Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme

Aurélien Alfonsi, Benjamin Jourdain, Arturo Kohatsu-Higa

Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme

A. Alfonsi, B. Jourdain*and A. Kohatsu-Higa †

September 3, 2012

Abstract

In the present paper, we prove that the Wasserstein distance on the space of continuous sample-paths equipped with the supremum norm between the laws of a uniformly elliptic one-dimensional diffusion process and its Euler discretization with N steps is smaller than $O(N^{-2/3+\varepsilon})$ where ε is an arbitrary positive constant. This rate is intermediate between the strong error estimation in $O(N^{-1/2})$ obtained when coupling the stochastic differential equation and the Euler scheme with the same Brownian motion and the weak error estimation $O(N^{-1})$ obtained when comparing the expectations of the same function of the diffusion and of the Euler scheme at the terminal time T. We also check that the supremum over $t \in [0,T]$ of the Wasserstein distance on the space of probability measures on the real line between the laws of the diffusion at time t and the Euler scheme at time t behaves like $O(\sqrt{\log(N)}N^{-1})$.

Keywords: Euler scheme, Wasserstein distance, weak trajectorial error, diffusion bridges
AMS (2010): 65C30, 60H35

For $\sigma : \mathbb{R} \to \mathbb{R}$ and $b : \mathbb{R} \to \mathbb{R}$, we are interested in the simulation of the stochastic differential equation

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt \tag{0.1}$$

where $X_0 = x_0 \in \mathbb{R}$ and $W = (W_t)_{t\geq 0}$ is a standard Brownian motion. We make the standard Lipschitz assumptions on the coefficients:

$$\exists K \in (0,+\infty),\ \forall x, y \in \mathbb{R},\ |\sigma(x) - \sigma(y)| + |b(x) - b(y)| \leq K|x - y|.$$

For $T > 0$, we are interested in the approximation of $X = (X_t)_{t\in[0,T]}$ by its Euler scheme $\bar{X} = (\bar{X}_t)_{t\in[0,T]}$ with $N \geq 1$ time-steps. We consider the regular grid $\{0 = t_0 < t_1 < t_2 < \ldots < t_N = T\}$ of the interval $[0,T]$ with $t_k = \frac{kT}{N}$ and define inductively $\bar{X}_0 = x_0$ and

$$\bar{X}_t = \bar{X}_{t_k} + \sigma(\bar{X}_{t_k})(W_t - W_{t_k}) + b(\bar{X}_{t_k})(t - t_k) \text{ for } t \in [t_k, t_{k+1}]. \tag{0.2}$$

*Université Paris-Est, CERMICS, Projet MathFi ENPC-INRIA-UMLV, 6 et 8 avenue Blaise Pascal, 77455 Marne La Vallée, Cedex 2, France, e-mails : alfonsi@cermics.enpc.fr, jourdain@cermics.enpc.fr. This research benefited from the support of the "Chaire Risques Financiers", Fondation du Risque and the French National Research Agency (ANR) under the program ANR-08-BLAN-0218 BigMC.

†Ritsumeikan University and Japan Science and Technology Agency, Department of Mathematical Sciences, 1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577, Japan. e-mail: arturokohatsu@gmail.com. This research was supported by grants of the Japanese goverment.

It is well known that the order of convergence of the strong error of discretization is $N^{-1/2}$. Indeed, we have (see [17])

$$\forall p \geq 1,\ \exists C < +\infty,\ \forall N \geq 1,\ \mathbb{E}^{1/p}\left[\sup_{t \leq T} |X_t - \bar{X}_t|^p\right] \leq \frac{C}{\sqrt{N}}. \tag{0.3}$$

See Section 1 for a more precise statement. This upper-bound gives the correct order of convergence since according to Remark 3.6 [20], when σ and b are continuously differentiable, $(\sqrt{N}(X_t - \bar{X}_t))_{t \leq T}$ converges in law as N goes to ∞ to some diffusion limit which is non zero as soon as σ is positive and non constant (see also [21] and [15] where stable convergence is also proved). When σ is constant, then the Euler scheme coincides with the Milstein scheme and the strong order of convergence is N^{-1}.
On the other hand, the order of convergence of the weak error of discretization is always N^{-1}. For example, according to [31], when σ and b are C^∞ with bounded derivatives of all orders and $f : \mathbb{R} \to \mathbb{R}$ is C^∞ with polynomial growth together with its derivatives then, for each integer $L \geq 1$, the expansion

$$\mathbb{E}[f(X_T)] - \mathbb{E}[f(\bar{X}_T)] = \sum_{l=1}^{L} \frac{a_l}{N^l} + \mathcal{O}(N^{-(L+1)}) \tag{0.4}$$

in powers of N^{-1} holds for the weak error. The bound $|\mathbb{E}[f(\bar{X}_T)] - \mathbb{E}[f(X_T)]| \leq \frac{C}{N}$ holds when σ, b and f are C^4 with the same growth assumptions. When f is only assumed to be measurable and bounded, it is proved in [2, 3] that the expansion (0.4) still holds for $L = 1$ if b and σ are smooth functions satisfying an hypoellipticity condition. Under uniform ellipticity, [13] even extends this expansion by only assuming that f is a tempered distribution acting on the densities of both X_T and $\bar{X}_T$.

In view of financial applications, the weak error analysis gives the convergence rate to 0 of the discretization bias introduced when replacing X by its Euler scheme $\bar{X}$ for the computation of the price $\mathbb{E}[f(X_T)]$ of a vanilla European option with payoff f and maturity T written on X. Let $\mathcal{C}$ denote the space $C([0,T], \mathbb{R})$ of continuous paths endowed with the sup norm. When dealing with exotic options with payoff $F : \mathcal{C} \to \mathbb{R}$ Lipschitz continuous,

$$\left|\mathbb{E}[F(X)] - \mathbb{E}[F(\bar{X})]\right| \leq \mathbb{E}|F(X) - F(\bar{X})| \leq \frac{C}{\sqrt{N}},$$

where the second inequality follows from the strong error estimate. But the first inequality is very rough and prevents from taking advantage of the cancellations in the mean which occur and permit to obtain the upper-bound $\frac{C}{N}$ for vanilla options. The weak error analysis has been performed for specific path-dependent payoffs, typically when $F(X) = f(X_T, Y_T)$ with Y_t a function of $(X_s)_{0 \leq s \leq t}$ such that $((X_t, Y_t))_{0 \leq t \leq T}$ is a Markov process. The cases $Y_t = \int_0^t X_s ds$ and $Y_t = \max_{0 \leq s \leq t} X_s$ respectively correspond to Asian [30] and barrier [9, 10, 11] or lookback options [27]. But no general theory has been developped so far to analyse the weak trajectorial error. The Wasserstein distance between the laws $\mathcal{L}(X)$ and $\mathcal{L}(\bar{X})$ of X and $\bar{X}$ defined by

$$\mathcal{W}_1(\mathcal{L}(X), \mathcal{L}(\bar{X})) = \sup_{F : \mathcal{C} \to \mathbb{R} : Lip(F) \leq 1} |\mathbb{E}[F(\bar{X})] - \mathbb{E}[F(X)]|,$$

where $Lip(F)$ denotes the Lipschitz constant of F is the appropriate measure to deal with the whole class of exotic Lipschitz payoffs. Notice that this distance has already been used in the context of discretization schemes for SDEs : in the multidimensional setting, by a clever rotation of the driving Brownian motion, Cruzeiro, Malliavin and Thalmaier [4] construct a modified Milstein scheme which does not involve the simulation of iterated Brownian integrals and with order of convergence N^{-1} for the Wasserstein distance. A simpler scheme with the same convergence properties is exhibited in [16] for usual stochastic volatility models.

The weak and strong error estimations recalled above imply that

$$\exists c, C < +\infty,\ \forall N \geq 1,\ \frac{c}{N} \leq \mathcal{W}_1(\mathcal{L}(X), \mathcal{L}(\bar{X})) \leq \frac{C}{\sqrt{N}}. \tag{0.5}$$

A very nice feature of the Wasserstein distance is its primal representation in the Kantorovitch duality theory. This representation is obtained by choosing $p = 1$, $E = \mathcal{C}$ and $(\mu, \nu) = (\mathcal{L}(X), \mathcal{L}(\bar{X}))$ in the general definition

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{E \times E} |x - y|^p \pi(dx, dy) \right)^{1/p} \tag{0.6}$$

where $p \in [1, +\infty)$, $(E, |\ |)$ is a normed vector space, μ and ν are two probability measures on E endowed with its Borel sigma-field and the infimum is computed on the set $\Pi(\mu, \nu)$ of probability measures on $E \times E$ with respective marginals μ and ν (see for instance Remark 6.5 p95 [29]). When one is able to exhibit some coupling $(Y, \bar{Y})$ with $Y \overset{\mathcal{L}}{=} X$ and $\bar{Y} \overset{\mathcal{L}}{=} \bar{X}$, then the law of $(Y, \bar{Y})$ belongs to $\Pi(\mathcal{L}(X), \mathcal{L}(\bar{X}))$ and necessarily $\mathcal{W}_p(\mathcal{L}(X), \mathcal{L}(\bar{X})) \leq \mathbb{E}^{1/p}\left[\sup_{t \in [0,T]} |Y_t - \bar{Y}_t|^p\right]$. For the obvious coupling $(Y, \bar{Y}) = (X, \bar{X})$ obtained by choosing the same driving Brownian motion for the diffusion and its Euler scheme, one recovers the upper-bound in (0.5) from the strong error analysis. The main result of the present paper is the construction of a better coupling which leads to the upper-bound

$$\forall p \geq 1,\ \forall \varepsilon > 0,\ \exists C < +\infty,\ \forall N \geq 1,\ \mathcal{W}_p(\mathcal{L}(X), \mathcal{L}(\bar{X})) \leq \frac{C}{N^{\frac{2}{3} - \varepsilon}}$$

proved in Section 3 under additional regularity assumptions on the coefficients and uniform ellipticity. To construct this coupling, we first obtain in Section 2 a time-uniform estimation of the Wasserstein distance between the respective laws $\mathcal{L}(X_t)$ and $\mathcal{L}(\bar{X}_t)$ of X_t and $\bar{X}_t$:

$$\forall p \geq 1,\ \exists C < +\infty,\ \forall N \geq 1,\ \sup_{t \in [0,T]} \mathcal{W}_p(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t)) \leq \frac{C\sqrt{\log(N)}}{N}.$$

Before, in Section 1, we recall well-known results concerning the moments and the dependence on the initial condition of the solution to the SDE (0.1) and its Euler scheme. Also, we explicit the dependence of the strong error estimations $\mathbb{E}[\sup_{s \leq t} |\bar{X}_s - X_s|]$ with respect to $t \in [0, T]$, which will play a key role in our analysis.

1 Basic estimates on the SDE and its Euler scheme

We recall some well-known results concerning the flow defined by (0.1) (see e.g. Karatzas and Shreve [18], p 306) and its Euler approximation.

Proposition 1.1 *Let us denote by $(X_t^x)_{t \in [0,T]}$ the solution of* (0.1)*, starting from $x \in \mathbb{R}$. One* *has that for any $p \geq 1$, the existence of a positive constant $C \equiv C(p,T)$ such that:*

$$\forall x \in \mathbb{R},\ \mathbb{E}\left[\sup_{t\in[0,T]} |X_t^x|^p\right] \leq C(1+|x|)^p \tag{1.1}$$

$$\forall x \in \mathbb{R},\ \forall s \leq t \leq T,\ \mathbb{E}\left[\sup_{u\in[s,t]} |X_u^x - X_s^x|^p\right] \leq C(1+|x|)^p(t-s)^{\frac{p}{2}} \tag{1.2}$$

$$\forall x,y \in \mathbb{R},\ \mathbb{E}\left[\sup_{t\in[0,T]} |X_t^x - X_t^y|^p\right] \leq C|y-x|^p \tag{1.3}$$

Proposition 1.2 *Let $(\bar{X}_t^x)_{t\in[0,T]}$ denote the Euler scheme* (0.2) *starting from x. For any $p \in [1,\infty)$, there exists a positive constant $C \equiv C(p,T)$ such that*

$$\forall N \geq 1,\ \forall x \in \mathbb{R},\ \mathbb{E}\left[\sup_{t\in[0,T]} |\bar{X}_t^x|^p\right] \leq C(1+|x|)^p \tag{1.4}$$

$$\forall N \geq 1,\ \forall x \in \mathbb{R},\ \forall t \in [0,T],\ \mathbb{E}\left[\sup_{r\in[0,t]} |\bar{X}_r^x - X_r^x|^p\right] \leq \frac{Ct^{\frac{p}{2}}(1+|x|)^p}{N^{\frac{p}{2}}}. \tag{1.5}$$

The moment bound (1.4) for the Euler scheme holds in fact as soon as the drift and the diffusion coefficients have a sublinear growth. The strong convergence order is established in Kanagawa [17] for Lipschitz and bounded coefficients. In fact, it is straightforward to extend Kanagawa's proof to merely Lipschitz coefficients by using the estimates (1.1) and (1.4) and obtain

$$\forall N \geq 1,\ \forall x \in \mathbb{R},\ \forall t \in [0,T],\ \mathbb{E}\left[\sup_{r\in[0,t]} |\bar{X}_r^x - X_r^x|^p\right] \leq \frac{C(1+|x|)^p}{N^{\frac{p}{2}}}. \tag{1.6}$$

The estimate (1.5) precises the dependence on t. This slight improvement will in fact play a crucial role to construct the coupling between the diffusion and the Euler scheme. We prove it for the sake of completeness even though the arguments are really standard.

