$\mathbb{E}[f(Z_k)] \geqslant \varepsilon_* \int_{\mathbb{R}^N} \varphi_{r_*/2}(\ z - z_{*,k}\ ) f(z)dz. \tag{32}$$

By a change of variable

$$\mathbb{E}[f(\frac{1}{\sqrt{n}}Z_k)] \geqslant \varepsilon_* \int_{\mathbb{R}^N} n^{N/2} \varphi_{r_*/2}\big(\sqrt{n}(z - \frac{z_{*,k}}{\sqrt{n}})\big) f(z)dz. \tag{33}$$

We denote

$$m_* = \varepsilon_* \int_{\mathbb{R}^N} \varphi_{r_*/2}(z)dz = \varepsilon_* \int_{\mathbb{R}^N} \varphi_{r_*/2}(z - z_{*,k})dz \tag{34}$$

and

$$\phi_n(z) = n^{N/2} \varphi_{r_*/2}(\sqrt{n}z) \tag{35}$$

and we notice that $\int \phi_n(z)dz = m_* \varepsilon_*^{-1}$.
We consider a sequence of independent random variables $\chi_k \in \{0, 1\}$, $U_k, V_k \in \mathbb{R}^N$, $k \in \{1, \cdots, n\}$ with laws given by

$$\begin{aligned}
\mathbb{P}(\chi_k &= 1) = m_*, \qquad \mathbb{P}(\chi_k = 0) = 1 - m_*, \\
\mathbb{P}(U_k &\in dz) = \frac{\varepsilon_*}{m_*} \phi_n(z - \frac{z_{*,k}}{\sqrt{n}})dz, \\
\mathbb{P}(V_k &\in dz) = \frac{1}{1 - m_*}(\mathbb{P}(\frac{1}{\sqrt{n}}Z_k \in dz) - \varepsilon_* \phi_n(z - \frac{z_{*,k}}{\sqrt{n}})dz).
\end{aligned} \tag{36}$$

Notice that (33) guarantees that $\mathbb{P}(V_k \in dz) \geqslant 0$. Then a direct computation shows that

$$\mathbb{P}(\chi_k U_k + (1-\chi_k)V_k \in dz) = \mathbb{P}(\frac{1}{\sqrt{n}} Z_k \in dz). \tag{37}$$

This is the splitting procedure for $\frac{1}{\sqrt{n}} Z_k$. Now on we will work with this representation of the law of $\frac{1}{\sqrt{n}} Z_k$. So, we always take

$$\frac{1}{\sqrt{n}} Z_k = \chi_k U_k + (1-\chi_k)V_k.$$

**Remark 2.2.** *The above splitting procedure has already been widely used in the litterature: in [28] and [23], it is used in order to prove convergence to equilibrium of Markov processes. In [7], [8] and [32], it is used to study the Central Limit Theorem. Besides, in [27], the above splitting method (with $\mathbb{1}_{B_{r_*}(z_{*,k})}$ instead of $\phi_n(z - \frac{z_{*,k}}{\sqrt{n}})$) is used in a framework which is similar to the one in this paper. Last ine [4], this exact framework is used to prov total variation convergence in a regular settings.*

## 2.3 Markov chains

In this section, $n \in \mathbb{N}$ will still be fixed and will be the number of time step between 0 and $T$ equipped with the time grid $t_k^n = kT/n$, $k \in \mathbb{N}$. We consider two sequences of independent random variables $Z_k \in \mathbb{R}^N, \kappa_k \in \mathbb{R}$, $k \in \mathbb{N}$ and we assume that $Z_k$ are centered verifies (71). We construct the $\mathbb{R}^d$ valued Markov chain

$$X^n_{t_{k+1}} = \psi(\kappa_k, X^n_{t_k}, \frac{Z_{k+1}}{\sqrt{n}}, t^n_{k+1} - t^n_k), \qquad k \in \mathbb{N} \tag{38}$$

where

$$\psi \in \mathcal{C}^\infty(\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^N \times \mathbb{R}_+; \mathbb{R}^d) \quad \text{and} \quad \psi(\kappa, x, 0, 0) = x. \tag{39}$$

We denote

$$\|\psi\|_{1,r,\infty} = 1 \vee \sum_{|\alpha|=0}^{r} \sum_{|\beta|+|\gamma|=1}^{r-|\alpha|} \|\partial_x^\alpha \partial_z^\beta \partial_t^\gamma \psi\|_\infty. \tag{40}$$

**Remark 2.3.** *The reason to consider the random variables $\kappa_k$ 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 $\psi$ according to the fact that $\kappa_k$ is equal to 1 or to $-1$.*

Since the function $\psi$ only needs to be measurable with respect to $\kappa$ and that all our estimates will be done in terms of $\|\psi\|_{1,r,\infty}$, then without loss of generality, we can simplify the notations and denote

$$\psi_k(x, z, t) = \psi(\kappa_k, x, z, t).$$

Then, we slightly modify the definition (40) and instead, in the sequel, we will consider the norm

$$\|\psi\|_{1,r,\infty} = \sup_{k \in \mathbb{N}} \|\psi_k\|_{1,r,\infty} = 1 \vee \sup_{k \in \mathbb{N}} \sum_{|\alpha|=0}^{r} \sum_{|\beta|+|\gamma|=1}^{r-|\alpha|} \|\partial_x^\alpha \partial_z^\beta \partial_t^\gamma \psi_k\|_\infty. \tag{41}$$

Finally for $r \in \mathbb{N}^*$, we denote

$$\mathfrak{K}_r(\psi) = (1 + \|\psi\|_{1,r,\infty}) \exp(\|\psi\|^2_{1,3,\infty}). \tag{42}$$

### 2.3.1 The regularization property

In the following, we will not work under $\mathbb{P}$, but under a localized probability measure defined as follows. We fix $S \leqslant T$ and we consider the set

$$\Lambda_S = \{\frac{1}{\lfloor Sn/T \rfloor} \sum_{k=1}^{\lfloor Sn/T \rfloor} \chi_k \geqslant \frac{m_*}{2}\}. \tag{43}$$

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

$$\mathbb{P}(\Lambda_S^c) \leqslant \exp(-m_*^2 \lfloor Sn/T \rfloor /2)) \tag{44}$$

We consider also the localization function $\varphi_{n^{1/4}/2}$, defined in (30), and we construct the random variable

$$\Theta = \Theta_{S,n} = \mathbb{1}_{\Lambda_S} \times \prod_{k=1}^{n} \varphi_{n^{1/4}/2}(Z_k). \tag{45}$$

Since $Z_k$ has finite moments of any order, the following inequality can be shown: for every $l \in \mathbb{N}$ there exists $C$ such that

$$\mathbb{P}(\Theta_{S,n} = 0) \leqslant \mathbb{P}(\Lambda_M^c) + \sum_{k=1}^{n} \mathbb{P}(|Z_k| \geqslant n^{1/4}) \leqslant \exp(m_*^2 \lfloor Sn/T \rfloor /2) + \frac{M_{4(l+1)}(Z)}{n^l}. \tag{46}$$

We define the probability measure

$$d\mathbb{P}_\Theta = \frac{1}{\mathbb{E}[\Theta]} \Theta d\mathbb{P}. \tag{47}$$

We still fix $T > 0$ and $n \in \mathbb{N}^*$ and we consider the Markov chain $(X_t^n)_{t \in \pi_{T,n}}$, defined in (38). We also recall that $\Theta_{S,n}$ is defined in (45) and we introduce $(Q_t^{\Theta,n})_{t \in \pi_{T,n}}$ such that,

$$\forall t \in \pi_T^n, \quad Q_t^{\Theta,n} f(x) := \mathbb{E}_{\Theta_{t,n}}[f(X_t^n(x))] = \frac{1}{\mathbb{E}[\Theta_{t,n}]} \mathbb{E}[\Theta_{t,n} f(X_t^n(x))]. \tag{48}$$

Notice that $(Q_t^{\Theta,n})_{t \in \pi_{T,n}}$, is not a semigroup, but this is not necessary. We are not be able to prove the regularization property for $Q^n$ but for its modification $Q^{\Theta,n}$. Considering the hypothesis of Theorem 2.1, this is sufficient to obtain total variation convergence.

