



**HAL**  
open science

# Total variation distance between two diffusions in small time with unbounded drift: application to the Euler-Maruyama scheme

Pierre Bras, Gilles Pagès, Fabien Panloup

► **To cite this version:**

Pierre Bras, Gilles Pagès, Fabien Panloup. Total variation distance between two diffusions in small time with unbounded drift: application to the Euler-Maruyama scheme. *Electronic Journal of Probability*, 2022, 27, pp.1-19. 10.1214/22-EJP881 . hal-03897789

**HAL Id: hal-03897789**

**<https://hal.science/hal-03897789>**

Submitted on 14 Dec 2022

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

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



Distributed under a Creative Commons Attribution 4.0 International License

# Total variation distance between two diffusions in small time with unbounded drift: application to the Euler-Maruyama scheme

Pierre Bras<sup>\*†</sup>, Gilles Pagès<sup>\*</sup> and Fabien Panloup<sup>‡</sup>

## Abstract

We give bounds for the total variation distance between the solutions to two stochastic differential equations starting at the same point and with close coefficients, which applies in particular to the distance between an exact solution and its Euler-Maruyama scheme in small time. We show that for small  $t$ , the total variation distance is of order  $t^{r/(2r+1)}$  if the noise coefficient  $\sigma$  of the SDE is elliptic and  $\mathcal{C}_b^{2r}$ ,  $r \in \mathbb{N}$  and if the drift is  $C^1$  with bounded derivatives, using multi-step Richardson-Romberg extrapolation. We do not require the drift to be bounded. Then we prove with a counterexample that we cannot achieve a bound better than  $t^{1/2}$  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, in this paper we consider the two following SDEs in  $\mathbb{R}^d$  starting at the same point:

$$\begin{aligned} X_0^x &= x \in \mathbb{R}^d, & dX_t^x &= b_1(t, X_t^x)dt + \sigma_1(t, X_t^x)dW_t, \\ Y_0^x &= x, & dY_t^x &= b_2(t, Y_t^x)dt + \sigma_2(t, Y_t^x)dW_t, \end{aligned}$$

where  $W$  is a Brownian motion. We generally assume that for  $i = 1, 2$ ,  $b_i$  is Lipschitz continuous and that  $\sigma_i$  is elliptic, bounded and Lipschitz continuous, but we do not assume that  $b_i$  is bounded. Our objective is to give bounds of the total variation distance between the law of  $X_t^x$  and the law of  $Y_t^x$ , denoted  $d_{\text{TV}}(X_t^x, Y_t^x)$ , as  $t \rightarrow 0$ . In particular, we apply our results to the case where  $Y^x = \bar{X}^x$  is the one-step Euler-Maruyama scheme associated to the SDE  $X$ , given by

$$dY_t^x = b_1(0, x)dt + \sigma_1(0, x)dW_t.$$

Such bounds are well known for  $L^p$  distances and their associated Wasserstein distances and are known to be of order  $t$  as  $t \rightarrow 0$ . Yet the literature seems to lack results as it comes to  $d_{\text{TV}}$ . If  $\sigma_1 = \sigma_2$  is constant, then it is classical background that  $d_{\text{TV}}(X_t^x, Y_t^x)$  is of order  $t$ , using a Girsanov

---

<sup>\*</sup>Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation, UMR 8001, case 158, 4 pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail: pierre.bras@sorbonne-universite.fr and gilles.pages@sorbonne-universite.fr.

<sup>†</sup>Corresponding author

<sup>‡</sup>LAREMA, Faculté des Sciences, 2 Boulevard Lavoisier, Université d’Angers, 49045 Angers, France. E-mail: fabien.panloup@univ-angers.fr

change of measure (see for example [PP20, Proposition 4.1]) 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  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is Lipschitz continuous, then we can bound  $|f(x) - f(y)|$  by  $[f]_{\text{Lip}}|x - y|$ ; whereas 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.

Results in the literature focus on the Euler-Maruyama scheme. In [BT96] is proved the convergence for a fixed time horizon  $T > 0$  and as  $N \rightarrow \infty$ , where  $N$  is the number of steps in the Euler-Maruyama scheme on a finite horizon. More precisely, if  $\sigma_1$  is elliptic and if  $b_1$  and  $\sigma_1$  are  $\mathcal{C}^\infty$  with bounded derivatives (but  $b_1$  and  $\sigma_1$  are not supposed bounded themselves), then ([BT96, Theorem 3.1])

$$\forall x \in \mathbb{R}^d, \text{d}_{\text{TV}}(X_T^x, \bar{X}_T^{x,N}) \leq \frac{K(T)(1 + |x|^Q)}{NT^q},$$

where  $\bar{X}_T^{x,N}$  stands for the Euler scheme with  $N$  steps, where  $Q$  and  $q$  are positive exponents and where  $K$  is a non-decreasing function depending on  $b_1$  and  $\sigma_1$ . 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  $T \rightarrow 0$  since we do not know whether  $K(T)/T^q \rightarrow 0$  as  $T \rightarrow 0$  in general. In [GL08] are given bounds in small time and as  $N \rightarrow \infty$ . Assuming that  $\sigma_1$  is uniformly elliptic and that  $b_1$  and  $\sigma_1$  are bounded with bounded derivatives up to order 3, then ([GL08, Theorem 2.3])

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

where  $p$  and  $\bar{p}^N$  denote the transition densities of  $X^x$  and  $\bar{X}^{x,N}$  respectively and where  $C$  is a positive constant depending on  $d$  and on the bounds on  $b_1$  and  $\sigma_1$  and on their derivatives. However, we cannot directly use this result for the total variation distance: taking  $N = 1$  yields

$$\text{d}_{\text{TV}}(X_t^x, \bar{X}_t^x) = \int_{\mathbb{R}^d} |p(t, x, y) - \bar{p}^1(t, x, y)| dy \leq K(T)Tt^{-1/2} \int_{\mathbb{R}^d} \frac{1}{t^{d/2}} e^{-C|x-y|^2/t},$$

giving a bound in  $t^{-1/2}$  which does not converge to 0 as  $t \rightarrow 0$ . Moreover, [GL08] assumes that  $b_1$  is bounded. [BJ22] focuses on the case where  $b_1$  is bounded and measurable but not necessarily regular and where  $\sigma_1$  is constant; it proves that the convergence in total variation of the Euler scheme on a finite horizon which is regularized with respect to the irregular drift  $b_1$  and with step  $h$ , is of order  $\sqrt{h}$ .

In the present paper, we first prove a convergence rate of order  $t^{1/3}$  for  $\text{d}_{\text{TV}}(X_t^x, Y_t^x)$ , provided that for  $i = 1, 2$ ,  $\sigma_i$  is elliptic,  $\sigma_i$  and  $b_i$  are Lipschitz-continuous with respect to their time variable and that  $\sigma_i$  is  $\mathcal{C}_b^2$  and  $b_i$  is  $\mathcal{C}^1$  and Lipschitz-continuous with respect to their space variable. More generally, if we furthermore assume that  $\sigma$  is  $\mathcal{C}_b^{2r}$ , then we obtain a convergence rate of order  $t^{r/(2r+1)}$ . Letting  $r \rightarrow \infty$ , we also prove that if  $\sigma \in \mathcal{C}_b^\infty$  with some technical condition on the derivatives of the densities of the random variables  $X_t^x$  and  $Y_t^x$ , 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  $t^{r/(2r+1)}$ , 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  $t^{r/(2r+1)}$ ).

Interestingly, the difference between the drift coefficients  $b_1 - b_2$  does not need to be small for our bounds to be valid. This is because the dominant term in  $\text{d}_{\text{TV}}(X_t^x, Y_t^x)$  comes from the the diffusion part.

Recent results (see [Cle21]) establish a convergence in small time at rate  $t^{1/2}$  for the Euler scheme of certain classes of diffusions driven by stable Lévy processes, not directly including the Brownian case. This approach relies on Malliavin calculus techniques. In this work the "standard" drift is replaced in the Euler scheme by the flow of the associated (noiseless) ODE. This seems to be specific to Lévy driven SDEs. Adapting this approach to our general continuous framework is not as straightforward as could be expected and would deserve further investigations for future work.

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 the drift is bounded, then the density and its partial derivatives admit sub-gaussian Aronson's bounds (see [Fri64] and Section 3.1). However, giving estimates and bounds for the solution of the PDE in the case of unbounded drift appears to be more difficult, see [Lun97], [Cer00], [BL05]. Recent improvements have been made in [MPZ21] using the parametrix method. Studying this case is useful to study the convergence in total variation of SDE's with unbounded drift, in particular for the Langevin equation, very popular in stochastic optimization, and which reads

$$dX_t = -\nabla V(X_t)dt + \sigma(X_t)dW_t,$$

where in many cases,  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  has quadratic growth and  $\nabla V$  has linear growth (see for example [BP21]).

In order to deal with unbounded  $b_i$ , we propose two different methods. First, we use a localization argument and "cut" the drift  $b_i$  into  $\tilde{b}_i$  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  $\|\tilde{b}_i\|_\infty$ . A second method consists in using the density estimates from [MPZ21, Section 4] to improve the dependency with respect to the bounds in  $x$ . However this second approach relies on advanced parametrix methods which require further regularity assumptions on the coefficients of the SDE and which are not fully detailed for higher order derivatives. Our first method is clearly much more elementary, starting from a quite general bound established for any pair of integrable random vectors (see Theorem 2.7) and calling upon a standard regularization strategy which combined with a multistep procedure, seems to be at least quasi-optimal in a very general framework.