Proof of (1.5)**.** Let $\tau_s = \sup\{t_i, t_i \leq s\}$ denote the last discretization time before s. We have $\bar{X}_t^x - X_t^x = \int_0^t b(\bar{X}_{\tau_s}^x) - b(X_s^x)ds + \int_0^t \sigma(\bar{X}_{\tau_s}^x) - \sigma(X_s^x)dW_s$. By Jensen's and Burkholder-Davis-Gundy inequalities,

$$\begin{aligned}\mathbb{E}\left[\sup_{r\in[0,t]} |\bar{X}_r^x - X_r^x|^p\right] &\leq 2^p\left(\mathbb{E}\left[\left(\int_0^t |b(\bar{X}_{\tau_s}^x) - b(X_s^x)|ds\right)^p\right] + C_p\mathbb{E}\left[\left(\int_0^t (\sigma(\bar{X}_{\tau_s}^x) - \sigma(X_s^x))^2 ds\right)^{\frac{p}{2}}\right]\right)\\ &\leq 2^p\left(t^{p-1}\int_0^t \mathbb{E}\left[|b(\bar{X}_{\tau_s}^x) - b(X_s^x)|^p\right]ds + C_p t^{\frac{p}{2}-1}\int_0^t \mathbb{E}\left[|\sigma(\bar{X}_{\tau_s}^x) - \sigma(X_s^x)|^p\right]ds\right)\end{aligned}$$

Denoting by $Lip(\sigma)$ the finite Lipschitz constant of σ, we have $|\sigma(\bar{X}_{\tau_s}^x) - \sigma(X_s^x)| \leq Lip(\sigma)(|\bar{X}_{\tau_s}^x - X_{\tau_s}^x| + |X_{\tau_s}^x - X_s^x|)$. Thus, (1.2) and (1.6) yield $\mathbb{E}[|\sigma(\bar{X}_{\tau_s}^x) - \sigma(X_s^x)|^p] \leq \frac{C(1+|x|)^p}{N^{\frac{p}{2}}}$, and the same bound holds for b replacing σ. Since $t^p \leq T^{p/2}t^{p/2}$, we easily conclude. ∎

2 The Wasserstein distance between the marginal laws

In this section, we are interested in finding an upper bound for the Wasserstein distance between the marginal laws of the SDE (0.1) and its Euler scheme. It is well known that the optimal coupling between two one-dimensional random variables is obtained by the inverse transform sampling. Thus, let F_t and $\bar{F}_t$ denote the respective cumulative distribution functions of X_t and $\bar{X}_t$. The p-Wasserstein distance between the time-marginals of the solution to the SDE and its Euler scheme is given by (see Theorem 3.1.2 in [24]):

$$\mathcal{W}_p(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t)) = \left(\int_0^1 |F_t^{-1}(u) - \bar{F}_t^{-1}(u)|^p du \right)^{1/p}. \tag{2.1}$$

Let us state now the main result of this Section. We set:

$$C_b^k = \{f : \mathbb{R} \to \mathbb{R} \ \ k \text{ times continuously differentiable } \ s.t. \ \|f^{(i)}\|_\infty < \infty, \ 0 \le i \le k\}.$$

Hypothesis 2.1 *Let $a = \sigma^2$. We assume that*

$$\begin{aligned}
&\exists \underline{a} > 0, \ \forall x \in \mathbb{R}, \ a(x) \ge \underline{a} \ \textit{(uniform ellipticity)},\\
&a \in C_b^2 \ \textit{and } a'' \textit{ is globally } \gamma\textit{-Hölder continuous with } \gamma > 0,\\
&b \in C_b^2.
\end{aligned}$$

Since σ is Lipschitz continuous, under Hypothesis 2.1, we have either $\sigma \equiv \sqrt{a}$ or $\sigma \equiv -\sqrt{a}$. From now on, we assume without loss of generality that $\sigma \equiv \sqrt{a}$ which is a C_b^2 function bounded from below by the positive constant $\underline{\sigma} = \sqrt{\underline{a}}$.

Theorem 2.2 *Under Hypothesis 2.1, we have for any $p \ge 1$,*

$$\forall N \ge 1, \ \sup_{t \in [0,T]} \mathcal{W}_p(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t)) \le \frac{C\sqrt{\log(N)}}{N},$$

where C is a positive constant that only depends on p, T, $\underline{a}$ and $(\|a^{(i)}\|_\infty, \|b^{(i)}\|_\infty, 0 \le i \le 2)$ and does not depend on the initial condition $x \in \mathbb{R}$.

Remark 2.3 *When $p = 1$, the slightly better bound $\sup_{t\in[0,T]} \mathcal{W}_1(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t)) \le \frac{C}{N}$ holds if σ is uniformly elliptic, according to [28] chapter 3. This is proved in a multidimensional setting for C^∞ coefficients σ and b with bounded derivatives by extending the results of [13] but can also be derived from a result of Gobet and Labart [12] only supposing that $b, \sigma \in C_b^3$. Let $p_t(x, y)$ and $\bar{p}_t(x, y)$ denote respectively the density of $X_t^{0,x}$ and $\bar{X}_t^{0,x}$. Then, Theorem 2.3 in [12] gives:*

$$\forall (t, x, y) \in (0, T] \times \mathbb{R}^2, \ |p_t(x, y) - \bar{p}_t(x, y)| \le \frac{TK(T)}{Nt} \exp\left(-\frac{c|x - y|^2}{t}\right).$$

As remarked in [28] chapter 3, for $f : \mathbb{R} \to \mathbb{R}$ a Lipschitz continuous function with Lipschitz constant not greater than one, one deduces that

$$\begin{aligned}
|\mathbb{E}[f(X_t)] - \mathbb{E}[f(\bar{X}_t)]| &= \left| \int_{\mathbb{R}} (f(y) - f(x))(p_t(x, y) - \bar{p}_t(x, y)) dy \right|\\
&\le \frac{K(T)T}{Nt} \int_{\mathbb{R}} |y - x| \exp\left(-\frac{c|x - y|^2}{t}\right) dy = \frac{K(T)T}{cN},
\end{aligned}$$

which gives $\sup_{t\le T} \mathcal{W}_1(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t)) \le \frac{CT}{N}$ by the dual formulation of the 1-Wasserstein distance.

Our approach consists in controlling the time evolution of the Wasserstein distance. To do so, we need to compute the evolution of both $F_t^{-1}(u)$ and $\bar{F}_t^{-1}(u)$. In the two next propositions, we derive partial differential equations satisfied by these functions by integrating in space the Fokker-Planck equations and then applying the implicit function theorem.

Proposition 2.4 *Assume that Hypothesis 2.1 holds.Then for any $t \in (0,T]$, the cumulative distribution function $x \mapsto F_t(x)$ is invertible with inverse denoted by $F_t^{-1}(u)$. Moreover, the function $(t,u) \mapsto F_t^{-1}(u)$ is $C^{1,2}$ on $(0,T] \times (0,1)$and satisfies*

$$\partial_t F_t^{-1}(u) = -\frac{1}{2}\partial_u \left(\frac{a(F_t^{-1}(u))}{\partial_u F_t^{-1}(u)} \right) + b(F_t^{-1}(u)). \tag{2.2}$$

Proposition 2.5 *Assume that σ and b have linear growth : $\exists C > 0,\ \forall x \in \mathbb{R},\ |\sigma(x)| + |b(x)| \leq C(1+|x|)$ and that uniform ellipticity holds : $\exists \underline{a} > 0,\ \forall x \in \mathbb{R},\ a(x) \geq \underline{a}$. Then for any $t \in (0,T]$, $\bar{X}_t$ admits a density $\bar{p}_t(x)$ with respect to the Lebesgue measure and its cumulative distribution function $x \mapsto \bar{F}_t(x)$ is invertible with inverse denoted by $\bar{F}_t^{-1}(u)$. Moreover, for each $k \in \{0,\dots,N-1\}$, the function $(t,u) \mapsto \bar{F}_t^{-1}(u)$ is $C^{1,2}$ on $(t_k, t_{k+1}] \times (0,1)$ and, on this set, it is a classical solution of*

$$\partial_t \bar{F}_t^{-1}(u) = -\frac{1}{2}\partial_u \left(\frac{\alpha_t(u)}{\partial_u \bar{F}_t^{-1}(u)} \right) + \beta_t(u). \tag{2.3}$$

where $\alpha_t(u) = \mathbb{E}[a(\bar{X}_{t_k})|\bar{X}_t = \bar{F}_t^{-1}(u)]$ and $\beta_t(u) = \mathbb{E}[b(\bar{X}_{t_k})|\bar{X}_t = \bar{F}_t^{-1}(u)]$.

The proofs of these two propositions are postponed to Appendix A. Let us mention here that Proposition 2.4 also holds when b' is only Hölder continuous: the Lipschitz assumption on b' is needed later to prove Theorem 2.2. The PDEs (2.2) and (2.3) enable us to compute the time derivative of the p-th power of the Wasserstein distance (2.1) and prove, again in Appendix A the following key lemma.

Lemma 2.6 *Under Hypothesis 2.1, for $p \geq 2$, the function $t \mapsto \mathcal{W}_p^p(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t))$ is continuous on $[0,T]$ and its first order distribution derivative $\partial_t \mathcal{W}_p^p(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t))$ is an integrable function on $[0,T]$. Moreover, dt a.e.,*

$$\begin{aligned}\partial_t \mathcal{W}_p^p(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t)) \leq C\Bigg(\mathcal{W}_p^p(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t)) + \int_0^1 |F_t^{-1}(u) - \bar{F}_t^{-1}(u)|^{p-1} |b(\bar{F}_t^{-1}(u)) - \beta_t(u)| du \\ + \int_0^1 |F_t^{-1}(u) - \bar{F}_t^{-1}(u)|^{p-2} \left(a(\bar{F}_t^{-1}(u)) - \alpha_t(u) \right)^2 du \Bigg),\end{aligned} \tag{2.4}$$

where C is a positive constant that only depends on p, $\underline{a}$, $\|a'\|_\infty$ and $\|b'\|_\infty$.

The last ingredient of the proof of Theorem 2.2 is the next Lemma, the proof of which is also postponed in Appendix A.

Lemma 2.7 *Let $\tau_t = \sup\{t_i, t_i \leq t\}$ denote the last discretization time before t. Under Hypothesis 2.1, we have for all $p \geq 1$:*

$$\exists C < +\infty,\ \forall N \geq 1,\ \forall t \in [0,T],\ \mathbb{E}\left[\left| \mathbb{E}\left[W_t - W_{\tau_t} | \bar{X}_t \right] \right|^p \right] \leq C \left(\frac{1}{N \vee (N^2 t)} \right)^{p/2}.$$

Proof of Theorem 2.2. Since $\mathcal{W}_p(\mathcal{L}(X_t),\mathcal{L}(\bar{X}_t)) \leq \mathcal{W}_{p'}(\mathcal{L}(X_t),\mathcal{L}(\bar{X}_t))$ for $p \leq p'$, it is enough to prove the estimation for $p \geq 2$. Therefore we suppose without loss of generality that $p \geq 2$. Let $\psi_p(t) = \mathcal{W}_p^2(\mathcal{L}(X_t),\mathcal{L}(\bar{X}_t))$ and

$$\text{for any integer } k \geq 1,\ h_k(x) = k^{-2/p}h(kx) \text{ where } h(x) = \begin{cases} x^{2/p} \text{ if } x \geq 1, \\ 1 + \frac{2}{p}(x-1) \text{ otherwise.} \end{cases}$$

Since h_k is C^1 and non-decreasing, Lemma 2.6 and Hölder's inequality imply that

$$\begin{aligned} h_k\left(\psi_p^{p/2}(t)\right) &= h_k\left(\mathcal{W}_p^p(\mathcal{L}(X_0),\mathcal{L}(\bar{X}_0))\right) + \int_0^t h_k'\left(\psi_p^{p/2}(s)\right)\partial_s \mathcal{W}_p^p(\mathcal{L}(X_s),\mathcal{L}(\bar{X}_s))ds \\ &\leq h_k(0) + C\int_0^t h_k'\left(\psi_p^{p/2}(s)\right)\left[\psi_p^{p/2}(s) + \psi_p^{(p-1)/2}(s)\left(\int_0^1 |b(\bar{F}_s^{-1}(u)) - \beta_s(u)|^p du\right)^{1/p}\right. \\ &\qquad \left. + \psi_p^{(p-2)/2}(s)\left(\int_0^1 |a(\bar{F}_s^{-1}(u)) - \alpha_s(u)|^p du\right)^{2/p}\right] ds. \end{aligned}$$

Since for fixed $x \geq 0$, the sequence $(h_k'(x))_k$ is non-decreasing and converges to $\frac{2}{p}x^{\frac{2}{p}-1}$ as $k \to \infty$, one may take the limit in this inequality thanks to the monotone convergence theorem and remark that the image of the Lebesgue measure on $[0,1]$ by $\bar{F}_s^{-1}$ is the distribution of $\bar{X}_s$ to deduce

$$\psi_p(t) \leq \frac{2C}{p}\int_0^t \psi_p(s) + \psi_p^{1/2}(s)\mathbb{E}^{1/p}\left(|b(\bar{X}_s) - \mathbb{E}(b(\bar{X}_{\tau_s})|\bar{X}_s)|^p\right) + \mathbb{E}^{2/p}\left(|a(\bar{X}_s) - \mathbb{E}(a(\bar{X}_{\tau_s})|\bar{X}_s)|^p\right)ds. \tag{2.5}$$

One has

$$\begin{aligned} a(\bar{X}_{\tau_s}) - a(\bar{X}_s) &= a'(\bar{X}_s)\sigma(\bar{X}_s)(W_{\tau_s} - W_s) - a'(\bar{X}_s)\left[(\sigma(\bar{X}_{\tau_s}) - \sigma(\bar{X}_s))(W_s - W_{\tau_s}) + b(\bar{X}_{\tau_s})(s-\tau_s)\right] \\ &\quad + (\bar{X}_{\tau_s} - \bar{X}_s)\int_0^1 a'(v\bar{X}_{\tau_s} + (1-v)\bar{X}_s) - a'(\bar{X}_s)dv. \end{aligned}$$

Using Jensen's inequality, the boundedness assumptions on a, b and their derivatives and Lemma 2.7, one gets

$$\begin{aligned} \mathbb{E}\left(|a(\bar{X}_s) - \mathbb{E}(a(\bar{X}_{\tau_s})|\bar{X}_s)|^p\right) &\leq C\mathbb{E}\left(|\sigma a'(\bar{X}_s)|^p|\mathbb{E}(W_s - W_{\tau_s})|\bar{X}_s)|^p\right) \\ &\quad + C\mathbb{E}\left((s-\tau_s)^p + |(\sigma(\bar{X}_{\tau_s}) - \sigma(\bar{X}_s))(W_s - W_{\tau_s})|^p + |\bar{X}_{\tau_s} - \bar{X}_s|^{2p}\right) \\ &\leq \frac{C}{N^{p/2} \vee (N^p s^{p/2})}. \end{aligned}$$

The same bound holds with a replaced by b. With (2.5) and Young's inequality, one deduces

$$\psi_p(t) \leq C\int_0^t \psi_p(s) + \frac{\psi_p^{1/2}(s)}{\sqrt{N} \vee (N\sqrt{s})} + \frac{1}{N \vee (N^2 s)}ds \leq C\int_0^t \psi_p(s) + \frac{1}{N \vee (N^2 s)}ds.$$

One concludes by Gronwall's lemma. ∎

Remark 2.8 *When $a(x) \equiv a$ is constant, the term $\mathbb{E}^{2/p}\left(|a(\bar{X}_s) - \mathbb{E}(a(\bar{X}_{\tau_s})|\bar{X}_s)|^p\right)$ in (2.5) vanishes and the above reasoning ensures that $\bar{\psi}_p(t)$ defined as $\sup_{s\in[0,T]}\psi_p(s)$ satisfies*

$$\bar{\psi}_p(t) \leq C\int_0^t \bar{\psi}_p(s)ds + C\bar{\psi}_p^{1/2}(t)\int_0^t \frac{1}{\sqrt{N} \vee (N\sqrt{s})}ds \leq C\int_0^t \bar{\psi}_p(s)ds + \frac{1}{2}\bar{\psi}_p(t) + \frac{C^2(T+1)^2}{2N}.$$

By Gronwall's lemma, we recover the estimation $\sup_{t\in[0,T]}\mathcal{W}_p(\mathcal{L}(X_t),\mathcal{L}(\bar{X}_t)) \leq \frac{C}{N}$ which is also a consequence of the strong order of convergence of the Euler scheme when the diffusion coefficient is constant.

