Total variation convergence of the Euler-Maruyama scheme in small time with unbounded drift

Pierre Bras∗†, Gilles Pagès∗ and Fabien Panloup‡

Abstract

We give bounds for the total variation distance between the solution of a stochastic differential equation in \( \mathbb{R}^d \) and its one-step Euler-Maruyama scheme in small time. We show that for small \( t \), the total variation distance is of order \( t^{1/3} \), and more generally of order \( t^{r/(2r+1)} \) if the noise coefficient \( \sigma \) of the SDE is elliptic and \( C^2_b, r \in \mathbb{N} \), using multi-step Richardson-Romberg extrapolation. We also extend our results to the case where the drift is not bounded. Then we prove with a counterexample that we cannot achieve better bounds in general.

Keywords– Stochastic Differential Equation, Euler scheme, Total Variation, Richardson-Romberg extrapolation, Aronson’s bounds

MSC Classification– 65C30, 60H35

1 Introduction

The convergence properties of Euler-Maruyama schemes to approximate the solution of a Stochastic Differential Equation (SDE) have been extensively studied, in particular for \( L^p \) distances. However, the literature seems to lack some results about the convergence in total variation in small time. More specifically, let us consider the following SDE in \( \mathbb{R}^d \):

\[
X_0^x = x \in \mathbb{R}^d, \quad dX_t^x = b(X_t^x)dt + \sigma(X_t^x)dW_t
\]

along with its one-step Euler-Maruyama scheme

\[
\tilde{X}_t^x = x + tb(x) + \sigma(x)W_t,
\]

where \( W \) is a Brownian motion. We generally assume that \( b \) is Lipschitz continuous and that \( \sigma \) is elliptic, bounded and Lipschitz continuous, but we do not assume that \( b \) is bounded. Our objective is to give bounds of the total variation distance between the law of \( X_t^x \) and the law of \( \tilde{X}_t^x \), denoted \( d_{TV}(X_t^x, \tilde{X}_t^x) \), as \( t \to 0 \). Such bounds are well known for \( L^p \) distances and their associated Wasserstein distances (see for example Lemma A.2) and are known to be of order \( t \) as \( t \to 0 \). Yet the literature seems to lack results as it comes to \( d_{TV} \). If \( \sigma \) is constant, then it is classical background that \( d_{TV}(X_t^x, \tilde{X}_t^x) \) is of order \( t \), using a Girsanov change of measure (see for example [PP20, Proposition 4.2]) but this strategy cannot be applied to non-constant \( \sigma \). The difficulty of the total variation distance in small time is the following: considering its representation formula and comparing it with the \( L^1 \)-Wasserstein...
distance, if \( x \) and \( y \in \mathbb{R}^d \) are close to each other and if \( f : \mathbb{R}^d \to \mathbb{R} \) is Lipschitz continuous, then we can bound \( |f(x) - f(y)| \) by \( \|f\|_{\text{Lip}} |x - y| \); however if \( f \) is simply measurable and bounded, then we cannot directly bound \( |f(x) - f(y)| \) in terms of \( |x - y| \). Moreover the regularizing properties of the semi-group cannot be used in small time for the total variation distance.

In [BT96] is proved the convergence for a fixed time horizon \( T > 0 \) and as \( N \to \infty \), where \( N \) is the number of steps in the multi-step Euler-Maruyama scheme. More precisely, if \( \sigma \) is elliptic and if \( b \) and \( \sigma \) are \( C^\infty \) with bounded derivatives (but \( b \) and \( \sigma \) are not supposed bounded themselves), then ([BT96, Theorem 3.1])

\[
\forall x \in \mathbb{R}^d, \space \text{d}_{TV}(X^x_T, \bar{X}^x_{T,N}) \leq \frac{K(T)(1 + |x|^Q)}{NT^q},
\]

where \( X^x_T \) stands for the \( N \)-step Euler scheme, where \( Q \) and \( q \) are positive exponents and where \( K \) is a non-decreasing function depending on \( b \) and \( \sigma \). The common strategy of proof for such bounds is to use Malliavin calculus in order to perform an integration by parts and to use bounds on the derivatives of the density. However, we cannot infer a bound as \( T \to 0 \) since we do not have \( K(T)/T^q \to 0 \) as \( T \to 0 \) in general. In [GL08] are given bounds in small time and as \( N \to \infty \). Assuming that \( \sigma \) is uniformly elliptic and that \( b \) and \( \sigma \) are bounded with bounded derivatives up to order 3, then ([GL08, Theorem 3])

\[
\forall t \in (0, T], \forall x, y \in \mathbb{R}^d, \space |p(t, x, y) - \bar{p}^N(t, x, y)| \leq \frac{K(T)T}{N(t+1)^2} e^{-C|x-y|^2/t},
\]

where \( p \) and \( \bar{p}^N \) denote the transition densities of \( X^x \) and \( X^x_{T,N} \) respectively and where \( C \) is a positive constant depending on \( d \) and on the bounds on \( b \) and \( \sigma \) and their derivatives. However, we cannot directly use this result for the total variation distance: taking \( N = 1 \) yields

\[
\text{d}_{TV}(X^x_T, \bar{X}^x_T) = \int_{\mathbb{R}^d} |p(t, x, y) - \bar{p}^N(t, x, y)| dy \leq K(T)Tt^{-1/2} \int_{\mathbb{R}^d} \frac{1}{t^{d/2}} e^{-C|x-y|^2/t},
\]
giving a bound in \( t^{-1/2} \) which does not converge to 0 as \( t \to 0 \). Moreover, [GL08] assumes that \( b \) is bounded. [BJ20] focuses on the case where \( b \) is bounded and measurable but not necessarily regular and where \( \sigma \) is constant; it proves that the convergence in total variation of the multi-step Euler scheme which is regularized with respect to the irregular drift \( b \) and with step \( h \), is of order \( \sqrt{h} \).

In the present paper, we prove a convergence rate of order \( t^{1/3} \) for \( \text{d}_{TV}(X^x_T, \bar{X}^x_T) \). More generally, if we assume that \( \sigma \) is \( C^2_b \) and that \( b \) is \( C^1 \) and Lipschitz continuous (but we do not assume that \( b \) is bounded itself), then we obtain a convergence rate of order \( t^{1/(2r+1)} \). Letting \( r \to \infty \), we also prove that if \( \sigma \in C^\infty_b \) with some technical condition on the derivatives of the density, then the convergence rate is of order \( t^{1/2} \exp(C\sqrt{\log(1/t)}) \) which is "almost" \( t^{1/2} \). Moreover, we provide an example using the geometric Brownian motion where the convergence rate is exactly \( t^{1/2} \), thus showing that we cannot achieve better bounds in general. To prove the bound in \( t^{1/(2r+1)} \), we use a multi-step Richardson-Romberg extrapolation [RG11] [LP17], which is a method imported from numerical analysis that we use in our case for theoretical purposes. It relies on a Taylor expansion with null coefficients up to some high order. Such method can be used in more general settings with regularization arguments in order to improve the convergence rates (in our case, we improve \( t^{1/3} \) into \( t^{1/(2r+1)} \)).

Extending our results to the case where \( b \) is not bounded is challenging. Indeed, the total variation distance is closely related to the estimation of the density of the solution to an SDE and this density satisfies a Fokker-Planck Partial Differential Equation (PDE) (3.2). If \( b \) is bounded, then the density \( p(t, x, y) \) and its partial derivatives admit sub-gaussian bounds (see [Fri64] and Section 3.1). However, giving estimates and bounds for the solution of the PDE in the case of unbounded \( b \) appears to be more difficult, see [Lun97], [Cer00], [BL05]. In [MPZ21] are given bounds in the unbounded drift case for the derivatives of the density with respect to the starting point \( x \), nonetheless finding bounds for the derivatives with respect to the final point \( y \) is still an open problem. Studying this case is useful to...
study the convergence in total variation of SDE’s with unbounded drift, in particular for the Langevin equation, which is popular in stochastic optimization and which reads
\[ dX_t = -\nabla V(X_t)dt + \sigma(X_t)dW_t, \]
where in many cases, \( V: \mathbb{R}^d \to \mathbb{R} \) has quadratic growth (see for example [BP21]) and \( \nabla V \) has linear growth.