**Proposition 2.4.** ***A.*** *Let $T > 0$ and $n \in \mathbb{N}^*$. We assume that $n$ and $t \in \pi_{T,n}^{0,T}$ are sufficiently large in order to have :*

$$\frac{3\|\psi\|_{1,3,\infty}}{n^{1/4}} + \frac{M_8(Z)}{n} + \exp(-m_*^2 nt/(2T)) \leqslant \frac{1}{2} \tag{49}$$

*and*

$$n^{1/2} \geqslant 3\frac{2^{N+2}}{\lambda_*}\|\psi\|^2_{1,3,\infty}. \tag{50}$$

*Moreover we assume that*

$$\inf_{\kappa\in\mathbb{R}}\inf_{x\in\mathbb{R}^d}\inf_{|\xi|=1}\sum_{i=1}^{N}\langle\partial_{z_i}\psi(\kappa,x,0,0),\xi\rangle^2 \geqslant \lambda_*. \tag{51}$$

*Then for every $q\in\mathbb{N}$ and multi index $\alpha,\beta$ with $|\alpha|+|\beta|\leqslant q$, there exists $l\in\mathbb{N}^*$ and $C\geqslant 1$ which depend on $m_*, r_*$ and the moments of $Z$ such that*

$$\|\partial_\alpha Q_t^{n,\Theta}\partial_\beta f\|_\infty \;\leqslant\; C\frac{\mathfrak{K}_{q+3}(\psi)^l}{(\lambda_* t)^{q(q+1)}}\|f\|_\infty \tag{52}$$

*with $\mathfrak{K}_r(\psi)$ defined in (42). In particular, $Q_t^{n,\Theta}(x,dy) = p_t^{n,\Theta}(x,y)dy$ and $(x,y)\mapsto p_t^{n,\Theta}(x,y)$ belongs to $\mathcal{C}^\infty(\mathbb{R}^d\times\mathbb{R}^d)$.*

***B.*** *There exists $C\geqslant 1$, such that for every $l\in\mathbb{N}$ and $t\in\pi_{T,n}^T$, we have*

$$\|Q_t^n f - Q_t^{n,\Theta} f\|_\infty \leqslant 4(\exp(-m_*^2 nt/(2T)) + \frac{M_{4(l+1)}(Z)}{n^l})\|f\|_\infty. \tag{53}$$

**Remark 2.4.** *(52) means that the strong regularization property $\overline{R}_{q,\eta}$, with $\eta(q)=q(q+1)$, holds for $Q^{\Theta,n}$.*

Now, we have obtained regularization property under regularity assumption on the scheme function (see hypothesis of Property 2.4), we can come back to our initial problem.

# 3 Markov diffusion processes with locally bounded coefficients

We consider the d-dimensional diffusion process

$$dX_t = V_0(X_t)dt + \sum_{i=1}^{N} V_i(X_t)\circ dW_t^i. \tag{54}$$

For now we assume that $V_i\in\mathcal{C}^\infty(\mathbb{R}^d;\mathbb{R}^d)$, $i=0,\dots,N$, $(W_t)_{t\geqslant 0}$ a standard Brownian motion and $\circ dW_t^i$ the Stratonovich integral with respect to $W^i$. In the same way we will denote $V_{0,\mathrm{Ito}}$ such that

$$dX_t = V_{0,\mathrm{Ito}}(X_t)dt + \sum_{i=1}^{N} V_i(X_t)dW_t^i. \tag{55}$$

The infinitesimal operator of this Markov process is

$$A = V_0 + \frac{1}{2}\sum_{i=1}^{N} V_i^2 \tag{56}$$

with the notation $Vf(x) = \langle V(x), \nabla f(x)\rangle$. In the latter, when it is relevant, we will denote by $X_t(x)$ the process starting at $x$.

## 3.1 Regularization of the coefficients of the diffusion

In this section, we assume that the coefficients $V_i \in \mathcal{C}_b^\infty(\mathcal{D};\mathbb{R}^d)$ for a subset $\mathcal{D}$ of $\mathbb{R}^d$. Moreover we denote $\mathcal{D}^v$ the biggest compact contained in $\{x \in \mathcal{D}, \inf_{y\in\mathbb{R}^d\setminus\mathcal{D}} |x-y| > v\}$ with the convention $\inf_{y\in\emptyset}|x-y| = +\infty$. Now for $v \geqslant 0$ and $\mathcal{D} \subset \mathbb{R}^d$, we introduce $\phi_v^{\mathcal{D}} \in \mathcal{C}_b^\infty(\mathbb{R}^d : \mathbb{R}^d)$ such that $\phi_v^{\mathcal{D}}(x) = x$ if $x \in \mathcal{D}^v$, is constant if $x \in \mathbb{R}^d \setminus \mathcal{D}$. Moreover we assume that $\phi_v^{\mathcal{D}} \in \mathcal{C}_b^\infty(\mathbb{R}^d \setminus \mathcal{D}^v;\mathbb{R}^d)$ and that for all $x \in \mathbb{R}^d \setminus \mathcal{D}^v$, we have

$$|\partial_\alpha \phi_v^{\mathcal{D}}(x)| \leqslant \frac{C}{v^{|\alpha|}}, \tag{57}$$

where $C$ does not depend on $v$. Now, we can introduce a regularization of (54). For $i = 0, \cdots, N$, we denote $V_i^{\mathcal{D},v} = V_i \circ \phi_v^{\mathcal{D}}$ and we define

$$dY_t^{\mathcal{D},v} = V_0^{\mathcal{D},v}(Y_t)dt + \sum_{i=1}^{N} V_i^{\mathcal{D},v}(Y_t^{\mathcal{D},v}) \circ dW_t^i. \tag{58}$$

In order to prove weak convergence for bounded measurable test functions, we will use this modification of the underlying process. For $\delta > 0$ we define

$$\overline{X}_t(x) = \begin{cases} X_t(x), & \text{if } t \leqslant T-\delta, \\ Y_{t-(T-\delta)}(X_{T-\delta}(x)), & \text{if } \quad T-\delta \leqslant t \leqslant T, \end{cases} \tag{59}$$

Indeed, let us assume that we are able to build a scheme for $Y$ that converges for the total variation distance when $T-\delta \leqslant t \leqslant T$ (for instance, see Theorem 4.2 for the Ninomiya Victoir scheme or [4] for a more detailed approach). Then if we are able to find a scheme which satisfies (12) and (15) for $t \leqslant T-\delta$, then Theorem 2.1 will ensure convergence for measurable test function between $\overline{X}$ and its scheme. The last step consists in estimating the distance between the underlying process (respectively scheme) and its modification (resp. modified scheme). This is the purpose of the next section.

## 3.2 Concentration inequalities

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*

$$\mathbb{P}(\sup_{t\in\mathcal{T}}|H_t-y| < v) = \mathbb{P}(\sup_{t\in\mathcal{T}}|K_t-y| < v) \tag{60}$$

*Proof.* We consider that $H$ and $K$ are non null processes. Otherwise th proof is immediate. 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 $\{\sup_{t\in\mathcal{T}}|H_t-y|<v\}\cap\{H=H\mathbb{1}_{\sup_{t\in\mathcal{T}}|H_t-y|\geqslant v}\}=\emptyset$ so we obtain

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

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.

□

**The continuous case - The Bernstein's inequality**

**Proposition 3.1.** *Let $(M_t)_{t\geqslant 0}$ a continuous $\mathcal{F}_t$-local martingale such that $M_0=0$ and $\langle M\rangle_\infty=\infty$, a.s., then, for all $v\geqslant 0$,*

$$\mathbb{P}(\sup_{0\leqslant s\leqslant t}|M_s|>v|\langle M\rangle_t\leqslant c)\leqslant 2\exp(-\frac{v^2}{2c}) \tag{61}$$

*Proof.* In order to prove (61), we will use the following result which corresponds to the specific case of the Brownian motion.

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

$$\forall v>0,\quad \mathbb{P}(\sup_{0\leqslant s\leqslant t}|W_s|>v)\leqslant 2\exp(-\frac{v^2}{2t}) \tag{62}$$

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

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

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 (62) follows. □

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

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

□

Applying this result to Markov diffusions we get the following result

**Corollary 3.1.** *Let $0 < \delta \leqslant T$. Assume that $(X_t)_{t\geqslant 0}$ follows a diffusions of the form (55) with $V_{0,\mathrm{Ito}}, V_i \in \mathcal{C}_b^0(\mathbb{R}^d;\mathbb{R}^d)$. Then, for all $v \geqslant 0$, we have*

$$\mathbb{P}(\sup_{T-\delta\leqslant t\leqslant T}|X_T - X_t| \geqslant v|X_T) \leqslant 2\exp(-\frac{v^2/2 - \delta^2\|V_{0,\mathrm{Ito}}\mathbb{1}_{\overline{B}_v(X_T)}\|_\infty^2}{\delta\sum_{i=1}^N \|V_i\mathbb{1}_{\overline{B}_v(X_T)}\|_\infty^2}). \tag{63}$$

*Proof.* Let $\epsilon > 0$. Using Lemma 3.1, we have

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

Now, (63) follows from Proposition 3.1 with $(M_t)_{0\leqslant t\leqslant\delta} = \sum_{i=1}^N\int_{T-t}^T V_i(X_s)\mathbb{1}_{\overline{B}_v(X_T)}(X_s)dW_s^i$ using that $\langle M\rangle_t \leqslant \delta\sum_{i=1}^N\|V_i\mathbb{1}_{\overline{B}_v(X_T)}\|_\infty^2$.