3 The Wasserstein distance between the pathwise laws

We now state the main result of the paper.

Hypothesis 3.1 *We assume that $a \in C_b^4$, $b \in C_b^3$, and*

$$\exists \underline{a} > 0,\ \forall x \in \mathbb{R},\ a(x) \geq \underline{a} \text{ (uniform ellipticity).}$$

Clearly, Hypothesis 3.1 implies Hypothesis 2.1.

Theorem 3.2 *Under Hypothesis 3.1, we have:*

$$\forall p \geq 1,\ \forall \varepsilon > 0,\ \exists C < +\infty,\ \forall N \geq 1,\ \mathcal{W}_p(\mathcal{L}(X), \mathcal{L}(\bar{X})) \leq \frac{C}{N^{\frac{2}{3}-\varepsilon}}.$$

Before proving the theorem, let us state some of its consequences for the pricing of lookback options. It is well-known that if $(U_k)_{0 \leq k \leq N-1}$ are independent random variables uniformly distributed on $[0,1]$ and independent from the Brownian increments $(W_{t_{k+1}} - W_{t_k})_{0 \leq k \leq N-1}$ then $\bar{\bar{X}} \stackrel{\text{def}}{=} \frac{1}{2} \max_{0 \leq k \leq N-1} \left(\bar{X}_{t_k} + \bar{X}_{t_{k+1}} + \sqrt{(\bar{X}_{t_{k+1}} - \bar{X}_{t_k})^2 - 2\sigma^2(\bar{X}_{t_k}) t_1 \ln(U_k)} \right)$ is such that $\left(\bar{X}_0, \bar{X}_{t_1}, \dots, \bar{X}_T, \bar{\bar{X}} \right) \stackrel{\mathcal{L}}{=} (\bar{X}_0, \bar{X}_{t_1}, \dots, \bar{X}_T, \max_{t \in [0,T]} \bar{X}_t)$.

Corollary 3.3 *If $f : \mathbb{R}^2 \to \mathbb{R}$ is Lipschitz continuous, then, under Hypothesis 3.1,*

$$\forall \varepsilon > 0,\ \exists C < +\infty,\ \forall N \geq 1,\ \left| \mathbb{E}\left[f\left(X_T, \max_{t \in [0,T]} X_t \right) \right] - \mathbb{E}[f(\bar{X}_T, \bar{\bar{X}})] \right| \leq \frac{C}{N^{\frac{2}{3}-\varepsilon}}. \tag{3.1}$$

To our knowledge, this result appears to be new. Of course, when f is also differentiable with respect to its second variable, one has

$$\mathbb{E}\left[f\left(X_T, \max_{t \in [0,T]} X_t \right) \right] = \mathbb{E}\left[f\left(X_T, x_0 \right) \right] + \int_{x_0}^{+\infty} \mathbb{E}\left[\partial_2 f(X_T, x) 1_{\{\max_{t \in [0,T]} X_t \geq x\}} \right] dx.$$

One could contemplate combining the weak error analysis for the first term in the right-hand-side with Theorem 2.3 [10] devoted to barrier options to obtain the order N^{-1} instead on $N^{-2/3+\varepsilon}$ in (3.1). In Theorem 2.3 [10], Gobet assumes C_b^5 regularity and uniform ellipticity on the coefficients σ and b and it is not clear whether the estimation is preserved by integration over $[x_0, +\infty)$. More importantly a structure condition on the payoff function implying that $\partial_2 f(x,x) = 0$ for all $x \geq x_0$ is needed.

Proof of Theorem 3.2. We first deduce from Theorem 2.2 some bound on the Wasserstein distance between the finite dimensional marginals of the diffusion X and its Euler scheme $\bar{X}$ on a coarse time-grid. For $m \in \{1, \dots, N-1\}$, we set $n = \lfloor N/m \rfloor$ and define

$$s_l = \frac{lmT}{N},\ \text{for } l \in \{0, \dots, n-1\},\ \text{and } s_n = T.$$

Combining the next proposition, the proof of which is postponed in Appendix B with Theorem 2.2, one obtains that

$$\mathcal{W}_p(\mathcal{L}(X_{s_1}, \dots, X_{s_n}), \mathcal{L}(\bar{X}_{s_1}, \dots, \bar{X}_{s_n})) \leq \frac{C\sqrt{\log N}}{m} \tag{3.2}$$

where the constant C does not depend on (m, N).

Proposition 3.4 *Let $\mathbb{R}^n$ be endowed with the norm $|(x_1,\dots,x_n)| = \max_{1\le l\le n}|x_l|$. For any $p\ge 1$, there is a constant C not depending on n such that*

$$\mathcal{W}_p(\mathcal{L}(X_{s_1},\dots,X_{s_n}),\mathcal{L}(\bar{X}_{s_1},\dots,\bar{X}_{s_n}))\le Cn\sup_{0\le t\le T,x\in\mathbb{R}}\mathcal{W}_p(\mathcal{L}(\bar{X}^x_t),\mathcal{L}(X^x_t)).$$

There is a probability measure $\pi(dx_1,\dots,dx_n,d\bar{x}_1,\dots,d\bar{x}_n)$ in $\Pi(\mathcal{L}(X_{s_1},\dots,X_{s_n}),\mathcal{L}(\bar{X}_{s_1},\dots,\bar{X}_{s_n}))$ which attains the Wasserstein distance in the left-hand-side of (3.2) (see for instance Theorem 3.3.11 [24]). Let $\tilde{\pi}(x_1,\dots,x_n,d\bar{x}_1,\dots,d\bar{x}_n)$ denote a regular conditional probability of $(\bar{x}_1,\dots,\bar{x}_n)$ given $(x_1,\dots,x_n)$ when $\mathbb{R}^{2n}$ is endowed with π and $(\bar{Y}_{s_1},\dots,\bar{Y}_{s_n})$ be distributed according to $\tilde{\pi}(X_{s_1},\dots,X_{s_n},d\bar{x}_1,\dots,d\bar{x}_n)$. The vector $(X_{s_1},\dots,X_{s_n},\bar{Y}_{s_1},\dots,\bar{Y}_{s_n})$ is distributed according to π so that

$$(\bar{Y}_{s_1},\dots,\bar{Y}_{s_n})\overset{\mathcal{L}}{=}(\bar{X}_{s_1},\dots,\bar{X}_{s_n})\text{ and }\mathbb{E}^{1/p}\left[\max_{1\le l\le n}|X_{s_l}-\bar{Y}_{s_l}|^p\right]\le\frac{C\sqrt{\log N}}{m}. \tag{3.3}$$

Let $p_t(x,y)$ denote the transition density of the SDE (0.1) and $\ell_t(x,y)=\log(p_t(x,y))$. According to Appendix C devoted to diffusion bridges, the processes

$$\left(W^l_t=\int_{s_l}^t\left(dW_s-\sigma(X_s)\partial_x\ell_{s_{l+1}-s}(X_s,X_{s_{l+1}})ds\right),t\in[s_l,s_{l+1})\right)_{0\le l\le n-1}$$

are independent Brownian motions independent from $(X_{s_1},\dots,X_{s_n})$. We suppose from now on that the vector $(\bar{Y}_{s_1},\dots,\bar{Y}_{s_n})$ has been generated independently from these processes and so will be all the random variables and processes needed in the remaining of the proof (see in particular the construction of β below). Moreover

$$\begin{cases}Z^{x,y}_t=x+\int_{s_l}^t\sigma(Z^{x,y}_s)dW^l_s+\int_{s_l}^t[b(Z^{x,y}_s)+\sigma^2(Z^{x,y}_s)\partial_x\ell_{s_{l+1}-s}(Z^{x,y}_s,y)]ds,\ t\in[s_l,s_{l+1})\\ Z^{x,y}_{s_{l+1}}=y\end{cases} \tag{3.4}$$

is distributed according to the conditional law of $(X_t)_{t\in[s_l,s_{l+1}]}$ given $(X_{s_l},X_{s_{l+1}})=(x,y)$ and for each $l\in\{0,\dots,n-1\}$, one has $(Z_t^{X_{s_l},X_{s_{l+1}}})_{t\in[s_l,s_{l+1}]}=(X_t)_{t\in[s_l,s_{l+1}]}$. In order to construct a good coupling between $\mathcal{L}(X)$ and $\mathcal{L}(\bar{X})$, a natural idea would be to extend $(\bar{Y}_{s_1},\dots,\bar{Y}_{s_n})$ to a process $(\bar{Y}_t)_{t\in[0,T]}$ with law $\mathcal{L}(\bar{X})$ by defining for each $l\in\{0,\dots,n-1\}$, $(\bar{Y}_t)_{t\in[s_l,s_{l+1}]}$ as an Euler scheme bridge driven by W^l and starting from $\bar{Y}_{s_l}$ and ending at $\bar{Y}_{s_{l+1}}$. Unfortunately, even if the Euler scheme bridge is deduced by a simple transformation of the Brownian bridge on a single time-step, it becomes a complicated process when the difference between the starting and ending times is larger than $\frac{T}{N}$ because of the lack of Markov property. We are finally going to choose the difference $s_{l+1}-s_l$ of order $\frac{T}{N^{1/3}}$ and therefore much larger than the time-step $\frac{T}{N}$. In addition, it is not clear how to compare the paths of the diffusion bridge and the Euler scheme bridge driven by the same Brownian motion. That is why we are going to introduce some new process $(\tilde{\chi}_t)_{t\in[0,T]}$ such that the comparison will be performed at the diffusion bridge level, which is not so easy yet.

To construct $\tilde{\chi}$, we are going to exhibit a Brownian motion $(\beta_t)_{t\in[0,T]}$ such that $(\bar{Y}_{s_1},\dots,\bar{Y}_{s_n})$ are the values on the coarse time-grid of the Euler scheme with time-step $\frac{T}{N}$ driven by β. The extension $(\bar{Y}_t)_{t\in[0,T]}$ with law $\mathcal{L}(\bar{X})$ is then simply defined as the whole Euler scheme driven by β :

$$\bar{Y}_t=\bar{Y}_{t_k}+\sigma(\bar{Y}_{t_k})(\beta_t-\beta_{t_k})+b(\bar{Y}_{t_k})(t-t_k),\ t\in[t_k,t_{k+1}],\ 0\le k\le N-1.$$

The construction of β is postponed at the end of the present proof. One then defines

$$\chi_t=\bar{Y}_{s_l}+\int_{s_l}^t\sigma(\chi_s)d\beta_s+\int_{s_l}^t b(\chi_s)ds,t\in[s_l,s_{l+1}),\ 0\le l\le n-1$$

Notice that the process $\chi = (\chi_t)_{t\in[0,T]}$ which evolves according to the SDE (0.1) with β replacing W on each time-interval $[s_l, s_{l+1})$ is càdlàg : discontinuities may arise at the points $\{s_{l+1},\ 0 \leq l \leq n-1\}$. We denote by $\chi_{s_{l+1}-}$ its left-hand limit at time s_{l+1} and set $\chi_T = \chi_{s_n-}$. The strong error estimation (1.5) will permit to estimate the difference between the processes $\bar{Y}$ and χ. Of course, there is no hope for the processes χ and X to be close. Nethertheless, the process $\tilde{\chi}$ obtained by setting

$$\forall l \in \{0,\ldots,n-1\},\ \forall t \in [s_l, s_{l+1}),\ \tilde{\chi}_t = Z_t^{\chi_{s_l},\chi_{s_{l+1}-}} \text{ and } \tilde{\chi}_T = \chi_T$$

where $Z^{x,y}$ is defined in (3.4) is such that $\mathcal{L}(\tilde{\chi}) = \mathcal{L}(\chi)$ by Propositions C.1 and C.3. On each coarse time-interval $[s_l, s_{l+1})$ the diffusion bridges associated with X and $\tilde{\chi}$ are driven by the same Brownian motion W^l. Moreover the differences $|X_{s_l} - Y_{s_l}|$ between the starting points and $|X_{s_{l+1}} - \chi_{s_{l+1}-}| \leq |X_{s_{l+1}} - Y_{s_{l+1}}| + |Y_{s_{l+1}} - \chi_{s_{l+1}-}|$ between the ending points is controlled by (3.3) and the above mentionned strong error estimation. That is why one may expect to obtain a good estimation of the difference between the processes X and $\tilde{\chi}$. By the triangle inequality and since $\mathcal{L}(\bar{X}) = \mathcal{L}(\bar{Y})$ and $\mathcal{L}(\tilde{\chi}) = \mathcal{L}(\chi)$,

$$\begin{aligned}\mathcal{W}_p(\mathcal{L}(\bar{X}),\mathcal{L}(X)) &\leq \mathcal{W}_p(\mathcal{L}(\bar{X}),\mathcal{L}(\chi)) + \mathcal{W}_p(\mathcal{L}(\chi),\mathcal{L}(X))\\ &\leq \mathbb{E}^{1/p}\left[\sup_{t\in[0,T]}|\bar{Y}_t - \chi_t|^p\right] + \mathbb{E}^{1/p}\left[\sup_{t\in[0,T]}|X_t - \tilde{\chi}_t|^p\right], \end{aligned} \tag{3.5}$$

where, for the definition of $\mathcal{W}_p(\mathcal{L}(\bar{X}),\mathcal{L}(\chi))$ and $\mathcal{W}_p(\mathcal{L}(\chi),\mathcal{L}(X))$, the space of càdlàg sample-paths from $[0,T]$ to $\mathbb{R}$ is endowed with the supremum norm. Let us first estimate the first term in the right-hand-side. Let $q \geq 1$. From (1.5), we get

$$\mathbb{E}\left[\sup_{t\in[s_l,s_{l+1})}|\bar{Y}_t - \chi_t|^{pq}\middle|\bar{Y}_{s_l}\right] \leq C\frac{m^{\frac{pq}{2}}(1+|\bar{Y}_{s_l}|)^{pq}}{N^{pq}}$$

where the constant C does not depend on (N,m). We deduce that

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}|\bar{Y}_t - \chi_t|^{pq}\right] &= \mathbb{E}\left[\max_{0\leq l\leq n-1}\sup_{t\in[s_l,s_{l+1})}|\bar{Y}_t - \chi_t|^{pq}\right] \leq \sum_{l=0}^{n-1}\mathbb{E}\left[\mathbb{E}\left[\sup_{t\in[s_l,s_{l+1})}|\bar{Y}_t - \chi_t|^{pq}\middle|\bar{Y}_{s_l}\right]\right]\\ &\leq C\frac{m^{\frac{pq}{2}}}{N^{pq}}\sum_{l=0}^{n-1}\mathbb{E}\left[(1+|\bar{Y}_{s_l}|)^{pq}\right] \leq C\frac{m^{\frac{pq}{2}-1}}{N^{pq-1}},\end{aligned}$$

where we used (1.4) for the last inequality. As a consequence,

$$\mathbb{E}^{1/p}\left[\sup_{t\in[0,T]}|\bar{Y}_t - \chi_t|^p\right] \leq \mathbb{E}^{1/pq}\left[\sup_{t\in[0,T]}|\bar{Y}_t - \chi_t|^{pq}\right] \leq C\frac{m^{\frac{1}{2}-\frac{1}{pq}}}{N^{1-\frac{1}{pq}}}. \tag{3.6}$$