In order to deal with unbounded \( b \), our strategy of proof is the following. Since we are mainly interested in the total variation distance, we do not state any differentiability property with respect to \( y \) for the density with unbounded \( b \). Instead, we use a localization argument and "cut" the drift \( b \) outside a compact set, so that we can use bounds from [Fri64] for the bounded drift case. We use the Girsanov formula to explicit the dependence of these bounds in \( \|\hat{b}\|_\infty \).

**Notations**

We endow the space \( \mathbb{R}^d \) with the canonical Euclidean norm denoted by \( |\cdot| \). For \( x \in \mathbb{R}^d \) and for \( R > 0 \), we denote \( \mathcal{B}(x, R) = \{ y \in \mathbb{R}^d : |y - x| \leq R \} \).

For \( M \in (\mathbb{R}^d)^{\otimes k} \), we denote by \( \|M\| \) its operator norm, i.e. \( \|M\| = \sup_{u \in \mathbb{R}^d, \|u\|=1} M \cdot u \). If \( M: \mathbb{R}^d \to (\mathbb{R}^d)^{\otimes k} \), we denote \( \|M\|_\infty = \sup_{x \in \mathbb{R}^d} \|M(x)\| \). We say that \( M \) is \( C^k \) for some \( r \in \mathbb{N} \cup \{0\} \) if \( M \) is bounded and has bounded derivatives up to the order \( r \).

For \( k \in \mathbb{N} \) and if \( f: \mathbb{R}^d \to \mathbb{R} \) is \( C^k \), we denote by \( \nabla^k f: \mathbb{R}^d \to (\mathbb{R}^d)^{\otimes k} \) its differential of order \( k \). If \( f \) is Lipschitz continuous, we denote by \( [f]_{\text{Lip}} \) its Lipschitz constant.

We denote the total variation distance between two distributions \( \pi_1 \) and \( \pi_2 \) on \( \mathbb{R}^d \):
\[ d_{TV}(\pi_1, \pi_2) = 2 \sup_{A \in \mathcal{B}(\mathbb{R}^d)} |\pi_1(A) - \pi_2(A)|. \]

Without ambiguity, if \( Z_1 \) and \( Z_2 \) are two \( \mathbb{R}^d \)-valued random vectors, we also write \( d_{TV}(Z_1, Z_2) \) to denote the total variation distance between the law of \( Z_1 \) and the law of \( Z_2 \). We have as well
\[ d_{TV}(\pi_1, \pi_2) = \sup \left\{ \int_{\mathbb{R}^d} \left| f d\pi_1 - \int_{\mathbb{R}^d} f d\pi_1, f: \mathbb{R}^d \to [-1, 1] \text{ measurable} \right| \right\}. \]

Moreover, we recall that if \( \pi_1 \) and \( \pi_2 \) admit densities with respect to some measure \( \lambda \), then
\[ d_{TV}(\pi_1, \pi_2) = \int_{\mathbb{R}^d} \left| \frac{d\pi_1}{d\lambda} - \frac{d\pi_2}{d\lambda} \right| d\lambda. \]

For \( x \in \mathbb{R}^d \), we denote by \( \delta_x \) the Dirac mass at \( x \).

If \( Z \) is a Markov time-homogeneous process with values in \( \mathbb{R}^d \), we denote, when it exists, its transition probability from \( x \) to \( y \) after time \( t \), \( p_x(t, x, y) \).

In this paper, we use the notation \( C \) and \( c \) to denote positive constants, which may change from line to line.

**2 Main results**

We consider the SDE in \( \mathbb{R}^d \):
\[ X_0^x = x \in \mathbb{R}^d, \quad dX_t^x = b(X_t^x)dt + \sigma(X_t^x)dW_t \quad (2.1) \]
along with its one-step Euler-Maruyama scheme
\[ X_t^x = x + tb(x) + \sigma(x)W_t, \quad (2.2) \]
where \( b: \mathbb{R}^d \to \mathbb{R}^d \), where \( \sigma: \mathbb{R}^d \to \mathcal{M}_d(\mathbb{R}) \) and where \( W \) is a standard \( \mathbb{R}^d \)-valued Brownian motion defined on a probability space \( (\Omega, \mathcal{A}, \mathbb{P}) \). We assume that \( \sigma \) is elliptic i.e.
\[ \exists \sigma_0 > 0, \forall x \in \mathbb{R}^d, \sigma(x)\sigma(x)^\top \geq \sigma_0^2 I_d. \quad (2.3) \]
Theorem 2.1. Let \( X \) be the solution of the SDE \((2.1)\). Assume that \( \sigma \) is \( C^2_b \) and elliptic \((2.3)\) and that \( b \) is \( C^1 \) with bounded derivatives (but we do not assume that \( b \) is bounded itself). Let \( T > 0 \) be a finite time horizon. Then

\[
\forall t \in [0,T], \forall x \in \mathbb{R}^d, \quad d_{TV}(X^r_t, X^r_t) \leq C e^{c|x|^2} t^{1/3},
\]

where the positive constants \( C \) and \( c \) only depend on \( d, T, \sigma_0, \|\sigma\|_\infty, [b]_{\text{Lip}}, [\sigma]_{\text{Lip}} \) and \( \|\nabla^2 \sigma\|_\infty \).

Theorem 2.2. Let \( X \) be the solution of the SDE \((2.1)\). Assume that \( \sigma \) is elliptic and is \( C^2_b \) for some \( r \in \mathbb{N} \) and that \( b \) is \( C^1 \) with bounded derivatives (but we do not assume that \( b \) is bounded itself). Let \( T > 0 \) be a finite time horizon. Then

\[
\forall t \in [0,T], \forall x \in \mathbb{R}^d, \quad d_{TV}(X^r_t, X^r_t) \leq C e^{c|x|^2} t^{1/3},
\]

where the positive constants \( C \) and \( c \) only depend on \( d, T, \sigma_0, \|\sigma\|_\infty \) and on the bounds on the derivatives of \( b \) and \( \sigma \).

To improve the rate of convergence from \( t^{1/3} \) in Theorem 2.1 to \( t^{r/(2r+1)} \) in Theorem 2.2, we rely on a Richardson-Romberg extrapolation [RG11] [LP17] that we apply in our case to an SDE and its Euler-Maruyama scheme, however this argument can also be applied in a more general framework as stated right after.

Proposition 2.3. Let \( Z_1 \) and \( Z_2 \) be two random vectors in \( L^1(\mathbb{R}^d) \) and admitting densities \( p_1 \) and \( p_2 \) respectively with respect to the Lebesgue measure. Assume furthermore that \( p_1 \) and \( p_2 \) are \( C^2r \) with \( r \in \mathbb{N} \) and that \( \nabla^k p_1 \in L^1(\mathbb{R}^d) \) for \( i = 1, 2 \) and for \( k = 1, \ldots, 2r \). Then for every \( \varepsilon > 0 \) we have

\[
d_{TV}(Z_1, Z_2) \leq C_{d,r} \varepsilon^r \int_{\mathbb{R}^d} (\|\nabla^2 r p_1(x)\| + \|\nabla^2 r p_2(x)\|) \, dx + C_{d,r} \varepsilon^{-1/2} W_1(Z_1, Z_2),
\]

where the constant \( C_{d,r} \) depends only on \( d \) and on \( r \).

If \( \sigma \in C^\infty_0 \), then we also prove that we can "almost" get a convergence rate of order \( t^{1/2} \).

Theorem 2.4. Let \( X \) be the solution of the SDE \((2.1)\). Assume that \( \sigma \) is elliptic and is \( C^\infty_0 \) and that \( b \) is \( C^1 \) with bounded derivatives (but we do not assume that \( b \) is bounded itself). Let \( T > 0 \) be a finite time horizon. Assume furthermore that if \( Z \) is the martingale \( dZ_t = \sigma(Z_t) dW_t \), then

\[
\forall t \in (0,T], \forall x, y \in \mathbb{R}^d, \quad \|\nabla^2 r p_x(t, x, y)\| \leq \frac{C_{2r}}{t^{(d+2r)/2}} e^{-c_2 t^{-1/2}},
\]

with \( \limsup_{r \to \infty} \left( C_{2r} e^{d/2} \right)^{1/(2r)} < \infty \).