□

**The discrete case**

**Proposition 3.2.** *The Hoeffding inequality. Let $(M_n)_{n\in\mathbb{N}}$ be a discrete centered Markov process such that there exists two sequences $(b_n^d)_{n\in\mathbb{N}} \leqslant (b_n^u)_{n\in\mathbb{N}}$ such that for all $k \in \mathbb{N}^*$, $\mathbb{P}(M_n - M_{n-1} \in [b_n^l, b_n^u]) = 1$, then*

$$\mathbb{P}(|M_n| \geqslant v) \leqslant 2\exp(-\frac{2v^2}{\sum_{k=1}^N(b_l^u - b_k^l)^2}). \tag{64}$$

**Corollary 3.2.** *We fix $T > 0$ and $n \in \mathbb{N}^*$. Let $2T/n \leqslant \delta \leqslant T$ and define $\tilde{\delta}_n = \inf\{t; t > \delta, T - t \in \pi_{T,n}\}$. We assume that $(X_{t_k}^n)_{k\in\mathbb{N}}$ is defined by (38) with $\psi_k \in \mathcal{C}_b^1(\mathbb{R}^d\times\mathbb{R}^N\times\mathbb{R}_+;\mathbb{R}^d)$ and $Z_k \in [b_k^l(Z), b_k^u(Z)]$ and that*

$$n \geqslant \sup_{k\in\mathbb{N}^*} T^2/(b_k^u(Z) - b_k^l(Z))^2. \tag{65}$$

*Then, for every $v > 0$ we have*

$$\mathbb{P}(\sup_{t\in\pi_{T,n}^T; t\geqslant T-\tilde{\delta}_n}|X_T^n - X_t^n| \geqslant v|X_T^n) \leqslant 2\exp(-\frac{C_Z v^2}{\delta\|\psi\mathbb{1}_{\overline{B}_v(X_T)}\|_{1,1,\infty}^2}), \tag{66}$$

*with $C_Z = \inf_{k\in\mathbb{N}^*} T/(12(b_k^u(Z) - b_k^l(Z))^2)$.*

*Proof.* Let $N = 1$ for sake of simplicity in the writing. We have

$$X^n_{t^n_{k+1}} - X^n_{t^n_k} = w^0_{k+1}\int_0^1 (1-\lambda)\partial_t\psi(\kappa_k, X^n_{t_k}, \lambda w^0_{k+1}, w^1_{k+1})d\lambda + w^1_{k+1}\int_0^1 (1-\lambda)\partial_z\psi(\kappa_k, X^n_{t_k}, 0, \lambda w^1_{k+1})d\lambda,$$

with $w^0_{k+1} = T/n$ and $w^1_{k+1} \in [b^l_{k+1}(Z)/\sqrt{n}, b^u_{k+1}(Z)/\sqrt{n}]$.We apply Lemma 3.1 and Proposition 3.2 in order to obtain

$$\mathbb{P}(\sup_{t\in\pi^T_{T,n};t\geqslant T-\tilde{\delta}_n} |X^n_T - X^n_t| \geqslant v|X^n_T) \leqslant$$
$$2\exp(-\frac{v^2}{4\sum_{k=k_\delta}^n (T/n\|\partial_t\psi\mathbb{1}_{\overline{B}_v(X^n_T)}\|_\infty)^2 + (1/\sqrt{n}(b^u_k(Z) - b^l_k(Z))\|\partial_z\psi\mathbb{1}_{\overline{B}_v(X^n_T)}\|_\infty)^2}),$$

with $k_\delta = n(T-\tilde{\delta}_n)/T$. 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 (65), we rearrange the terms and the proof of (66) is completed. □

## 3.3 Convergence results for Markov diffusion with locally bounded coefficients.

Let $n \in \mathbb{N}^*$, $q \in \mathbb{N}$, $v > 0$ and $\tilde{\psi} \in \mathcal{C}^q_b(\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^N\times\mathbb{R}_+)$. We introduce the hypothesis concerning the existence of a regular modification of the Markov discrete process defined in (38). We will exploit this modification in the neighborhood of the terminal date $T$ and when the terminal value of the Markov process belong to a compact set.

(**Reg**$(\psi, \tilde{\psi}, v, q, n, \delta)$) For all $m \geqslant n$, we assume that there exists a discrete process $(\overline{X}^m_t)_{t\in\pi_{T,m}}$ defined by

$$\forall k \in \mathbb{N}^*, \quad \overline{X}^m_{t^m_{k+1}} = \tilde{\psi}(\kappa_k, \overline{X}^m_{t^m_k}, Z_{k+1}/\sqrt{m}, t^m_{k+1} - t^m_k),$$

and such that,

$$(X^m_{t^m_k}\mathbb{1}_{\sup_{T-\tilde{\delta}_m\leqslant t^m_j\leqslant T}|X^m_{t^m_j} - X^m_T|\leqslant v})_{T-\tilde{\delta}_m\leqslant t^m_k\leqslant T} \overset{Law}{=} (\overline{X}^m_{t^m_k}\mathbb{1}_{\sup_{T-\tilde{\delta}_m\leqslant t^m_j\leqslant T}|\overline{X}^m_{t^m_j} - \overline{X}^m_T|\leqslant v})_{T-\tilde{\delta}_m\leqslant t^m_k\leqslant T}$$
$$\text{on the event} \quad \{X^m_{T-\tilde{\delta}_m} = \overline{X}^m_{T-\tilde{\delta}_m}\} \cap \{X^m_T = \overline{X}^m_T \in \mathcal{D}^{2v}\},$$

with $\tilde{\delta}_m$ defined in (25). Moreover, we also suppose that (49) (50) and (51) hold with $\psi$ replaced by $\tilde{\psi}$. Finally, we denote $(P^{\mathcal{D},v}_t)_{t\geqslant 0}$ the semigroup of the process $(Y^{\mathcal{D},v}_t)_t$ defined in (58) and $(Q^{\mathcal{D},v,m})_{t\in\pi_{T,m}}$ the one associated to $(\overline{X}^m_t(x))_{t\in\pi_{T,m}}$. and we assume that $E_m(h,q)$ (see (13)) and $E^*_m(h,q)$ (see (20)) hold between $(P^{\mathcal{D},v,m}_t)_{t\in\pi_{T,m}} = (P^{\mathcal{D},v}_t)_{t\in\pi_{T,m}}$ and $(Q^{\mathcal{D},v,m}_t)_{t\in\pi_{T,m}}$.

**Theorem 3.1.** *We recall that $T > 0$ is fixed. Let $q \in \mathbb{N}$, $h > 0$, $\delta \in (0,T)$ and $\tilde{\psi} \in \mathcal{C}^q_b(\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^N\times\mathbb{R}_+)$. For a given $n \in \mathbb{N}^*$, we consider the Markov semigroup $(P_t)_{t\geqslant 0}$ (see (54)), and the approximation Markov chain $(Q^n_t)_{t\in\pi_{T,n}}$ (see (38)), defined above. Moreover, we assume that there exists $n_0 \in \mathbb{N}^*$ such that $T/n_0 \leqslant 2\delta$ and, (49), (50) and (65) hold with $n = n_0$ and*

*$t = \delta/2$. Then, for all $n \geqslant n_0$, we have the following property.*

*Assume that $(\mathbf{Reg}(\psi, \tilde{\psi}, v, q, n, \delta))$ holds and that, on the interval $[0, T-\delta]$, $P^n$ and $Q^n$ satisfy respectively (11) for $q = 0$ and (12) for this $q$ and that $E'_{n,\mathrm{pol}}(h, q)$ (see (15)) holds. Then, there exists $l \in \mathbb{N}^*$, $C, \beta \geqslant 1$ such that we have for every bounded and measurable test function $f$ on $\mathbb{R}^d$ with $\mathrm{supp}(f) \subset \mathcal{D}^{2v}$,*

$$\begin{aligned}\mathbb{E}[f(X_T(x)) - f(X_T^n(x))] \leqslant & C(\exp(-\frac{v^2/2 - \delta^2 \|V_{0,\mathrm{Ito}} \mathbb{1}_{\mathcal{D}}\|_\infty^2}{\delta \sum_{i=1}^N \|V_i \mathbb{1}_{\mathcal{D}}\|_\infty^2}) + \exp(-\frac{C_Z v^2}{\delta \|\tilde{\psi} \mathbb{1}_{\mathcal{D}}\|_{1,1,\infty}^2})) \|f\|_\infty \\ & + C(1 + |x|^\beta) \frac{\mathfrak{K}_{q+3}(\tilde{\psi})^l}{(\lambda_* S)^{\eta(q)}} \|f\|_\infty / n^h \end{aligned} \tag{67}$$

*with $\eta(q) = q(q+1)$ and $C_Z = \inf_{k \in \mathbb{N}^*} T/(12(b_k^u(Z) - b_k^l(Z))^2)$.*

**Remark 3.1.** *We can obtain the same result as (67) for measurable and bounded test functions $f$ with $\mathrm{supp}(f) \subset \mathbb{R}^d$. In order to do it, we simply have to assume that $\mathbb{P}(X_T \notin \mathcal{D})$ and $\mathbb{P}(X_T^n \notin \mathcal{D})$ are small enough. In this case we can rigoursly speak of total variation distance estimation. However, there is not particular interest in doing it if we can not describre $\mathbb{P}(X_T \notin \mathcal{D})$ or $\mathbb{P}(X_T^n \notin \mathcal{D})$. That is why, we simply provide (67) and consider this result as a total variation distance estimation between the process and its scheme.*

**Remark 3.2.** *Notice that from Lemma 3.1, we deduce that the previous result remains true if we replace $V$ by $V^{\mathcal{D},v}$ in (67) or if we replace $\|\tilde{\psi} \mathbb{1}_{\mathcal{D}}\|_{1,1,\infty}$ by $\|\psi \mathbb{1}_{\mathcal{D}}\|_{1,1,\infty}$, $\|\tilde{\psi}\|_{1,1,\infty}$ or $\|\psi\|_{1,1,\infty}$ if thos quantities are finite.*