Let us now estimate the second term in the right-hand-side of (3.5). By Proposition C.3 and since for $l \in \{0,\ldots,n-1\}$, $\chi_{s_l} = \bar{Y}_{s_l}$,

$$\sup_{t\leq T}|X_t - \tilde{\chi}_t| = \max_{0\leq l\leq n-1}\sup_{t\in[s_l,s_{l+1})}|Z_t^{X_{s_l},X_{s_{l+1}}} - Z_t^{\chi_{s_l},\chi_{s_{l+1}-}}| \leq C\max_{0\leq l\leq n-1}|X_{s_l} - \bar{Y}_{s_l}| \vee |X_{s_{l+1}} - \chi_{s_{l+1}-}|.$$

Since, by the triangle inequality and the continuity of $\bar{Y}$,

$$|X_{s_{l+1}} - \chi_{s_{l+1}-}| \leq |X_{s_{l+1}} - \bar{Y}_{s_{l+1}}| + |\bar{Y}_{s_{l+1}} - \chi_{s_{l+1}-}| \leq |X_{s_{l+1}} - \bar{Y}_{s_{l+1}}| + \sup_{t\in[0,T]}|\bar{Y}_t - \chi_t|,$$

one deduces that

$$\sup_{t\le T}|X_t-\tilde{\chi}_t|\le C\left(\max_{1\le l\le n}|X_{s_l}-\bar{Y}_{s_l}|+\sup_{t\in[0,T]}|\bar{Y}_t-\chi_t|\right).$$

Combined with (3.3) and (3.6), this implies

$$\mathbb{E}^{1/p}\left[\sup_{t\le T}|X_t-\tilde{\chi}_t|^p\right]\le C\mathbb{E}^{1/p}\left[\max_{1\le l\le n}|X_{s_l}-\bar{Y}_{s_l}|^p\right]+C\mathbb{E}^{1/p}\left[\sup_{t\in[0,T]}|\bar{Y}_t-\chi_t|^p\right]\le C\left(\frac{\sqrt{\log N}}{m}+\frac{m^{\frac{1}{2}-\frac{1}{pq}}}{N^{1-\frac{1}{pq}}}\right).$$

Plugging this inequality together with (3.6) in (3.5), we deduce that

$$\mathcal{W}_p(\mathcal{L}(X),\mathcal{L}(\bar{X}))\le C\left(\frac{\sqrt{\log N}}{m}+\frac{m^{\frac{1}{2}-\frac{1}{pq}}}{N^{1-\frac{1}{pq}}}\right)$$

and conclude by choosing $m=\lfloor N^{\frac{2}{3}}\rfloor$ and $q\ge\frac{1}{3p\varepsilon}$.

To end the proof, we still have to construct the Brownian motion β. We first reconstruct on the fine time grid $(t_k)_{1\le k\le N}$ an Euler scheme $(\bar{Y}_{t_k},0\le k\le N)$ interpolating the values on the coarse grid $(s_l)_{1\le l\le n}$. Let us denote by $\bar{p}(x,y)$ the density of the law $\mathcal{N}(x+b(x)T/N,\sigma(x)^2T/N)$ of the Euler scheme starting from x after one time step T/N. Thanks to the ellipticity assumption, we have $\bar{p}(x,y)>0$ for any $x,y\in\mathbb{R}$. Conditionally on $(\bar{Y}_{s_1},\dots,\bar{Y}_{s_n})$, we generate independent random vectors

$$(\bar{Y}_{s_{l-1}+t_1},\dots,\bar{Y}_{s_{l-1}+t_{m-1}})_{1\le l\le n-1}\text{ and }(\bar{Y}_{s_{n-1}+t_1},\dots,\bar{Y}_{t_{N-1}})$$

with respective densities

$$\frac{\bar{p}(\bar{Y}_{s_{l-1}},x_1)\bar{p}(x_1,x_2)\dots\bar{p}(x_{n-1},\bar{Y}_{s_l})}{\int_{\mathbb{R}^{n-1}}\bar{p}(\bar{Y}_{s_{l-1}},y_1)\bar{p}(y_1,y_2)\dots\bar{p}(y_{n-1},\bar{Y}_{s_l})dy_1\dots dy_{n-1}}$$
$$\text{and }\frac{\bar{p}(\bar{Y}_{s_{n-1}},x_1)\bar{p}(x_1,x_2)\dots\bar{p}(x_{N-1-m(n-1)},\bar{Y}_{s_n})}{\int_{\mathbb{R}^{N-1-m(n-1)}}\bar{p}(\bar{Y}_{s_{n-1}},y_1)\bar{p}(y_1,y_2)\dots\bar{p}(y_{N-1-m(n-1)},\bar{Y}_{s_n})dy_1\dots dy_{N-1-m(n-1)}}$$

and get immediately $(\bar{Y}_{t_k})_{0\le k\le n}\overset{\mathcal{L}}{=}(\bar{X}_{t_k})_{0\le k\le n}$. Then, thanks to the ellipticity condition, $\left(\frac{1}{\sigma(\bar{Y}_{t_{k-1}})}(\bar{Y}_{t_k}-\bar{Y}_{t_{k-1}}-b(\bar{Y}_{t_{k-1}}))\right)_{1\le k\le N}$ are independent centered Gaussian variables with variance T/N. By using independent Brownian bridges, we can then construct a Brownian motion $(\beta_t)_{t\in[0,T]}$ such that

$$\beta_{t_k}-\beta_{t_{k-1}}=\frac{1}{\sigma(\bar{Y}_{t_{k-1}})}(\bar{Y}_{t_k}-\bar{Y}_{t_{k-1}}-b(\bar{Y}_{t_{k-1}})),$$

which ends the construction. ∎

Conclusion

In this paper, we prove that the order of convergence of the Wasserstein distance $\mathcal{W}_p$ on the space of continuous paths between the laws of a uniformly elliptic one-dimensional diffusion and its Euler scheme with N-steps is not worse that $N^{-2/3+\varepsilon}$. In view of a possible extension to multidimensional settings, two main difficulties have to be overcomed. First, we took advantage of the optimality of the inverse transform coupling in dimension one to obtain a uniform bound on

the Wasserstein distance between the marginal laws with optimal rate N^{-1} up to a logarithmic factor. In dimension $d > 1$, the optimal coupling between two probability measures on $\mathbb{R}^d$ is not available, which makes the estimation of the Wasserstein distance between the marginal laws much more complicated even if, for $\mathcal{W}_1$, the order N^{-1} may be deduced from the results of [12] (see Remark 2.3). In the second place, one has to generalize the estimation on diffusion bridges given by Proposition C.3 which we deduce from the Lamperti transform in dimension $d = 1$.
In the perspective of the multi-level Monte Carlo method introduced by Giles [8], coupling with order of convergence $N^{-2/3+\varepsilon}$ the Euler schemes with N and $2N$ steps would also be of great interest for variance reduction, especially in multidimensional situations where the Milstein scheme is not feasible (see [16] for the implementation of this idea in the example of a discretization scheme devoted to usual stochastic volatility models). But this does not seem obvious from our non constructive coupling between the Euler scheme and its diffusion limit.
For both the derivation of the order of convergence of the Wasserstein distance on the path space and the explicitation of the coupling, the limiting step in our approach is Proposition 3.4. In this proposition, we bound the dual formulation of the Wasserstein distance between n-dimensional marginals by the Wasserstein distance between one-dimensional marginals multiplied by n.
Even if the order of convergence of the Wasserstein distance on the path space obtained in the present paper may not be optimal, it provides the first significant step from the order $N^{1/2}$ obtained with the trivial coupling where the diffusion and the Euler scheme are driven by the same Brownian motion.

A Proofs of Section 2

Proof of Proposition 2.4. According to [6], Theorems 5.4 and 4.7, for any $t \in (0, T]$, X_t admits a density $p_t(x)$ w.r.t. the Lebesgue measure on the real line, the function $(t, x) \mapsto p_t(x)$ is $C^{1,2}$ on $(0, T] \times \mathbb{R}$ and, on this set, it is a classical solution of the Fokker-Planck equation

$$\partial_t p_t(x) = \frac{1}{2}\partial_{xx}(a(x)p_t(x)) - \partial_x(b(x)p_t(x)). \tag{A.1}$$

Moreover, the following Gaussian bounds hold

$$\exists C > 0,\ \forall t \in (0, T],\ \forall x \in \mathbb{R},\ |p_t(x)| + \sqrt{t}|\partial_x p_t(x)| \le \frac{C}{\sqrt{t}} e^{-\frac{(x-x_0)^2}{Ct}} \tag{A.2}$$

The partial derivatives $\partial_x F_t(x) = p_t(x)$ and $\partial_{xx} F_t(x) = \partial_x p_t(x)$ exist and are continuous on $(0, T] \times \mathbb{R}$. For $0 < s < t \le T$ and $y \le x$, integrating (A.1) over $[s, t] \times [y, x]$, then letting $y \to -\infty$ thanks to (A.2), one obtains $F_t(x) - F_s(x) = \int_s^t \frac{1}{2}\partial_x(a(x)p_r(x)) - b(x)p_r(x)dr$. By continuity of the integrand w.r.t. (r, x) one deduces that the partial derivative $\partial_t F_t(x)$ exists and is continuous on $(0, T] \times \mathbb{R}$. So, $(t, x) \mapsto F_t(x)$ is $C^{1,2}$ on $(0, T] \times \mathbb{R}$ and solves

$$\partial_t F_t(x) = \frac{1}{2}\partial_x(a(x)\partial_x F_t(x)) - b(x)\partial_x F_t(x). \tag{A.3}$$

According to [1], the density is also bounded from below by some Gaussian kernel : $\exists c > 0,\ \forall (t, x) \in (0, T] \times \mathbb{R},\ |p_t(x)| \ge \frac{c}{\sqrt{t}} e^{-\frac{(x-x_0)^2}{ct}}$. This enables us to apply the implicit function theorem to $(t, x, u) \mapsto F_t(x) - u$ to deduce that the inverse $u \mapsto F_t^{-1}(u)$ of $x \mapsto F_t(x)$ is $C^{1,2}$ in the variables $(t,u)\in(0,T]\times(0,1)$ and solves

$$\begin{aligned}\partial_t F_t^{-1}(u) &= -\frac{\partial_t F_t}{\partial_x F_t}(F_t^{-1}(u))\\ &= -\frac{1}{2}\partial_x(a(x)\partial_x F_t(x))|_{x=F_t^{-1}(u)}\partial_u F_t^{-1}(u)+b(F_t^{-1}(u))\\ &= -\frac{1}{2}\partial_u\left(\frac{a(F_t^{-1}(u))}{\partial_u F_t^{-1}(u)}\right)+b(F_t^{-1}(u))\end{aligned}$$

where we used (A.3) for the second equality and $\partial_u F_t^{-1}(u)=\frac{1}{\partial_x F_t(F_t^{-1}(u))}$ for both the second and the third equalities. ∎

Proof of Proposition 2.5. For $t\in(0,t_1]$, $\bar{X}_t$ admits the gaussian density with mean $x_0+b(x_0)t$ and variance $a(x_0)t$. By induction on k and independence of $W_t-W_{t_k}$ and $\bar{X}_{t_k}$ in (0.2), one checks that for $k\in\{1,\dots,n-1\}$, $\bar{X}_{t_k}$ admits a density $\bar{p}_{t_k}(x)$ and that for $t\in(t_k,t_{k+1}]$, $(\bar{X}_{t_k},\bar{X}_t)$ admits the density

$$\rho(t_k,t,y,x)=\bar{p}_{t_k}(y)\frac{e^{-\frac{(x-y-b(y)(t-t_k))^2}{2a(y)(t-t_k)}}}{\sqrt{2\pi a(y)(t-t_k)}}.$$

The marginal density $\bar{p}_t(x)=\int_{\mathbb{R}}\bar{p}_{t_k}(y)\frac{e^{-\frac{(x-y-b(y)(t-t_k))^2}{2a(y)(t-t_k)}}}{\sqrt{2\pi a(y)(t-t_k)}}dy$ of $\bar{X}_t$ is continuous on $(t_k,t_{k+1}]\times\mathbb{R}$ by Lebesgue's theorem and positive.

Let $N(x)=\int_{-\infty}^x e^{-\frac{y^2}{2}}\frac{dy}{\sqrt{2\pi}}$ denote the cumulative distribution function of the standard Gaussian law and $k\in\{0,\dots,N-1\}$. Again by the independence structure in (0.2), for $(t,x)\in(t_k,t_{k+1}]\times\mathbb{R}$, $\bar{F}_t(x)=\mathbb{E}\left(N\left(\frac{x-\bar{X}_{t_k}-b(\bar{X}_{t_k})(t-t_k)}{\sqrt{a(\bar{X}_{t_k})(t-t_k)}}\right)\right)$. One has

$$\partial_t N\left(\frac{x-y-b(y)(t-t_k)}{\sqrt{a(y)(t-t_k)}}\right)=-\left(\frac{x-y-b(y)(t-t_k)}{2\sqrt{2\pi a(y)(t-t_k)^3}}+\frac{b(y)}{\sqrt{2\pi a(y)(t-t_k)}}\right)e^{-\frac{(x-y-b(y)(t-t_k))^2}{2a(y)(t-t_k)}}.$$

By the growth assumption on σ and b, one easily checks that $\forall k\in\{0,\dots,N\}$, $\mathbb{E}(\bar{X}_{t_k}^2)<+\infty$. With the uniform ellipticity assumption and Lebesgue's theorem, one deduces that $\bar{F}_t(x)$ is differentiable w.r.t. t with partial derivative

$$\partial_t\bar{F}_t(x)=-\mathbb{E}\left[\left(\frac{x-\bar{X}_{t_k}-b(\bar{X}_{t_k})(t-t_k)}{2\sqrt{2\pi a(\bar{X}_{t_k})(t-t_k)^3}}+\frac{b(\bar{X}_{t_k})}{\sqrt{2\pi a(\bar{X}_{t_k})(t-t_k)}}\right)e^{-\frac{(x-\bar{X}_{t_k}-b(\bar{X}_{t_k})(t-t_k))^2}{2a(\bar{X}_{t_k})(t-t_k)}}\right]\tag{A.4}$$

continuous in $(t,x)\in(t_k,t_{k+1}]\times\mathbb{R}$. In the same way, one checks smoothness of $\bar{F}_t(x)$ in the spatial variable x and obtains that this function is $C^{1,2}$ on $(t_k,t_{k+1}]\times\mathbb{R}$.
When $k\geq1$,

$$\mathbb{E}\left[b(\bar{X}_{t_k})\frac{e^{-\frac{(x-\bar{X}_{t_k}-b(\bar{X}_{t_k})(t-t_k))^2}{2a(\bar{X}_{t_k})(t-t_k)}}}{\sqrt{2\pi a(\bar{X}_{t_k})(t-t_k)}}\right]=\int_{\mathbb{R}}b(y)\rho(t_k,t,y,x)dy=\mathbb{E}[b(\bar{X}_{t_k})|\bar{X}_t=x]\bar{p}_t(x).$$

For $k=0$, even if $(\bar{X}_0,\bar{X}_t)$ has no density, the equality between the opposite sides of this equation remains true.