(see Theorem 3.1). Then

\[
\forall t \in [0,T], \forall x \in \mathbb{R}^d, \quad d_{TV}(X^r_t, X^r_t) \leq C e^{c|x|^2} t^{1/2} e^{\sqrt{\log(1/t)}},
\]

where the positive constants \( C \) and \( c \) only depend on \( d, T, \sigma_0, \|\sigma\|_\infty \) and on the bounds on the derivatives of \( b \) and \( \sigma \).

Remark 2.5. Assumption \((2.6)\) is satisfied in the case of a Brownian motion, which suggests that this assumption is satisfied in general provided that \( \sigma \) is "regular enough". Indeed, if \( Z^r_t = x + \sigma W_t \) then with \( \Sigma := \sigma \sigma^T \) we have

\[
p_x(t, x, y) = \frac{1}{\sqrt{\det(\Sigma) t^{d/2}}} \Phi \left( \Sigma^{-1/2} y - x / \sqrt{t} \right), \quad \Phi(u) = \frac{1}{(2\pi)^{d/2}} e^{-|u|^2/2}.
\]
Moreover for every \( r \in \mathbb{N} \) we have
\[
\left\| \frac{d^r}{du^r} \Phi(u) \right\| \leq \frac{1}{(2\pi)^{d/2}} |\text{He}_r(|u|)| e^{-|u|^2/2}
\]
where \( \text{He}_r \) is the \( r \)-th probabilist Hermite polynomial. Moreover following [Kra04] we have
\[
\forall u \in \mathbb{R}^+, \quad |\text{He}_{2r}(u)| e^{-u^2/2} \leq C 2^{-r} \sqrt{r} \frac{(2r)!^2}{r^r} \leq C 2^r \sqrt{r},
\]
using the Stirling formula for the last inequality. On the other hand, using \([\text{AS64, 22.14.15}]\) we have
\[
\forall u \in \mathbb{R}^+, \quad |\text{He}_{2r}(u)| e^{-u^2/4} \leq 2^{r+1} r!.
\]
Then, for every \( \varepsilon \in (0,1) \),
\[
|\text{He}_{2r}(u)| e^{-u^2/2} = |\text{He}_{2r}(u)| e^{-u^2/4} \left|\text{He}_{2r}(u)| e^{-u^2/2}\right|^{1-\varepsilon} e^{-\varepsilon u^2/4} \leq C (2^r r!)^{\varepsilon} \left(2^r r^{1/2}\right)^{1-\varepsilon} e^{-\varepsilon u^2/2}.
\]
Then if we choose \( \varepsilon_r = \log^{-1}(r) \), we have
\[
(2^r r!)^{\varepsilon} \leq e^{e_r r \log(r)} = e^r
\]
so that
\[
\left\| \frac{d^r}{du^r} \Phi(u) \right\| \leq C 2^{r^2} e^r 2^r r^{1/2} e^{-r^2 u^2/2} =: A_r e^{-e_r |u|^2/2}
\]
and then
\[
\left\| \nabla_y^{2r} p_x(t,x,y) \right\| \leq \frac{\|\Sigma^{-1/2}\|^{2r}}{\sqrt{\det(\Sigma)}} \frac{d^r}{du^r} \Phi \left( \frac{\Sigma^{-1/2} y - x}{\sqrt{t}} \right) \leq \frac{\|\Sigma^{-1/2}\|^{2r}}{\sqrt{\det(\Sigma)}} A_r e^{-e_r |x-y|^2/(2t)}
\]
where \( \left(\|\Sigma^{-1/2}\|^{2r} A_r e^{-d/2}\right)^{1/(2r)} \) is bounded. Thus Assumption (2.6) is satisfied.

3 Proof of the Theorems

3.1 Recalls on density estimates for SDEs with bounded drift

We recall results on the bounds for the density of the solution of the SDE using the theory of partial differential equations. Let us consider a generic SDE:
\[
Z_0^x = x \in \mathbb{R}^d, \quad dZ_t^x = b(Z_t^x)dt + \sigma(Z_t^x) dW_t,
\]
where \( \sigma \) is elliptic. Then the transition probability \( p_x \) is solution of the backward Kolmogorov PDE:
\[
p_x(0,x,) = \delta_x, \quad \partial_t p_x = \langle b(Z_t^x), \nabla p_x \rangle + \frac{1}{2} \text{Tr}(\nabla^2 Z_t^x p_x \cdot \sigma(Z_t^x)).
\]
Moreover, \( p_x \) and its derivatives satisfy sub-gaussian bounds:

**Theorem 3.1** ([Fri64], Chapter 9, Theorem 7). Let \( Z \) be the solution of (3.1) and let \( T > 0 \). Assume that \( \sigma \) is elliptic and that \( b \) and \( \sigma \) are \( C^r \) for some \( r \in \mathbb{N} \). Then for every \( m_0 = 0,1 \) and for every \( 0 \leq m_1 + m_2 \leq r \),
\[
\forall t \in (0,T), \quad \forall x,y \in \mathbb{R}^d, \quad \left\| \nabla_x^{m_0} \nabla_y^{m_1} p_x(t,x,y) \right\| \leq \frac{C}{t^{(d+m_0+m_1+m_2)/2}} e^{-c|y-x|^2/t},
\]
where the constants \( C \) and \( c \) only depend on the bounds on \( b \) and \( \sigma \) and on their derivatives, on the modulus of ellipticity of \( \sigma \), on \( d \) and on \( T \).
3.2 Preliminary results

In order to apply the bounds from Theorem 3.1, we first "cut" the drift $b$ on a compact set. That is, we instead consider the process $Y$ defined by

$$Y^x_0 = x, \quad dY^x_t = \tilde{b}^x(Y^x_t)dt + \sigma(Y^x_t)dW_t,$$

(3.4)

where $\tilde{b}$ is defined as follows. We choose $R > 0$ and we consider a $C^\infty$ decreasing function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\psi = 1$ on $[0, R^2]$ and $\psi = 0$ on $[(R + 1)^2, \infty)$ and we define $\tilde{b}^x(y) := b(y)\psi(|y - x|^2)$, so that $\tilde{b}^x$ is bounded:

$$\forall y \in \mathbb{R}^d, \quad |\tilde{b}^x(y)| \leq \sup_{z \in \mathbb{B}(x, R+1)} |b(z)| \leq C(1 + |x|),$$

(3.5)

because $b$ is Lipschitz continuous.

Lemma 3.2. We have

$$d_{TV}(Y^x_t, X^x_t) \leq C(1 + |b(x)|^2)t.$$

(3.6)

Proof. Let $f : \mathbb{R}^d \to \mathbb{R}$ be measurable and bounded. We remark that on the event $\{\sup_{s \in [0,t]} |X^x_s - x|^2 \leq R^2\}$, we have $Y^x_t = X^x_t$, so that

$$|\mathbb{E}f(Y^x_t) - \mathbb{E}f(X^x_t)| \leq 2\|f\|_\infty \mathbb{P}\left(\sup_{s \in [0,t]} |X^x_s - x|^2 > R^2\right).$$

But using the inequality $|u + v|^2 \leq 2|u|^2 + 2|v|^2$ we have

$$|X^x_t - x|^2 \leq 2\left|\int_0^t b(X^x_s)ds\right|^2 + 2\left|\int_0^t \sigma(X^x_s)dW_s\right|^2 \leq 2t \int_0^t |b(X^x_s)|^2ds + 2\left|\int_0^t \sigma(X^x_s)dW_s\right|^2 \leq 4t|b|_{Lip}^2 \int_0^t |X^x_s - x|^2ds + 4t^2|b(x)|^2 + 2\left|\int_0^t \sigma(X^x_s)dW_s\right|^2,$$

so that

$$\mathbb{E}\sup_{s \in [0,t]} |X^x_s - x|^2 \leq 4t|b|_{Lip}^2 \int_0^t \left(\mathbb{E}\sup_{u \in [0,s]} |X^x_u - x|^2\right)ds + 4t^2|b(x)|^2 + 2\mathbb{E}\sup_{s \in [0,t]} \left|\int_0^t \sigma(X^x_u)dW_u\right|^2.$$