*Proof of Theorem 3.1.* First we define

$$\overline{X}_t(x) = \begin{cases} X_t(x), & \text{if } t < T - \tilde{\delta}_n, \\ Y_{t-(T-\tilde{\delta}_n)}(X_{T-\delta}(x)), & \text{if } \quad T - \tilde{\delta}_n \leqslant t \leqslant T, \end{cases}$$

$$\overline{X}^n_{t^n_{k+1}}(x) = \begin{cases} X^n_{t^n_{k+1}}(x), & \text{if } \quad t_k^n < T - \tilde{\delta}_n, \\ \tilde{\psi}_k(\overline{X}^n_{t^n_k}(x), w^1_{k+1}, w^0_{k+1}) & \text{if } \quad T - \tilde{\delta}_n \leqslant t_k^n \leqslant T, \end{cases}$$

with $Y$ defined in (58) and $\tilde{\delta}_n$ defined in (25). First,from Lemma 3.1, we have

$$\mathbb{P}(\sup_{T-\tilde{\delta}_n \leqslant t \leqslant T} |X_t - X_T| \geqslant v | X_T \in \mathcal{D}^{2v}) = \mathbb{P}(\sup_{T-\tilde{\delta}_n \leqslant t \leqslant T} |\overline{X}_t - \overline{X}_T| \geqslant v | \overline{X}_T \in \mathcal{D}^{2v})$$

because $(X_t \mathbb{1}_{\sup_{T-\tilde{\delta}_n \leqslant t \leqslant T} |X_t - X_T| \leqslant v})_{T-\tilde{\delta}_n \leqslant t \leqslant T}$ and $(\overline{X}_t \mathbb{1}_{\sup_{T-\tilde{\delta}_n \leqslant t \leqslant T} |\overline{X}_t - \overline{X}_T| \leqslant v})_{T-\tilde{\delta}_n \leqslant t \leqslant T}$ follow the same law as soon as $X_{T-\tilde{\delta}_n} = \overline{X}_{T-\tilde{\delta}_n}$ and $X_T, \overline{X}_T \in \mathcal{D}^{2v}$. Moreover, we deduce from Lemma 3.1 again and $(\mathbf{Reg}(\psi, \tilde{\psi}, v, q, n, \delta))$ that

$$\mathbb{P}(\sup_{T-\tilde{\delta}_n \leqslant t_k^n \leqslant T} |X^n_{t^n_k} - X^n_T| \geqslant v | X^n_T \in \mathcal{D}^{2v}) = \mathbb{P}(\sup_{T-\tilde{\delta}_n \leqslant t_k^n \leqslant T} |\overline{X}^n_{t^n_k} - \overline{X}^n_T| \geqslant v | \overline{X}^n_T \in \mathcal{D}^{2v}).$$

Then from the Corollary 3.1 and the Corollary 3.2, we deduce that for every bounded and measurable test function $f$ with $\mathrm{supp}(f) \subset \mathcal{D}^{2v}$, we have

$$\begin{aligned}
\mathbb{E}[f(X_T) - f(X_T^n)] &\leqslant |\mathbb{E}[f(\overline{X}_T) - f(\overline{X}_T^n)]| \\
&\quad + 4\|f\|_\infty \Big( \mathbb{P}( \sup_{T-\tilde{\delta}_n \leqslant t \leqslant T} |X_t - X_T| \geqslant v) + \mathbb{P}( \sup_{T-\tilde{\delta}_n \leqslant t \leqslant T} |\overline{X}_t - \overline{X}_T| \geqslant v) \\
&\quad + \mathbb{P}( \sup_{T-\tilde{\delta}_n \leqslant t_k^n \leqslant T} |X_{t_k^n}^n - X_T^n| \geqslant v) + \mathbb{P}( \sup_{T-\tilde{\delta}_n \leqslant t_k^n \leqslant T} |\overline{X}_{t_k^n}^n - \overline{X}_T^n| \geqslant v) \Big) \\
&\leqslant |\mathbb{E}[f(\overline{X}_T(x)) - f(\overline{X}_T^n(x))]| + C\|f\|_\infty \Big( \exp(-\frac{v^2/2 - \delta^2 \|V_{0,\mathrm{Ito}} \mathbb{1}_{\mathcal{D}}\|_\infty^2}{\delta \sum_{i=1}^N \|V_i \mathbb{1}_{\mathcal{D}}\|_\infty^2}) \\
&\quad + \exp(-\frac{C_Z v^2}{\delta \|\tilde{\psi}\|_{1,1,\infty}^2}) \Big)
\end{aligned}$$

Now, we use Theorem 4.2 and Theorem 2.1 in order to estimate $|\mathbb{E}[f(\overline{X}_T(x)) - f(\overline{X}_T^n(x))]|$ and we obtain (67). □

**Remark 3.3.** *Notice that from Lemma 3.1, we deduce that the previous result remains true if we replace $V$ by $V^{\mathcal{D},v}$ in (67) or if we replace $\|\tilde{\psi}\mathbb{1}_{\mathcal{D}}\|_{1,1,\infty}$ by $\|\psi\mathbb{1}_{\mathcal{D}}\|_{1,1,\infty}$, $\|\tilde{\psi}\|_{1,1,\infty}$ or $\|\psi\|_{1,1,\infty}$ if thos quantities are finite.*

# 4 Second order total variation convergence towards CIR processes

In this section, we are going to apply the results we have just 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 method. Initially those methods were used in [26] to build the so called Ninomiya Victoir schemes for SDE with smooth coefficients. Then, Alfonsi [2], inspired by this approach, built a second weak order scheme for the CIR process. Finally, in [4], the author have shown that the total variation convergence takes place for those cubature scheme as soon as the coefficients of the SDE are smooth. In this section, our purpose to exploit and extend those result in order to obtain total variation convergence results for the CIR which has singular coefficients in the neighborhood of zero.

## 4.1 The Ninomiya Victoir scheme

We begin by presenting the Ninomiya Victoir scheme [26]. Let us define $\exp(V)(x) := \Phi_V(x,1)$ where $\Phi_V$ solves the deterministic equation

$$\Phi_V(x,t) = x + \int_0^t V(\Phi_V(x,s))ds. \tag{68}$$

By a change of variables one obtains $\Phi_{\varepsilon V}(x,t) = \Phi_V(x,\varepsilon t)$ 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 operator $A_V f(x) = Vf(x)$. 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

$$f(\Phi_V(x,t))) = f(x) + \sum_{r=1}^{m} \frac{t^r}{r!} V^r f(x) + \frac{1}{m!} \int_0^t (t-s)^m V^{m+1} P_s^V f(x) ds. \tag{69}$$

We present now a second order scheme introduced defined as in [2]. We consider a sequence $\rho_k,\ k \in \mathbb{N}$ of independent Bernoulli random variables and we define $\psi_k : \mathbb{R}^d \times \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}^d$ using a splitting procedure. Let us define

$$\psi(\rho, x, w^1, w^0) = \begin{cases} \exp(w^0 V_0) \circ \exp(w^{1,1} V_1) \circ \cdot \circ \exp(w^{1,N} V_N) \circ \exp(w^0 V_0)(x), & \text{if } \rho = 1, \\ \exp(w^0 V_0) \circ \exp(w^{1,N} V_N) \circ \cdot \circ \exp(w^{1,1} V_1) \circ \exp(w^0 V_0)(x), & \text{if } \rho = -1. \end{cases} \tag{70}$$

Moreover, we denote $w_k^0 = T/n$ and $w_k^1 = (w_k^{1,i})_{i=1,\cdot,N}$ with $w_k^i = \sqrt{T} Z_k^i / \sqrt{n},\ i = 1, \cdots, N$ and we assume that $Z_k,\ k \in \mathbb{N}$ are independent random variables which are lower bounded by the Lebesgue measure: 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 $k \in \mathbb{N}$

$$L_{z_*}(\varepsilon_*, r_*) \qquad \mathbb{P}(Z_k \in A) \geqslant \varepsilon_* \lambda(A \cap B_{r_*}(z_{*,k})). \tag{71}$$

Finally, we assume that the sequence $Z_k$ satisfies the following moment conditions:

$$\begin{aligned} \mathbb{E}[Z_k^i] = \mathbb{E}[(Z_k^i)^3] = \mathbb{E}[(Z_k^i)^5] = 0, \qquad \mathbb{E}[(Z_k^i)^2] = 1, \qquad \mathbb{E}[(Z_k^i)^4] = 3, \\ \forall p \geqslant 1, \qquad \mathbb{E}[|Z_k|^p] < \infty. \end{aligned} \tag{72}$$

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

$$X_{t_{k+1}^n}^n = \psi(\rho_k, X_{t_k^n}^n, w_{k+1}, w_{k+1}^0). \tag{73}$$

### 4.1.1 Convergence results for diffusion with smooth coefficients

**Smooth test functions**

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

**Theorem 4.1.** *Let $(X_t)_{t \geqslant 0}$ the process defined by (54) and $(X_t^n)_{t \in \pi_{T,n}}$ given by (73). We also assume that (72) holds.*

**A.** *We assume that for all* $l \leqslant 3$, *we have* $V_0^l : \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^d, \mathbb{R}^d) \to \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^d, \mathbb{R}^d)$ *and for all* $i = 1, \cdots N$, *we have* $V_i^{2l}, A^l : \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^d, \mathbb{R}^d) \to \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^d, \mathbb{R}^d)$. *We also assume that* $\sum_{i=0}^N V_i(x) \leqslant C(1 + |x|)$ *and that for all* $t \geqslant 0$, $x \mapsto \Phi_{V_i}(x,t) \in \mathcal{C}^0_{\text{pol}}(\mathbb{R}^d)$. *Then,* $E_{n,\text{pol}}(2,6)$ *(see (14)) and* $E'_{n,\text{pol}}(2,6)$ *(see (15)) are satisfied for* $(X_t)_{t\pi_{T,n}}$ *and* $(X_t^n)_{t\in\pi_{T,n}}$.

