ESAIM: PS 27 (2023) 841–866
https://doi.org/10.1051/ps/2023017

ESAIM: Probability and Statistics
www.esaim-ps.org

APPROXIMATION OF THE INVARIANT DISTRIBUTION FOR A CLASS OF ERGODIC SDES WITH ONE-SIDED LIPSCHITZ CONTINUOUS DRIFT COEFFICIENT USING AN EXPLICIT TAMED EULER SCHEME

Charles-Edouard Brehier*

Abstract. We study the behavior in a large time regime of an explicit tamed Euler–Maruyama scheme applied to a class of ergodic Itô stochastic differential equations with one-sided Lipschitz continuous drift coefficient and bounded globally Lipschitz diffusion coefficient. Our first main contribution is to prove moments for the numerical scheme, which, on the one hand, are uniform with respect to the time-step size, and which, on the other hand, may not be uniform but have at most polynomial growth with respect to time. Our second main contribution is to apply this result to obtain weak error estimates to quantify the error to approximate averages with respect to the invariant distribution of the continuous-time process, as a function of the time-step size and of the time horizon. The explicit tamed Euler scheme is shown to be computationally effective for the approximation of the invariant distribution: even if the moment bounds and error estimates are not proved to be uniform with respect to time, the obtained polynomial growth results in a marginal increase in the upper bound of the computational cost. To the best of our knowledge, this is the first result in the literature concerning the approximation of the invariant distribution for stochastic differential equations with non-globally Lipschitz coefficients using an explicit tamed Euler–Maruyama scheme.

Mathematics Subject Classification. 60H35, 65C30.

Received October 9, 2020. Accepted September 22, 2023.

1. Introduction

Understanding the long-time behavior of deterministic and stochastic processes and of their discrete-time approximations has been an active research area in the last decades, with applications to all field of science. In this work, we consider Itô stochastic differential equations (SDEs) of the type

$$\mathrm{d}X(t) = f(X(t))\mathrm{d}t + \sum_{k=1}^{K} \sigma_k(X(t))\mathrm{d}\beta^k(t),$$

Keywords and phrases: Stochastic differential equations, tamed Euler scheme, invariant distribution.

Université de Pau et des Pays de l'Adour, E2S UPPA, CNRS, LMAP, Pau, France.

* Corresponding author: charles-edouard.brehier@univ-pau.fr

where $X(t) \in \mathbb{R}^d$, $d \in \mathbb{N}$, and β^k are independent standard real-valued Wiener processes with $k = 1, \ldots, K$, $K \in \mathbb{N}$. The drift coefficient $f : \mathbb{R}^d \to \mathbb{R}^d$ is assumed to be a non-globally Lipschitz continuous mapping, precisely it has polynomial growth and satisfies a one-sided Lipschitz continuous condition, see Assumptions 2.1 and 2.2. The diffusion coefficients $\sigma_k : \mathbb{R}^d \to \mathbb{R}^d$, for $k = 1, \ldots, K$, are assumed to be bounded and globally Lipschitz continuous, see Assumption 2.3. Moreover, the nonlinearities f and $\sigma_1, \ldots, \sigma_K$ are assumed to be of class $\mathcal{C}^\infty$. Under the above regularity conditions, for any initial value $x_0 \in \mathbb{R}^d$ there exists a unique strong solution of this SDE, which is defined for all times $t \geq 0$. In addition, under an appropriate sufficient ergodicity condition (see Asm. 2.2), the SDE admits a unique invariant probability distribution $\mu_\star$, and one has the exponentially fast convergence (see Eq. (2.7))

$$\mathbb{E}[\varphi(X(T))] \underset{T\to\infty}{\to} \int \varphi \mathrm{d}\mu_\star,$$

for any initial condition $X(0)$ and any real-valued Lipschitz continuous function φ. In general, computing $\int \varphi \mathrm{d}\mu_\star$ by deterministic approaches is not feasible, because no explicit expression for $\mu_\star$ is known, or when the dimension d is large. In this work, we approximate $\int \varphi \mathrm{d}\mu_\star$ for arbitrary sufficiently smooth functions $\varphi : \mathbb{R}^d \to \mathbb{R}$ using an Euler type temporal discretization scheme applied to the SDE. The objective of this article is to show that a well-chosen explicit scheme can be used, without loss of computational efficiency, even if the nonlinearity f is not globally Lipschitz continuous.