Moreover using Doob’s martingale inequality we have

$$\mathbb{E}\sup_{s \in [0,t]} \left|\int_0^s \sigma(X^x_u)dW_u\right|^2 \leq 4\mathbb{E}\left|\int_0^t \sigma(X^x_u)dW_u\right|^2 = 4\mathbb{E}\int_0^t \sigma^2(X^x_u)du \leq 4\|\sigma\|_{Lip}^2 t.$$

Then we define the non-decreasing deterministic process

$$S_t := \mathbb{E}\sup_{s \in [0,t]} |X^x_s - x|^2$$

and we get the differential inequality (using $t^2 \leq tT$)

$$S_t \leq 4t(T|b(x)|^2 + 2\|\sigma\|_{Lip}^2) + 4t|b|_{Lip}^2 \int_0^t S_udu,$$

implying that

$$S_t \leq 4t(T|b(x)|^2 + 2\|\sigma\|_{Lip}^2)e^{2t^2|b|_{Lip}^2} \leq C(1 + |b(x)|^2)t.$$
Using Markov’s inequality, we have then
\[
\mathbb{P} \left( \sup_{s \in [0,t]} |X_s^x - x|^2 > R^2 \right) \leq \frac{C(1 + |b(x)|^2)t}{R^2}
\]
and then
\[
|\mathbb{E} f(Y_s^x) - \mathbb{E} f(X_t^x)| \leq 2\|f\|_{\infty} \frac{C(1 + |b(x)|^2)t}{R^2}.
\]

We can now apply Theorem 3.1 to $Y$ however the constants arising depend on the bound on $\|\tilde{b}^x\|_{\infty}$ and thus on $x$. In order to deal with the dependency in $\|\tilde{b}^x\|_{\infty}$, we apply the Girsanov formula and reduce to the null drift case.

**Proposition 3.3.** Let $Z$ be the solution of
\[
Z_0^x = x, \quad dZ_t^x = \sigma(Z_t^x)dW_t.
\]
Then we have for every $t \in [0,T]$, $x, y \in \mathbb{R}^d$,
\[
p_x(t,x,y) = p_x(t,x,y) + \int_0^t \mathbb{E} \left[ U_s^x (\tilde{b}^x(Z_s^x), \nabla_x p_x(t-s, Z_s^x, y)) \right] ds,
\]
where $Y$ is defined in (3.4) and $U$ is defined as
\[
U_s^x = \exp \left( \int_0^s \langle g(Z_u^x)\tilde{b}^x(Z_u^x), dZ_u^x \rangle - \frac{1}{2} \int_0^s \langle g(Z_u^x)\tilde{b}^x(Z_u^x), \tilde{b}^x(Z_u^x) \rangle du \right),
\]
\[
g = (\sigma \sigma^\top)^{-1}.
\]

**Proof.** First, note that since $\sigma$ is elliptic (2.3) and since $\tilde{b}^x$ and $\sigma$ are bounded, $C^1$ and Lipschitz continuous, then $p_x$ and $p_x^y$ exist as well as $\nabla_x p_x$ (Theorem 3.1). We then use [QZ04, Theorem 2.4]. Following [QZ04, Remark 2.5], since $\sigma$ is elliptic and bounded, then the assumptions of [QZ04, Theorem 2.4] hold.

We also have the following bounds on the process $U$.

**Lemma 3.4.** For every $p \geq 2$ we have
\[
\mathbb{E} \left[ \sup_{s \in [0,t]} |U_s^x|^p \right] \leq |x|^p e^{Cp^2\|\tilde{b}^x\|_{\infty}^2 t}.
\]

**Proof.** We recall that for every $q \geq 1$, the process $U^q$ is a martingale with
\[
d(U_s^x)^q = q(U_s^x)^{q-1} \langle g(Z_s^x)\tilde{b}^x(Z_s^x), \sigma(Z_s^x) dW_s \rangle.
\]
Thus, Doob’s martingale inequality yields
\[
\mathbb{E} \left[ \sup_{s \in [0,t]} |U_s^x|^{2q} \right] \leq Cq^2\|\tilde{b}^x\|_{\infty}^2 \mathbb{E} \int_0^t |U_s^x|^{2q} ds \leq Cq^2\|\tilde{b}^x\|_{\infty}^2 \int_0^t \mathbb{E} \left[ \sup_{u \in [0,s]} |U_u^x|^{2q} \right] ds.
\]
Using the Gronwall inequality we obtain
\[
\mathbb{E} \left[ \sup_{s \in [0,t]} |U_s^x|^{2q} \right] \leq |x|^{2q} e^{Cq^2\|\tilde{b}^x\|_{\infty}^2 t}.
\]
Lemma 3.5. We have
\[
\left| \int_0^t \mathbb{E} \left[ U_s^x (\tilde{b}^x(Z_s^x), \nabla_x p_Z(t - s, Z_s^x, y)) \right] ds \right| \leq C e^{C\|\tilde{b}^x\|_2^2 t} e^{-c|y-x|^2/t} \frac{1}{t^{(d-1)/2}} \tag{3.9}
\]

Proof. Using Theorem 3.1 on the process $Z$, which yields bounds with constants depending on $\sigma$ but not on $b$, for every $q \geq 1$ and for every $s \in [0, t]$: \[
\mathbb{E} |\nabla_x p_Z(t - s, Z_s^x, y)|^q = \int_{\mathbb{R}^d} |\nabla_x p_Z(t - s, \xi, y)|^q p_Z(s, x, \xi) d\xi
\]
\[
\leq \frac{C_q}{(t - s)^{(q+1)d)/2}} \int_{\mathbb{R}^d} \frac{1}{s(t - s)^{d/2}} \exp \left( -c_q \left( \frac{|y - \xi|^2}{t - s} + \frac{|\xi - x|^2}{s} \right) \right) d\xi
\]
\[
\leq \frac{C_q}{t^{d/2} (t - s)^{(q+1)d)/2}} e^{-c_q|y-x|^2/t},
\]
where we used Lemma A.1 in the appendix. Then for $p^{-1} + q^{-1} = 1$ and $p \geq 2$, using the Hölder inequality we have
\[
\left| \int_0^t \mathbb{E} \left[ U_s^x (\tilde{b}^x(Z_s^x), \nabla_x p_Z(t - s, Z_s^x, y)) \right] ds \right|
\]
\[
\leq ||\tilde{b}^x||_\infty \left( \sup_{s \in [0, t]} \mathbb{E} |U_s^x|^{p} \right)^{1/p} \int_0^t \mathbb{E} |\nabla_x p_Z(t - s, Z_s^x, y)|^{q} ds
\]
\[
\leq ||\tilde{b}^x||_\infty |x| e^{Cp||\tilde{b}^x||_\infty^2 t} \frac{C_q e^{-c_q|y-x|^2/t}}{t^{d/2}} \int_0^t (t - s)^{(1+(1-q^{-1})d)/2} ds.
\]
The integral in $ds$ converges under the condition $q < d/(d-1)$ if $d > 1$, and for any value of $q > 1$ if $d = 1$. Then performing the change of variable $s = tu$ we obtain \[
\left| \int_0^t \mathbb{E} \left[ U_s^x (\tilde{b}^x(Z_s^x), \nabla_x p_Z(t - s, Z_s^x, y)) \right] ds \right| \leq ||\tilde{b}^x||_\infty |x| e^{Cp||\tilde{b}^x||_\infty^2 T} \frac{C_q e^{-c_q|y-x|^2/t}}{t^{(d-1)/2}} \leq C e^{C\|\tilde{b}^x\|_2^2 t} e^{-c|y-x|^2/t} \frac{1}{t^{(d-1)/2}}.
\]

\[\square\]

3.3 Proof of Theorem 2.1

Proof. Let us introduce an artificial regularization. For $\varepsilon > 0$ we write \[
\begin{align*}
d_{TV}(\tilde{X}_t^\varepsilon, X_t^\varepsilon) &\leq d_{TV}(\tilde{X}_t^\varepsilon, Z_t^\varepsilon) + d_{TV}(Z_t^\varepsilon, Z_t^\varepsilon + \sqrt{\varepsilon} \zeta) + d_{TV}(Z_t^\varepsilon + \sqrt{\varepsilon} \zeta, Z_t^\varepsilon + \sqrt{\varepsilon} \zeta) \\
&\quad + d_{TV}(Z_t^\varepsilon + \sqrt{\varepsilon} \zeta, Z_t^\varepsilon) + d_{TV}(Z_t^\varepsilon, Y_t^\varepsilon) + d_{TV}(Y_t^\varepsilon, X_t^\varepsilon) \\
&=: D_1 + D_2 + D_3 + D_4 + D_5 + D_6 \tag{3.10}
\end{align*}
\]
where $\zeta \sim \mathcal{N}(0, I_d)$ and is independent of the Brownian motion $W$, where $Y$ is defined in (3.4), where $Z$ is defined in (3.7) and where $\tilde{Z}$ denotes the one-step Euler-Maruyama scheme associated to $Z$. Let us now investigate each term.