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

$$\sup_{t_k^n \leqslant T} |\mathbb{E}[f(X_{t_k^n})) - \mathbb{E}[f(X^n_{t_k^n})]| \leqslant CC_6(V)^l \|f\|_{6,\infty}/n^2, \tag{74}$$

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

*Proof.* We will prove only point **A**. The proof of (15) for **B.** is very similar. Then, it simply remains to show (12) which is done in the literature (see [4]). We focus on the proof of (15). We will assume that $N = 1$ for sake of simplicity. It is sufficient to prove that the schemes with transition probability laws $\Phi_{V_0}(t_{k+1} - t_k, .)$ and $\Phi_{V_i}(\frac{Z_k^i}{\sqrt{n}}, .)$ are weak second order schemes. We will prove that they are $\nu$-order schemes, for all integer $\nu \in \mathbb{N}^*$ as soon as $Z_k$ matches the $2\nu + 1$ 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 the moments functions $x \mapsto \mathbb{E}[|X_t(x)|^q]$ belong to $\mathcal{C}^0_{\text{pol}}(\mathbb{R}^d)$ (see [1]). Let us consider $f \in \mathcal{C}^\infty_b(\mathbb{R}^d, \mathbb{R})$. According to the definition of $V_0$, we have $\forall l \in \mathbb{N}, V_0^l f \in \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^d)$. Writing as expansion (69), we get :

$$\forall \nu \in \mathbb{N}, \quad f(\Phi_{V_0}(t,x)) = f(x_0) + \textstyle\sum_{l=1}^\nu \frac{t^l}{l!} V_0^l f(x) + R_{0,t}^{\nu+1} f(x_0)$$

with

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

We have already precised that $V_0^{\nu+1} f \in \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^d)$. Besides we have $\Phi_{V_0} \in \mathcal{C}^0_{\text{pol}}(\mathbb{R}^d, \mathbb{R}; \mathbb{R}^d)$, then for all $t \in [0,1]$, there exists $\beta \in \mathbb{N}$ such that

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

so the scheme with transition probability $\Phi_{V_0}$ is a $\nu$-order scheme for the operator $V_0$.

Let $H_k = \sqrt{t} Z_k^1$. We have

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

with

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

Moreover, $V_1^{2\nu+2} f, \in \mathcal{C}^\infty_{\text{pol}}(\mathbb{R}^d, \mathbb{R})$ and $\Phi_{V_1} \in \mathcal{C}^0_{\text{pol}}(\mathbb{R}^d, \mathbb{R}; \mathbb{R}^d)$ and again, there exists $\beta \in \mathbb{N}^*$, $C > 0$, such that

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

and the scheme with transition probability $\Phi_{V_1}$ is a $\nu$-order scheme for the operator $V_1$.

Finally, for all $f \in \mathcal{C}^\infty_{\text{pol}}$ then $Af \in \mathcal{C}^\infty_{\text{pol}}$ and we obtain (15) using the Ninomiya-Victoir composition and the polynomial control of the moments of the diffusion and $\Phi_{V_i}$. The proof of (14) is very similar so we leave it out. □

**Remark 4.1.** *Notice that property (12) has already been studied in [1] for the CIR and since $\Phi_{V_i} \in \mathcal{C}^0_{\text{pol}}(\mathbb{R}^d)$, 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 are going to obtain an estimate of the total variation distance between a diffusion process of the form (54) and its second order scheme (73).

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

$$\inf_{|\xi|=1} \sum_{i=1}^N \langle V_i(x), \xi \rangle^2 \geqslant \lambda_* > 0 \qquad \forall x \in \mathbb{R}^d. \tag{75}$$

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

$$\sup_{t \in \pi^{2S,T}_{T,n}} |\mathbb{E}[f(X_t)] - \mathbb{E}[f(X^n_t)]| \leqslant C \frac{C_6(V)^l \mathfrak{K}_9(\psi)^l}{(\lambda_* S)^{42}} \|f\|_\infty / n^2. \tag{76}$$

**Remark 4.2.** *This result has already been obtained in [4]. The result (76) 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 (70) 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 (75). Consequently a similar procedure could be used in order to prove the convergence in total variation for even higher order scheme as soon as we control this error for smooth test function. Finally, it is important to notice that the generic property (71) 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 variable $(Z_k)_{k \in \mathbb{N}^*}$ do not have a specific law but only satisfy the Doeblin condition (71) and (72), the result can be seen as an invariance priciple.*

## 4.2 The CIR model

The CIR model is a $\mathbb{R}_+$-valued random process defined by the following SDE,

$$dX_t = (a - kX_t)dt + \sigma\sqrt{X_t}dW_t. \tag{77}$$

This model was first presented in 1985 inspired from Vasicek (1977), by modifying the volatility term introducing a "square root" term, among others in order to guarantee non-negativity. We suppose in this paper $a, k, \sigma > 0$. In this case, it is impotant 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 (54), we have

$$\forall x \in \mathbb{R}, \quad V_{0,cir}f(x) = (a - kx - \frac{\sigma^2}{4})\partial_x f(x) \tag{78}$$

$$\forall x \in \mathbb{R}_+, \quad V_{1,cir}f(x) = \sigma\sqrt{x}\partial_x f(x) \tag{79}$$

Solving the PDE (68) brings 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 (73) is not well defined anymore. In this case, we will introduce another scheme in the neighborhood of zero and we will use a switching procedure. Otherwise, we will prove that,as soon as the scheme (73) is well defined, then it is a second weak order scheme. As a consequence, if $4a \geqslant \sigma^2$, we define

$$\psi_{cir}(x, w^1, w^0) = \exp(w^0 V_{0,cir}) \circ \exp(w^1 V_{1,cir}) \circ \exp(w^0 V_{0,cir})(x). \tag{80}$$

Besides, we denote $w_k^0 = T/n$ and $w_k^1 = \sqrt{T}Z_k/\sqrt{n}$, where $Z_k$, $k \in \mathbb{N}^*$ are independent random variables which are lower bounded by the Lebesgue measure (see (71)). Finally, we assume that the sequence $Z_k$ satisfies the moment conditions (72).
One step of the scheme for the CIR diffusion (between times $t_k$ and $t_{k+1}$) is given by

$$X_{t_{k+1}}^n = \psi_{cir}(\rho_k, X_{t_k}^n, w_{k+1}^1, w_{k+1}^0). \tag{81}$$

### 4.2.2 Second weak order scheme in the neighborhood of zero

It is easy to show that

$$\mathbb{E}[X_t] = xe^{-kt} + \frac{a}{k}(1-e^{-kt}) \tag{82}$$

$$Var[X_t] = x\frac{\sigma^2}{k}(e^{-kt}-e^{-2kt}) + a\frac{\sigma^2}{2k^2}(1-e^{-kt})^2 \tag{83}$$

Now, we consider the case $\sigma^2 > 4a$. First of all we have to identify the bound such that, for every time step we use the scheme (81) or we introduce another scheme. Since we will consider some bounded random variables $Z_k$, we will use the following result in order to chose if we switch schemes.

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

$$K(t,A) = e^{\frac{kt}{2}}\left[\left(\frac{\sigma^2}{4}-a\right)\frac{1-e^{\frac{-kt}{2}}}{k} + \left(\sqrt{e^{\frac{kt}{2}}\left(\frac{\sigma^2}{4}-a\right)\frac{1-e^{-\frac{kt}{2}}}{k}} + \frac{\sigma}{2}A\right)^2\right]. \tag{84}$$

*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 if $K(t,A) \geqslant x^*(\frac{t}{2})$ and then for all $t \in [0,1]$, the function $x \mapsto \phi_{0,cir}(\frac{t}{2},.) \circ \phi_{1,cir}(w^1,.) \circ \phi_{0,cir}(\frac{t}{2},x)$ is well defined. Moreover $\phi_{0,cir}$ is increasing with respect to $x$ and :

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

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

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

and the prof is complete. □

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

**Moments matching approach** Our purpose is to verify for each step if the Ninomiya Victoir scheme remains positive. Otherwise, we switch with a scheme inspired from Andersen [3] and introduced in [2]. We recall that the first two moments of the CIR :are given by

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

Let us introduce the discrete random process $(\zeta_t)_{t\geqslant 0} \in \{y_1, y_2\} \in \mathbb{R}_+^2$ and denote $u_q(x,t) = \mathbb{E}[X_t^q]$. Then, one step of the scheme in the neighborhood of zero for the CIR will be given by $Y_{t_{k+1}-t_k}$ with $(y_1, y_2)$ given by the solution of the following equation

$$\left\{\begin{array}{ll} P(Y_t = y_1)y_1 + P(\zeta_t = y_2)y_2 & = u_1(x,t) \\ P(Y_t = y_1)y_1^2 + P(\zeta_t = y_2)y_2^2 & = u_2(x,t). \end{array}\right. \tag{85}$$

In this equation the only fixed parameters are $u_1(t,x)$ and $u_2(t,x)$, and obviously $P(Y_t = y_2) = 1 - P(Y_t = y_1) \in ]0,1[$. Thus we can fix values for $y_1$ and $y_2$ to solve the first equation and then the second will give a second order equation to solve to find $P(Y_t = y_1)$. For instance, if $v \in ]0,1[$, let us choose :

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

The second equation gives :