Let us recall that applying a crude explicit Euler scheme for SDEs with non-globally Lipschitz continuous coefficients, may lead to important issues due to the lack of suitable moment bounds for the approximation, see for instance [9]. In the last two decades, many strategies to design more advanced explicit schemes have been explored: see for instance the monograph [8] and references therein, and the articles [10, 11, 20, 23], among many other contributions. When one is interested in the approximation of the invariant distribution, using the standard explicit Euler-Maruyama scheme for such SDEs is non appropriate, as demonstrated in [18].

In this article, we study the explicit tamed Euler-Maruyama method (see also Eq. (2.9))

$$\tilde{X}_{n+1} = \tilde{X}_n + \frac{\Delta t f(\tilde{X}_n)}{1 + \Delta t \|f(\tilde{X}_n)\|} + \sum_{k=1}^{K} \sigma_k(\tilde{X}_n) \Delta\beta_n^k,$$

with time-step size Δt, $t_n = n\Delta t$ and with Wiener increments $\Delta\beta_n^k = \beta^k(t_{n+1}) - \beta^k(t_n)$. Explicit tamed Euler schemes for SDEs have been studied extensively in the recent years, see for instance the article [23], the monograph [8] and references therein. It has been shown that it satisfies moment bounds of the type

$$\sup_{0 \leq n\Delta t \leq T} \left(\mathbb{E}[\|\tilde{X}_n\|^m]\right)^{\frac{1}{m}} \leq C(T, m, \|\tilde{X}_0\|)$$

for all $m \in \mathbb{N}$, $T \in (0, \infty)$ and any deterministic initial value $\tilde{X}_0$, where $C(T, m, \|\tilde{X}_0\|) \in (0, \infty)$ is independent of the time-step size Δt. The moment bounds above are fundamental to ensure convergence of the numerical scheme and to prove strong and weak error estimates. To the best of our knowledge, the question of the dependence of the upper bound above with respect to time T, when the scheme is applied to ergodic SDEs, in the large time regime $T \to \infty$, has not been studied in the literature yet. More precisely, whether moment bounds can be uniform or have some (polynomial or exponential) dependence with respect to time T has not been considered so far in the literature, to the best of our knowledge. The first contribution of this work (see Thm. 3.1) is to provide a first answer to this open question: we prove that moment bounds where the dependence with respect to T is at most polynomial are satisfied. More precisely, we prove that, under appropriate assumptions, for arbitrary $T \in (0,\infty)$ one has moment bounds

$$\sup_{0\le n\Delta t\le T} \big(\mathbb{E}[\|\tilde{X}_n\|^m]\big)^{\frac{1}{m}} \le C(\|x_0\|)(1+T^M),$$

where $\tilde{X}_0 = x_0$, for some $M \ge 1$, where the positive real number $C(\|x_0\|) \in (0,\infty)$ the right-hand side is independent of the time-step size Δt and of $T \in (0,\infty)$. In the moment bounds above, the growth with respect to T is thus at most polynomial.

Let us mention that the moment bounds above, which are one of the main results of this work, may be pessimistic: it might be possible to retrieve moment bounds which are uniform with respect to T. Note that this would correspond to choosing $M = 0$ instead of $M \ge 1$. The arguments employed in the proof of Theorem 3.1 below only give values $M \ge 1$.

The second contribution of this work is to apply the moment bounds on the tamed explicit Euler–Maruyama scheme for the analysis of the error between $\mathbb{E}[\varphi(\tilde{X}_N)]$ and $\int \varphi \mathrm{d}\mu_\star$, see Theorem 3.3. Under appropriate regularity assumptions (stated more precisely below), one obtains weak error estimates of the type

$$\big|\mathbb{E}[\varphi(\tilde{X}_N)] - \int \varphi \mathrm{d}\mu_\star\big| \le C(\|x_0\|,\varphi)\Big(\exp(-\lambda N\Delta t) + \big(1+(N\Delta t)^M\big)\Delta t\Big),$$