Combining Lebesgue's theorem and a similar reasoning, one checks that

$$-\mathbb{E}\left[\frac{x-\bar{X}_{t_k}-b(\bar{X}_{t_k})(t-t_k)}{\sqrt{2\pi a(\bar{X}_{t_k})(t-t_k)^3}}e^{-\frac{(x-\bar{X}_{t_k}-b(\bar{X}_{t_k})(t-t_k))^2}{2a(\bar{X}_{t_k})(t-t_k)}}\right]=\partial_x\mathbb{E}\left[a(\bar{X}_{t_k})\frac{e^{-\frac{(x-\bar{X}_{t_k}-b(\bar{X}_{t_k})(t-t_k))^2}{2a(\bar{X}_{t_k})(t-t_k)}}}{\sqrt{2\pi a(\bar{X}_{t_k})(t-t_k)}}\right]$$
$$=\partial_x\left[\mathbb{E}(a(\bar{X}_{t_k})|\bar{X}_t=x)\bar{p}_t(x)\right].$$

With (A.4), one deduces that

$$\partial_t\bar{F}_t(x)=\frac{1}{2}\partial_x\left(\mathbb{E}[a(\bar{X}_{t_k})|\bar{X}_t=x]\partial_x\bar{F}_t(x)\right)-\mathbb{E}[b(\bar{X}_{t_k})|\bar{X}_t=x]\partial_x\bar{F}_t(x). \tag{A.5}$$

One checks that the function $(t,u)\mapsto\bar{F}_t^{-1}(u)$ is smooth and satisfies the partial differential equation (2.3) by arguments similar to the ones given at the end of the proof of Proposition 2.4.
■

Remark A.1 *In the same way, for $k\in\{0,\dots,N-1\}$, one could prove that on $(t_k,t_{k+1}]\times\mathbb{R}$, $(t,x)\mapsto\bar{p}_t(x)$ is $C^{1,2}$ and satisfies the partial differential*

$$\partial_t\bar{p}_t(x)=\frac{1}{2}\partial_{xx}\left(\mathbb{E}[a(\bar{X}_{t_k})|\bar{X}_t=x]\bar{p}_t(x)\right)-\partial_x\left(\mathbb{E}[b(\bar{X}_{t_k})|\bar{X}_t=x]\bar{p}_t(x)\right).$$

obtained by spatial derivation of (A.5)*. This is related to [14].*

Proof of Lemma 2.6. By the continuity of the paths of X and $\bar{X}$ and the finiteness of $\mathbb{E}\left[\sup_{t\le T}(|X_t|^{p+1}+|\bar{X}_t|^{p+1})\right]$, one easily checks that $t\mapsto\mathcal{W}_p^p(\mathcal{L}(X_t),\mathcal{L}(\bar{X}_t))$ is continuous.
Let $k\in\{0,\dots,N-1\}$ and $s,t\in(t_k,t_{k+1}]$ with $s\le t$. Combining Propositions 2.4 and 2.5 with a spatial integration by parts, one obtains for $\varepsilon\in(0,1/2)$

$$\begin{aligned}
&\int_\varepsilon^{1-\varepsilon}|F_t^{-1}(u)-\bar{F}_t^{-1}(u)|^pdu=\int_\varepsilon^{1-\varepsilon}|F_s^{-1}(u)-\bar{F}_s^{-1}(u)|^pdu\\
&+p\int_s^t\int_\varepsilon^{1-\varepsilon}|F_r^{-1}(u)-\bar{F}_r^{-1}(u)|^{p-2}(F_r^{-1}(u)-\bar{F}_r^{-1}(u))(b(F_r^{-1}(u))-\beta_r(u))dudr\\
&+\frac{p(p-1)}{2}\int_s^t\int_\varepsilon^{1-\varepsilon}|F_r^{-1}(u)-\bar{F}_r^{-1}(u)|^{p-2}(\partial_uF_r^{-1}(u)-\partial_u\bar{F}_r^{-1}(u))\left(\frac{a(F_r^{-1}(u))}{\partial_uF_r^{-1}(u)}-\frac{\alpha_r(u)}{\partial_u\bar{F}_r^{-1}(u)}\right)dudr\\
&+\frac{p}{2}\int_s^t|F_r^{-1}(1-\varepsilon)-\bar{F}_r^{-1}(1-\varepsilon)|^{p-2}(F_r^{-1}(1-\varepsilon)-\bar{F}_r^{-1}(1-\varepsilon))\left(\frac{\alpha_r(1-\varepsilon)}{\partial_u\bar{F}_r^{-1}(1-\varepsilon)}-\frac{a(F_r^{-1}(1-\varepsilon))}{\partial_uF_r^{-1}(1-\varepsilon)}\right)dr\\
&-\frac{p}{2}\int_s^t|F_r^{-1}(\varepsilon)-\bar{F}_r^{-1}(\varepsilon)|^{p-2}(F_r^{-1}(\varepsilon)-\bar{F}_r^{-1}(\varepsilon))\left(\frac{\alpha_r(\varepsilon)}{\partial_u\bar{F}_r^{-1}(\varepsilon)}-\frac{a(F_r^{-1}(\varepsilon))}{\partial_uF_r^{-1}(\varepsilon)}\right)dr
\end{aligned} \tag{A.6}$$

We are now going to take the limit as $\varepsilon\to 0$. We will check at the end of the proof that

$$\lim_{u\to 0^+\text{ or }1^-}\sup_{r\in[s,t]}\frac{a(F_t^{-1}(u))}{\partial_uF_t^{-1}(u)}|F_t^{-1}(u)-\bar{F}_t^{-1}(u)|^{p-1}+\sup_{r\in[s,t]}\frac{\alpha_t(u)}{\partial_u\bar{F}_t^{-1}(u)}|F_t^{-1}(u)-\bar{F}_t^{-1}(u)|^{p-1}=0. \tag{A.7}$$

which enables us to get rid of the two last boundary terms.

Combining Young's inequality with the uniform ellipticity assumption and the positivity of $\partial_u F_t^{-1}(u)$ and $\partial_u \bar{F}_t^{-1}(u)$, one obtains

$$\begin{aligned}
&(\partial_u F_r^{-1}(u) - \partial_u \bar{F}_r^{-1}(u)) \left(\frac{a(F_r^{-1}(u))}{\partial_u F_r^{-1}(u)} - \frac{\alpha_r(u)}{\partial_u \bar{F}_r^{-1}(u)} \right) \\
&= \big(a(F_r^{-1}(u)) - \alpha_r(u)\big) \frac{\partial_u F_r^{-1}(u) - \partial_u \bar{F}_r^{-1}(u)}{\partial_u F_r^{-1}(u) \vee \partial_u \bar{F}_r^{-1}(u)} \\
&- a(F_r^{-1}(u)) \frac{((\partial_u \bar{F}_r^{-1}(u) - \partial_u F_r^{-1}(u))^+)^2}{\partial_u F_r^{-1}(u) \partial_u \bar{F}_r^{-1}(u)} - \alpha_r(u) \frac{((\partial_u F_r^{-1}(u) - \partial_u \bar{F}_r^{-1}(u))^+)^2}{\partial_u F_r^{-1}(u) \partial_u \bar{F}_r^{-1}(u)} \\
&\leq \frac{1}{4\underline{a}} \big(a(F_r^{-1}(u)) - \alpha_r(u)\big)^2 + \underline{a} \frac{(\partial_u F_r^{-1}(u) - \partial_u \bar{F}_r^{-1}(u))^2}{(\partial_u F_r^{-1}(u) \vee \partial_u \bar{F}_r^{-1}(u))^2} - (a(F_r^{-1}(u)) \wedge \alpha_r(u)) \frac{(\partial_u \bar{F}_r^{-1}(u) - \partial_u F_r^{-1}(u))^2}{\partial_u F_r^{-1}(u) \partial_u \bar{F}_r^{-1}(u)} \\
&\leq \frac{1}{4\underline{a}} \big(a(F_r^{-1}(u)) - \alpha_r(u)\big)^2.
\end{aligned}$$

Hence, up to the factor $\frac{p(p-1)}{2}$, the third term of the right-hand-side of (A.6) is equal to

$$\begin{aligned}
&\int_s^t \int_\varepsilon^{1-\varepsilon} |F_r^{-1}(u) - \bar{F}_r^{-1}(u)|^{p-2} \Bigg[(\partial_u F_r^{-1}(u) - \partial_u \bar{F}_r^{-1}(u)) \left(\frac{a(F_r^{-1}(u))}{\partial_u F_r^{-1}(u)} - \frac{\alpha_r(u)}{\partial_u \bar{F}_r^{-1}(u)} \right) \\
&- \frac{\big(a(F_r^{-1}(u)) - \alpha_r(u)\big)^2}{4\underline{a}} \Bigg] dudr + \frac{1}{4\underline{a}} \int_s^t \int_\varepsilon^{1-\varepsilon} |F_r^{-1}(u) - \bar{F}_r^{-1}(u)|^{p-2} \big(a(F_r^{-1}(u)) - \alpha_r(u)\big)^2 dudr.
\end{aligned}$$

where the integrand in the first integral is non positive. Since

$$\begin{aligned}
&\int_s^t \int_0^1 |F_r^{-1}(u) - \bar{F}_r^{-1}(u)|^{p-2} \left(|F_r^{-1}(u) - \bar{F}_r^{-1}(u)| |b(F_r^{-1}(u)) - \beta_r(u)| + (a(F_r^{-1}(u)) - \alpha_r(u))^2 \right) dudr \\
&\qquad \leq 2\|b\|_\infty \int_s^t \mathcal{W}_p^{p-1}(\mathcal{L}(X_r), \mathcal{L}(\bar{X}_r)) dr + 4\|a\|_\infty^2 \int_s^t \mathcal{W}_p^{p-2}(\mathcal{L}(X_r), \mathcal{L}(\bar{X}_r)) dr < +\infty,
\end{aligned}$$

one can take the limit $\varepsilon \to 0$ in (A.6) using Lebesgue's theorem for the second term of the right-hand-side and combining Lebesgue's theorem with monotone convergence for the third term to obtain

$$\begin{aligned}
&\mathcal{W}_p^p(\mathcal{L}(X_t), \mathcal{L}(\bar{X}_t)) = \mathcal{W}_p^p(\mathcal{L}(X_s), \mathcal{L}(\bar{X}_s)) \\
&\quad + p \int_s^t \int_0^1 |F_r^{-1}(u) - \bar{F}_r^{-1}(u)|^{p-2} (F_r^{-1}(u) - \bar{F}_r^{-1}(u))(b(F_r^{-1}(u)) - \beta_r(u)) dudr \\
&\quad + \frac{p(p-1)}{2} \int_s^t \int_0^1 |F_r^{-1}(u) - \bar{F}_r^{-1}(u)|^{p-2} (\partial_u F_r^{-1}(u) - \partial_u \bar{F}_r^{-1}(u)) \left(\frac{a(F_r^{-1}(u))}{\partial_u F_r^{-1}(u)} - \frac{\alpha_r(u)}{\partial_u \bar{F}_r^{-1}(u)} \right) dudr.
\end{aligned} \tag{A.8}$$

The last term which belongs to $[-\infty, +\infty)$ is finite since so are all the other terms. We deduce integrability of

$$(r, u) \mapsto |F_r^{-1}(u) - \bar{F}_r^{-1}(u)|^{p-2} (\partial_u F_r^{-1}(u) - \partial_u \bar{F}_r^{-1}(u)) \left(\frac{a(F_r^{-1}(u))}{\partial_u F_r^{-1}(u)} - \frac{\alpha_r(u)}{\partial_u \bar{F}_r^{-1}(u)} \right)$$

on $[s, t] \times (0, 1)$. Similar arguments show that the integrability property and (A.8) remain true for $s = t_k$. By summation, they remain true for $0 \leq s \leq t \leq T$. So integrability holds on $[0, T]$

for the distribution derivative

$$\partial_t \mathcal{W}_p^p(\mathcal{L}(X_t),\mathcal{L}(\bar X_t)) = p\int_0^1 |F_t^{-1}(u)-\bar F_t^{-1}(u)|^{p-2}(F_t^{-1}(u)-\bar F_t^{-1}(u))(b(F_t^{-1}(u))-\beta_t(u))du$$
$$+\frac{p(p-1)}{2}\int_0^1 |F_t^{-1}(u)-\bar F_t^{-1}(u)|^{p-2}(\partial_u F_t^{-1}(u)-\partial_u \bar F_t^{-1}(u))\left(\frac{a(F_t^{-1}(u))}{\partial_u F_t^{-1}(u)}-\frac{\alpha_t(u)}{\partial_u \bar F_t^{-1}(u)}\right)du$$
$$\leq p\int_0^1 |F_t^{-1}(u)-\bar F_t^{-1}(u)|^{p-2}\left[(F_t^{-1}(u)-\bar F_t^{-1}(u))(b(F_t^{-1}(u))-\beta_t(u))+\frac{(p-1)\left(a(F_t^{-1}(u))-\alpha_t(u)\right)^2}{8\underline{a}}\right]du.$$

Equation (2.4) follows by remarking that

$$\left(a(F_r^{-1}(u))-\alpha_r(u)\right)^2 \leq 2\left(\|a'\|_\infty^2|F_t^{-1}(u)-\bar F_t^{-1}(u)|^2+\left(a(\bar F_t^{-1}(u))-\alpha_t(u)\right)^2\right)$$

and using a similar idea for $|b(F_t^{-1}(u))-\beta_t(u)|$.