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

Thus, if we consider the second order equation which depends on the parameter v :

$$u_2(t,x)\kappa^2 + \left[(1-2v)u_1(t,x) - u_2(t,x)\right]\kappa + v^2 u_1(t,x)^2 = 0 \tag{86}$$

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

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

For sake of simplicity we set $v = \frac{1}{2}$ that simplifies the calculus and as shown below, fulfill the desired conditions. We obtain the following solution for (86) :

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

Now, we set

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

Since $u_1(t,x) \geqslant \max(a^2\left(\frac{1-e^{-kt}}{k}\right), 2x\frac{a}{k}(e^{-kt}-e^{-2kt}))$, we are able to bound :

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

Then, we have the following crucial property in order to prove the second order convergence. Assume that $0 < K(t) < Ct$, then, for all $q \in \mathbb{N}$, there exists $C > 0$ such that

$$\forall t \in ]0,1], x \in [0, K(t)], \quad \exists C > 1, \mathbb{E}[Y_t^q] < Ct^q. \tag{87}$$

We define the transition function in the neighborhood of zero by

$$\hat{\psi}_{cir}(\rho, x, w^0) = \begin{cases} \dfrac{u_1(w^0,x)}{2p(x,w^0)}, & \text{if } \rho = 1, \\ \dfrac{u_1(w^0,x)}{2(1-p(x,w^0))}, & \text{if } \rho = -1. \end{cases} \tag{88}$$

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

$$X^n_{t_{k+1}} = \hat{\psi}_{cir}(\rho_k, X^n_{t_k}, w^0_{k+1}).,\tag{89}$$

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

$$X^n_{t_{k+1}} = \begin{cases} \psi_{cir}(\rho_k, X^n_{t_k}, w^1_{k+1}, w^0_{k+1}), & \text{if } X^n_k \geqslant K(T/n, 3/\sqrt{n}), \\ \hat{\psi}_{cir}(\rho_k, X^n_{t_k}, w^0_{k+1}), & \text{if } X^n_k < K(T/n, 3/\sqrt{n}). \end{cases} \tag{90}$$

## 4.3 Convergence results

### 4.3.1 Smooth test functions

We focus on the convergence of the CIR schemes defined in (81) and (90). We will use Property 2.1. In order to apply this result we we need to establish the following straightforward property.

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

$$V^{2l}_{1,cir} : \mathcal{C}^\infty_p(\mathbb{R}^*_+, \mathbb{R}) \to \mathcal{C}^\infty_p(\mathbb{R}^*_+.\mathbb{R}) \tag{91}$$

*Proof.* It is sufficient to show the result for $l = 1$. We have :

$$\begin{aligned} \forall f \in \mathcal{C}^\infty_p(\mathbb{R}^*_+, \mathbb{R}), V^2_{1,cir}f(x) &= \sigma^2\sqrt{x}\left(\frac{1}{2\sqrt{x}}\partial_x f(x) + \sqrt{x}\partial^2_x f(x)\right) \\ &= \frac{\sigma^2}{2}\partial_x f(x) + \sigma^2 x\partial^2_x f(x) \in \mathcal{C}^\infty(\mathbb{R}^*_+, \mathbb{R}) \end{aligned}$$

□

The property (91) enables to obtain the following short time estimate.

**Theorem 4.3.** *let us assume that $(X_t)_{t\geqslant 0}$ is the process given by (77) and $(X^n_{t_k})$ is defined by (81) if $4a \geqslant \sigma^2$ and (90) otherwise. We denote by $\mu_k$ and $\nu^n_k$ their probability transition measures. Then, there exists $l \in \mathbb{N}^*$, $C, \beta > 1$, such for all $f \in \mathcal{C}^6_b(\mathbb{R}_+;\mathbb{R})$, we have $E'_{n,\mathrm{pol}}(2,6)$ (see (15)) and*

$$\forall x \in \mathbb{R}_+, \quad |\mathbb{E}[f(X_T(x)] - \mathbb{E}[f(X^n_T(x)]| \leqslant C(1+|x|^\beta)\|f\|_{6,\infty}/n^2. \tag{92}$$

*Proof.* First we recall that the proof of (12) for the CIR diffusion and its scheme are given in [2]. If $4a \geqslant \sigma^2$, it is sufficient to use Theorem 4.1 **A.** to obtain (15) and then (74) follows from Property 2.1. Now, let $\sigma^2 > 4a$. Here, the only thing we still have to check is that (15) is satisfied for the moment matching scheme as soon as $x \in [0, K(T/n, 3/\sqrt{n})]$ with $K$ defined in (84). This is a consequence of the two following results.

**Lemma 4.3.** *Let us consider the case $\sigma^2 > 4a$. For $K$ defined in (84), $\lambda \in \mathbb{R}$, we have, for all $q \in \mathbb{N}$,*

$$\exists C_q > 0, t_q \in [0,1], \quad \forall t \in [0,t_q], x \in [0, K(t, \lambda\sqrt{t})[, \quad \mathbb{E}[X^q_t] \leqslant C_q t^q \tag{93}$$

*Proof.* Let us consider the function $g(x) = x^{\nu+1}$ for $x \geqslant 0$. Applying Ito's formula to $g$ for the CIR process gives :

$$X^{\nu+1}_t = x^{\nu+1} + \int_0^t \left[(\nu+1)X^\nu_s(a-kX_s) + \frac{1}{2}\nu(\nu+1)\sigma^2 X^\nu_s\right] ds + \int_0^t (\nu+1)X^{\nu+\frac{1}{2}}_s \sigma dW_s$$

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

$$\begin{aligned}
\mathbb{E}[|X^{\nu+1}_{t\wedge\mu_m}|] &\leqslant x^{\nu+1} + \mathbb{E}\Big[\Big|\int_0^{t\wedge\mu_m} \left[(\nu+1)X^\nu_s(a-kX_s) + \frac{1}{2}\nu(\nu+1)\sigma^2 X^\nu_s\right] ds\Big|\Big] \\
&\leqslant x^{\nu+1} + \int_0^t \left[(\nu+1)a + \frac{1}{2}\nu(\nu+1)\sigma^2\right] \mathbb{E}[|X^\nu_{s\wedge\mu_m}|] - k\mathbb{E}[X^{\nu+1}_{s\wedge\mu_m}]ds
\end{aligned}$$

Reasoning by induction, we assume that $\exists C_\nu > 0, t_\nu \in [0,1], \mathbb{E}[X^\nu_{s\wedge\mu_m}] \leqslant C_\nu t^\nu$. Since $x \in [0, K(t,\lambda\sqrt{t})[$ with $K(t) \underset{t\to 0}{=} O(t)$, $\exists t_{\nu+1} \in [0,1]$ such that :

$$\forall t \in [0,t_{\nu+1}], \mathbb{E}[|X^{\nu+1}_{t\wedge\mu_m}|] \leqslant \; K(t,\lambda\sqrt{t})^{\nu+1} + C_\nu\left((\nu+1)a + \tfrac{1}{2}\nu(\nu+1)\sigma^2\right)t^{\nu+1} + \int_0^t k\mathbb{E}[|X^{\nu+1}_{s\wedge\mu_m}|]ds$$

Applying Gronwall lemma and the fact that $0 \leqslant t \leqslant t_{\nu+1}$, and that there exists $C > 0$ such that $K(t,\lambda\sqrt{t}) < Ct$ we deduce :

$$\mathbb{E}[|X^{\nu+1}_{t\wedge\mu_m}|] \leqslant \left(C_\nu\left[(\nu+1)a + \frac{1}{2}\nu(\nu+1)\sigma^2\right] + C\right)e^k t^{\nu+1}$$

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

Using this result (15) will be a consequence of the following theorem

**Lemma 4.4.** *Let $\nu, \mu \in \mathbb{N}$, $x \in [0, C^*(t)[$ with $C^*(t) = \underset{t\to 0}{O}(t)$ and $(Y_t)_{t\geqslant 0} \in \mathbb{R}_+$ a random process such that $\forall q = 1, ..., \nu, \mathbb{E}[Y_t^q] = \mathbb{E}[X_t^q]$ with $(X_t)_{t\geqslant 0}$ the CIR process defined in (77) and $\forall q \in \leqslant \nu + 1$, there exists $C_q > 0, t_q^Y \in [0,1]$, $\forall t \in [0, t_q^Y], \mathbb{E}[|Y_t|^q] \leqslant C_q t^q$. Then the scheme with transition probability $\mathbb{P}(Y_{T/n} \in dx)$ satisfies (15) with $h = p = \nu + 1$.*

*Proof.* Let us write the Taylor expansion of $f$ in point $Y_t$ :

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

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

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

Moreover, $X_t$ has uniformly bounded moments and $f \in \mathcal{C}_b^\infty(\mathbb{R}_+)$, so using the Lemma 4.3, we have

$$\begin{aligned} \mathbb{E}[f(X_t) - f(Y_t)] &\leqslant C\|f\|_{\nu+1,\infty}\mathbb{E}\left|\frac{X_t^{\nu+1}}{\nu!} + \frac{Y_t^{\nu+1}}{\nu!}\right| \\ &\leqslant C\|f\|_{\nu+1,\infty} t^{\nu+1} \end{aligned}$$

and the proof is completed.