where $\lambda > 0$ and $M \ge 1$. Contrary to existing works in the literature concerning the numerical approximation of the invariant distribution for SDEs (this is a classical problem, see for instance [18, 19, 26, 28]), the weak error $|\mathbb{E}[\varphi(\tilde{X}_N)] - \mathbb{E}[\varphi(X(N\Delta t))]$ is not of size Δt uniformly in time (however sufficient conditions for this to hold have been identified, at least for the Euler scheme, in [7]), and one cannot take the limit $T = N\Delta t \to \infty$ in the weak error estimate above. Note that, for a given time-step size Δt, it is not necessary to assume or prove that the numerical scheme admits at least one invariant probability distribution.

Observe that having at most polynomial dependence in the moment bounds and then in the weak error estimates with respect to $T = N\Delta t$ is a non-trivial result and has important practical consequences. Analyzing the computational cost (see Cor. 3.4) shows that, in order to ensure that the approximation weak error is less than ε, it is sufficient to choose the time horizon T and the time-step size Δt such that the required number $N(\varepsilon) = T/\Delta t$ of iterations of the scheme is chosen such that

$$N(\varepsilon) \le C\varepsilon^{-1}|\log(\varepsilon)|^{1+M}.$$

Observe that having non-uniform moment bounds ($M \ge 1$) instead of uniform moment bounds ($M = 0$) thus results only in a marginal increase of the computational cost (supplementary polynomial dependence with respect to $|\log(\varepsilon)|$). Having for instance an exponential dependence would result in a substantial increase of the cost (supplementary polynomial dependence with respect to ε^{-1}). As a consequence, this supports the statement that the considered tamed explicit Euler–Maruyama scheme is effective for the approximation of the invariant distribution.

In this article, we thus prove moment bounds and weak error estimates which are not uniform with respect to time. The numerical scheme considered for an arbitrary time-step size may even be non-ergodic. Since the moments of the exact solution are bounded uniformly with respect to time, and since the considered SDE is ergodic, our results may be non optimal. It is worth mentioning two recent works where the problem studied in this manuscript have been treated, and where authors have proposed other explicit numerical methods which seem to provide uniform moment bounds with respect to the time horizon. First, the authors of the preprint [3] propose truncated tamed Euler schemes and prove uniform second order moment bounds, see [3], Proposition 5.3. In [3], Theorem 5.5, the approximation error for the invariant distribution is of order 1/2 with respect to the time-step size. Second, the authors of the preprint [22] proposed linear implicit numerical schemes with a projection procedure in the explicit treatment of the nonlinearity, see [22], Lemma 4.3 for uniform moment bounds and [22], Theorem 2.5 for their main result, following time-independent weak error estimates.

Thus modifying the scheme considered in this work may provide uniform moment bounds, and whether this is possible for the explicit tamed Euler scheme remains an open question.

One of the objectives of this article is to present in a simplified framework the strategy and the arguments used in the recent article [5] where a tamed explicit exponential Euler scheme is applied for the approximation of the invariant distribution of semilinear parabolic stochastic partial differential equations with one-sided Lipschitz continuous nonlinearities and driven by additive noise.

The literature concerning the numerical approximation of the invariant distribution for ergodic SDEs has been and still is an active research area. Two standard references in the globally Lipschitz continuous situation are [26] and [19], see also [25]. The difficulties to study the non-globally Lipschitz continuous situation have been raised and studied in [18]. The case of Hamiltonian systems has been considered in [27]. The weak error of ergodic schemes has been studied using modified equations and weak backward error analysis in [2] and [12] respectively. A general analysis of order conditions has been developed recently in [14]. Innovative algorithms have been introduced and studied in [1, 4, 15–17, 29]. Finally, algorithms using decreasing-time step sequences have been introduced and studied in [13], and revisited in [21]. Let us also mention the recent article [7]. The list of references above is not exhaustive. Note that the numerical approximation of the invariant distribution for stochastic partial differential equations has also been the subject of works in the last decade: see the recent work [5] and references therein.