To prove (A.7) for $0<s\leq t\leq T$, we use the Aronson estimates recalled in the proof of Proposition 2.4 for X_t and deduced from Theorem 2.1 [22] for the Euler scheme.

$$\frac{c}{\sqrt{r}}\exp\left(-\frac{(x-x_0)^2}{cr}\right)\leq p_r(x)\wedge \bar p_r(x)\leq p_r(x)\vee \bar p_r(x)\leq \frac{C}{\sqrt{r}}\exp\left(-\frac{(x-x_0)^2}{Cr}\right), \tag{A.9}$$

Setting $K_1=\frac{c}{\sqrt{t}}$, $c_1=cs/2$, $K_2=\frac{C}{\sqrt{s}}$ and $c_2=Ct/2$, one has

$$\forall r\in[s,t],\ \forall x\in\mathbb{R},\ K_1\exp\left(-\frac{(x-x_0)^2}{2c_1}\right)\leq \rho_r(x)\leq K_2\exp\left(-\frac{(x-x_0)^2}{2c_2}\right). \tag{A.10}$$

where ρ_r denotes either p_r or $\bar p_r$. The four limits in (A.7) can be obtained similarly, and we focus on the one of $\sup_{r\in[s,t]}\frac{a(F_r^{-1}(u))}{\partial_u F_r^{-1}(u)}|F_r^{-1}(u)-\bar F_r^{-1}(u)|^{p-1}$. Up to modifying $K_1>0$ and decreasing $c_1>0$, we get from (A.10) that

$$\forall r\in[s,t],\ \forall x\leq x_0-1,\ K_1(x_0-x)\exp\left(-\frac{(x-x_0)^2}{2c_1}\right)\leq \rho_r(x)\leq K_2(x_0-x)\exp\left(-\frac{(x-x_0)^2}{2c_2}\right),$$

which leads to

$$\forall x\leq x_0-1,\ K_1c_1\exp\left(-\frac{(x-x_0)^2}{2c_1}\right)\leq G_r(x)\leq K_2c_2\exp\left(-\frac{(x-x_0)^2}{2c_2}\right),$$

where G_r denotes either F_r or $\bar F_r$. Thus, the inverse function satisfies

$$x_0-\sqrt{-2c_2\log(\frac{u}{K_2c_2})}\leq \bar F_r^{-1}(u)\leq x_0-\sqrt{-2c_1\log(\frac{u}{K_1c_1})} \tag{A.11}$$

for u small enough. The two last inequalities imply that when $x\to-\infty$,

$$\forall r\in[s,t],\ \bar F_r^{-1}(F_r(x))\geq x_0-\sqrt{-2c_2\left[\log(\frac{K_1c_1}{K_2c_2})-\frac{(x-x_0)^2}{2c_1}\right]}$$

and $\sup_{r\in[s,t]}|x-\bar F_r^{-1}(F_r(x))|\underset{x\to-\infty}{=}O(x)$. With the boundedness of a and (A.10), we easily deduce that

$$\lim_{x\to-\infty}\sup_{r\in[s,t]}a(x)p_r(x)|x-\bar F_r^{-1}(F_r(x))|^{p-1}=0.$$

Since, by (A.11), $\bar{F}_r^{-1}(u)$ converges to $-\infty$ uniformly in $r \in [s,t]$ as u tends to 0, we conclude that

$$\lim_{u\to 0^+} \sup_{r\in[s,t]} \frac{a(F_r^{-1}(u))}{\partial_u F_r^{-1}(u)} |F_r^{-1}(u) - \bar{F}_r^{-1}(u)|^{p-1} = 0.$$

∎

Proof of Lemma 2.7. By Jensen's inequality,

$$\mathbb{E}\left[|\mathbb{E}(W_t - W_{\tau_t}|\bar{X}_t)|^p\right] \le \mathbb{E}\left[|W_t - W_{\tau_t}|^p\right] \le \frac{C}{N^{p/2}}.$$

Let us now check that the left-hand-side is also smaller than $\frac{C}{t^{p/2}N^p}$. To do this, we will study

$$\mathbb{E}\left[(W_t - W_{\tau_t})g(\bar{X}_t)\right]$$

where g is any smooth real valued function.

In order to continue, we need to do various estimations on the Euler scheme and its stochastic derivatives. Let $\eta_t = \min\{t_i; t \le t_i\}$ denote the discretization time just after t. We have $D_u\bar{X}_t = 0$ for $u > t$, and

$$D_u\bar{X}_t = 1_{\{t\le\eta_u\}}\sigma(\bar{X}_{\tau_t}) + 1_{\{t>\eta_u\}}\left(1 + \sigma'(\bar{X}_{\tau_t})(W_t - W_{\tau_t}) + b'(\bar{X}_{\tau_t})(t-\tau_t)\right) D_u\bar{X}_{\tau_t} \text{ for } u \le t.$$

Then by induction, one clearly obtains that for $u \le t$,

$$\begin{aligned} &D_u\bar{X}_t = \sigma(\bar{X}_{\tau_u})\bar{\mathcal{E}}_{u,t}, \\ &\bar{\mathcal{E}}_{u,t} = \begin{cases} 1 & \text{if } \tau_t \le \eta_u \\ \left(1 + b'(\bar{X}_{\tau_t})(t-\tau_t) + \sigma'(\bar{X}_{\tau_t})(W_t - W_{\tau_t})\right) & \text{if } \eta_u = \tau_t \\ \prod_{i=\frac{N\eta_u}{T}}^{\frac{N\tau_t}{T}-1}\left(1 + b'(\bar{X}_{t_i})(t_{i+1}-t_i) + \sigma'(\bar{X}_{t_i})(W_{t_{i+1}} - W_{t_i})\right) & \text{if } \eta_u < \tau_t \\ \quad \times\left(1 + b'(\bar{X}_{\tau_t})(t-\tau_t) + \sigma'(\bar{X}_{\tau_t})(W_t - W_{\tau_t})\right). & \end{cases} \end{aligned}$$

Note that $\bar{\mathcal{E}}$ satisfies the following properties: 1. $\bar{\mathcal{E}}_{u,t} = \bar{\mathcal{E}}_{\eta(u),t}$ and 2. $\bar{\mathcal{E}}_{t_i,t_j}\bar{\mathcal{E}}_{t_j,t} = \bar{\mathcal{E}}_{t_i,t}$ for $t_i \le t_j \le t$. We also introduce the process $\mathcal{E}$ defined by

$$\mathcal{E}_{u,t} = \exp\left(\int_u^t b'(X_s) - \frac{1}{2}\sigma'(X_s)^2 ds + \int_u^t \sigma'(X_s)dW_s\right).$$

The next lemma, the proof of which is postponed at the end of the present proof states some useful properties of the processes $\mathcal{E}$ and $\bar{\mathcal{E}}$.

Lemma A.2 *Let us assume that $b, \sigma \in C_b^2$. Then, we have:*

$$\sup_{0\le s\le t\le T} \mathbb{E}\left[\mathcal{E}_{s,t}^{-p}\right] + \mathbb{E}\left[\mathcal{E}_{s,t}^p\right] \le C, \quad \sup_{0\le s\le t\le T} \mathbb{E}\left[\bar{\mathcal{E}}_{s,t}^p\right] \le C, \tag{A.12}$$

$$\sup_{0\le s,u\le t\le T} \mathbb{E}\left[|D_u\bar{\mathcal{E}}_{s,t}|^p + |D_u\mathcal{E}_{s,t}|^p\right] \le C, \tag{A.13}$$

$$\sup_{0\le t\le T} \mathbb{E}\left[\left|\mathcal{E}_{0,t} - \bar{\mathcal{E}}_{0,t}\right|^p\right] \le \frac{C}{N^{\frac{p}{2}}}, \tag{A.14}$$

where C is a positive constant depending only on p and T.

We next define the localization given by

$$\psi = \varphi\left(\mathcal{E}_{0,t}^{-1}\left(\mathcal{E}_{0,t} - \bar{\mathcal{E}}_{0,t}\right)\right).$$

Here $\varphi : \mathbb{R} \to [0,1]$ is a C^∞ symmetric function so that

$$\varphi(x) = \begin{cases} 0, & \text{if} \quad |x| > \frac{1}{2} \\ 1, & \text{if} \quad |x| < \frac{1}{4}. \end{cases}$$

One has

$$\begin{aligned}
\mathbb{E}\left[(W_t - W_{\tau_t})g(\bar{X}_t)\right] &= \mathbb{E}\left[(W_t - W_{\tau_t})g(\bar{X}_t)\psi\right] + \mathbb{E}\left[(W_t - W_{\tau_t})g(\bar{X}_t)(1-\psi)\right] \\
&= \int_{\tau_t}^t \mathbb{E}\left[\psi g'(\bar{X}_t)D_u\bar{X}_t\right]du + \mathbb{E}\left[g(\bar{X}_t)\int_{\tau_t}^t D_u\psi du\right] + \mathbb{E}\left[(W_t - W_{\tau_t})g(\bar{X}_t)(1-\psi)\right]
\end{aligned}$$

where the second equality follows from the duality formula (see e.g. Definition 1.3.1 in [23]). Since for $\tau_t \le u \le t$

$$\begin{aligned}
\mathbb{E}\left[\psi g'(\bar{X}_t)D_u\bar{X}_t\right] = \mathbb{E}\left[\psi g'(\bar{X}_t)\sigma(\bar{X}_{\tau_t})\right] &= t^{-1}\mathbb{E}\left[\int_0^t D_s g(\bar{X}_t)\frac{\psi\sigma(\bar{X}_{\tau_t})}{D_s\bar{X}_t}ds\right] \\
&= t^{-1}\mathbb{E}\left[g(\bar{X}_t)\int_0^t \psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1}\delta W_s\right],
\end{aligned}$$

one deduces

$$\begin{aligned}
\mathbb{E}\left[W_t - W_{\tau_t}\middle|\,\bar{X}_t\right] = t^{-1}\int_{\tau_t}^t \mathbb{E}\left[\int_0^t \psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1}\delta W_s\middle|\,\bar{X}_t\right]du \\
+ \mathbb{E}\left[\int_{\tau_t}^t D_u\psi du\middle|\bar{X}_t\right] + \mathbb{E}\left[\left(W_t - W_{\tau_t}\right)(1-\psi)\middle|\,\bar{X}_t\right].
\end{aligned} \tag{A.15}$$

Here δW denotes the Skorohod integral. In order to obtain the conclusion of the Lemma, we need to bound the L^p-norm of each term on the right-hand-side of (A.15). In particular, we will use the following estimate (which also proves the existence of the Skorohod integral on the left side below) which can be found in Proposition 1.5.4 in [23]:

$$\left\|\int_0^t \psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1}\delta W_s\right\|_p \le C(p)\left\|\psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_\cdot}\right)\bar{\mathcal{E}}_{\cdot,t}^{-1}\right\|_{1,p}, \tag{A.16}$$

where $\|F_\cdot\|_{1,p}^p = \mathbb{E}\left[\left(\int_0^t F_s^2 ds\right)^{p/2} + \left(\int_0^t\int_0^t (D_uF_s)^2 dsdu\right)^{p/2}\right]$. By Jensen's inequality for $p \ge 2$, we have

$$\|F_\cdot\|_{1,p}^p \le t^{p/2-1}\int_0^t \mathbb{E}[|F_s|^p]ds + t^{p-2}\int_0^t\int_0^t \mathbb{E}[|D_uF_s|^p]dsdu, \tag{A.17}$$

and we will use this inequality to upper bound (A.16). When $1 \le p \le 2$, we will use alternatively the following upper bound $\|F_\cdot\|_{1,p}^p \le \left(\int_0^t \mathbb{E}[F_s^2]ds\right)^{p/2} + \left(\int_0^t\int_0^t \mathbb{E}[(D_uF_s)^2]dsdu\right)^{p/2}$ that comes from Hölder's inequality.

For $\psi > 0$, $\bar{\mathcal{E}}_{0,t} \ge \frac{1}{2}\mathcal{E}_{0,t} > 0$. From Hypothesis 3.1, there are constants $0 < \underline{\sigma} \le \overline{\sigma} < \infty$ such that $0 < \underline{\sigma} \le \sigma \le \overline{\sigma}$, and one has

$$\begin{aligned}
\int_0^t \mathbb{E}\left[\left(\psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1}\right)^p\right]ds &\le \left(\frac{\overline{\sigma}}{\underline{\sigma}}\right)^p\int_0^t \mathbb{E}\left[\psi^p\bar{\mathcal{E}}_{0,t}^{-p}\bar{\mathcal{E}}_{0,\eta(s)}^p\right]ds \\
&\le \left(\frac{2\overline{\sigma}}{\underline{\sigma}}\right)^p\sqrt{\mathbb{E}[\mathcal{E}_{0,t}^{-2p}]}\int_0^t\sqrt{\mathbb{E}[|\bar{\mathcal{E}}_{0,\eta(s)}|^{2p}]}ds \le Ct,
\end{aligned}$$

by using the estimates (A.12).

Next, we focus on getting an upper bound for

$$\int_0^t \int_0^t \mathbb{E}\left[\left|D_u\left(\psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1}\right)\right|^p\right] dsdu. \tag{A.18}$$

To do so, we compute the derivative using basic derivation rules, which gives

$$\begin{aligned} D_u\left(\psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1}\right) &= D_u\psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1} + \psi\sigma'(\bar{X}_{\tau_t})D_u\bar{X}_{\tau_t}\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1} \\ &\quad - \psi\sigma(\bar{X}_{\tau_t})\sigma^{-2}\sigma'\left(\bar{X}_{\tau_s}\right)\sigma(\bar{X}_{\tau_u})\bar{\mathcal{E}}_{u,\tau_s}\bar{\mathcal{E}}_{s,t}^{-1}\mathbf{1}_{u\le\tau_s} \\ &\quad - \psi\sigma(\bar{X}_{\tau_u})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-2}D_u\bar{\mathcal{E}}_{s,t}. \end{aligned} \tag{A.19}$$

One has then to get an upper bound for the L^p-norm of each term. As many of the arguments are repetitive, we show the reader only some of the arguments that are involved. Let us start with the first term. We have

$$D_u\psi = \varphi'\left(\mathcal{E}_{0,t}^{-1}\left(\mathcal{E}_{0,t} - \bar{\mathcal{E}}_{0,t}\right)\right) D_u\left[\mathcal{E}_{0,t}^{-1}\left(\mathcal{E}_{0,t} - \bar{\mathcal{E}}_{0,t}\right)\right],$$

and $D_u\left[\mathcal{E}_{0,t}^{-1}\left(\mathcal{E}_{0,t} - \bar{\mathcal{E}}_{0,t}\right)\right] = \mathcal{E}_{0,t}^{-2}D_u\mathcal{E}_{0,t}\bar{\mathcal{E}}_{0,t} - \mathcal{E}_{0,t}^{-1}D_u\bar{\mathcal{E}}_{0,t}$. From the estimates in (A.12) and (A.13), we obtain

$$\sup_{u\in[0,t]} \|D_u\psi\|_p \le \|\varphi'\|_\infty C(p). \tag{A.20}$$

Since $\bar{\mathcal{E}}_{s,t}^{-1} = \bar{\mathcal{E}}_{0,\eta(s)}\bar{\mathcal{E}}_{0,t}^{-1}$ and $\bar{\mathcal{E}}_{0,t} \ge \frac{1}{2}\mathcal{E}_{0,t} > 0$ if $\varphi'\left(\mathcal{E}_{0,t}^{-1}\left(\mathcal{E}_{0,t} - \bar{\mathcal{E}}_{0,t}\right)\right) \ne 0$, we have

$$\mathbb{E}\left[\left|D_u\psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1}\right|^p\right] \le \left(\frac{2\overline{\sigma}}{\underline{\sigma}}\right)^p \|D_u\psi\|_{2p}^p \,\mathbb{E}\left[\left|\mathcal{E}_{0,t}^{-1}\bar{\mathcal{E}}_{0,\eta(s)}\right|^{2p}\right]^{1/2}.$$

Similar bounds hold for the three other terms. Note that the highest requirements on the derivatives of b and σ will come from the terms involving $D_u\bar{\mathcal{E}}$ in (A.19). Gathering all the upper bounds, we get that $\left\|\psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_\cdot}\right)\bar{\mathcal{E}}_{\cdot,t}^{-1}\right\|_{1,p}^p \le C(t^{p/2} + t^p) \le Ct^{p/2}$ since $0 \le t \le T$. From (A.16), we finally obtain

$$\left\|\int_0^t \psi\sigma(\bar{X}_{\tau_u})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1}\delta W_s\right\|_p \le C(p)t^{1/2}.$$