□

The proof of (92) is a direct application of this theorem for $\nu = 3$ and $Y$ given by (90). □

### 4.3.2 Convergence for measurable test function

Using the results of the previous section we are now able to study the total variation convergence of the CIR scheme with weak order 2. We will be able to prove the convergence for bounded measurable function but with support strictly contained in $\mathbb{R}_+$. We now introduce this space. Let $d_2 \geqslant d_1 > 0$ and define

$$\mathcal{D}_{cir} = [d_1, d_2] \tag{94}$$

Now, we provide a way to estimate the norms $\|\tilde{\psi}\|_{1,r,\infty}$, $r \in \mathbb{N}^*$. The first lemma treats several cases of interest given the form of the coefficients $V_i$.

**Lemma 4.5.** *Let $\mathcal{D}$ a compact subset of $\mathbb{R}^d$, $\mathcal{T}$ a compact subset of $\mathbb{R}$, and $v \in (0,1)$. We assume that $\phi_v^{\mathcal{D}}$ defined in (57) satisfy $|\phi_v^{\mathcal{D}}(x)| \leqslant |2x|$.*

***A.*** *We take $a_V, b_V \in \mathbb{R}^d$ and $V(x) = a_V + \langle b_V, x\rangle$. We denote by $\Phi_{V^{\mathcal{D},v}}$ the solution of (68) with $V$ replaced by $V^{\mathcal{D},v}$. Let $q \in \mathbb{N}$ and $\alpha$, $\beta$ two multi indexes such that $|\alpha| + |\beta| \leqslant q$. Then, there exists $C \geqslant 1$ such that, for all $x \in \mathbb{R}^d$,*

$$\sup_{t\in\mathcal{T}} |\partial_x^\alpha \partial_t^\beta \Phi_{V^{\mathcal{D},v}}(x,t)| \leqslant \begin{cases} C(1 + v^{-|\alpha|+1}) \exp(C|b_V|/v^{|\alpha|})(a_V^{|\beta|} + b_V^{|\beta|}), & \text{if } \alpha \geqslant 1 \\ C(|a_V|^q + |b_V|^q)(|x| + \|V^{\mathcal{D},v}\|_\infty), & \text{if } \alpha = 0 \end{cases} \tag{95}$$

***B.*** *Let $\zeta \in (0,1)$ such that $1/(1-\zeta) \in \mathbb{N}$, and assume that $V(x) = x^\zeta$. We denote by $\Phi_{V^{\mathcal{D},u}}$ the solution of (68) with $V$ replaced by $V^{\mathcal{D},u}$. Let $q \in \mathbb{N}$ and $\alpha$, $\beta$ two multi indexes such that $|\alpha| + |\beta| \leqslant q$. Then, there exists $C \geqslant 1$ such that, for all $x \in \mathbb{R}^d$,*

$$\sup_{t\in\mathcal{T}} |\partial_x^\alpha \partial_t^\beta \Phi_{V^{\mathcal{D},v}}(x,t)| \leqslant C(1 + v^{-\zeta/(1-\zeta)-|\alpha|})(1+|x|^{1-\zeta-|\alpha|}) \exp(Cv^{-\zeta/(1-\zeta)}) \tag{96}$$

*Proof.* In order to simplify the notations, we consider that $d = 1$. The proof for the multi dimensional case follows the same line so we will leave it out. We focus on **A.** first. Using (68), we have

$$\Phi_{V^{\mathcal{D},v}}(x,t) = x + \int_0^t V^{\mathcal{D},v}(\Phi_{V^{\mathcal{D},v}}(x,s))ds = x + a_V t + \int_0^t b_V \phi_v^{\mathcal{D}}(\Phi_{V^{\mathcal{D},v}}(x,s))ds$$

First we study the derivatives with respect to $x$. If we assume that. It follows that

$$\partial_x \Phi_{V^{\mathcal{D},u}}(x,t) = 1 + \int_0^t b_V \partial_x \phi_v^{\mathcal{D}}(\Phi_{V^{\mathcal{D},v}}(x,s)) \partial_x \Phi_{V^{\mathcal{D},v}}(x,s)ds$$

Now, we use (57) and it follows that

$$|\partial_x \Phi_{V^{\mathcal{D},v}}(x,t)| = 1 + \frac{C|b_V|}{v} \int_0^t |\partial_x \Phi_{V^{\mathcal{D},v}}(x,s)|ds.$$

Now, using the Gronwall inequality, we conclude $|\partial_x \Phi_{V^{\mathcal{D},u}}(x,t)| \leqslant \exp(C|b_V|t/u)$. Using a recursive approach we obtain for higher orders $|\partial_x^\alpha \Phi_{V^{\mathcal{D},u}}(x,t)| \leqslant C(1+u^{-|\alpha|+1}) \exp(C|b_V|t/v^{|\alpha|})$. Now, we study the derivative with respect to $t$. First, we notice that $\Phi_{V^{\mathcal{D},v}}(x,t) \leqslant |x| + t\|V^{\mathcal{D},v}\|_\infty$. From (68), we have $\partial_t \Phi_{V^{\mathcal{D},v}}(x,t) = a_V + b_V \phi_v^{\mathcal{D}}(\Phi_{V^{\mathcal{D},v}}(x,t))$ and we use these formula several times together with $|\phi_v^{\mathcal{D}}(x)| \leqslant |2x|$, in order to obtain (95).
Now, we prove **B.** Using (68) up to order $l = 1/(1-\zeta)$, if we denote $I_d$ the identity function (that is $I_d(x) = x$), we obtain

$$\Phi_{V^{\mathcal{D},v}}(x,t) = \sum_{i=0}^{l-1} \frac{t^i}{i!} (V^{\mathcal{D},v})^i I_d(x) + \frac{1}{l!} \int_0^t (t-s)^{l-1} (V^{\mathcal{D},v})^l (\Phi_{V^{\mathcal{D},v}}(x,s))ds$$

Now we remark that for an operator $V$, if $l \geqslant 2$, $V^i I_d(x)$ can be written has

$$V^i I_d(x) = \sum_{j=1}^{i-1} \sum_{\substack{|\gamma|=i-1 \\ \gamma\in\{1,\cdot,l\}^{i-j}}} C_{j,\gamma} V(x)^j \prod_{m=1}^{i-j} \partial_x^{\gamma_m} V(x),$$

with $C_{j,\gamma} \in \mathbb{R}$. Applying this decomposition to $V(x) = (\phi_v^{\mathcal{D}}(x))^\gamma$ and $i = l$, we show that $(V^{\mathcal{D},v})^l I_d(x)$ is a linear function of $\phi_v^{\mathcal{D}}(x)$ and moreover $(V^{\mathcal{D},v})^l I_d(x) \leqslant Cv^{-l+1}(1 + |(\phi_v^{\mathcal{D}}(x)|)$. Finally, we have for $i < l$,

$$|\partial_x^\alpha (V^{\mathcal{D},v})^i I_d(x)| = |\partial_x^\alpha \sum_{j=1}^{i-1} \sum_{\substack{|\gamma|=i-1 \\ \gamma\in\{1,\cdot,l\}^{i-j}}} C_{j,\gamma} V^{\mathcal{D},v}(x)^j \prod_{m=1}^{i-j} \partial_x^{\gamma_m} V^{\mathcal{D},v}(x)|$$
$$\leqslant C(1+v^{-i-|\alpha|+1})(1+|\phi_v^{\mathcal{D}}(x)|^{i(\zeta-1)+1-|\alpha|}) \leqslant C(1+v^{-i-|\alpha|+1})(1+|x|^{i(\zeta-1)+1-|\alpha|})$$

It remains to use the Gronwall lemma and we conclude

$$|\partial_x^\alpha \Phi_{V^{\mathcal{D},v}}(x,t)| \leqslant C(1+t^{l-1}v^{-l-|\alpha|+1})(1+|x|^{(l-1)(\zeta-1)+1-|\alpha|}) \exp(Ct^l v^{-l+1}),$$

and (96) follows using the same approach as in the proof of **A**. □

From Lemma 4.5, we immediatly obtain the following estimates

**Lemma 4.6.** *Let $v \in (0,1]$, $r \in \mathbb{N}^*$ and $\tilde{\psi}_{cir}$ the function defined as in (80) with $V_{0,cir}$ and $V_{1,cir}$ replaced by $V_{0,cir}^{\mathcal{D}_{cir},v}$ and $V_{1,cir}^{\mathcal{D}_{cir},v}$. Then, there exists $\mathfrak{C}_r \geqslant 1$ such that we have*

$$\|\tilde{\psi}_{cir}\|_{1,r,\infty} \leqslant \mathfrak{C}_r(1+d_1^{1/2-r}+d_2)(1+v^{-r-1})\exp(\mathfrak{C}_r/v^r) \tag{97}$$

with $\tilde{\psi}$ defined in Lemma 4.6. Moreover, we have $C_q(V_{cir}^{\mathcal{D}_{cir},2v}) \leqslant Cv^{-q}(1+d_1^{-q/2}+d_2)$, and the following ellipticity condition holds:

$$|V_{1,cir}^{\mathcal{D}_{cir},v}(x)|^2 \geqslant \sigma^2 d_1 > 0. \tag{98}$$

Now, we have all these estimates, we can state our main result for the second weak order convergence of the CIR process for bounded measurable functions.