This manuscript is organized as follows. The setting is presented in Section 2, including the statement of precise assumptions (Asms. 2.1 and 2.2), properties of the solutions of the SDE (2.4), and the definition of the numerical scheme (Eq. (2.9)). The main results are stated and discussed in Section 3. The proof of Theorem 3.1 is provided in Section 4, whereas Section 5 presents the proof of Theorem 3.3.

2. Setting

Let $\mathbb{N} = \{1, 2, \dots\}$ denote the set of integers.

Let $d \in \mathbb{N}$. The standard inner product and norm in the space $\mathbb{R}^d$ are denoted by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ respectively. Let $d' \in \mathbb{N}$, if $\phi : \mathbb{R}^d \to \mathbb{R}^{d'}$, is a mapping of class $\mathcal{C}^\infty$, for all $j \in \mathbb{N}$ and $x, h_1, \dots, h_j \in \mathbb{R}^d$, the derivative of order j of ϕ at $x \in \mathbb{R}^d$ in directions $h_1, \dots, h_j \in \mathbb{R}^d$ is denoted by $D^j\phi(x).(h_1, \dots, h_j) \in \mathbb{R}^{d'}$. For all $j \in \mathbb{N}$ and $r \geq 0$, set

$$\mathcal{N}_{j,r}(\phi) = \sup_{x\in\mathbb{R}^d} \frac{\|\varphi(x)\|}{1 + \|x\|^r} + \sum_{i=1}^{j} \sup_{x\in\mathbb{R}^d} \sup_{\|h_1\|,\dots,\|h_i\|\leq 1} \frac{\|D^i\phi(x).(h_1, \dots, h_i)\|}{1 + \|x\|^r} < \infty. \tag{2.1}$$

In the proofs, the values of real-valued constants $C \in (0, \infty)$ and of polynomial functions $\mathcal{P}$ may vary from line to line. Dependence with respect to relevant parameters is indicated.

For all $p \in [1, \infty)$ and all $x \in \mathbb{R}^d$, let $\psi_p(x) = \|x\|^{2p}$. Then the mapping ψ_p is of class $\mathcal{C}^2$ and for all $x, h, h_1, h_2 \in \mathbb{R}^d$ one has

$$\begin{aligned}
&D\psi_p(x).h = 2p\langle x, h\rangle\|x\|^{2(p-1)}\\
&D^2\psi_p(x).(h_1, h_2) = 2p\langle h_1, h_2\rangle\|x\|^{2(p-1)} + 4p(p-1)\langle x, h_1\rangle\langle x, h_2\rangle\|x\|^{2(p-2)}.
\end{aligned}$$

The auxiliary mappings ψ_p are used below for the application of Itô's formula in order to prove moment bounds.

Let us first state regularity assumptions concerning the nonlinear drift coefficient f: it is assumed to be of class $\mathcal{C}^\infty$, with at most polynomial growth (Asm. 2.1).

Assumption 2.1. The mapping $f : \mathbb{R}^d \to \mathbb{R}^d$ is of class $\mathcal{C}^\infty$. Moreover, f and all its derivatives D^jf, $j \in \mathbb{N}$, have at most polynomial growth in the following sense: there exists $q_0 \geq 0$ such that one has

$$\sup_{x\in\mathbb{R}^d} \frac{\|f(x)\|}{1 + \|x\|^{q_0}} < \infty$$

and for all $j \in \mathbb{N}$ there exists $q_j \geq 0$ such that

$$\sup_{x\in\mathbb{R}^d} \sup_{\|h_1\|,\dots,\|h_j\|\leq 1} \frac{\|D^j f(x).(h_1,\dots,h_j)\|}{1+\|x\|^{q_j}} < \infty.$$

Using (2.1), an equivalent formulation of Assumption 2.1 is obtained: for all $j \in \mathbb{N}$, there exists $q_j \geq 0$ such that

$$\mathcal{N}_{j,q_j}(f) < \infty.$$

Note that the moment bounds and error estimates below will depend only on $q = \max(q_0, q_1, q_2, q_3)$ and on $\mathcal{N}_{3,q}(f)$. The assumptions for higher-order derivatives are imposed only to ensure regularity properties for solutions of Kolmogorov equations (see Sect. 5.1).

One of the key assumptions in this work is that the nonlinearity f satisfies a one-sided Lipschitz continuity condition.