#### 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  $\mathbf{B}(x, R) = \{y \in \mathbb{R}^d : |y - x| \leq R\}$ .

We denote  $\mathcal{M}_d(\mathbb{R})$  the set of  $d \times d$  matrices with real coefficients.

For  $M \in (\mathbb{R}^d)^{\otimes k}$ , we denote by  $\|M\|$  its operator norm, i.e.  $\|M\| = \sup_{u \in \mathbb{R}^d \times k, |u|=1} M \cdot u$ . If  $M : \mathbb{R}^d \rightarrow (\mathbb{R}^d)^{\otimes k}$ , we denote  $\|M\|_\infty = \sup_{x \in \mathbb{R}^d} \|M(x)\|$ .

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

We denote the total variation distance between two distributions  $\pi_1$  and  $\pi_2$  on  $\mathbb{R}^d$ :

$$d_{\text{TV}}(\pi_1, \pi_2) = 2 \sup_{A \in \mathcal{B}(\mathbb{R}^d)} |\pi_1(A) - \pi_2(A)|.$$

Without ambiguity, if  $Z_1$  and  $Z_2$  are two  $\mathbb{R}^d$ -valued random vectors, we also write  $d_{\text{TV}}(Z_1, Z_2)$  to denote the total variation distance between the law of  $Z_1$  and the law of  $Z_2$ . We have as well

$$d_{\text{TV}}(\pi_1, \pi_2) = \sup \left\{ \int_{\mathbb{R}^d} f d\pi_1 - \int_{\mathbb{R}^d} f d\pi_2, f : \mathbb{R}^d \rightarrow [-1, 1] \text{ measurable} \right\}.$$

Moreover, we recall that if  $\pi_1$  and  $\pi_2$  admit densities with respect to some measure  $\lambda$ , then

$$d_{\text{TV}}(\pi_1, \pi_2) = \int_{\mathbb{R}^d} \left| \frac{d\pi_1}{d\lambda} - \frac{d\pi_2}{d\lambda} \right| d\lambda.$$

We denote by  $\mathcal{W}_1$  the  $L^1$ -Wasserstein distance.

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

If  $Z$  is a Markov process with values in  $\mathbb{R}^d$ , we denote, when it exists, its transition probability from  $x$  to  $y \in \mathbb{R}^d$  between times  $s < t$ ,  $p_Z(s, t, x, y)$ .

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 two following SDEs in  $\mathbb{R}^d$ :

$$X_0^x = x \in \mathbb{R}^d, \quad dX_t^x = b_1(t, X_t^x)dt + \sigma_1(t, X_t^x)dW_t, \quad t \in [0, T], \quad (2.1)$$

$$Y_0^x = x, \quad dY_t^x = b_2(t, Y_t^x)dt + \sigma_2(t, Y_t^x)dW_t, \quad t \in [0, T], \quad (2.2)$$

where  $T$  is a finite time horizon,  $b_i : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma_i : \mathbb{R}^d \rightarrow \mathcal{M}_d(\mathbb{R})$ ,  $i = 1, 2$ , are Borel functions and  $W$  is a standard  $\mathbb{R}^d$ -valued Brownian motion defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . The one-step Euler-Maruyama scheme of  $X$ , denoted  $\bar{X}$ , is defined by  $\bar{X}^x = Y^x$  when  $Y^x$  reads

$$dY_t^x = b_1(0, x)dt + \sigma_1(0, x)dW_t, \quad t \in [0, T]. \quad (2.3)$$

To alleviate notations, we also define

$$\Delta b(x) := |b_1(0, x) - b_2(0, x)|, \quad \Delta \sigma(x) := |\sigma_1(0, x) - \sigma_2(0, x)|. \quad (2.4)$$

Let us remark that if  $Y = \bar{X}$ , then  $\Delta b(x) = 0$  and  $\Delta \sigma(x) = 0$ . For  $g : (t, x) \in [0, T] \times \mathbb{R}^d \mapsto g(t, x) \in \mathbb{R}^q$  and  $r \in \mathbb{N}$ , let us define the following assumptions:

- $\text{Lip}_t(g)$ :  $g$  is Lipschitz continuous with respect to  $t$ , uniformly in  $x$ .
- $g \in \mathcal{C}^r$ :  $g$  is differentiable with respect to  $x$  with continuous partial derivatives up to the order  $r$ .
- $g \in \mathcal{C}_b^r$ :  $g \in \mathcal{C}^r$  and is bounded with bounded partial derivatives up to the order  $r$ .
- $g \in \tilde{\mathcal{C}}_b^r$ :  $g \in \mathcal{C}^r$  and has partial bounded derivatives up to the order  $r$ , but we do not assume that  $g$  is bounded itself.
- For  $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathcal{M}_d(\mathbb{R})$ , we say that  $\sigma$  is (uniformly) elliptic if

$$\exists \varrho_0 > 0, \quad \forall x \in \mathbb{R}^d, \quad \forall t \in [0, T], \quad \sigma(t, x)\sigma(t, x)^\top \geq \varrho_0^2 I_d. \quad (2.5)$$

**Theorem 2.1.** *Let  $X$  and  $Y$  be the solutions of the SDEs (2.1) and (2.2). For  $i = 1, 2$ , assume  $\text{Lip}_t(b_i)$ ,  $\text{Lip}_t(\sigma_i)$ ,  $\sigma_i \in \mathcal{C}_b^2$ ,  $b_i \in \tilde{\mathcal{C}}_b^1$  and  $\sigma_i$  is elliptic. Then*

$$\forall t \in [0, T], \quad \forall x \in \mathbb{R}^d, \quad d_{\text{TV}}(X_t^x, Y_t^x) \leq C(t^{1/2} + \Delta \sigma(x))^{2/3} + Ce^{c|x|^2}t^{1/2}, \quad (2.6)$$

where the positive constants  $C$  and  $c$  only depend on  $d, T, \varrho_0, \|\sigma_i\|_\infty, [b_i]_{\text{Lip}}, [\sigma_i]_{\text{Lip}}$  and  $\|\nabla^2 \sigma_i\|_\infty$ . In particular, if  $Y = \bar{X}$  we have

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq Ct^{1/3} + Ce^{c|x|^2}t^{1/2}.$$

**Theorem 2.2.** *Let  $X$  and  $Y$  be the solutions of the SDEs (2.1) and (2.2). For  $i = 1, 2$ , assume  $\text{Lip}_t(b_i)$ ,  $\text{Lip}_t(\sigma_i)$ , that  $\sigma_i \in \mathcal{C}_b^{2r}$ ,  $b_i \in \tilde{\mathcal{C}}_b^1$  and that  $\sigma_i$  is elliptic. Then*

$$\forall t \in [0, T], \quad \forall x \in \mathbb{R}^d, \quad d_{\text{TV}}(X_t^x, Y_t^x) \leq C(t^{1/2} + \Delta \sigma(x))^{2r/(2r+1)} + Ce^{c|x|^2}t^{1/2}, \quad (2.7)$$

where the positive constants  $C$  and  $c$  only depend on  $d, T, \varrho_0$ , on  $\|\sigma_i\|_\infty$ , on the bounds on the derivatives of  $b_i$  and  $\sigma_i$  and on their Lipschitz constants. In particular, if  $Y = \bar{X}$  we have

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq Ct^{r/(2r+1)} + Ce^{c|x|^2}t^{1/2}.$$

*Remark 2.3.* In Theorems 2.1 and 2.2, we can actually improve the dependency of the constants in  $x$  in small time since we have more precisely:

$$\forall t \in [0, Ce^{-2(2r+1)c|x|^2}], \quad d_{\text{TV}}(X_t^x, Y_t^x) \leq C(t^{1/2} + \Delta\sigma(x))^{2r/(2r+1)}.$$

*Remark 2.4.* We can adapt the framework of Theorem 2.2 to the study of SDEs with homogeneous vanishing noise, i.e. where for  $a > 0$ ,

$$dX_t^x = b_1(t, X_t^x)dt + a\sigma_1(X_t^x), \quad dY_t^x = b_2(t, Y_t^x) + a\sigma_2(Y_t^x)$$

and we identify the dependency in  $a$  as  $a \rightarrow 0$ . Namely, with the same assumptions, for  $a > 0$  small enough we have

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq Ce^{ca^{-1}|x|^2}t^{1/2} + Ca^{-(d+r)}(t^{1/2} + \Delta\sigma(x))^{2r/(2r+1)}, \quad (2.8)$$

where the constant  $C$  does not depend on  $a$ . This bound is obtained adapting the proof of Theorem 2.2 in Section 3.4 and using that if  $Z^x$  is the martingale  $dZ_t^x = a\sigma(Z_t^x)dW_t$  then  $(Z_t^x) \sim (\tilde{Z}_{a^2t}^x)$  where  $\tilde{Z}_t^x = \sigma(\tilde{Z}_t^x)dW_t$  which does not depend on  $a$ .

We can also improve the dependency in the initial condition  $x \in \mathbb{R}^d$  using [MPZ21], however at the expense of further regularity assumptions on  $b_i$  and  $\sigma_i$ ,  $i = 1, 2$ .