**Theorem 4.4.** *Let $v \in (0,1]$ and $0 < d_1 \leqslant d_2$, $\delta \in (0,T)$. We us assume that $(X_t)_{t\geqslant 0}$ is the process given by (77) and $(X_t^n)_{t\in\pi_{T,n}}$ is defined by (81) if $4a \geqslant \sigma^2$ and (90) otherwise. Now, let $n_0 \in \mathbb{N}^*$ such that $2T/n_0 \leqslant \delta$, $(K(T/n_0, 3/\sqrt{n_0}) - 2v)_+ < d_1$ and that assumptions (49) and (50) hold with $n = n_0$ and $\psi$ replaced by $\tilde{\psi}_{cir}$, that is:*

$$\frac{3\|\tilde{\psi}_{cir}\|_{1,3,\infty}}{n_0^{1/4}} + \frac{M_8(Z)}{n_0} + \exp(-m_*^2 n_0 \delta/(2T)) \leqslant \frac{1}{2}, \quad n_0^{1/2} \geqslant 3\frac{2^{N+2}}{\lambda_*}\|\tilde{\psi}_{cir}\|_{1,3,\infty}^2.$$

*We also assume that (65) holds with $n = n_0$.*

*Then, there exists $C, \beta \geqslant 0$, $\mathfrak{l}_* \in \mathbb{N}$ for all bounded measurable test function $f$ on $\mathbb{R}^d$ with $\mathrm{supp}(f) \subset \mathcal{D}_{cir}^{2v} = [d_1+2v, d_2-2v]$, and $n \geqslant n_0$, then*

$$\mathbb{E}[f(X_T(x)) - f(X_T^n(x))] \leqslant C(\exp(-\frac{v^2/2 - \delta^2 a^2 \vee (kd_2)^2}{\delta\sigma^2 d_2}) + \exp(-\frac{C_Z \mathfrak{C}_1^{-1} v^4 \exp(\mathfrak{C}_1/v)}{2\delta(1+d_1^{-1/2}+d_2)^2}))\|f\|_\infty$$
$$+C(1+|x|^\beta)\frac{\Upsilon(v,d_1,d_2)^{\mathfrak{l}_*}}{(\delta d_1)^{42}}\|f\|_\infty/n^2. \tag{99}$$

*with* $\mathfrak{C}_r$, $r \in \mathbb{N}^*$, *is introduced in Lemma 4.6,* $C_Z$ *is defined in (66) and,*

$$\Upsilon(v, d_1, d_2) = v^{-6}(1 + d_2 + d_1^{-3})\mathfrak{K}_{0,9}(v, d_1, d_2), \tag{100}$$

*with*

$$\begin{aligned}\mathfrak{K}_{0,r}(v, d_1, d_2) =& \mathfrak{C}_r(1 + (1 + v^{-r-1})(1 + d_1^{1/2-r} + d_2)\exp(\mathfrak{C}_r/v^r)) \\ &\times \exp\left(\mathfrak{C}_3(1 + v^{-4})(1 + d_1^{-5/2} + d_2)\exp(\mathfrak{C}_3/v^3)\right).\end{aligned} \tag{101}$$

*Proof.* The proof follows from Theorem 3.1 (with Remark 3.3) and Lemma 4.6 since one can easily verifies that $(\mathbf{Reg}(\psi_{cir}, \tilde{\psi}_{cir}, v, q, n_0, \delta))$ holds. □

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

**Corollary 4.1.** *Let* $v \in (0, 1]$ *and* $0 < d_1 \leqslant d_2$, $\delta > 0$, $n_0 \in \mathbb{N}$. *We assume that the hypothesis from Theorem 4.4 are fulfilled with those parameters and that there exists a sequence* $(\rho_n)_{n\in\mathbb{N}}$ *taking positive values such that for all* $n \geqslant n_0$,

$$\delta \in [\underline{\delta}(n, v, d_1, d_2), \overline{\delta}(n, v, d_1, d_2)] \tag{102}$$

*with*

$$\underline{\delta}(n, v, d_1, d_2) = \frac{\exp\left(21^{-1}\mathfrak{l}_*\mathfrak{C}_3(1 + v^{-4})(1 + d_1^{-5/2} + d_2)\exp(\mathfrak{C}_3/v^3)\right)}{d_1\rho_n^{1/42}}$$

*and*

$$\overline{\delta}(n, v, d_1, d_2) = \frac{v^2}{4\sigma^2 d_2 \ln(n)} \wedge \frac{\ln(\rho_n)\sigma^2 d_2}{(a \vee k d_2)^2} \wedge \frac{v^4 C_Z}{4\mathfrak{C}_1(1 + d_1^{-1/2} + d_2)^2 \ln(n)}$$

*Then, there exists* $C, \beta \geqslant 0$, *such that for all bounded measurable test function* $f$ *with* $\mathrm{supp}(f) \subset \mathcal{D}_{cir}^{2v} = [d_1 + 2v, d_2 - 2v]$, *and* $n \geqslant n_0$,

$$\mathbb{E}[f(X_T(x)) - f(X_T^n(x))] \leqslant C(1 + |x|^\beta)\|f\|_\infty \rho_n / n^2. \tag{103}$$

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 $\delta$ which satisfies the hypothesis from Theorem 4.4 and (102) and for all bounded and measurable test function $f$ with $\mathrm{supp}(f) \subset [d_1 + 2v, d_2 - 2v]$

$$\mathbb{E}[f(X_T(x)) - f(X_T^n(x))] \leqslant C(1 + |x|^\beta)\|f\|_\infty \ln(n)^\zeta / n^2. \tag{104}$$

Moreover, through the sequence $(\rho_n)_{n\in\mathbb{N}}$, this result show 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 $\overline{\delta}$, one can identify the speeds of convergence of $v(n) \to 0$, $d_1(n) \to 0$ or $d_2(n) \to \infty$, with respect to $n$, which enable to obtain (102) and the hypothesis of Theorem 4.4 and consequently obtain a similar result to (104).

In particular, let $\epsilon > 0$ and define $\rho_n = n^\epsilon$. We fix $d_1$ and $v$ and put

$$d_2(n) = \zeta \frac{\epsilon C_v}{42} \ln(n) + v \quad \text{with} \quad C_v = \exp(-\mathfrak{C}_3/v^3)/(21^{-1}\mathfrak{l}_*\mathfrak{C}_3(1+v^{-4})) \quad \text{and} \quad \zeta \in [0,1)$$

We observe that in this case $\underline{\delta}(n, v, d_1, d_2(n)) = O_{n\to\infty}(n^{-(1-\zeta)\epsilon/42})$ and $\overline{\delta}(n, v, d_1, d_2(n)) = O_{n\to\infty}(\ln(n)^{-3})$ so for $n$ large enough we can find $\delta > 0$ wich satisfies the hypothesis from Theorem 4.4 and (102). 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 $\lambda \leqslant 2(1-\exp(-kt))/(k\sigma^2)$. Therefore, if the following hypothesis is satisfied:

$$\lambda_{v,\epsilon} := \frac{42(2-\epsilon)}{\zeta\epsilon C_v} \leqslant 2(1-\exp(-kT))/(k\sigma^2), \tag{105}$$

then $\mathbb{P}(X_t \geqslant d_2(n) - v) \leqslant C\exp(Cx)\exp(-\lambda_{v,\epsilon}(d_2(n) - v)) \leqslant C\exp(Cx)/n^{2-\epsilon}$. The reader may notice that we can always find $\epsilon \in (0,2)$ such that (105) holds. We conclude that there exists $n_0 \in \mathbb{N}$, $C, \beta \geqslant 0$, such that for all bounded and measurable test function $f$ with $\mathrm{supp}(f) \subset [d_1 + 2v, \infty)$, and $n \geqslant n_0$, we obtain

$$\mathbb{E}[f(X_T(x)) - Y_T^n(x)] \leqslant C(1+|x|^\beta)\|f\|_\infty/n^{2-\epsilon} + C\exp(\beta x)\|f\|_\infty/n^{2-\epsilon} \tag{106}$$

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.

*Proof of Corrolary 4.1.* The result is a consequence of Theorem 4.4. First, we remark that there exists $C \geqslant 1$ such that

$$d_1^{-42}\Upsilon(v, d_1, d_2)^{\mathfrak{l}_*} \leqslant C\exp\left(2\mathfrak{l}_*\mathfrak{C}_3(1+v^{-4})(1+d_1^{-5/2}+d_2)\exp(\mathfrak{C}_3/v^3)\right),$$

where $\Upsilon$ is defined in (100). Since $\delta \geqslant \underline{\delta}(n, v, d_1, d_2)$, it follows that

$$\Upsilon(v, d_1, d_2)/(d_1\delta)^{42} \leqslant C\rho_n$$

We study the other term in (99). Since $\delta \leqslant \overline{\delta}(n_0, v, d_1, d_2)$, it follows that

$$\begin{aligned}\exp(-\frac{v^2/2 - \delta^2 a^2 \vee (kd_2)^2}{\delta\sigma^2 d_2}) + \exp(-\frac{C_Z\mathfrak{C}_1^{-1}v^4\exp(\mathfrak{C}_1/v)}{2\delta(1+d_1^{-1/2}+d_2)^2}) \\ \leqslant C\rho_n/n^2 + C/n^{2\exp(\mathfrak{C}_1/v)} \leqslant C\rho_n/n^2,\end{aligned}$$

which completes the proof. □

# 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 semigroups in total variation distance. *Electron. J. Probab.*, 21:44 pp., 2016.

[5] 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.

[6] 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.

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

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

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

[10] 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.

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

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

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

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

[15] 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.

[16] 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.

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

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

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

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

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

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

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

[31] 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.

[32] 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.