Assumption 2.2. There exists $\gamma \in (0, \infty)$ such that for all $x_1, x_2 \in \mathbb{R}^d$ one has

$$\langle f(x_2) - f(x_1), x_2 - x_1 \rangle \leq -\gamma \|x_2 - x_1\|^2. \tag{2.2}$$

Assumptions 2.1 and 2.2 are satisfied for instance when $d = 1$ for polynomial functions

$$f(x) = -a_{2p+1}x^{2p+1} - \epsilon \sum_{j=2}^{2p} a_j x^j - a_1 x - a_0$$

with $p \in \mathbb{N}$ and $a_{2p+1}, a_1 > 0$ and sufficiently small $\epsilon > 0$ (to ensure that $\sup\limits_{x\in\mathbb{R}} f(x) < 0$).

Let us now describe the assumptions concerning the diffusion coefficients. Let $K \in \mathbb{N}$.

Assumption 2.3. For all $k \in \{1, \dots, K\}$, the mapping $\sigma_k : \mathbb{R}^d \to \mathbb{R}^d$ is of class $\mathcal{C}^\infty$. Moreover, σ_k and all its derivatives $D^j\sigma_k$, $j \in \mathbb{N}$, are bounded in the following sense: for all $j \in \mathbb{N}$, one has

$$\mathcal{N}_{j,0}(\sigma_k) < \infty.$$

An equivalent formulation of Assumption 2.3 is the following: one has $\sum_{k=1}^K \sup\limits_{x\in\mathbb{R}^d} \|\sigma_k(x)\| < \infty$, and for all $j \in \mathbb{N}$, one has $\sum_{k=1}^K \sup\limits_{x\in\mathbb{R}^d} \sup\limits_{\|h_1\|,\dots,\|h_j\|\leq 1} \|D^j\sigma_k(x).(h_1,\dots,h_j)\| < \infty$.

Like for the drift coefficient f, only the value of $\max\limits_{k=1,\dots,K} \mathcal{N}_{3,0}(\sigma_k)$ will appear in error estimates, however assumptions on higher-order derivatives are needed in the analysis of regularity properties for solutions of Kolmogorov equations (see Sect. 5.1).

Define the real number

$$\mathbf{C}_\sigma = \frac{1}{2}\sum_{k=1}^K \sup_{x_1\neq x_2\in\mathbb{R}^d} \frac{\|\sigma_k(x_2) - \sigma_k(x_1)\|^2}{\|x_2 - x_1\|^2} = \frac{1}{2}\sum_{k=1}^K \sup_{x,h\in\mathbb{R}^d} \|D\sigma_k(x).h\|^2, \tag{2.3}$$

and note that one has $\mathbf{C}_\sigma \in [0, \infty)$ owing to Assumption 2.3. In addition, observe that $\mathbf{C}_\sigma = 0$ when $\sigma_1, \dots, \sigma_K$ are constant (corresponding to additive noise below).

Let $\big(\beta^1(t)\big)_{t\geq 0}, \dots \big(\beta^K(t)\big)_{t\geq 0}$ be independent standard real-valued Wiener processes, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the usual conditions. The expectation operator is denoted by $\mathbb{E}[\cdot]$. In the sequel, the

$$\sigma(\cdot)\mathrm{d}B(t) = \sum_{k=1}^{K} \sigma_k(\cdot)\mathrm{d}\beta^k(t)$$

is used.

In this work, we consider the following SDE with values in $\mathbb{R}^d$:

$$\mathrm{d}X(t) = f(X(t))\mathrm{d}t + \sigma(X(t))\mathrm{d}B(t). \tag{2.4}$$

Owing to the locally Lipschitz and the one-sided Lipschitz continuity properties of the function f (Asms. 2.1 and 2.2), and to the global Lipschitz continuity of the functions $\sigma_1, \dots, \sigma_K$ (Asm. 2.3), for any initial value $x_0 \in \mathbb{R}^d$, there exists a unique global strong solution $\big(X_{x_0}(t)\big)_{t\geq 0}$ to (2.4) with $X_{x_0}(0) = x_0$. In the sequel, instead of writing $\mathbb{E}[\phi(X_{x_0}(t))]$, the notation $\mathbb{E}_{x_0}[\phi(X(t))]$ is often used, and sometimes the dependence with respect to the initial condition x_0 is completely omitted if it is not relevant.