**Theorem 2.5.** *Let  $X$  and  $Y$  be the solutions of the SDEs (2.1) and (2.2). For  $i = 1, 2$ , assume  $\text{Lip}_t(b_i)$ ,  $\text{Lip}_t(\sigma_i)$ ,  $\sigma_i \in \mathcal{C}_b^{2r+1}$ ,  $b_i \in \tilde{\mathcal{C}}_b^{2r}$  and  $\sigma_i$  is elliptic. Then*

$$\forall t \in [0, T], \quad \forall x \in \mathbb{R}^d, \quad d_{\text{TV}}(X_t^x, Y_t^x) \leq C \left( t^{1/2}(1 + \Delta b(x)) + \Delta\sigma(x) + t(|b_1| + |b_2|)(0, x) \right)^{2r/(2r+1)}, \quad (2.9)$$

where the positive constants  $C$  and  $c$  only depend on  $d$ ,  $T$ ,  $\sigma_0$ , on  $\|\sigma_i\|_\infty$ , on the bounds on the derivatives of  $b_i$  and  $\sigma_i$  and on their Lipschitz constants. In particular, if  $Y = \bar{X}$ , we have

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq Ct^{r/(2r+1)} \left( 1 + t^{1/2}(|b_1| + |b_2|)(0, x) \right)^{2r/(2r+1)}.$$

*Remark 2.6.* Choosing  $Y$  not to be the Euler-Maruyama scheme of  $X$  but a general SDE and expressing the bounds in the Theorems in terms of  $\Delta b(x)$  and  $\Delta\sigma(x)$  allows to extend our results to more general couples of diffusions with "close" coefficients, for example to SDE solvers other than the genuine Euler-Maruyama scheme. Also, it is helpful to study perturbed SDEs, for example if we consider

$$dX_t^x = b_1(t, X_t^x)dt + a_1(t)\sigma(X_t^x)dW_t, \quad dY_t^x = b_2(t, Y_t^x)dt + a_2(t)\sigma(Y_t^x)dW_t \quad (2.10)$$

where  $|a_1(t) - a_2(t)| \rightarrow 0$  as  $t \rightarrow \infty$ . Then we have

$$d_{\text{TV}}(X_{t+s}^x, Y_{t+s}^x) \leq Ce^{c|x|^2}s^{1/2} + C(s^{1/2} + |a_1(t) - a_2(t)|)^{2r/(2r+1)}$$

and we obtain different convergence rates as  $t \rightarrow \infty$  and  $s \rightarrow 0$ , depending on  $(t, s)$ . A noticeable example of this is the Langevin-simulated annealing SDE [BP21].

Furthermore we remark that in Theorems 2.1 and 2.2, the bounds do not depend on  $\Delta b$ , enhancing that the dominant term in the total variation comes from the diffusion part.

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]; this argument can also be applied in a more general framework. The following proposition gives bounds on the total variation between two random vectors, knowing bounds on the  $L^1$ -Wasserstein distance and bounds on the partial derivatives of the densities up to some order.

**Theorem 2.7.** 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  $\mathcal{C}^{2r}$  with  $r \in \mathbb{N}$  and that  $\nabla^k p_i \in L^1(\mathbb{R}^d)$  for  $i = 1, 2$  and  $k = 1, \dots, 2r$ . Then we have

$$d_{\text{TV}}(Z_1, Z_2) \leq C_{d,r} \mathcal{W}_1(Z_1, Z_2)^{2r/(2r+1)} \left( \int_{\mathbb{R}^d} (\|\nabla^{2r} p_1(\xi)\| + \|\nabla^{2r} p_2(\xi)\|) d\xi \right)^{1/(2r+1)} \quad (2.11)$$

where the constant  $C_{d,r}$  depends only on  $d$  and on  $r$ .

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

**Theorem 2.8.** Let  $X$  and  $Y$  be the solutions of the SDEs (2.1) and (2.2). For  $i = 1, 2$ , assume  $\text{Lip}_t(b_i)$ ,  $\text{Lip}_t(\sigma_i)$ , that  $\sigma_i \in \mathcal{C}_b^{2r}$  for every  $r \in \mathbb{N}$ ,  $b_i \in \tilde{\mathcal{C}}_b^1$ , that  $\sigma_i$  is elliptic and that  $\Delta\sigma(x) = 0$ . Assume furthermore that if  $Z$  and  $V$  are the martingales  $dZ_t = \sigma_1(t, Z_t)dW_t$  and  $dV_t = \sigma_2(t, Z_t)dW_t$ , then

$$\forall r \in \mathbb{N}, \forall t \in (0, T], \forall x, y \in \mathbb{R}^d, \|\nabla_y^{2r} p_Z(0, t, x, y)\| + \|\nabla_y^{2r} p_V(0, t, x, y)\| \leq \frac{C_{2r}}{t^{(d+2r)/2}} e^{-c_{2r}|y-x|^2/t}$$

with  $\limsup_{r \rightarrow \infty} \left( C_{2r} c_{2r}^{-d/2} \right)^{1/(2r)} < \infty.$  (2.12)

(see Theorem 3.1). Then

$$\forall t \in (0, T], \forall x \in \mathbb{R}^d, d_{\text{TV}}(X_t^x, Y_t^x) \leq C e^{c|x|^2} t^{1/2} + C t^{1/2} e^{c\sqrt{\log(1/t)}}, \quad (2.13)$$

where the positive constants  $C$  and  $c$  only depend on  $d$ ,  $T$ ,  $\underline{\sigma}_0$ , on  $\|\sigma_i\|_\infty$ , on the bounds on the derivatives of  $b_i$  and  $\sigma_i$  and on their Lipschitz constants.

*Remark 2.9.* Assumption (2.12) 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  $dZ_t = \sigma dW_t$  with  $\sigma \in \mathcal{M}_d(\mathbb{R})$  being non degenerate, then with  $\Sigma := \sigma\sigma^\top$  we have for  $t > 0$  and  $x, y \in \mathbb{R}^d$ :

$$p_Z(0, t, x, y) = \frac{1}{\sqrt{\det(\Sigma)t^{d/2}}} \Phi\left(\Sigma^{-1/2} \frac{y-x}{\sqrt{t}}\right), \quad \Phi(u) := \frac{1}{(2\pi)^{d/2}} e^{-|u|^2/2}.$$

Moreover for every  $r \in \mathbb{N}$  and  $u \in \mathbb{R}^d$  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  $r^{\text{th}}$  probabilist Hermite polynomial. Following [Kra04] we have

$$\forall u \geq 0, |\text{He}_{2r}(u)| e^{-u^2/2} \leq C 2^{-r} \sqrt{r} \frac{(2r)!^2}{r!^2} \leq C 2^r \sqrt{r},$$

using the Stirling formula for the last inequality. On the other hand, using [AS64, 22.14.15] we have

$$\forall u \geq 0, |\text{He}_{2r}(u)| e^{-u^2/4} \leq 2^{r+1} r!.$$

Then, for for every  $\varepsilon \in (0, 1]$ ,

$$\begin{aligned} |\text{He}_{2r}(u)| e^{-u^2/2} &= \left| \text{He}_{2r}(u) e^{-u^2/4} \right|^\varepsilon \left| \text{He}_{2r}(u) e^{-u^2/2} \right|^{1-\varepsilon} e^{-\varepsilon u^2/4} \\ &\leq C (2^r r!)^\varepsilon \left( 2^r r^{1/2} \right)^{1-\varepsilon} e^{-\varepsilon u^2/2}. \end{aligned}$$

Then if we choose  $\varepsilon_r = \log^{-1}(r)$  for  $r \geq 3$ , we have  $(r!)^{\varepsilon_r} \leq e^{\varepsilon_r r \log(r)} = e^r$  so that

$$\left\| \frac{d^r}{du^r} \Phi(u) \right\| \leq C 2^{r\varepsilon_r} e^r 2^r r^{1/2} e^{-\varepsilon_r |u|^2/2} =: A_r e^{-\varepsilon_r |u|^2/2}$$

and then for  $r \geq 3$  we have

$$\begin{aligned} \|\nabla_y^{2r} p_Z(0, t, x, y)\| &\leq \frac{\|\Sigma^{-1/2}\|^{2r}}{\sqrt{\det(\Sigma)} t^{(d+2r)/2}} \frac{d^{2r}}{du^{2r}} \Phi\left(\Sigma^{-1/2} \frac{y-x}{\sqrt{t}}\right) \\ &\leq \frac{\|\Sigma^{-1/2}\|^{2r}}{\sqrt{\det(\Sigma)} t^{(d+2r)/2}} A_r e^{-\varepsilon_r \|\Sigma^{-1}\| \|y-x\|^2/(2t)} \end{aligned}$$

where  $(\|\Sigma^{-1/2}\|^{2r} A_r \varepsilon_r^{-d/2})^{1/(2r)}$  is bounded. Thus Assumption (2.12) is satisfied.

*Remark 2.10.* For the Euler-Maruyama scheme (2.3), with a slight abuse of notation,  $x$  is used both for the starting point and in the definition of the drift and diffusion coefficients. The transition density should be considered for constant drift and diffusion coefficients in this case. However the results remain valid as the Euler scheme is simply a Brownian process.