• Term $D_5$: Using the formula (3.8) and the inequality (3.9), we have \[
\begin{align*}
d_{TV}(Z_t^\varepsilon, Y_t^\varepsilon) &= \int_{\mathbb{R}^d} |p_z(t, x, y) - p_{Y}(t, x, y)| dy = \int_{\mathbb{R}^d} \left| \int_0^t \mathbb{E} \left[ U_s^x (\tilde{b}^x(Z_s^x), \nabla_x p_Z(t - s, Z_s^x, y)) \right] ds \right| dy \\
&\leq C e^{C||\tilde{b}^x||_2^2 t^{1/2}} \int_{\mathbb{R}^d} e^{-c|x-y|^2/t} \frac{1}{t^{d/2}} dy \leq C e^{C|x|^2 t^{1/2}},
\end{align*}
\]
where we used (3.5). The term \( D_1 \) is treated the same way.

- **Term** \( D_3 \): Let \( f : \mathbb{R}^d \to \mathbb{R} \) be measurable and bounded and let us define
  \[
  \varphi : y \in \mathbb{R}^d \mapsto \mathbb{E}_f(Z_t^x + y) = \int_{\mathbb{R}^d} f(\xi + y)p_z(t, x, \xi)d\xi = \int_{\mathbb{R}^d} f(\xi)p_z(t, x, \xi - y)d\xi.
  \] (3.11)

Then \( \varphi \) is \( C^2 \) with
\[
\nabla^2 \varphi(y) = \int_{\mathbb{R}^d} f(\xi)\nabla^2 p_z(t, x, \xi - y)d\xi.
\]

Moreover, using Theorem 3.1, we have
\[
\|\nabla^2 \varphi(y)\| \leq C\|f\|\infty t^{-1} \int_{\mathbb{R}^d} \frac{1}{\epsilon^{(d+2)/2}} e^{-c|x-\epsilon|^2/\epsilon^t} d\xi \leq C\|f\|\infty t^{-1}.
\]

Then using the Taylor formula, for every \( y \in \mathbb{R}^d \) there exists \( \tilde{y} \in (0, y) \) such that
\[
\varphi(y) = \varphi(0) + \nabla \varphi(0) \cdot y + \frac{1}{2} \nabla^2 \varphi(\tilde{y}) \cdot y^{\otimes 2}
\]
and then for some random \( \tilde{\zeta} \in (0, \zeta) \) we have
\[
|\mathbb{E}_f(Z_t^x + \sqrt{\epsilon}\zeta) - \mathbb{E}_f(Z_t^x)| = |\mathbb{E}_f(\sqrt{\epsilon}\zeta) - \varphi(0)| = |\sqrt{\epsilon}\mathbb{E}[\nabla \varphi(0) \cdot \zeta] + \frac{\epsilon}{2} \mathbb{E}[\nabla^2 \varphi(\sqrt{\epsilon}\zeta) \cdot \zeta^{\otimes 2}]|
\leq C\epsilon \|f\|\infty t^{-1},
\]

where we used that \( \mathbb{E}[\nabla \varphi(0) \cdot \zeta] = \nabla \varphi(0) \cdot \mathbb{E}[\zeta] = 0 \). This way we obtain
\[
D_3 \leq C\epsilon t^{-1}.
\]

The term \( D_2 \) is treated likewise.

- **Term** \( D_4 \): Using Lemma 3.2 with (3.5), we have \( D_4 \leq C(1 + |x|^2)t \).

- **Term** \( D_5 \): Let \( f : \mathbb{R}^d \to \mathbb{R} \) be measurable and bounded and let us define
\[
f_{\epsilon} : y \mapsto \mathbb{E}_f(y + \sqrt{\epsilon} \zeta) = \frac{1}{(2\pi\epsilon)^{d/2}} \int_{\mathbb{R}^d} f(y + \sqrt{\epsilon} \zeta)e^{-|\xi|^2/2\epsilon}d\xi = \frac{1}{(2\pi\epsilon)^{d/2}} \int_{\mathbb{R}^d} f(\xi)e^{-|\xi - y|^2/(2\epsilon)}d\xi.
\] (3.12)

Then \( f_{\epsilon} \) is \( C^1 \) with
\[
\nabla f_{\epsilon}(y) = \frac{1}{(2\pi\epsilon)^{d/2}} \int_{\mathbb{R}^d} f(\xi)\frac{\xi - y}{\epsilon}e^{-|\xi - y|^2/(2\epsilon)}d\xi = \frac{\epsilon^{-1/2}}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(y + \sqrt{\epsilon} \zeta)e^{-|\xi|^2/2\epsilon}d\xi
\]
and then
\[
[f_{\epsilon}]_{\text{Lip}} \leq \|f\|\infty \epsilon^{-1/2}\mathbb{E}[\zeta] \leq C\|f\|\infty \epsilon^{-1/2}.
\] (3.13)

So that
\[
|\mathbb{E}_f(Z_t^x + \sqrt{\epsilon}\zeta) - \mathbb{E}_f(Z_t^x + \sqrt{\epsilon}\zeta)| = |\mathbb{E}_f(Z_t^x) - \mathbb{E}_f(Z_t^x)| \leq \frac{C\|f\|\infty}{\sqrt{\epsilon}} \|Z_t^x - Z_t^x\|_1
\]
\[
\leq \frac{C\|f\|\infty}{\sqrt{\epsilon}} t(1 + |x|),
\] (3.14)
where we used Lemma A.2 in the appendix. This implies that
\[
dTV(Z_t^x + \sqrt{\varepsilon} \zeta, Z_t^x + \sqrt{\varepsilon} \zeta) \leq C\varepsilon^{-1/2} t(1 + |x|).
\]

**Conclusion**: Considering (3.10), we have
\[
dTV(X_t^x, X_t^x) \leq Ce^{C|z|^2 t^{1/2} + C\varepsilon t^{-1} + C(1 + |x|^2)t + C\varepsilon^{-1/2} t(1 + |x|)}.
\]

We now choose \( \varepsilon = t^{4/3} \), so that
\[
dTV(X_t^x, X_t^x) \leq Ce^{C|z|^2 t^{1/3}}.
\]

### 3.4 Proof of Theorem 2.2 using Proposition 2.3

We first prove Proposition 2.3.

**Proof.** Let \( f : \mathbb{R}^d \to \mathbb{R} \) be measurable and bounded, let \( \varepsilon > 0 \) and let \( \zeta \sim \mathcal{N}(0, I_d) \) be independent of \( (Z_1, Z_2) \). We have
\[
|E f(Z_1) - E f(Z_2)| \leq \left| E f(Z_1) - \sum_{i=1}^r w_i E f_{\varepsilon/n_i}(Z_1) \right| + \sum_{i=1}^r w_i E f_{\varepsilon/n_i}(Z_1) - E f(Z_1)
\]
\[
+ \sum_{i=1}^r w_i E f_{\varepsilon/n_i}(Z_2) - E f(Z_2)
\]
where \( f_\varepsilon \) is defined as in (3.12) and where the \( n_i \)’s and the \( w_i \)’s will be defined later.