One has moment bounds for solutions, which are uniform with respect to time: for all $p \in [1, \infty)$, there exists $C_p \in (0, \infty)$ such that for all $x_0 \in \mathbb{R}^d$, one has

$$\sup_{t\geq 0} \mathbb{E}[\|X_{x_0}(t)\|^{2p}] \leq C_p(1 + \|x_0\|^{2p}). \tag{2.5}$$

Proof of (2.5). Using Itô's formula, writing $f(X(t)) = f(X(t)) - f(0) + f(0)$, and using the one-sided Lipschitz continuity of f (Asm. 2.2) and the boundedness of σ (Asm. 2.3), one obtains

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[\|X(t)\|^{2p}]}{\mathrm{d}t} &= \frac{\mathrm{d}\mathbb{E}[\psi_p(X(t))]}{\mathrm{d}t} \\
&= 2p\mathbb{E}[\langle f(X(t)), X(t)\rangle \|X(t)\|^{2(p-1)}] \\
&+ 2p\sum_{k=1}^{K} \mathbb{E}[\|\sigma_k(X(t))\|^2\|X(t)\|^{2(p-1)}] + 2p(p-1)\sum_{k=1}^{K} \mathbb{E}[\langle X(t), \sigma_k(X(t))\rangle^2 \|X(t)\|^{2(p-2)}] \\
&\leq -2p\gamma\mathbb{E}[\|X(t)\|^{2p}] + 2p\mathbb{E}[\langle f(0), X(t)\rangle \|X(t)\|^{2(p-1)}] + C_p\mathbb{E}[\|X(t)\|^{2(p-1)}].
\end{aligned}$$

Using Young's inequality, one then obtains

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[\|X(t)\|^{2p}]}{\mathrm{d}t} &\leq -2p\gamma\mathbb{E}[\|X(t)\|^{2p}] + C_p\mathbb{E}[\|X(t)\|^{2p-1}] \\
&\leq -p\gamma\mathbb{E}[\|X(t)\|^{2p}] + C_p,
\end{aligned}$$

and the conclusion follows from the application of Gronwall's lemma. □

Owing to the uniform moment bounds (2.5) and to the Krylov–Bogoliubov criterion, there exists invariant probability distributions associated with the SDE (2.4). To ensure the uniqueness of the invariant distribution and an exponential convergence to equilibrium result, the following condition is imposed in the sequel.

Assumption 2.4. The condition

$$\mathbf{C}_\sigma < \gamma \tag{2.6}$$

is satisfied.

Under Assumption 2.4, there exists a unique invariant distribution $\mu_\star$, such that for any Lipschitz continuous function $\varphi : \mathbb{R}^d \to \mathbb{R}$, for all $T \in (0,\infty)$ and $x_0 \in \mathbb{R}^d$, one has

$$\big|\mathbb{E}[\varphi(X_{x_0}(T))] - \int \varphi \mathrm{d}\mu_\star\big| \le \mathrm{e}^{-(\gamma-\mathbf{C}_\sigma)T}\mathrm{Lip}(\varphi)(1+\|x_0\|) \tag{2.7}$$

with $\mathrm{Lip}(\varphi) = \sup\limits_{x_1\neq x_2 \in \mathbb{R}^d} \frac{\|\varphi(x_2)-\varphi(x_1)\|}{\|x_2-x_1\|}$. In addition, for all $m \in [1,\infty)$ one has $\int \|x\|^m \mathrm{d}\mu_\star(x) < \infty$.