### 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:

$$Z_0^x = x \in \mathbb{R}^d, \quad dZ_t^x = b_Z(t, Z_t^x) dt + \sigma_Z(t, Z_t^x) dW_t, \quad t \in [0, T]. \quad (3.1)$$

Then under regularity assumptions on  $b_Z$  and on  $\sigma_Z$ , the transition probability  $p_Z$  exists and is solution of the backward Kolmogorov PDE:

$$\begin{aligned} p_Z(t, t, x, \cdot) &= \delta_x, \quad t \in [0, T], \\ \partial_s p_Z(s, t, x, y) &= \langle b_Z(s, x), \nabla_x p_Z(s, t, x, y) \rangle + \frac{1}{2} \text{Tr} \left( \sigma_Z^\top(s, x) \nabla_x^2 p_Z(s, t, x, y) \sigma_Z(s, x) \right), \quad s < t \in [0, T]. \end{aligned} \quad (3.2)$$

Moreover,  $p_Z$  and its derivatives satisfy sub-gaussian Aronson's bounds:

**Theorem 3.1** ([Fri64], Chapter 9, Theorem 7). *Let  $Z$  be the solution of (3.1) and let  $T > 0$ . Assume  $\text{Lip}_t(b_Z)$  and  $\text{Lip}_t(\sigma_Z)$ , that  $b_Z, \sigma_Z \in \mathcal{C}_b^r$  and that  $\sigma_Z$  is elliptic. Then for every  $m_0 = 0, 1$  and for every  $0 \leq m_1 + m_2 \leq r$ ,  $\nabla_x^{m_0+m_1} \nabla_y^{m_2} p_Z$  exists and*

$$\forall s < t \in [0, T], \quad \forall x, y \in \mathbb{R}^d, \quad \|\nabla_x^{m_0+m_1} \nabla_y^{m_2} p_Z(s, t, x, y)\| \leq \frac{C}{(t-s)^{(d+m_0+m_1+m_2)/2}} e^{-c|y-x|^2/(t-s)}, \quad (3.3)$$

where the constants  $C$  and  $c$  only depend on the bounds on  $b_Z$  and  $\sigma_Z$  and on their derivatives and their Lipschitz constants, on the modulus of ellipticity of  $\sigma_Z$ , on  $d$  and on  $T$ .

Let us also recall the recent result from [MPZ21] giving Aronson's bounds of the partial derivatives with respect to  $y$  in the case where  $b_Z$  is unbounded. Considering [MPZ21, Section 4] with [MPZ21, (3.1)], we have the following result.



**Theorem 3.2.** Let  $Z$  be the solution of (3.1) and let  $T > 0$ . Assume  $\text{Lip}_t(b_Z)$  and  $\text{Lip}_t(\sigma_Z)$ , that  $b_Z \in \tilde{\mathcal{C}}_b^r$ ,  $\sigma_Z \in \mathcal{C}_b^{r+1}$  and that  $\sigma_Z$  is elliptic. Then for every  $0 \leq m \leq r$ ,  $\nabla_y^m p_Z$  exists and

$$\forall s < t \in [0, T], \forall x, y \in \mathbb{R}^d, \|\nabla_y^m p_Z(s, t, x, y)\| \leq \frac{C}{(t-s)^{(d+m)/2}} e^{-c|y-x|^2/(t-s)}, \quad (3.4)$$

where the constants  $C$  and  $c$  only depend on the bounds on  $b_Z$  and  $\sigma_Z$  and on their derivatives and on their Lipschitz constants, on the modulus of ellipticity of  $\sigma_Z$ , on  $d$  and on  $T$ .

### 3.2 Preliminary results

In order to apply the bounds on the densities from Theorem 3.1 to Theorem 2.7, we first "cut" the drifts  $b_1$  and  $b_2$  on a compact set. That is, we instead consider the processes  $\tilde{X}$  and  $\tilde{Y}$  defined by

$$d\tilde{X}_t^x = \tilde{b}_1^x(t, \tilde{X}_t^x)dt + \sigma_1(t, \tilde{X}_t^x)dW_t, \quad t \in [0, T], \quad (3.5)$$

$$d\tilde{Y}_t^x = \tilde{b}_2^x(t, \tilde{Y}_t^x)dt + \sigma_2(t, \tilde{Y}_t^x)dW_t, \quad t \in [0, T], \quad (3.6)$$

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

$$\forall y \in \mathbb{R}^d, \forall t \in [0, T], |\tilde{b}_i^x(t, y)| \leq \sup_{z \in \mathbf{B}(x, R+1)} |b_i(t, z)| \leq C(1 + |x|), \quad (3.7)$$

because  $b_i$  is Lipschitz continuous.

**Lemma 3.3.** Assume  $\text{Lip}_t(b_1)$ ,  $\text{Lip}_t(\sigma_1)$ ,  $b_1 \in \tilde{\mathcal{C}}_b^1$ ,  $\sigma_1 \in \mathcal{C}_b^1$ . Then for every  $x \in \mathbb{R}^d$  and  $t \in [0, T]$ ,

$$d_{\text{TV}}(X_t^x, \tilde{X}_t^x) \leq C(1 + |b_1(0, x)|^2)t. \quad (3.8)$$

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

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

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

$$\begin{aligned} |X_t^x - x|^2 &\leq 2\left|\int_0^t b_1(s, X_s^x)ds\right|^2 + 2\left|\int_0^t \sigma_1(s, X_s^x)dW_s\right|^2 \leq 2t\int_0^t |b_1(s, X_s^x)|^2 ds + 2\left|\int_0^t \sigma_1(s, X_s^x)dW_s\right|^2 \\ &\leq 4t[b_1]_{\text{Lip}}^2 \left(\int_0^t |X_s^x - x|^2 ds + \frac{1}{3}t^3\right) + 4t^2|b_1(0, x)|^2 + 2\left|\int_0^t \sigma_1(s, X_s^x)dW_s\right|^2 \end{aligned}$$

so that

$$\begin{aligned} \mathbb{E} \sup_{s \in [0, t]} |X_s^x - x|^2 &\leq 4t[b_1]_{\text{Lip}}^2 \int_0^t \left(\mathbb{E} \sup_{u \in [0, s]} |X_u^x - x|^2\right) ds + \frac{4}{3}[b_1]_{\text{Lip}}^2 t^4 + 4t^2|b_1(0, x)|^2 \\ &\quad + 2\mathbb{E} \sup_{s \in [0, t]} \left|\int_0^s \sigma_1(u, X_u^x)dW_u\right|^2. \end{aligned}$$

Moreover using Doob's martingale inequality we have

$$\mathbb{E} \sup_{s \in [0, t]} \left|\int_0^s \sigma_1(u, X_u^x)dW_u\right|^2 \leq 4\mathbb{E} \left|\int_0^t \sigma_1(u, X_u^x)dW_u\right|^2 = 4\mathbb{E} \int_0^t \sigma_1^2(u, X_u^x)du \leq 4\|\sigma_1\|_\infty^2 t.$$

Then we define the non-decreasing deterministic process  $S_t := \mathbb{E} \sup_{s \in [0, t]} |X_s^x - x|^2$  and we get the differential inequality (using  $t \leq T$ )

$$S_t \leq 4t \left( T|b_1(0, x)|^2 + \frac{1}{3}[b_1]_{\text{Lip}}^2 T^3 + 2\|\sigma_1\|_\infty^2 \right) + 4t[b_1]_{\text{Lip}}^2 \int_0^t S_s ds,$$

so the Gronwall lemma yields

$$S_t \leq 4t \left( T|b_1(0, x)|^2 + \frac{1}{3}[b_1]_{\text{Lip}}^2 T^3 + 2\|\sigma_1\|_\infty^2 \right) e^{2t^2[b_1]_{\text{Lip}}^2} \leq C(1 + |b_1(0, x)|^2)t.$$

Using Markov's inequality, we have then

$$|\mathbb{E}f(\tilde{X}_t^x) - \mathbb{E}f(X_t^x)| \leq 2\|f\|_\infty \mathbb{P} \left( \sup_{s \in [0, t]} |X_s^x - x|^2 > R^2 \right) \leq 2\|f\|_\infty \frac{C(1 + |b_1(0, x)|^2)t}{R^2}.$$

□

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

**Proposition 3.4.** *Let  $Z^x$  be the solution of*

$$Z_0^x = x, \quad dZ_t^x = \sigma_1(t, Z_t^x) dW_t, \quad t \in [0, T]. \quad (3.9)$$

*Assume  $\text{Lip}_t(b_1)$ ,  $\text{Lip}_t(\sigma_1)$ ,  $b_1 \in \tilde{C}_b^1$ ,  $\sigma_1 \in C_b^1$  and  $\sigma_1$  is elliptic. Then we have for every  $t \in [0, T]$ ,  $x, y \in \mathbb{R}^d$ ,*

$$p_{\tilde{X}}(0, t, x, y) = p_Z(0, t, x, y) + \int_0^t \mathbb{E} \left[ U_s^x \langle \tilde{b}_1^x(s, Z_s^x), \nabla_x p_Z(s, t, Z_s^x, y) \rangle \right] ds, \quad (3.10)$$

where  $\tilde{X}$  is defined in (3.5) and  $U^x$  is defined as

$$U_s^x = \exp \left( \int_0^s \langle g(u, Z_u^x) \tilde{b}_1^x(u, Z_u^x), dZ_u^x \rangle - \frac{1}{2} \int_0^s \langle g(u, Z_u^x) \tilde{b}_1^x(u, Z_u^x), \tilde{b}_1^x(u, Z_u^x) \rangle du \right), \quad (3.11)$$

$$g = (\sigma_1 \sigma_1^\top)^{-1}. \quad (3.12)$$

*Proof.* First, note that since  $\sigma_1$  is elliptic and since  $\tilde{b}_1^x, \sigma_1 \in C_b^1$ , then  $p_{\tilde{X}}$  and  $p_Z$  exist as well as  $\nabla_x p_Z$  (Theorem 3.1). We then use [QZ04, Theorem 2.4] extended to non-homogeneous diffusion processes. Following [QZ04, Remark 2.5], since  $\sigma_1$  is elliptic and bounded, the assumptions of [QZ04, Theorem 2.4] hold. □