Let \( \varphi \) be as defined in (3.11) replacing \( Z_t^x \) by \( Z_1 \). Then, \( \varphi \) is differentiable up to the order \( 2r \) and for all \( k = 0, 1, \ldots, 2r \):
\[
\nabla^k \varphi(y) = (-1)^k \int_{\mathbb{R}^d} f(\xi) \nabla^k p_1(\xi - y) d\xi.
\]

Using the Taylor formula up to order \( 2r \), for every \( y \in \mathbb{R}^d \) there exits \( \tilde{y} \in (0, y) \) such that
\[
\varphi(y) = \varphi(0) + \sum_{k=1}^{2r-1} \frac{\nabla^k \varphi(0)}{k!} \cdot y^\otimes k + \frac{\nabla^{2r} \varphi(\tilde{y})}{(2r)!} \cdot y^\otimes 2r.
\]

Moreover, we have
\[
|\nabla^{2r} \varphi(\tilde{y}) \cdot y^\otimes 2r| \leq C\|f\|_\infty |y|^{2r} \int_{\mathbb{R}^d} \|\nabla^{2r} p_1\|.
\]

Then there exists a random \( \tilde{\zeta} \in (0, \zeta) \) such that
\[
E f(Z_1 + \sqrt{\varepsilon} \zeta) - E f(Z_1) = E \varphi(\sqrt{\varepsilon} \zeta) - \varphi(0) = \sum_{k=1}^{2r-1} \frac{\nabla^k \varphi(0)}{k!} \varepsilon^{k/2} \cdot E[\zeta^\otimes k] + \frac{E[\nabla^{2r} \varphi(\sqrt{\varepsilon} \zeta) \cdot \zeta^\otimes 2r]}{(2r)!} \varepsilon^r
\]
\[
= \sum_{k=1}^{r-1} \frac{\nabla^{2k} \varphi(0)}{(2k)!} \varepsilon^k \cdot E[\zeta^\otimes 2k] + \frac{E[\nabla^{2r} \varphi(\sqrt{\varepsilon} \zeta) \cdot \zeta^\otimes 2r]}{(2r)!} \varepsilon^r
\]
\[
=: \sum_{k=1}^{r-1} \beta_k(t)e^k + \tilde{\beta}_r(t, \varepsilon)e^r,
\]
because if \( k \) is odd, then \( E[\zeta^{\otimes k}] = 0 \). We now rely on a multi-step Richardson-Romberg extrapolation [LP17, Appendix A]. Let us denote the refiners \( n_i = 2^{r-1} \) and the auxiliary sequences

\[
\begin{align*}
  u_k := & \left( \prod_{\ell=1}^{k-1} (1 - 2^{-\ell}) \right)^{-1}, \\
  v_k := & (-1)^k 2^{-k(k+1)/2} u_{k+1}
\end{align*}
\]

and the weights

\[
  w_k = u_k v_{r-k}, \quad k = 1, \ldots, r.
\]

These weights are the unique solution to the \( r \times r \) Vandermonde system

\[
  \sum_{i=1}^{r} w_i n_i^{-k} = \delta_{k,0}, \quad k = 0, 1, \ldots, r - 1.
\]

Then we have

\[
\begin{align*}
  \sum_{i=1}^{r} w_i \left( E(f(Z_1 + \sqrt{\varepsilon/n_i} \zeta) - E(f(Z_1)) \right) & = \sum_{i=1}^{r} w_i \sum_{k=1}^{r-1} \beta_k(t) \varepsilon^k n_i^{-k} + \sum_{i=1}^{r} w_i \tilde{\beta}_r(t, \varepsilon/n_i) \varepsilon^r n_i^{-r} \\
  & = \sum_{k=1}^{r-1} \varepsilon^k \beta_k(t) \sum_{i=1}^{r} w_i n_i^{-k} + \varepsilon^r \sum_{i=1}^{r} \tilde{\beta}_r(t, \varepsilon/n_i) w_i n_i^{-r} \\
  & = \varepsilon^r \sum_{i=1}^{r} \tilde{\beta}_r(t, \varepsilon/n_i) w_i n_i^{-r}. \quad (3.17)
\end{align*}
\]

Now, using (3.15) we have

\[
\left| \sum_{i=1}^{r} \tilde{\beta}_r(t, \varepsilon/n_i) w_i n_i^{-r} \right| \leq C \|f\|_\infty \left( \int_{\mathbb{R}^d} \|\nabla^{2r} p_1\| \right) \sum_{i=1}^{r} |w_i| n_i^{-r}.
\]

Since \( u_k \to u_\infty = \prod_{\ell=1}^{\infty} (1 - 2^{-\ell})^{-1} < \infty \), the weights satisfy

\[
|w_i| \leq u_\infty^2 2^{1-r-i} 2^{-i} / (r-i+1)/2, \quad i = 1, \ldots, r,
\]

so that

\[
\sum_{i=1}^{r} \frac{|w_i|}{n_i^r} \leq u_\infty^2 \sum_{i=1}^{r} 2^{1-r-i} 2^{-i} / (r-i+1)/2 \leq u_\infty^2 \sum_{i=1}^{r} 2^{r-i}/2 = u_\infty^2 \sum_{i=0}^{r-1} 2^{-i}/2 \leq C. \quad (3.18)
\]

As a consequence and since \( \sum_{i=1}^{r} v_i = 1 \), we may write from (3.17)

\[
\left| E(f(Z_1)) - \sum_{i=1}^{r} w_i E f_{n_i}(Z_1) \right| \leq C \|f\|_\infty \varepsilon^r \int_{\mathbb{R}^d} \|\nabla^{2r} p_1\|. \quad (3.19)
\]

The same way, we obtain

\[
\left| E(f(Z_2)) - \sum_{i=1}^{r} w_i E f_{n_i}(Z_2) \right| \leq C \|f\|_\infty \varepsilon^r \int_{\mathbb{R}^d} \|\nabla^{2r} p_2\|.
\]

On the other side, using (3.13) we have

\[
\left| \sum_{i=1}^{r} w_i E f_{n_i}(Z_1) - \sum_{i=1}^{r} w_i E f_{n_i}(Z_2) \right| \leq \frac{C \|f\|_\infty}{\sqrt{\varepsilon}} W_1(Z_1, Z_2) \left( \sum_{i=1}^{r} |w_i|^2 (i-1)/2 \right). \quad (3.20)
\]

11
Moreover, for every \( i = 1, \ldots, r, \)
\[
|w_i|^2(i-1)/2 \leq u_{\infty}^22^{-(r-i)(r-i+1)/2+(i-1)/2}
\]
and then
\[
\sum_{i=1}^{r} |w_i|^2(i-1)/2 \leq u_{\infty}^2 \sum_{i=1}^{r} 2^{(i-1)/2} \leq u_{\infty}^22^r. \tag{3.21}
\]

We now prove Theorem 2.2.

**Proof.** The strategy of the proof is the same as for the proof of Theorem 2.1 for the terms \( D_1, D_5 \) and \( D_6, \) but is different for the terms \( D_2, D_3 \) and \( D_4 \) in (3.10). More precisely, let us write
\[
d_{TV}(\bar{X}_t^x, X_t^x) \leq d_{TV}(\bar{X}_t^x, Z_t^x) + d_{TV}(Z_t^x, Z_t^y) + d_{TV}(Z_t^y, Y_t^x) + d_{TV}(Y_t^x, X_t^x)
\]
\[
= : D_1 + D_{234} + D_5 + D_6. \tag{3.22}
\]
The bounds for the terms \( D_1, D_5 \) and \( D_6 \) are the same as in Section 3.3. For the term \( D_{234}, \) we apply Proposition 2.3 with the random vectors \( Z_1 = Z_t^x \) and \( Z_2 = Z_t^y. \) Assuming that \( \sigma \) is \( C_5^{2r} \) and using Theorem 3.1, \( \nabla^k\partial_x \) exists for \( k = 0, 1, \ldots, 2r \) and
\[
\forall k = 0, 1, \ldots, 2r, \forall t \in (0, T], \forall x, y \in \mathbb{R}^d, \|\nabla^k\partial_x(t, x, y)\| \leq \frac{C}{t^{(d+k)/2}}e^{-c|x-y|^2/t}.
\]
Then we have
\[
\int_{\mathbb{R}^d} \nabla^k\partial_x(t, x, \xi)d\xi \leq Ct^{-r} \int_{\mathbb{R}^d} \frac{1}{t^{d/2}}e^{-c|x-\xi+y|^2/t}d\xi \leq Ct^{-r}.
\]
The same way we have
\[
\int_{\mathbb{R}^d} \nabla^k\partial_x(t, x, \xi)d\xi \leq Ct^{-r}.
\]
Applying Proposition 2.3 with Lemma A.2 yields
\[
D_{234} \leq C\varepsilon t^{-r} + C\varepsilon^{-1/2}t(1 + |x|).
\]