Proof of (2.7). Let $x_0^1, x_0^2 \in \mathbb{R}^d$ be two initial values and consider the solutions $\big(X_{x_0^1}(t)\big)_{t\ge 0}$ and $\big(X_{x_0^2}(t)\big)_{t\ge 0}$ of (2.4) driven by the same Wiener processes. Using Itô's formula, one obtains

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}\mathbb{E}[\|X_{x_0^2}(t)-X_{x_0^1}(t)\|^2]}{\mathrm{d}t} &= \frac{1}{2}\frac{\mathrm{d}\mathbb{E}[\psi_1\big(X_{x_0^2}(t)-X_{x_0^1}(t)\big)]}{\mathrm{d}t}\\
&= \mathbb{E}[\langle f(X_{x_0^2}(t)) - f(X_{x_0^1}(t)), X_{x_0^2}(t)-X_{x_0^1}(t)\rangle]\\
&+ \frac{1}{2}\sum_{k=1}^K \mathbb{E}[\|\sigma_k(X_{x_0^2}(t)) - \sigma_k(X_{x_0^1}(t))\|^2]\\
&\le -\big(\gamma - \mathbf{C}_\sigma\big)\mathbb{E}[\|X_{x_0^2}(t)-X_{x_0^1}(t)\|^2].
\end{aligned}$$

Using Gronwall's lemma gives for all $t \ge 0$

$$\big(\mathbb{E}[\|X_{x_0^2}(t)-X_{x_0^1}(t)\|^2]\big)^{\frac{1}{2}} \le \mathrm{e}^{-(\gamma-\mathbf{C}_\sigma)t}\|x_0^2 - x_0^1\|, \tag{2.8}$$

which then implies the uniqueness of $\mu_\star$ and the inequality (2.7) by straightforward arguments. □

The objective of this work is to define and analyze estimators of $\int \varphi \mathrm{d}\mu_\star$, using an explicit integrator for the discretization of the SDE (2.4). Note that in general, no explicit expression of the invariant distribution $\mu_\star$ is known, and if dimension d is large quadrature rules are not efficient and Monte-Carlo methods are usually employed. In addition, since f has polynomial growth, using the standard explicit Euler–Maruyama scheme is not allowed, see the discussion in Section 1.

Let Δt denote the time-step size, and without loss of generality assume that $\Delta t \in (0, \Delta t_0]$ for some arbitrary fixed $\Delta t_0 > 0$ which plays no important role in this article. Let $t_n = n\Delta t$, and define the Wiener increments as $\Delta\beta_n^k = \beta^k(t_{n+1}) - \beta^k(t_n)$, and $\sigma(\cdot)\Delta B_n = \sum_{k=1}^K \sigma_k(\cdot)\Delta\beta_n^k$, for all $n \ge 0$. For all $t \ge 0$, let $\ell(t) = n$ be the unique integer $n \ge 0$ such that $t_n \le t < t_{n+1}$.

In this work, we consider the explicit tamed Euler-Maruyama method

$$\tilde{X}_{n+1} = \tilde{X}_n + \frac{\Delta t}{1+\alpha\Delta t\|f(\tilde{X}_n)\|} f(\tilde{X}_n) + \sigma(\tilde{X}_n)\Delta B_n, \tag{2.9}$$

where $\alpha \in (0,\infty)$ is an arbitrary parameter and $X_0 = x_0$. The parameter α does not play an important role in the sequel, for simplicity one may choose $\alpha = 1$. The notation $M_n = \alpha\|f(\tilde{X}_n)\|$ is used in the sequel. Introduce a continuous-time auxiliary process $\big(\tilde{X}(t)\big)_{t\ge 0}$ as follows: for all $t \ge 0$, set

$$\tilde{X}(t) = x_0 + \int_0^t \frac{1}{1+\Delta t M_{\ell(s)}} f(\tilde{X}_{\ell(s)})\mathrm{d}s + \int_0^t \sigma(\tilde{X}_{\ell(s)})\mathrm{d}B(s) \tag{2.10}$$

where $\ell(t)$ is defined above and $\big(\tilde{X}_n\big)_{n\ge 0}$ is defined by the scheme (2.9). Observe that by construction one has $\tilde{X}(t_n) = \tilde{X}_n$ for all $n \ge 0$.

The processes $\big(\tilde{X}_n\big)_{n\geq 0}$ and $\big(\tilde{X}(t)\big)_{t\geq 0}$ depend on the time-step size Δt and on the initial condition x_0, however the dependence is omitted to simplify notation.