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

**Lemma 3.5.** *With the same assumptions as in Proposition 3.4, for every  $p \geq 2$ ,  $x \in \mathbb{R}^d$  and  $t \in [0, T]$  we have*

$$\mathbb{E} \left[ \sup_{s \in [0, t]} |U_s^x|^p \right] \leq e^{Cp^2 \|\tilde{b}_1^x\|_\infty^2 t}. \quad (3.13)$$

*Proof.* We recall that for every  $q \geq 1$ , the process  $(U^x)^q$  is a martingale with

$$d(U_s^x)^q = q(U_s^x)^{q-1} \langle g(s, Z_s^x) \tilde{b}_1^x(s, Z_s^x), \sigma_1(s, Z_s^x) dW_s \rangle.$$

Thus, Doob's martingale inequality yields

$$\mathbb{E} \left[ \sup_{s \in [0, t]} |U_s^x|^{2q} \right] \leq Cq^2 \|\tilde{b}_1^x\|_\infty^2 \mathbb{E} \int_0^t |U_s^x|^{2q} ds \leq Cq^2 \|\tilde{b}_1^x\|_\infty^2 \int_0^t \mathbb{E} \left[ \sup_{u \in [0, s]} |U_s^x|^{2q} \right] ds.$$

So with  $U_0^x = 1$  we obtain

$$\mathbb{E} \left[ \sup_{s \in [0, t]} |U_s^x|^{2q} \right] \leq e^{Cq^2 \|\tilde{b}_1^x\|_\infty^2 t}.$$

□

**Lemma 3.6.** *With the same assumptions as in Proposition 3.4, we have for every  $x, y \in \mathbb{R}^d$  and  $t \in [0, T]$ ,*

$$\left| \int_0^t \mathbb{E} \left[ U_s^x \langle \tilde{b}_1^x(s, Z_s^x), \nabla_x p_Z(s, t, Z_s^x, y) \rangle \right] ds \right| \leq C e^{C \|\tilde{b}_1^x\|_\infty^2} \frac{e^{-c|y-x|^2/t}}{t^{(d-1)/2}}. \quad (3.14)$$

*Proof.* We use Theorem 3.1 on the process  $Z^x$ , which yields bounds with constants depending on  $\sigma_1$  but not on  $\tilde{b}_1^x$ . We obtain for every  $q \geq 1$  and for every  $s \in [0, t]$ :

$$\begin{aligned} \mathbb{E} |\nabla_x p_Z(s, t, Z_s^x, y)|^q &= \int_{\mathbb{R}^d} |\nabla_x p_Z(s, t, \xi, y)|^q p_Z(0, s, x, \xi) d\xi \\ &\leq \frac{C_q}{(t-s)^{(q+(q-1)d)/2}} \int_{\mathbb{R}^d} \frac{1}{(s(t-s))^{d/2}} \exp\left(-c_q \left(\frac{|y-\xi|^2}{t-s} + \frac{|\xi-x|^2}{s}\right)\right) d\xi \\ &\leq \frac{C_q}{t^{d/2}} \frac{1}{(t-s)^{(q+(q-1)d)/2}} e^{-c_q |y-x|^2/t}, \end{aligned}$$

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

$$\begin{aligned} &\left| \int_0^t \mathbb{E} \left[ U_s^x \langle \tilde{b}_1^x(s, Z_s^x), \nabla_x p_Z(s, t, Z_s^x, y) \rangle \right] ds \right| \\ &\leq \|\tilde{b}_1^x\|_\infty \left( \sup_{s \in [0, t]} \mathbb{E} |U_s^x|^p \right)^{1/p} \int_0^t (\mathbb{E} |\nabla_x p_Z(s, t, Z_s^x, y)|^q)^{1/q} ds \\ &\leq \|\tilde{b}_1^x\|_\infty e^{Cp \|\tilde{b}_1^x\|_\infty^2 t} \frac{C_q e^{-c_q |y-x|^2/t}}{t^{d/(2q)}} \int_0^t \frac{ds}{(t-s)^{(1+(1-q^{-1})d)/2}}. \end{aligned}$$

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

$$\left| \int_0^t \mathbb{E} \left[ U_s^x \langle \tilde{b}_1^x(s, Z_s^x), \nabla_x p_Z(s, t, Z_s^x, y) \rangle \right] ds \right| \leq \|\tilde{b}_1^x\|_\infty e^{Cp \|\tilde{b}_1^x\|_\infty^2 T} \frac{C_q e^{-c_q |y-x|^2/t}}{t^{(d-1)/2}} \leq C e^{C \|\tilde{b}_1^x\|_\infty^2} \frac{e^{-c|y-x|^2/t}}{t^{(d-1)/2}}.$$

□

### 3.3 Proof of Theorem 2.1

**Lemma 3.7.** *We have for every  $x \in \mathbb{R}^d$  and  $t \in [0, T]$ :*

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq d_{\text{TV}}(Z_t^x, V_t^x) + C e^{C|x|^2} t^{1/2},$$

where  $dZ_t^x = \sigma_1(t, Z_t^x) dW_t$  and  $dV_t^x = \sigma_2(t, V_t^x) dW_t$ .

*Proof.* Let us write

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq d_{\text{TV}}(X_t^x, \tilde{X}_t^x) + d_{\text{TV}}(\tilde{X}_t^x, Z_t^x) + d_{\text{TV}}(Z_t^x, V_t^x) + d_{\text{TV}}(V_t^x, \tilde{Y}_t^x) + d_{\text{TV}}(\tilde{Y}_t^x, Y_t^x),$$

where  $\tilde{X}$  and  $\tilde{Y}$  are defined in (3.5) and (3.6). Using Lemma 3.3 with (3.7), we have

$$d_{\text{TV}}(X_t^x, \tilde{X}_t^x) + d_{\text{TV}}(\tilde{Y}_t^x, Y_t^x) \leq C(1 + |x|^2)t.$$

Using the formula (3.10) and the inequality (3.14), we have

$$\begin{aligned} d_{\text{TV}}(\tilde{X}_t^x, Z_t^x) &= \int_{\mathbb{R}^d} |p_{\tilde{X}}(0, t, x, y) - p_Z(0, t, x, y)| dy = \int_{\mathbb{R}^d} \left| \int_0^t \mathbb{E} \left[ U_s^x \langle \tilde{b}_1^x(s, Z_s^x), \nabla_x p_Z(s, t, Z_s^x, y) \rangle \right] ds \right| dy \\ &\leq C e^{C \|\tilde{b}_1^x\|_\infty^2} t^{1/2} \int_{\mathbb{R}^d} \frac{e^{-c|x-y|^2/t}}{t^{d/2}} dy \leq C e^{C|x|^2} t^{1/2}, \end{aligned}$$

where we used (3.7). The term  $d_{\text{TV}}(\tilde{Y}_t^x, V_t^x)$  is treated likewise.  $\square$

We now prove Theorem 2.1.

*Proof.* Let us introduce an artificial regularization. For  $\varepsilon > 0$  and using Lemma 3.7 we have

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq C e^{C|x|^2} t^{1/2} + d_{\text{TV}}(Z_t^x, Z_t^x + \sqrt{\varepsilon}\zeta) + d_{\text{TV}}(Z_t^x + \sqrt{\varepsilon}\zeta, V_t^x + \sqrt{\varepsilon}\zeta) + d_{\text{TV}}(V_t^x + \sqrt{\varepsilon}\zeta, V_t^x) \quad (3.15)$$

where  $\zeta \sim \mathcal{N}(0, I_d)$  and is independent of the Brownian motion  $W$ .

- Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be measurable and bounded and let us define

$$\varphi : y \in \mathbb{R}^d \mapsto \mathbb{E} f(Z_t^x + y) = \int_{\mathbb{R}^d} f(\xi + y) p_Z(0, t, x, \xi) d\xi = \int_{\mathbb{R}^d} f(\xi) p_Z(0, t, x, \xi - y) d\xi. \quad (3.16)$$

Then  $\varphi$  is  $C^2$  with

$$\nabla^2 \varphi(y) = \int_{\mathbb{R}^d} f(\xi) \nabla_y^2 p_Z(0, t, x, \xi - y) d\xi.$$

Moreover, using Theorem 3.1, we have

$$\|\nabla_y^2 p_Z(0, t, x, \xi - y)\| \leq \frac{C}{t^{(d+2)/2}} e^{-c|x-\xi+y|^2/t},$$

where the constants  $C$  and  $c$  do not depend on  $\tilde{b}_1^x$ . This implies that for every  $y \in \mathbb{R}^d$ ,

$$\|\nabla^2 \varphi(y)\|_\infty \leq C \|f\|_\infty t^{-1} \int_{\mathbb{R}^d} \frac{1}{t^{d/2}} e^{-c|x-\xi+y|^2/t} d\xi \leq C \|f\|_\infty t^{-1}.$$

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

$$\begin{aligned} |\mathbb{E} f(Z_t^x + \sqrt{\varepsilon}\zeta) - \mathbb{E} f(Z_t^x)| &= |\mathbb{E} \varphi(\sqrt{\varepsilon}\zeta) - \varphi(0)| = \left| \sqrt{\varepsilon} \mathbb{E} [\nabla \varphi(0) \cdot \zeta] + \frac{\varepsilon}{2} \mathbb{E} [\nabla^2 \varphi(\sqrt{\varepsilon}\tilde{\zeta}) \cdot \zeta^{\otimes 2}] \right| \\ &\leq C \varepsilon \|f\|_\infty t^{-1}, \end{aligned}$$

where we used that  $\mathbb{E}[\nabla\varphi(0) \cdot \zeta] = \nabla\varphi(0) \cdot \mathbb{E}[\zeta] = 0$ . This way we obtain

$$d_{\text{TV}}(Z_t^x, Z_t^x + \sqrt{\varepsilon}\zeta) \leq C\varepsilon t^{-1}.$$

The term  $d_{\text{TV}}(V_t^x, V_t^x + \sqrt{\varepsilon}\zeta)$  is treated likewise.

- Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be measurable and bounded and let us define

$$f_\varepsilon : y \mapsto \mathbb{E}f(y + \sqrt{\varepsilon}\zeta) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(y + \sqrt{\varepsilon}\xi) e^{-|\xi|^2/2} d\xi = \frac{1}{(2\pi\varepsilon)^{d/2}} \int_{\mathbb{R}^d} f(\xi) e^{-|\xi-y|^2/(2\varepsilon)} d\xi. \quad (3.17)$$

Then  $f_\varepsilon$  is  $C^1$  with

$$\begin{aligned} \nabla f_\varepsilon(y) &= \frac{1}{(2\pi\varepsilon)^{d/2}} \int_{\mathbb{R}^d} f(\xi) \frac{\xi - y}{\varepsilon} e^{-|\xi-y|^2/(2\varepsilon)} d\xi = \frac{\varepsilon^{-1/2}}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(y + \sqrt{\varepsilon}\xi) \xi e^{-|\xi|^2/2} d\xi \\ &= \varepsilon^{-1/2} \mathbb{E}[f(y + \sqrt{\varepsilon}\zeta)\zeta] \end{aligned}$$

and then

$$[f_\varepsilon]_{\text{Lip}} \leq \|f\|_\infty \varepsilon^{-1/2} \mathbb{E}|\zeta| \leq C\|f\|_\infty \varepsilon^{-1/2}. \quad (3.18)$$

So that

$$\begin{aligned} |\mathbb{E}f(Z_t^x + \sqrt{\varepsilon}\zeta) - \mathbb{E}f(V_t^x + \sqrt{\varepsilon}\zeta)| &= |\mathbb{E}f_\varepsilon(Z_t^x) - \mathbb{E}f_\varepsilon(V_t^x)| \\ &\leq \frac{C\|f\|_\infty}{\sqrt{\varepsilon}} \|Z_t^x - V_t^x\|_1 \leq \frac{C\|f\|_\infty}{\sqrt{\varepsilon}} (t + t^{1/2} \Delta\sigma(x)), \end{aligned} \quad (3.19)$$

where we used Lemma A.3 in the Appendix. This implies that

$$d_{\text{TV}}(Z_t^x + \sqrt{\varepsilon}\zeta, V_t^x + \sqrt{\varepsilon}\zeta) \leq C\varepsilon^{-1/2} (t + t^{1/2} \Delta\sigma(x)).$$

- **Conclusion :** Considering (3.15), we have

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq C e^{C|x|^2} t^{1/2} + C\varepsilon t^{-1} + C\varepsilon^{-1/2} (t + t^{1/2} \Delta\sigma(x)).$$

We now choose  $\varepsilon = [t(t + t^{1/2} \Delta\sigma(x))]^{2/3}$ , so that

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq C e^{C|x|^2} t^{1/2} + C(t^{1/2} + \Delta\sigma(x))^{2/3}.$$

□

### 3.4 Proof of Theorem 2.2 using Theorem 2.7

We first prove Theorem 2.7.

*Proof.* Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  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

$$\begin{aligned} |\mathbb{E}f(Z_1) - \mathbb{E}f(Z_2)| &\leq \left| \mathbb{E}f(Z_1) - \sum_{i=1}^r w_i \mathbb{E}f_{\varepsilon/n_i}(Z_1) \right| + \left| \sum_{i=1}^r w_i \mathbb{E}f_{\varepsilon/n_i}(Z_1) - \sum_{i=1}^r w_i \mathbb{E}f_{\varepsilon/n_i}(Z_2) \right| \\ &\quad + \left| \sum_{i=1}^r w_i \mathbb{E}f_{\varepsilon/n_i}(Z_2) - \mathbb{E}f(Z_2) \right|, \end{aligned} \quad (3.20)$$

where  $f_\varepsilon$  is defined as in (3.17) and where the  $n_i$ 's and the  $w_i$ 's will be defined later.

Let  $\varphi$  be as defined in (3.16) replacing  $Z_t^x$  by  $Z_1$ . Then,  $\varphi$  is differentiable up to the order  $2r$  and for all  $k = 0, 1, \dots, 2r$ :

$$\nabla^k \varphi(y) = (-1)^k \int_{\mathbb{R}^d} f(\xi) \nabla^k p_1(\xi - y) d\xi.$$

Using the Taylor formula up to order  $2r$ , for every  $y \in \mathbb{R}^d$  there exists  $\tilde{y} \in (0, y)$  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

$$|\nabla^{2r} \varphi(\tilde{y}) \cdot y^{\otimes 2r}| \leq C \|f\|_\infty |y|^{2r} \int_{\mathbb{R}^d} \|\nabla^{2r} p_1(\xi)\| d\xi. \quad (3.21)$$

Then there exists a random  $\tilde{\zeta} \in (0, \zeta)$  such that

$$\begin{aligned} \mathbb{E}f(Z_1 + \sqrt{\varepsilon}\zeta) - \mathbb{E}f(Z_1) &= \mathbb{E}\varphi(\sqrt{\varepsilon}\zeta) - \varphi(0) = \sum_{k=1}^{2r-1} \frac{\nabla^k \varphi(0)}{k!} \varepsilon^{k/2} \cdot \mathbb{E}[\zeta^{\otimes k}] + \frac{\mathbb{E}[\nabla^{2r} \varphi(\sqrt{\varepsilon}\tilde{\zeta}) \cdot \zeta^{\otimes 2r}]}{(2r)!} \varepsilon^r \\ &= \sum_{k=1}^{r-1} \frac{\nabla^{2k} \varphi(0)}{(2k)!} \varepsilon^k \cdot \mathbb{E}[\zeta^{\otimes 2k}] + \frac{\mathbb{E}[\nabla^{2r} \varphi(\sqrt{\varepsilon}\tilde{\zeta}) \cdot \zeta^{\otimes 2r}]}{(2r)!} \varepsilon^r =: \sum_{k=1}^{r-1} \beta_k(t) \varepsilon^k + \tilde{\beta}_r(t, \varepsilon) \varepsilon^r, \end{aligned} \quad (3.22)$$

because if  $k$  is odd, then  $\mathbb{E}[\zeta^{\otimes k}] = 0$ . We now rely on a multi-step Richardson-Romberg extrapolation [LP17, Appendix A]. Let us denote the refiners  $n_i = 2^{i-1}$  and the auxiliary sequences and weights

$$u_k := \left( \prod_{\ell=1}^{k-1} (1 - 2^{-\ell}) \right)^{-1}, \quad v_k := (-1)^k 2^{-k(k+1)/2} u_{k+1}, \quad w_k := u_k v_{r-k}, \quad k = 1, \dots, r. \quad (3.23)$$

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

$$\sum_{i=1}^r w_i n_i^{-k} = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{else.} \end{cases}, \quad k = 0, 1, \dots, r-1. \quad (3.24)$$

Then we have

$$\begin{aligned} \sum_{i=1}^r w_i \left( \mathbb{E}f(Z_1 + \sqrt{\varepsilon/n_i}\zeta) - \mathbb{E}f(Z_1) \right) &= \sum_{i=1}^r w_i \sum_{k=1}^{r-1} \beta_k(t) \varepsilon^k n_i^{-k} + \sum_{i=1}^r w_i \tilde{\beta}_r(t, \varepsilon/n_i) \varepsilon^r n_i^{-r} \\ &= \sum_{k=1}^{r-1} \varepsilon^k \beta_k(t) \sum_{i=1}^r w_i n_i^{-k} + \varepsilon^r \sum_{i=1}^r \tilde{\beta}_r(t, \varepsilon/n_i) w_i n_i^{-r} \\ &= \varepsilon^r \sum_{i=1}^r \tilde{\beta}_r(t, \varepsilon/n_i) w_i n_i^{-r}, \end{aligned} \quad (3.25)$$

where we used (3.24) in the last equation. Now, using (3.21) we have

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

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

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

so that

$$\sum_{i=1}^r \frac{|w_i|}{n_i^r} \leq u_\infty^2 \sum_{i=1}^r 2^{-(r-i)(r-i+1)/2} \leq u_\infty^2 \sum_{i=1}^r 2^{(r-i)/2} = u_\infty^2 \sum_{i=0}^{r-1} 2^{-i/2} \leq C. \quad (3.26)$$