We are now in position to conclude. Using Jensen's inequality, the results (A.15), (A.12), (A.14), (A.20) and the definition of φ together with Chebyshev's inequality, we have for any $k > 0$ that

$$\begin{aligned} &\mathbb{E}\left[\left|\mathbb{E}\left[W_t - W_{\tau_t}|\,\bar{X}_t\right]\right|^p\right] \\ &\le C\left(t^{-p}(t-\tau_t)^p\left\|\int_0^t \psi\sigma(\bar{X}_{\tau_t})\sigma^{-1}\left(\bar{X}_{\tau_s}\right)\bar{\mathcal{E}}_{s,t}^{-1}\delta W_s\right\|_p^p + (t-\tau_t)^{p-1}\int_{\tau_t}^t \|D_u\psi\|_p^p\,du \right.\\ &\left. +\sqrt{\mathbb{E}(|W_t - W_{\tau_t}|^{2p})}4^{k/2}\left(\mathbb{E}(|\mathcal{E}_{0,t} - \bar{\mathcal{E}}_{0,t}|^{2k})\mathbb{E}(\mathcal{E}_{0,t}^{-2k})\right)^{1/4}\right) \\ &\le C\left(t^{-p/2}(t-\tau_t)^p + (t-\tau_t)^p + \left(\frac{1}{N}\right)^{(2p+k)/4}\right) \le C\left(\frac{1}{t^{p/2}N^p} + \frac{1}{N^{\frac{p}{2}+\frac{k}{4}}}\right). \end{aligned}$$

∎

Proof of Lemma A.2. The upper bounds (A.12) on $\mathcal{E}$ and $\bar{\mathcal{E}}$ are obvious since b' and σ' are bounded. Now, let us remark that $\bar{\mathcal{E}}$ and $\mathcal{E}$ satisfy

$$\mathcal{E}_{u,t} = 1 + \int_u^t \sigma'(X_s)\mathcal{E}_{u,s} dW_s + \int_u^t b'(X_s)\mathcal{E}_{u,s} ds,$$
$$\bar{\mathcal{E}}_{\eta_u,t} = 1 + \int_{\eta_u}^t \sigma'(\bar{X}_{\tau_s})\bar{\mathcal{E}}_{\eta_u,\tau_s} dW_s + \int_{\eta_u}^t b'(\bar{X}_{\tau_s})\bar{\mathcal{E}}_{\eta_u,\tau_s} ds.$$

Thus, (A.14) can be easily obtained by noticing that $(\bar{X}_t, \bar{\mathcal{E}}_{0,t})$ is the Euler scheme for the SDE $(X_t, \mathcal{E}_{0,t})$ which has Lipschitz coefficients, and by using the strong convergence order of $1/2$ (see e.g. [17]).

The estimate (A.13) on $D_u\mathcal{E}$ is given, for example, by Theorem 2.2.1 in [23]. On the other hand, we have for $\eta(s) \le u \le t$

$$D_u\bar{\mathcal{E}}_{\eta_s,t} = \sigma'(\bar{X}_{\tau_u})\bar{\mathcal{E}}_{\eta_s,\tau_u} + \int_{\eta_u}^t \left[\sigma''(\bar{X}_{\tau_r})\sigma(\bar{X}_{\tau_u})\bar{\mathcal{E}}_{\eta_u,\tau_r}\bar{\mathcal{E}}_{\eta_s,\tau_r} + \sigma'(\bar{X}_{\tau_r})D_u\bar{\mathcal{E}}_{\eta_s,\tau_r}\right] dW_r$$
$$+ \int_{\eta_u}^t \left[b''(\bar{X}_{\tau_r})\sigma(\bar{X}_{\tau_u})\bar{\mathcal{E}}_{\eta_u,\tau_r}\bar{\mathcal{E}}_{\eta_s,\tau_r} + b'(\bar{X}_{\tau_r})D_u\bar{\mathcal{E}}_{\eta_s,\tau_r}\right] dr.$$

In order to obtain a $L^p(\Omega)$ estimate, we then use (A.12), $b, \sigma \in C_b^2$ and Gronwall's lemma. ∎

B Proofs of Section 3

Proof of Proposition 3.4. We use the dual representation of the Wasserstein distance (0.6) deduced from Kantorovitch duality theorem (see for instance Theorem 5.10 p58 [29]) :

$$\mathcal{W}_p^p(\mu,\nu) = \sup_{\phi \in L^1(\nu)} \left(\int_E \tilde{\phi}(x)\mu(dx) - \int_E \phi(x)\nu(dx) \right)$$

where $\tilde{\phi}(x) = \inf_{y\in E} (\phi(y) + |y-x|^p)$.

We also denote by $(X_t^{s,x})_{t\in[s,T]}$ the solution to (0.1) starting from $x \in \mathbb{R}$ at time $s \in [0,T]$ and by $(\bar{X}_t^{t_j,x})_{t\in[t_j,T]}$ the Euler scheme starting from x at time t_j with $j \in \{0,\dots,N\}$. It is enough to check that

$$w_k \overset{\text{def}}{=} \mathcal{W}_p(\mathcal{L}(\bar{X}_{s_1},\dots,\bar{X}_{s_k}, X_{s_{k+1}}^{s_k,\bar{X}_{s_k}},\dots,X_{s_n}^{s_k,\bar{X}_{s_k}}), \mathcal{L}(\bar{X}_{s_1},\dots,\bar{X}_{s_{k-1}}, X_{s_k}^{s_{k-1},\bar{X}_{s_{k-1}}},\dots,X_{s_n}^{s_{k-1},\bar{X}_{s_{k-1}}}))$$

is smaller than $C\sup_{0\le t\le T, x\in\mathbb{R}} \mathcal{W}_p(\mathcal{L}(\bar{X}_t^x), \mathcal{L}(X_t^x))$ since $\mathcal{W}_p(\mathcal{L}(\bar{X}_{s_1},\dots,\bar{X}_{s_n}), \mathcal{L}(X_{s_1},\dots,X_{s_n})) \le \sum_{k=1}^n w_k$. For $f : \mathbb{R}^n \to \mathbb{R}$ a bounded measurable function and

$$\tilde{f}(x_1,\dots,x_n) = \inf_{(y_1,\dots,y_n)\in\mathbb{R}^n} \{f(y_1,\dots,y_n) + \max_{1\le j\le n} |y_j - x_j|^p\},$$

we set $f_k(x_1,\dots,x_k) = \mathbb{E}(f(x_1,\dots,x_k, X_{s_{k+1}}^{s_k,x_k},\dots,X_{s_n}^{s_k,x_k}))$. First choosing

$$(y_1,\dots,y_{k-1},y_{k+1},\dots,y_n) = (\bar{X}_{s_1},\dots,\bar{X}_{s_{k-1}}, X_{s_{k+1}}^{s_k,y_k},\dots,X_{s_n}^{s_k,y_k}),$$

then conditioning to $\sigma(W_s, s \le s_k)$ and using (1.3), next conditioning to $\sigma(W_s, s \le s_{k-1})$ and using the dual formulation of the Wasserstein distance, one gets

$$\begin{aligned}
&\mathbb{E}\left(\tilde{f}(\bar{X}_{s_1}, \dots, \bar{X}_{s_k}, X^{s_k,\bar{X}_{s_k}}_{s_{k+1}}, \dots, X^{s_k,\bar{X}_{s_k}}_{s_n}) - f(\bar{X}_{s_1}, \dots, \bar{X}_{s_{k-1}}, X^{s_{k-1},\bar{X}_{s_{k-1}}}_{s_k}, \dots, X^{s_{k-1},\bar{X}_{s_{k-1}}}_{s_n})\right) \\
&\le \mathbb{E}\Big(\inf_{y_k\in\mathbb{R}} \{ f(\bar{X}_{s_1}, \dots, \bar{X}_{s_{k-1}}, y_k, X^{s_k,y_k}_{s_{k+1}}, \dots, X^{s_k,y_k}_{s_n}) + \max_{k\le j\le n} |X^{s_k,y_k}_{s_j} - X^{s_k,\bar{X}_{s_k}}_{s_j}|^p \} \\
&\qquad - f(\bar{X}_{s_1}, \dots, \bar{X}_{s_{k-1}}, X^{s_{k-1},\bar{X}_{s_{k-1}}}_{s_k}, \dots, X^{s_{k-1},\bar{X}_{s_{k-1}}}_{s_n})\Big) \\
&\le \mathbb{E}\left(\inf_{y_k\in\mathbb{R}} \{ f_k(\bar{X}_{s_1}, \dots, \bar{X}_{s_{k-1}}, y_k) + C|y_k - \bar{X}_{s_k}|^p\} - f_k(\bar{X}_{s_1}, \dots, \bar{X}_{s_{k-1}}, X^{s_{k-1},\bar{X}_{s_{k-1}}}_{s_k})\right) \\
&\le C\mathbb{E}\left(W_p^p(\mathcal{L}(X^{s_{k-1},x}_{s_k}), \mathcal{L}(\bar{X}^{s_{k-1},x}_{s_k}))|_{x=\bar{X}_{s_{k-1}}}\right) \le C \sup_{x\in\mathbb{R}} \mathcal{W}_p^p(\mathcal{L}(\bar{X}^x_{s_k-s_{k-1}}), \mathcal{L}(X^x_{s_k-s_{k-1}})) \\
&\le C \sup_{0\le t\le \mathcal{T}, x\in\mathbb{R}} \mathcal{W}_p^p(\mathcal{L}(\bar{X}^x_t), \mathcal{L}(X^x_t)).
\end{aligned}$$

∎

C Some properties of diffusion bridges

Let us suppose that the SDE $dX_t = b(X_t)dt + \sigma(X_t)dW_t$, $X_0 = x$ has a transition density $p_t(x,y)$ which is positive and of class $\mathcal{C}^{1,2}$ with respect to $(t,x) \in \mathbb{R}_+^* \times \mathbb{R}$. We check later in this section that this holds under Hypothesis 3.1. Then, the law of the diffusion bridge with time horizon $\mathcal{T}$ is given by (see for instance Fitzsimmons, Pitman and Yor [5])

$$\mathbb{E}[F(X_u, 0 \le u \le t)|X_{\mathcal{T}} = y] = \mathbb{E}\left[F(X_u, 0\le u \le t)\frac{p_{\mathcal{T}-t}(X_t,y)}{p_{\mathcal{T}}(x,y)}\right], \; 0 \le t < \mathcal{T},$$

where $F : C([0,t],\mathbb{R}) \to \mathbb{R}$ is a bounded measurable function. Indeed for $g : \mathbb{R} \to \mathbb{R}$ measurable and bounded, using that $X_{\mathcal{T}}$ has the density $p_{\mathcal{T}}(x,y)$ then the Markov property at time t, one checks that

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[F(X_u, 0 \le u \le t)\frac{p_{\mathcal{T}-t}(X_t,y)}{p_{\mathcal{T}}(x,y)}\right]\Big|_{y=X_{\mathcal{T}}} g(X_{\mathcal{T}})\right] = \mathbb{E}\left[F(X_u, 0\le u\le t)\int_{\mathbb{R}} g(y)p_{\mathcal{T}-t}(X_t,y)dy\right] \\
&= \mathbb{E}\left[F(X_u, 0\le u \le t)\mathbb{E}[g(X_{\mathcal{T}})|X_t]\right] = \mathbb{E}\left[F(X_u, 0 \le u \le t) g(X_{\mathcal{T}})\right].
\end{aligned}$$

We thus focus on the change of probability measure

$$\frac{d\mathbb{P}^y}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = \frac{p_{\mathcal{T}-t}(X_t,y)}{p_{\mathcal{T}}(x,y)} =: M_t,$$

so that $\mathbb{E}[F(X_u, 0 \le u \le t)|X_{\mathcal{T}} = y] = \mathbb{E}^y[F(X_u, 0 \le u \le t)]$ where $\mathbb{E}^y$ denotes the expectation with respect to $\mathbb{P}^y$. We define $\ell_t(x,y) = \log p_t(x,y)$. The process $(M_t)_{t\in[0,\mathcal{T})}$ is a martingale, and by Itô's formula, we get $dM_t = M_t \partial_x \ell_{\mathcal{T}-t}(X_t,y)\sigma(X_t)dW_t$, which gives

$$M_t = \exp\left(\int_0^t \partial_x \ell_{\mathcal{T}-s}(X_s,y)\sigma(X_s)dW_s - \frac{1}{2}\int_0^t \partial_x\ell_{\mathcal{T}-s}(X_s,y)^2\sigma(X_s)^2 ds\right).$$

Girsanov Theorem then gives that for all $y \in \mathbb{R}$, $(W_t^y = W_t - \int_0^t \partial_x \ell_{\mathcal{T}-s}(X_s,y)\sigma(X_s)ds)_{t\in[0,\mathcal{T})}$ is a Brownian motion under $\mathbb{P}^y$, so that $(W_t^{X_{\mathcal{T}}})_{t\in[0,\mathcal{T})}$ is a Brownian motion independent of $X_{\mathcal{T}}$. Moreover, we have

$$dX_t = [b(X_t) + \partial_x \ell_{\mathcal{T}-t}(X_t,y)\sigma(X_t)^2]dt + \sigma(X_t)dW_t^y, \tag{C.1}$$

which gives precisely the diffusion bridge dynamics.

Conversely, we would like now to reconstruct the diffusion from the initial and the final value by using diffusion bridges. We have the following result.