**Conclusion:** Considering (3.22), we have
\[
d_{TV}(\bar{X}_t^x, X_t^x) \leq Ce^{C|x|^2}t^{1/2} + C\varepsilon^{-1/2}t(1 + |x|) + C(1 + |x|^2)t.
\]
We now choose \( \varepsilon = t^{(2r+2)/(2r+1)} \) so that
\[
d_{TV}(\bar{X}_t^x, X_t^x) \leq Ce^{C|x|^2}t^{r/(2r+1)}.
\]

### 3.5 Proof of Theorem 2.4

**Proof.** We write (3.22) again; while the other bounds remain the same, we rework the bound on the term \( D_{234} \) by paying attention to the dependency of the constants in \( r \) in the proof of Proposition 2.3 with \( Z_1 := Z_t^x \) and \( Z_2 := Z_t^y. \) Since \( \sigma \in \mathcal{C}_b^{2r} \) for every \( r \in \mathbb{N}, \) we write (3.16) for any \( r \in \mathbb{N} \) and we have
\[
|\tilde{\beta}_t(t, \varepsilon)| \leq \tilde{C}_2r \|f\|_{\infty}t^{-r}E[|\xi|^{2r}]/(2r)!,
\]
\[
\tilde{C}_2r := C_2r^{d/2}.
\]
where $C_{2r}$ and $c_{2r}$ are defined in (2.6) and where

$$\mathbb{E}[|\zeta|^{2r}] = \frac{2^r \Gamma(d/2 + r)}{\Gamma(d/2)} = \prod_{i=0}^{r-1} (d + 2i).$$

Using (3.18) we get

$$\left| \sum_{i=1}^{r} \tilde{\beta}_r(t, \varepsilon/n_i) w_i n_{i}^{-r} \right| \leq C \widetilde{C}_{2r} \|f\|_{\infty} t^{-r} \prod_{i=0}^{r-1} (d + 2i) \left(2r\right)!$$

and we obtain as in (3.19):

$$\left| \mathbb{E}f(Z_t^\varepsilon) - \sum_{i=1}^{r} w_i \mathbb{E}f_{\varepsilon/n_i}(Z_t^\varepsilon) \right| \leq \frac{1}{2} \kappa_1 \|f\|_{\infty} \varepsilon^r t^{-r}, \quad \kappa_1 := C \widetilde{C}_{2r} \prod_{i=0}^{r-1} (d + 2i) \left(2r\right)!$$

$$\left| \mathbb{E}f(Z_t^\varepsilon) - \sum_{i=1}^{r} w_i \mathbb{E}f_{\varepsilon/n_i}(Z_t^\varepsilon) \right| \leq \frac{1}{2} \kappa_1 \|f\|_{\infty} \varepsilon^r t^{-r}.$$

On the other hand, considering (3.20) and (3.21) with Lemma A.2 we have

$$\left| \sum_{i=1}^{r} w_i \mathbb{E}f_{\varepsilon/n_i}(Z_t^\varepsilon) \right| \leq \kappa_2 \|f\|_{\infty} t(1 + |x|), \quad \kappa_2 := C 2^r.$$

We now minimize $\kappa_1 \varepsilon^r t^{-r} + \kappa_2 \varepsilon^{-1/2} t (1 + |x|)$ in $\varepsilon$, giving

$$\varepsilon_* = \frac{(1 + |x|)^{2/(2r+1)} \Gamma(2r+3)/(2r+1)}{2 \kappa_1 \kappa_2^{2/(2r+1)}}$$

and then

$$\kappa_1 \varepsilon_*^r t^{-r} + \kappa_2 \varepsilon_*^{-1/2} t (1 + |x|) \leq C \kappa_2^{2r/(2r+1)} \kappa_1^{1/(2r+1)} (1 + |x|) t^{r/(2r+1)}$$

with as $r \to \infty$:

$$\kappa_2^{2r/(2r+1)} \kappa_1^{1/(2r+1)} \sim \tilde{C}_{2r}^{1/(2r+1)} \left(\prod_{i=0}^{r-1} (d + 2i)\right)^{1/(2r+1)} \left(2r\right)!^{1/(2r+1)}$$

with

$$\left(\prod_{i=0}^{r-1} (d + 2i)\right)^{1/(2r+1)} \sim \frac{c}{2r}, \quad \limsup_{r \to \infty} \tilde{C}_{2r}^{1/(2r+1)} < \infty$$

where we used Assumption (2.6), so that

$$\kappa_2^{2r/(2r+1)} \kappa_1^{1/(2r+1)} \leq C \sqrt{d + r - 1} \frac{c}{2r} 2^r.$$

Then we have $D_{234} \leq C 2^r r^{-1/2} t^{r/(2r+1)}$ and we choose $r(t) = \left[\log^{1/2}(1/t)\right]$ so that as $t \to 0$,

$$D_{234} \leq C t^{1/2} \exp\left(C \sqrt{\log(1/t)}\right).$$

$\square$
4 Counterexample

In this section we give a counter-example showing that we cannot achieve a bound better than $t^{1/2}$ in general. For $x > 0$ and $\sigma > 0$, let us consider the one-dimensional process

$$Y^x_t = xe^{\sigma W_t},$$

(4.1)

where $W$ is a standard Brownian motion. The process $Y$ is solution of the SDE $dY_t = (\sigma^2/2)Y_t^2 dt + \sigma Y_t dW_t$ and its associated Euler-Maruyama schemes reads

$$\tilde{Y}^x_t = x + (\sigma^2/2)xt + \sigma x W_t \sim N(x(1 + t\sigma^2/2), \sigma^2 x^2 t).$$

(4.2)

Proposition 4.1. Let $Y$ be the process defined in (4.1). Then for small enough $t$ we have

$$d_{TV}(Y^x_t, \tilde{Y}^x_t) \geq C_x t^{1/2}.$$  

(4.3)

Proof. We have

$$p_y(t, x, y) = \frac{1}{2\pi \sigma^2 t} \exp \left( -\frac{\log^2(y/x)}{2\sigma^2 t} \right) \mathbf{1}_{y \geq 0}$$

(4.4)

so that

$$d_{TV}(Y^x_t, \tilde{Y}^x_t) = \frac{1}{\sqrt{2\pi \sigma^2 t}} \int_{\mathbb{R}} \left| \exp \left( -\frac{\log^2(y/x)}{2\sigma^2 t} \right) y^{-1} \mathbf{1}_{y \geq 0} - \exp \left( -\frac{(y - x - x t \sigma^2/2)^2}{2\sigma^2 x^2 t} \right) x^{-1} \right| dy$$

$$\geq \frac{1}{\sqrt{2\pi \sigma^2}} \int_{-x/\sqrt{t}}^{\infty} \frac{1}{x + \sqrt{ty}} \left[ \frac{\log^2(1 + \sqrt{ty}/x)}{2\sigma^2 t} - 1 \times \exp \left( -\frac{(y - x - x t \sigma^2/2)^2}{2\sigma^2 x^2} \right) \right] dy.$$

But we have as $(t, y) \to 0$:

$$\frac{1}{1 + \sqrt{ty}/x} \exp \left( -\frac{\log^2(1 + \sqrt{ty}/x)}{2\sigma^2 t} \right) - \exp \left( -\frac{(y - x t \sigma^2/2)^2}{2\sigma^2 x^2} \right)$$

$$= \left( 1 - \sqrt{ty}/x + O(t^2) \right) \exp \left( -\frac{1}{2\sigma^2 t} \left( \frac{ty^2}{x^2} - \frac{t^3/2 y^3}{x^3} + O(t^4) \right) \right) - \exp \left( -\frac{y^2}{2\sigma^2 x^2} - \frac{t \sigma^2}{8} + \frac{\sqrt{ty}}{2 \sigma^2 x} \right)$$

$$= e^{-\frac{y^2}{2\sigma^2 x^2}} \left[ \left( 1 - \sqrt{ty} + O(t^2) \right) \left( 1 + \frac{\sqrt{ty}^3}{2 \sigma^2 x^3} + O(t^4) \right) - \left( 1 + \frac{\sqrt{ty}}{2 \sigma^2 x} - \frac{t \sigma^2}{8} + O(t^2) + O(t^2) \right) \right]$$

$$= e^{-\frac{y^2}{2\sigma^2 x^2}} \left[ -\frac{\sqrt{ty}}{x} - \frac{\sqrt{ty}}{2 \sigma^2 x^3} + \frac{\sqrt{ty}^3}{2 \sigma^2 x^3} + \frac{t \sigma^2}{8} + O(t^2) \right].$$