As a consequence and since  $\sum_{i=1}^r w_i = 1$ , we may write from (3.25)

$$\left| \mathbb{E}f(Z_1) - \sum_{i=1}^r w_i \mathbb{E}f_{\varepsilon/n_i}(Z_1) \right| \leq C \|f\|_\infty \varepsilon^r \int_{\mathbb{R}^d} \|\nabla^{2r} p_1(\xi)\| d\xi. \quad (3.27)$$

The same way, we obtain

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

On the other side, using (3.18) we have

$$\left| \sum_{i=1}^r w_i \mathbb{E}f_{\varepsilon/n_i}(Z_1) - \sum_{i=1}^r w_i \mathbb{E}f_{\varepsilon/n_i}(Z_2) \right| \leq \frac{C \|f\|_\infty}{\sqrt{\varepsilon}} \mathcal{W}_1(Z_1, Z_2) \left( \sum_{i=1}^r |w_i| 2^{(i-1)/2} \right). \quad (3.28)$$

Moreover, for every  $i = 1, \dots, r$ ,

$$|w_i| 2^{(i-1)/2} \leq u_\infty^2 2^{-(r-i)(r-i+1)/2 + (i-1)/2}$$

and then

$$\sum_{i=1}^r |w_i| 2^{(i-1)/2} \leq u_\infty^2 \sum_{i=1}^r 2^{(i-1)/2} \leq u_\infty^2 2^r. \quad (3.29)$$

Thus considering (3.20), we obtain for every  $\varepsilon > 0$ ,

$$d_{\text{TV}}(Z_1, Z_2) \leq C \varepsilon^r \int_{\mathbb{R}^d} (\|\nabla^{2r} p_1(\xi)\| + \|\nabla^{2r} p_2(\xi)\|) d\xi + C \varepsilon^{-1/2} \mathcal{W}_1(Z_1, Z_2).$$

Optimizing in  $\varepsilon$  gives

$$\varepsilon_\star = \left( \mathcal{W}_1(Z_1, Z_2) / (2r \int_{\mathbb{R}^d} (\|\nabla^{2r} p_1(\xi)\| + \|\nabla^{2r} p_2(\xi)\|) d\xi) \right)^{2/(2r+1)}$$

and then

$$d_{\text{TV}}(Z_1, Z_2) \leq C_{d,r} \mathcal{W}_1(Z_1, Z_2)^{2r/(2r+1)} \left( \int_{\mathbb{R}^d} (\|\nabla^{2r} p_1(\xi)\| + \|\nabla^{2r} p_2(\xi)\|) d\xi \right)^{1/(2r+1)}.$$

□

We now prove Theorem 2.2.

*Proof.* Using Lemma 3.7, we have

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq C e^{C|x|^2} t^{1/2} + d_{\text{TV}}(Z_t^x, V_t^x) \quad (3.30)$$

We now apply Theorem 2.7 with the random vectors  $Z_1 = Z_t^x$  and  $Z_2 = V_t^x$ . Assuming that  $\sigma_1$  is  $C_b^{2r}$  and using Theorem 3.1,  $\nabla_y^k p_Z$  exists for  $k = 0, 1, \dots, 2r$  and

$$\forall k = 0, 1, \dots, 2r, \forall t \in (0, T], \forall x, y \in \mathbb{R}^d, \|\nabla_y^k p_Z(0, t, x, y)\| \leq \frac{C}{t^{(d+k)/2}} e^{-c|y-x|^2/t}.$$

Then we have

$$\int_{\mathbb{R}^d} \nabla_y^{2r} p_Z(0, t, x, \xi) d\xi \leq Ct^{-r} \int_{\mathbb{R}^d} \frac{1}{t^{d/2}} e^{-c|x-\xi+y|^2/t} d\xi \leq Ct^{-r}.$$

The same way we have

$$\int_{\mathbb{R}^d} \nabla_y^{2r} p_V(0, t, x, \xi) d\xi \leq Ct^{-r}.$$

Applying Theorem 2.7 with Lemma A.3 yields

$$d_{\text{TV}}(Z_t^x, V_t^x) \leq C(\sqrt{t} + \Delta\sigma(x))^{2r/(2r+1)}.$$

□

### 3.5 Proof of Theorem 2.5

For the proof of Theorem 2.5, we do not use Lemma 3.7; instead we directly apply Theorem 2.7. Using Theorem 3.2,  $\nabla_y^k p_X$  and  $\nabla_y^k p_Y$  exist for  $k = 0, 1, \dots, 2r$  and satisfy the same bounds as previously. Then using Theorem 2.7 with Lemma A.3 we obtain

$$d_{\text{TV}}(X_t^x, Y_t^x) \leq C(\sqrt{t}(1 + \Delta b(x)) + \Delta\sigma(x) + t|b(x)|)^{2r/(2r+1)}.$$

### 3.6 Proof of Theorem 2.8

*Proof.* We use Lemma 3.7 again and rework the bound on  $d_{\text{TV}}(Z_t^x, V_t^x)$  by paying attention to the dependency of the constants in  $r$  in the proof of Theorem 2.7 with  $Z_1 := Z_t^x$  and  $Z_2 := V_t^x$ . Since  $\sigma_1 \in \mathcal{C}_b^{2r}$  for every  $r \in \mathbb{N}$ , we write (3.22) for any  $r \in \mathbb{N}$  and we have

$$|\tilde{\beta}_r(t, \varepsilon)| \leq \tilde{C}_{2r} \|f\|_{\infty} t^{-r} \frac{\mathbb{E}[|\zeta|^{2r}]}{(2r)!}, \quad \tilde{C}_{2r} := C_{2r} c_{2r}^{-d/2},$$

where  $C_{2r}$  and  $c_{2r}$  are defined in (2.12) and where

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

Using (3.26) we get

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

and we obtain as in (3.27):

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

On the other hand, considering (3.28) and (3.29) with Lemma A.3 with  $\Delta\sigma(x) = 0$  we have

$$\left| \sum_{i=1}^r w_i \mathbb{E}f_{\varepsilon/n_i}(V_t^x) - \sum_{i=1}^r w_i \mathbb{E}f_{\varepsilon/n_i}(Z_t^x) \right| \leq \kappa_2 \frac{\|f\|_{\infty}}{\sqrt{\varepsilon}} t, \quad \kappa_2 := C2^r. \quad (3.31)$$



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

$$\varepsilon_\star = \frac{t^{(2r+2)/(2r+1)}}{(2r\kappa_1)^{2/(2r+1)}} \kappa_2^{2/(2r+1)}$$

and then

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

with as  $r \rightarrow \infty$ :

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

with

$$\begin{aligned} \left( \prod_{i=0}^{r-1} (d+2i) \right)^{\frac{1}{(2r+1)}} &= \exp \left( \frac{r}{2r+1} \frac{1}{r} \sum_{i=0}^{r-1} \log(d+2i) \right) \leq \exp \left( \frac{r}{2r+1} \log(d+(r-1)) \right) \leq \sqrt{d+r-1}, \\ \frac{1}{(2r)!^{1/(2r+1)}} &\sim \frac{e}{2r}, \quad \limsup_{r \rightarrow \infty} \tilde{C}_{2r}^{1/(2r+1)} < \infty \end{aligned}$$

where we used Assumption (2.12), so that

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

Then we have  $d_{\text{TV}}(Z_t^x, V_t^x) \leq C 2^r r^{-1/2} t^{r/(2r+1)}$  and we choose  $r(t) = \lceil \log^{1/2}(1/t) \rceil$  so that as  $t \rightarrow 0$ ,

$$d_{\text{TV}}(Z_t^x, V_t^x) \leq C t^{1/2} \exp \left( C \sqrt{\log(1/t)} \right).$$

□

## 4 Counterexample

In this section we give a counter-example showing that we cannot achieve a bound better than  $t^{1/2}$  in general. More specifically, we show that we cannot achieve a bound better than  $t^{1/2}$  for the total variation between an SDE and its Euler-Maruyama-scheme in general. For  $x > 0$  and  $\sigma > 0$ , let us consider the one-dimensional process

$$Y_t^x = x e^{\sigma W_t}, \tag{4.1}$$

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

$$\bar{Y}_t^x = x + (\sigma^2/2)xt + \sigma x W_t \sim \mathcal{N}(x(1 + t\sigma^2/2), \sigma^2 x^2 t). \tag{4.2}$$

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

$$d_{\text{TV}}(Y_t^x, \bar{Y}_t^x) \geq C_x t^{1/2}. \tag{4.3}$$

*Proof.* We have

$$p_Y(t, x, y) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \frac{\exp\left(-\frac{1}{2\sigma^2 t} \log^2(y/x)\right)}{y} \mathbf{1}_{y \geq 0} \tag{4.4}$$

so that

$$\begin{aligned} d_{\text{TV}}(Y_t^x, \bar{Y}_t^x) &= \frac{1}{\sqrt{2\pi\sigma^2 t}} \int_{\mathbb{R}} \left| \exp\left(-\frac{\log^2(y/x)}{2\sigma^2 t}\right) y^{-1} \mathbb{1}_{y \geq 0} - \exp\left(-\frac{(y-x-xt\sigma^2/2)^2}{2\sigma^2 x^2 t}\right) x^{-1} \right| dy \\ &\geq \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-x/\sqrt{t}}^{\infty} \left| \frac{1}{x + \sqrt{t}y} \exp\left(-\frac{\log^2(1 + \sqrt{t}y/x)}{2\sigma^2 t}\right) - \frac{1}{x} \exp\left(-\frac{(y-x\sqrt{t}\sigma^2/2)^2}{2\sigma^2 x^2}\right) \right| dy. \end{aligned}$$