Proposition C.1 *We consider an SDE* $dX_t = b(X_t)dt + \sigma(X_t)dW_t$, $X_0 = x$ *with a transition density* $p_t(x, y)$ *positive and of class* $\mathcal{C}^{1,2}$ *on* $(t, x) \in \mathbb{R}_+^* \times \mathbb{R}$. *Let* $(B_t, t \geq 0)$ *be a standard Brownian motion and* $Z_\mathcal{T}$ *be a random variable with density* $p_\mathcal{T}(x, y)$ *drawn independently from* B. *We assume that pathwise uniqueness holds for the SDE*

$$dZ_t^{x,y} = [b(Z_t^{x,y}) + \partial_x \ell_{\mathcal{T}-t}(Z_t^{x,y}, y)\sigma(Z_t^{x,y})^2]dt + \sigma(Z_t^{x,y})dB_t,\ Z_0^{x,y} = x,\ t \in [0, \mathcal{T}), \tag{C.2}$$

for any $x, y \in \mathbb{R}$, *and set* $Z_t = Z_t^{x,Z_\mathcal{T}}$ *for* $t \in [0, \mathcal{T})$. *Then,* $(Z_t)_{t\in[0,\mathcal{T}]}$ *and* $(X_t)_{t\in[0,\mathcal{T}]}$ *have the same law.*

A consequence of this result is that $(Z_t, t \in [0, \mathcal{T}])$ has continuous paths, which gives that $\lim_{t\to\mathcal{T}-} Z_t^{x,y} = y$ a.s., dy-a.e.
Proof . Let $t \in [0, \mathcal{T})$ and $F : C([0, t], \mathbb{R}) \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be bounded and measurable functions. Since pathwise uniqueness for the SDE (C.2) implies weak uniqueness, we get

$$\mathbb{E}\left[F(Z_u^{x,y}, 0 \leq u \leq t)\right] = \mathbb{E}^y\left[F(X_u, 0 \leq u \leq t)\right] = \mathbb{E}\left[F(X_u, 0 \leq u \leq t)\frac{p_{\mathcal{T}-t}(X_t, y)}{p_\mathcal{T}(x, y)}\right].$$

Thus, we have

$$\mathbb{E}\left[F(Z_u, 0 \leq u \leq t)g(Z_\mathcal{T})\right] = \mathbb{E}\left[F(X_u, 0 \leq u \leq t)\int_\mathbb{R} p_{\mathcal{T}-t}(X_t, y)g(y)dy\right] = \mathbb{E}\left[F(X_u, 0 \leq u \leq t)g(X_\mathcal{T})\right].$$

Hence the finite-dimensional marginals of the two processes are equal. Since $(X_t)_{t\in[0,\mathcal{T}]}$ has continuous paths and $(Z_t)_{t\in[0,\mathcal{T}]}$ has càdlàg paths (continuous on $[0, \mathcal{T})$ with a possible jump at $\mathcal{T}$), this concludes the proof. ■

From now on, we assume that Hypothesis 3.1 holds. We introduce the Lamperti transformation of the stochastic process $(X_t, t \geq 0)$. We define $\varphi(x) = \int_0^x \frac{dy}{\sigma(y)}$ and $\alpha(y) = \left(\frac{b}{\sigma} - \frac{\sigma'}{2}\right) \circ \varphi^{-1}(y)$, $\hat{X}_t \overset{\text{def}}{=} \varphi(X_t)$ so that we have

$$d\hat{X}_t = \alpha(\hat{X}_t)dt + dW_t, t \in [0, T]. \tag{C.3}$$

By Hypothesis 3.1, φ is a C^5 bijection, $\alpha \in C_b^3$ and both φ and φ^{-1} are Lipschitz continuous. We denote by $\hat{p}_t(\hat{x}, \hat{y})$ the transition density of $\hat{X}$ and $\hat{\ell}_t(\hat{x}, \hat{y}) = \log(\hat{p}_t(\hat{x}, \hat{y}))$.

Lemma C.2 *The density* $\hat{p}_t(\hat{x}, \hat{y})$ *is* $C^{1,2}$ *with respect to* $(t, \hat{x}) \in \mathbb{R}_+^* \times \mathbb{R}$. *Besides, we have*

$$\partial_x \hat{\ell}_t(\hat{x}, \hat{y}) = \frac{\hat{y} - \hat{x}}{t} - \alpha(\hat{x}) + g_t(\hat{x}, \hat{y}),$$

where $g_t(\hat{x}, \hat{y})$ *is a continuous function on* $\mathbb{R}_+ \times \mathbb{R}^2$ *such that* $\partial_{\hat{x}} g_t(\hat{x}, \hat{y})$ *and* $\partial_{\hat{y}} g_t(\hat{x}, \hat{y})$ *exist and*

$$\forall T > 0, \quad \sup_{t\in[0,T],\ \hat{x},\hat{y}\in\mathbb{R}} |\partial_x g_t(\hat{x}, \hat{y})| + |\partial_y g_t(\hat{x}, \hat{y})| < \infty.$$

Proof . It is well-known that we can express the transition density $\hat{p}_t(\hat{x},\hat{y})$ by using Girsanov theorem as an expectation on a Brownian bridge between $\hat{x}$ and $\hat{y}$. Namely, since α and its derivatives are bounded, we can apply a result stated in Gihman and Skorohod [7] (Theorem 1, Chapter 3, § 13) to get that $\hat{p}_t(\hat{x},\hat{y})$ is positive and

$$\hat{\ell}_t(\hat{x},\hat{y}) = -\frac{(\hat{x}-\hat{y})^2}{2t} + \int_{\hat{x}}^{\hat{y}} \alpha(z)dz + \log \mathbb{E}\left(e^{-\frac{1}{2}\int_0^t (\alpha'+\alpha^2)(\hat{x}+W_s+\frac{s}{t}(\hat{y}-\hat{x}-W_t))ds}\right) - \frac{1}{2}\log(2\pi t).$$

Clearly, $\hat{\ell}_t(\hat{x},\hat{y})$ is $C^{1,2}$ in $(t,\hat{x}) \in \mathbb{R}_+^* \times \mathbb{R}$ (we can use carefree the dominated convergence theorem for the third term since $\alpha \in C_b^3$), and we have

$$g_t(\hat{x},\hat{y}) = -\frac{1}{2}\frac{\mathbb{E}\left[e^{-\frac{1}{2}\int_0^t (\alpha'+\alpha^2)(\hat{x}+W_s+\frac{s}{t}(\hat{y}-\hat{x}-W_t))ds}\int_0^t \frac{t-s}{t}(\alpha''+2\alpha\alpha')(\hat{x}+W_s+\frac{s}{t}(\hat{y}-\hat{x}-W_t))ds\right]}{\mathbb{E}\left[e^{-\frac{1}{2}\int_0^t (\alpha'+\alpha^2)(\hat{x}+W_s+\frac{s}{t}(\hat{y}-\hat{x}-W_t))ds}\right]}.$$

This is a continuous function on $\mathbb{R}_+ \times \mathbb{R}^2$, and we easily conclude by using the dominated convergence theorem and $\alpha \in C_b^3$. ∎

By straightforward calculations, we have

$$p_t(x,y) = \frac{1}{\sigma(y)}\hat{p}_t(\varphi(x),\varphi(y)),$$

and $p_t(x,y)$ is thus positive and $C^{1,2}$ with respect to (t,x). The diffusion bridge (C.1) is thus well defined. Since $\partial_x \ell_t(x,y) = \frac{1}{\sigma(x)}\partial_{\hat{x}}\hat{\ell}_t(\varphi(x),\varphi(y))$, we get by Itô formula from (C.1)

$$d\hat{X}_t = [\alpha(\hat{X}_t) + \partial_{\hat{x}}\hat{\ell}_{\mathcal{T}-t}(\hat{X}_t,\varphi(y))]dt + dW_t^y,\ dW_t^y = dW_t - \partial_{\hat{x}}\hat{\ell}_{\mathcal{T}-t}(\hat{X}_t,\varphi(y))dt.$$

Therefore, as one could expect, the Lamperti transform on the diffusion bridge coincides with the diffusion bridge on the Lamperti transform.

Proposition C.3 *Let Hypothesis 3.1 hold. There exists a deterministic constant C such that*

$$\forall \mathcal{T} \in (0,T],\ x,x',y,y' \in \mathbb{R},\ \sup_{t\in[0,\mathcal{T})} |Z_t^{x,y} - Z_t^{x',y'}| \le C(|x-x'| \vee |y-y'|),$$

and in particular, pathwise uniqueness holds for (C.2).

Proof . For $\hat{x},\hat{y} \in \mathbb{R}$, we consider the following SDE

$$d\hat{Z}_t^{\hat{x},\hat{y}} = dB_t + \left[\frac{\hat{y}-\hat{Z}_t^{\hat{x},\hat{y}}}{\mathcal{T}-t} + g_{\mathcal{T}-t}(\hat{Z}_t^{\hat{x},\hat{y}},\hat{y})\right]dt,\ \hat{Z}_0^{\hat{x},\hat{y}} = \hat{x},\ t \in [0,\mathcal{T}) \quad \text{(C.4)}$$

that corresponds to the diffusion bridge on the Lamperti transform $\hat{X}$. We set $\Delta_t = \hat{Z}_t^{\hat{x},\hat{y}} - \hat{Z}_t^{\hat{x}',\hat{y}'}$ for $t \in [0,\mathcal{T})$ and $\hat{x}',\hat{y}' \in \mathbb{R}$. We have

$$d\Delta_t = \left[\frac{\hat{y}-\hat{y}'-\Delta_t}{\mathcal{T}-t} + g_{\mathcal{T}-t}(\hat{Z}_t^{\hat{x},\hat{y}},\hat{y}) - g_{\mathcal{T}-t}(\hat{Z}_t^{\hat{x}',\hat{y}'},\hat{y}')\right]dt,$$

and thus $d(|\Delta_t| \vee |\hat{y}-\hat{y}'|) = \text{sign}(\Delta_t)\mathbf{1}_{|\Delta_t|\ge|\hat{y}-\hat{y}'|}d\Delta_t$. On the one hand, we observe that $\mathbf{1}_{|\Delta_t|\ge|\hat{y}-\hat{y}'|}[\text{sign}(\Delta_t)(\hat{y}-\hat{y}') - |\Delta_t|] \le 0$. On the other hand, g_t is uniformly Lipschitz w.r.t $(\hat{x},\hat{y})$ on $t \in [0,T]$ by Lemma C.2, which leads to:

$$d(|\Delta_t| \vee |\hat{y}-\hat{y}'|) \le C(|\Delta_t| \vee |\hat{y}-\hat{y}'|),$$

for some positive constant C. Gronwall's lemma gives then $|\Delta_t| \leq e^{CT}(|\hat{x}-\hat{x}'| \vee |\hat{y}-\hat{y}'|)$. This gives in particular pathwise uniqueness for (C.4).

Now, let us assume that $(Z_t^{x,y})_{t\in[0,\mathcal{T})}$ solves (C.2). Then, $\varphi(Z_t^{x,y})$ solves (C.4) with $\hat{x} = \varphi(x)$ and $\hat{y} = \varphi(y)$, and we necessarily have $Z_t^{x,y} = \varphi^{-1}(\hat{Z}_t^{\varphi(x),\varphi(y)})$ by pathwise uniqueness. Both φ and φ^{-1} are Lipschitz, and we denote by K a common Lipschitz constant. Then, we get

$$|Z_t^{x,y} - Z_t^{x',y'}| = |\varphi^{-1}(\hat{Z}_t^{\varphi(x),\varphi(y)}) - \varphi^{-1}(\hat{Z}_t^{\varphi(x'),\varphi(y')})| \leq K^2 e^{CT}(|x-x'| \vee |y-y'|),$$

which gives the desired result. ∎

References

[1] Aronson, D.G. Bounds for the fundamental solution of a parabolic equation. Bulletin of the American Mathematical Society, 73, (1967), 890–903.

[2] Bally, V. and Talay, D. The law of the Euler scheme for stochastic differential equations (II) : convergence rate of the density. Monte Carlo Methods and Applications, 2, (1996), 93–128.

[3] Bally, V. and Talay, D. The law of the Euler scheme for stochastic differential equations (I) : convergence rate of the distribution function. Probability Theory and Related Fields, 104, (1995), 43–60.

[4] Cruzeiro, A.B., Malliavin, P. and Thalmaier, A.: Geometrization of Monte-Carlo numerical analysis of an elliptic operator: strong approximation. C. R. Math. Acad. Sci. Paris 338(6), (2004), 481–486.

[5] Fitzsimmons, P., Pitman, J. and Yor, M., Markovian bridges: construction, Palm interpretation, and splicing, Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992), 101–134, Birkhäuser Boston, 1993.

[6] Friedman, A. Stochastic differential equations and applications vol 1, Probability and mathematical statistics 28, Academic Press, 1975.

[7] Gihman, I. I. and Skorohod, A. V. Stochastic differential equations. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 72. Springer-Verlag, New York-Heidelberg, 1972.

[8] Giles, M. B. Multi-level Monte Carlo path simulation. Operations Research, 56(3), (2008), 607–617.

[9] Gobet, E. Weak approximation of killed diffusion using Euler schemes. Stochastic Processes and their Applications, 87, (2000), 167–197.

[10] Gobet, E. Euler Schemes and Half-Space Approximation for the Simulation of Diffusion in a Domain, ESAIM P&S, 5, (2001), 261–297.

[11] Gobet, E. Menozzi, S. Exact approximation rate of killed hypoelliptic diffusions using the discrete Euler scheme, Stochastic Processes and their Applications, 112(2), (2004), 201–223.

[12] Gobet, E. and Labart, C. Sharp estimates for the convergence of the density of the Euler scheme in small time. Electron. Commun. Probab., 13 , (2008), 352–363.

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

[14] Gyöngy, I. Mimicking the one-dimensional marginal distributions of processes having an Itô's differential. Probab. Theory Relat. Fields, 71(4), (1986), 501–516.

[15] Jacod, J. and Protter, P. Asymptotic error distributions for the Euler method for stochastic differential equations. Ann. Probab., 26(1), (1998), 267–307.

[16] Jourdain, B. and Sbai, M. High order discretization schemes for stochastic volatility models, To appear in Journal of Computational Finance.

[17] Kanagawa, S. On the rate of convergence for Maruyama's approximate solutions of stochastic differential equations. Yokohama Math. J., 36(1), (1988), 79–86.

[18] Karatzas, Ioannis; Shreve, Steven E. Brownian motion and stochastic calculus. Second edition. Graduate Texts in Mathematics, 113. Springer-Verlag, New York, 1991.

[19] Kohatsu-Higa, A. Weak approximations: A Malliavin calculus approach. Mathematics of Computation, 70, 135-172, 2001.

[20] Kurtz, T. and Protter, P. Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. Stochastic analysis, Academic Press, Boston, MA, (1991), 331-346.

[21] Kurtz, T. and Protter, P. Weak limit theorems for stochastic integrals and stochastic differential equations. Ann. Probab., 19(3), (1991), 1035-1070.

[22] Lemaire, V. and Menozzi, S. On some non asymptotic bounds for the Euler scheme, Electronic Journal of Probability, 15, (2010), 1645–1681.

[23] Nualart, D. The Malliavin Calculus and Related Topics. Second Edition. Springer-Verlag, 2006.

[24] Rachev, S.T. and Rüschendorf, L. Mass Transportation problems. Springer-Verlag, 1998.

[25] Rogers, L. C. G. Smooth transition densities for one-dimensional diffusions. Bull. London Math. Soc., 17(2), (1985), 157–161.

[26] Stroock, D.W. and Varadhan, S.R.S. Multidimensional Diffusion Processes. Springer-Verlag, 1979.

[27] Seumen Tonou, P. Méthodes numériques probabilistes pour la résolution d'équations du transport et pour l'évaluation d'options exotiques. PhD thesis, Université Aix-Marseille 1, 1997.

[28] Sbai, M. Modélisation de la dépendance et simulation de processus en finance. PhD thesis, Université Paris-Est (2009), http://tel.archives-ouvertes.fr/tel-00451008/en/.

[29] Villani, C. Optimal transport. Old and new. Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009.

[30] Temam, E. Couverture approchée d'options exotiques. Pricing des options asiatiques. PhD thesis, Université Paris 6 (2001).

[31] Talay, D. and Tubaro, L. Expansion of the global error for numerical schemes solving stochastic differential equations. Stochastic Anal. Appl., 8(4), (1990), 483–509.