Thus there exists $\epsilon > 0$ and $t_0$ such that for every $t \leq t_0$:

$$d_{TV}(Y^x_t, \tilde{Y}^x_t) \geq \frac{1}{\sqrt{2\pi \sigma^2 x^2}} e^{-\frac{\epsilon^2}{2\sigma^2 x^2}} \sqrt{\frac{t}{2}} \left| -\frac{y^2}{2\sigma^2 x} - \frac{y^3}{2 \sigma^2 x^3} \right| dy,$$

so that $d_{TV}(Y^x_t, \tilde{Y}^x_t)$ is of order $t^{1/2}$ as $t \to 0$.

However, the process $Y$ does not satisfy the assumptions of Theorem 2.2 as its noise coefficient is not elliptic neither bounded on $(0, \infty)$. We then prove the following result.

Proposition 4.2. There exists a diffusion process on $\mathbb{R}$ with $C^1$ and Lipschitz continuous drift, with $C^0$ elliptic diffusion coefficient $\sigma$ and there exists $T > 0$ and $\epsilon > 0$ such that

$$\forall t \in [0, T], \forall x \in (\epsilon, \epsilon^{-1}), \quad d_{TV}(X^x_t, \tilde{X}^x_t) \geq C_x t^{1/2}$$

where the positive constant $C_x$ depends on $x$.  

14
Proof. We construct from the geometric Brownian motion $Y$ defined in (4.2), a process $X$ with elliptic and bounded drift and such that $d_{TV}(X^x_T, X^x_t) \geq C_x^{1/2}$. For $\varepsilon > 0$, let us consider $\psi : \mathbb{R} \to \mathbb{R}^+$ a $C^\infty$ approximation of

$$
\tilde{\psi} : x \in \mathbb{R} \mapsto \begin{cases} 
  x & \text{if } x \in [\varepsilon, \varepsilon^{-1}], \\
  \varepsilon & \text{if } x \leq \varepsilon \\
  \varepsilon^{-1} & \text{if } x \in [\varepsilon^{-1}, \infty).
\end{cases}
$$

Then we define the process with elliptic and bounded noise coefficient

$$
dX^x_t = -\frac{\sigma^2}{2} X^x_t dt + \sigma \psi(X^x_t) dW_t.
$$

Then for $x \in (\varepsilon, \varepsilon^{-1})$ we have $\tilde{X}^x_t = \tilde{Y}^x_t$ and

$$
P(Y^x_t \neq X^x_t) \leq \mathbb{P} \left( \sup_{s \in [0,t]} Y^x_s \geq \varepsilon^{-1} \right) + \mathbb{P} \left( \inf_{s \in [0,t]} Y^x_s \leq \varepsilon \right).
$$

With a proof similar to the proof of Lemma 3.2, we show that

$$
P \left( \sup_{s \in [0,t]} Y^x_s \leq \varepsilon^{-1} \right) \geq C_{x,\varepsilon} t.
$$

Moreover, we remark that $(Y^x)^{-1} \sim x^{-2} Y^x$ in law so

$$
P \left( \inf_{s \in [0,t]} Y^x_s \geq \varepsilon \right) = \mathbb{P} \left( \sup_{s \in [0,t]} (Y^x_s)^{-1} \leq \varepsilon^{-1} \right) = \mathbb{P} \left( \sup_{s \in [0,t]} Y^x_s \leq x^2 \varepsilon^{-1} \right) \leq C_{x,\varepsilon} t.
$$

Then we obtain

$$
d_{TV}(X^x_T, X^x_t) \geq d_{TV}(Y^x_t, \tilde{Y}^x_t) - d_{TV}(X^x_t, Y^x_t) \geq C_x \sqrt{t}.
$$

Remark 4.3. We could also consider the process $X$ with "cut" bounded drift $\bar{b}$ and get the same bounds, proving that the bounds established in Theorem 2.2 are optimal also if we assume that $b$ is bounded.

A Appendix

Lemma A.1 ([Fri64], Chapter 9, Lemma 7). For $a > 0$, $0 < u < t \leq T$, $x \in \mathbb{R}^d$, $\xi \in \mathbb{R}^d$, let

$$
I_a := \int_{\mathbb{R}^d} \frac{1}{u(t-u)^{d/2}} \exp \left( -a \left( \frac{|x-y|^2}{t-u} + \frac{|y-\xi|^2}{u} \right) \right) dy.
$$

Then there exists a constant $C > 0$ depending only on $d$ and $T$ such that for every $0 < \varepsilon < 1$,

$$
I_a \leq \frac{C}{(\varepsilon at)^{d/2}} \exp \left( -a(1-\varepsilon) \frac{|x-\xi|^2}{t} \right).
$$

Lemma A.2 ([PP20], Lemma 3.4). Let $Z$ be the solution of (3.1) and assume that $b_x$ and $\sigma_x$ are Lipschitz continuous and that $\sigma_x$ is bounded. Then for every $p \geq 1$,

$$
\forall t \in [0,T], \forall x \in \mathbb{R}^d, \|Z^x_t - \tilde{Z}^x_t\|_p \leq C_p(1 + |x|)t.
$$
References

[AS64] Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. National Bureau of Standards Applied Mathematics Series, No. 55. U. S. Government Printing Office, Washington, D. C., 1964.

[BJ20] Oumaima Bencheikh and Benjamin Jourdain. Convergence in total variation of the Euler-Maruyama scheme applied to diffusion processes with measurable drift coefficient and additive noise. *arXiv e-prints*, page arXiv:2005.09354, May 2020.

[BL05] Marcello Bertoldi and Luca Lorenzi. Estimates of the derivatives for parabolic operators with unbounded coefficients. *Trans. Amer. Math. Soc.*, 357(7):2627–2664, 2005.

[BP21] Pierre Bras and Gilles Pagès. Convergence of langevin-simulated annealing algorithms with multiplicative noise, 2021.

[BT96] 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.

[Cer00] Sandra Cerrai. Analytic semigroups and degenerate elliptic operators with unbounded coefficients: a probabilistic approach. *J. Differential Equations*, 166(1):151–174, 2000.

[Fri64] Avner Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964.

[GL08] Emmanuel Gobet and Céline Labart. Sharp estimates for the convergence of the density of the Euler scheme in small time. *Electron. Commun. Probab.*, 13:352–363, 2008.

[Kra04] Ilia Krasikov. New bounds on the Hermite polynomials. *East J. Approx.*, 10(3):355–362, 2004.

[LP17] Vincent Lemaire and Gilles Pagès. Multilevel Richardson-Romberg extrapolation. *Bernoulli*, 23(4A):2643–2692, 2017.

[Lun97] Alessandra Lunardi. Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in \( \mathbb{R}^n \). *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 24(1):133–164, 1997.

[MPZ21] Stéphane Menozzi, Antonello Pesce, and Xicheng Zhang. Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift. *J. Differential Equations*, 272:330–369, 2021.

[PP20] Gilles Pagès and Fabien Panloup. Unadjusted Langevin algorithm with multiplicative noise: Total variation and Wasserstein bounds. *arXiv e-prints*, page arXiv:2012.14310, 2020.

[QZ04] Zhongmin Qian and Weian Zheng. A representation formula for transition probability densities of diffusions and applications. *Stochastic Process. Appl.*, 111(1):57–76, 2004.

[RG11] Lewis Fry Richardson and Richard Tetley Glazebrook. IX. the approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 210(459-470):307–357, 1911.