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

$$\begin{aligned} &\frac{1}{1 + \sqrt{t}y/x} \exp\left(-\frac{\log^2(1 + \sqrt{t}y/x)}{2\sigma^2 t}\right) - \exp\left(-\frac{(y-x\sqrt{t}\sigma^2/2)^2}{2\sigma^2 x^2}\right) \\ &= (1 - \sqrt{t}y/x + O(ty^2)) \exp\left(-\frac{1}{2\sigma^2 t} \left(\frac{ty^2}{x^2} - \frac{t^{3/2}y^3}{x^3} + O(t^2y^4)\right)\right) - \exp\left(-\frac{y^2}{2\sigma^2 x^2} - \frac{t\sigma^2}{8} + \frac{\sqrt{t}y}{2\sigma^2 x}\right) \\ &= e^{-\frac{y^2}{2\sigma^2 x^2}} \left[ (1 - \sqrt{t}y/x + O(ty^2)) \left(1 + \frac{\sqrt{t}y^3}{2\sigma^2 x^3} + O(ty^4)\right) - \left(1 + \frac{\sqrt{t}y}{2\sigma^2 x} - \frac{t\sigma^2}{8} + O(t^2) + O(ty^2)\right) \right] \\ &= e^{-\frac{y^2}{2\sigma^2 x^2}} \left[ -\frac{\sqrt{t}y}{x} - \frac{\sqrt{t}y}{2\sigma^2 x} + \frac{\sqrt{t}y^3}{2\sigma^2 x^3} + \frac{t\sigma^2}{8} + O(ty^2) + O(t^2) \right]. \end{aligned}$$

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

$$d_{\text{TV}}(Y_t^x, \bar{Y}_t^x) \geq \frac{1}{\sqrt{2\pi\sigma^2 x^2}} e^{-\frac{\epsilon^2}{2\sigma^2 x^2}} \frac{\sqrt{t}}{2} \int_{-\epsilon}^{\epsilon} \left| -\frac{y}{x} - \frac{y}{2\sigma^2 x} + \frac{y^3}{2\sigma^2 x^3} \right| dy,$$

so that  $d_{\text{TV}}(Y_t^x, \bar{Y}_t^x)$  is of order  $t^{1/2}$  as  $t \rightarrow 0$ .  $\square$

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  $X$  on  $\mathbb{R}$  with  $\mathcal{C}^1$  and Lipschitz continuous drift, with  $\mathcal{C}_b^\infty$  and elliptic diffusion coefficient  $\sigma$  and there exists  $T > 0$  and  $\epsilon \in (0, 1)$  such that*

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

where  $\bar{X}$  is the Euler-Maruyama scheme of  $X$  and where the positive constant  $C_x$  depends on  $x$ .

*Proof.* We construct from the geometric Brownian motion  $Y$  defined in (4.1), a process  $X$  with elliptic and bounded drift and such that  $d_{\text{TV}}(X_t^x, \bar{X}_t^x) \geq C_x t^{1/2}$ . For  $\epsilon \in (0, 1/2)$ , let us consider  $\psi : \mathbb{R} \rightarrow \mathbb{R}^+$  a  $\mathcal{C}_b^\infty$  approximation of

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

such that  $\psi = \tilde{\psi}$  on  $[2\epsilon, \epsilon^{-1}/2] \cup (-\infty, \epsilon/2] \cup [2\epsilon^{-1}, \infty)$ . Then we define the process with elliptic and bounded noise coefficient

$$dX_t^x = -\frac{\sigma^2}{2} X_t^x dt + \sigma \psi(X_t^x) dW_t.$$

Then for  $x \in (2\epsilon, \epsilon^{-1}/2)$  we have  $\bar{X}_t^x = \bar{Y}_t^x$  and

$$\mathbb{P}(Y_t^x \neq X_t^x) \leq \mathbb{P}\left(\sup_{s \in [0, t]} Y_s^x \geq \epsilon^{-1}/2\right) + \mathbb{P}\left(\inf_{s \in [0, t]} Y_s^x \leq 2\epsilon\right).$$

With a proof similar to the proof of Lemma 3.3, we show that

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

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

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

Then we obtain

$$d_{\text{TV}}(X_t^x, \bar{X}_t^x) \geq d_{\text{TV}}(Y_t^x, \bar{Y}_t^x) - d_{\text{TV}}(X_t^x, Y_t^x) \geq C_x \sqrt{t}.$$

□

*Remark 4.3.* We could also consider the process  $X$  with "cut" bounded drift  $\tilde{b}$  and get the same bounds, proving then that we cannot achieve better bounds in general than the ones established in Theorem 2.8 even 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  $0 < \varepsilon < 1$ ,

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

Let us recall [PP20, Lemma 3.4(a)], with an immediate adaptation to the non-homogeneous case.

**Lemma A.2.** Let  $Z$  be solution to the generic SDE:

$$Z_0^x = x \in \mathbb{R}^d, \quad dZ_t^x = b(t, Z_t^x)dt + \sigma(t, Z_t^x)dW_t, \quad t \in [0, T],$$

where  $b$  and  $\sigma$  are Lipschitz continuous in  $(t, x)$  and where  $\sigma$  is bounded. Then for  $p \geq 1$ ,

$$\forall t \in [0, T], \quad \forall x \in \mathbb{R}^d, \quad \|Z_t^x - x\|_p \leq C(p, T, [b]_{\text{Lip}}, [\sigma]_{\text{Lip}}, \|\sigma\|_\infty) \left(t|b(0, x)| + t^{1/2}\right).$$

**Lemma A.3.** Let  $X$  and  $Y$  be the solution to the two general SDEs (2.1) and (2.2). Assume that  $b$  and  $\sigma$  are Lipschitz continuous in  $(t, x)$  and that  $\sigma$  is bounded. Then for every  $p \geq 1$ ,

$$\forall t \in [0, T], \quad \forall x \in \mathbb{R}^d, \quad \|X_t^x - Y_t^x\|_p \leq C \left(t(1 + \Delta b(x)) + t^{3/2}(|b_1| + |b_2|)(0, x) + \Delta \sigma(x)t^{1/2}\right) \quad (\text{A.1})$$

*Proof.* We first deal with the case  $p \geq 2$ . We have

$$\begin{aligned} \|X_t^x - Y_t^x\|_p &\leq \left\| \int_0^t (b_1(s, X_s^x) - b_2(s, Y_s^x)) ds \right\|_p + \left\| \int_0^t (\sigma_1(s, X_s^x) - \sigma_2(s, Y_s^x)) dW_s \right\|_p \\ &\leq \left\| \int_0^t (b_1(s, X_s^x) - b_1(0, x)) ds \right\|_p + t \Delta b(x) + \left\| \int_0^t (b_2(s, Y_s^x) - b_2(0, x)) ds \right\|_p \\ &\quad + \left\| \int_0^t (\sigma_1(s, X_s^x) - \sigma_1(0, x)) dW_s \right\|_p + \left\| \int_0^t \Delta \sigma(x) dW_s \right\|_p + \left\| \int_0^t (\sigma_2(s, Y_s^x) - \sigma_2(0, x)) dW_s \right\|_p \end{aligned}$$

But using the Burkholder-Davis-Gundy and the generalized Minkowski inequalities, we have

$$\begin{aligned} \left\| \int_0^t (\sigma_1(s, X_s^x) - \sigma_1(0, x)) dW_s \right\|_p &\leq C_p^{\text{BDG}} [\sigma_1]_{\text{Lip}} \left\| \int_0^t |(s, X_s^x) - (0, x)|^2 ds \right\|_{p/2}^{1/2} \\ &\leq C_p^{\text{BDG}} [\sigma_1]_{\text{Lip}} \left( \int_0^t \|(s, X_s^x) - (0, x)\|_p^2 ds \right)^{1/2} \leq C(t + t^{3/2}|b_1(0, x)|), \end{aligned}$$

where  $C_p^{\text{BDG}}$  is a constant which only depends on  $p$  and where we used Lemma A.2. So that

$$\begin{aligned} \|X_t^x - Y_t^x\|_p &\leq [b_1]_{\text{Lip}} \int_0^t \|(s, X_s^x) - (0, x)\|_p ds + [b_2]_{\text{Lip}} \int_0^t \|(s, Y_s^x) - (0, x)\|_p + t\Delta b(x) \\ &\quad + C(t + t^{3/2}(|b_1| + |b_2|)(0, x)) + \Delta\sigma(x)\sqrt{t}\|W_1\|_p \\ &\leq C \left( t(\Delta b(x) + 1) + t^{3/2}(|b_1| + |b_2|)(0, x) + \Delta\sigma(x)\sqrt{t} \right) \end{aligned}$$

which completes the proof for  $p \geq 2$ . For  $p \in [1, 2)$ , the inequality is still true remarking that  $\|\cdot\|_p \leq \|\cdot\|_2$ .  $\square$

## 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.
- [BJ22] 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. *SIAM J. Numer. Anal.*, 60(4):1701–1740, 2022.
- [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. *arXiv e-prints*, page arXiv:2109.11669, September 2021.
- [BT96] Vlad Bally and Denis 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.
- [Cle21] Emmanuelle Clement. Hellinger and total variation distance in approximating Levy driven SDEs. *arXiv e-prints*, page arXiv:2103.09648, March 2021.
- [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  $\mathbf{R}^n$ . *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, to appear in Annals of Applied Probability*, 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.