Remark 2.5. The explicit tamed Euler-Maruyama scheme (2.9) can be interpreted as coming from the application of the standard explicit Euler-Maruyama scheme for the SDE

$$\mathrm{d}X^{\Delta t}(t) = f_{\Delta t}(X^{\Delta t}(t))\mathrm{d}t + \sigma(X^{\Delta t}(t))\mathrm{d}B(t)$$

with modified drift coefficient

$$f_{\Delta t}(x) = \frac{f(x)}{1+\alpha\|f(x)\|}.$$

However, the function $f^{\Delta t}$ does not satisfy a one-sided Lipschitz continuity condition as in Assumption 2.2, with $\gamma > 0$. Indeed, let $d = 1$ and assume that f is a polynomial function, for instance $f(x) = -x - x^3$, then one checks that $\sup_{x\in\mathbb{R}} f'(x) = -1$, however one has

$$f'_{\Delta t}(x) \underset{x\to\pm\infty}{\to} 0.$$

As a consequence, the long-time behavior (uniform in time moment bounds, existence and uniqueness of invariant distriutions) of the process $X^{\Delta t}$ may be non trivial. It may even be the case that $\sup_{x\in\mathbb{R}} f'_{\Delta t}(x) > 0$, in which case uniform in time moment bounds are not expected to hold. In turn no information on the behavior of the scheme (2.9) is obtained by using the above interpretation of the tamed scheme.

3. Main results

The objective of this section is to state the two main results of this article: Theorem 3.1 concerning moment bounds and Theorem 3.3 concerning weak error estimates. The most important feature in these two results is the polynomial dependence with respect to the time T (and with respect to the norm of the initial condition). The consequences in terms of computational cost for the approximation of the invariant distributions are also discussed, see Corollary 3.4.

Theorem 3.1. *Let Assumptions 2.1, 2.2 and 2.3 be satisfied. For all $m \in [1,\infty)$, there exists a polynomial function $\mathcal{P}_m : \mathbb{R}^2 \to \mathbb{R}$, such that for all $T \in (0,\infty)$ and $x_0 \in \mathbb{R}^d$, one has*

$$\sup_{\Delta t\in(0,\Delta t_0]} \sup_{0\leq n\Delta t\leq T} \mathbb{E}[\|\tilde{X}_n\|^m] \leq \mathcal{P}_m(T, \|x_0\|). \tag{3.1}$$

Assumption 2.4 ensures the ergodicity of the SDE. It is not required for the proof of Theorem 3.1. In addition, regularity properties on f and σ are also not required. However, (2.6) from Assumption 2.4 is not sufficient for the proof of Theorem 3.3. In the sequel the following stronger condition is imposed.

Assumption 3.2. The condition

$$5\mathbf{C}_\sigma < \gamma. \tag{3.2}$$

is satisfied.

Theorem 3.3. *Let Assumptions 2.1, 2.2, 2.3 and 3.2 be satisfied. For all $r \geq 0$, there exist polynomial functions $\mathcal{P}_r : \mathbb{R}^2 \to \mathbb{R}$ and $\mathcal{Q}_r : \mathbb{R} \to \mathbb{R}$, such that the following holds: if $\varphi : \mathbb{R}^d \to \mathbb{R}$ is a function of class $\mathcal{C}^\infty$ such that*

$\mathcal{N}_{j,r}(\varphi) < \infty$ for all $j \in \mathbb{N}$ (see Eq. (2.1)), then for all $N \in \mathbb{N}$, $x_0 \in \mathbb{R}^d$ and $\Delta t \in (0, \Delta t_0]$, one has

$$\left|\mathbb{E}[\varphi(\tilde{X}_N)] - \int \varphi \mathrm{d}\mu_\star\right| \leq \mathcal{N}_{3,r}(\varphi)\Big(\mathrm{e}^{-(\gamma - \mathbf{C}_\sigma)N\Delta t} Q_r(\|x_0\|) + \Delta t \mathcal{P}_r(N\Delta t, \|x_0\|)\Big), \tag{3.3}$$

where $\big(\tilde{X}_n\big)_{n\geq 0}$ is defined by the tamed Euler-Maruyama scheme (2.9) with $\tilde{X}_0 = x_0$.

The value of Δt_0 in the statements above is arbitrary, the polynomial functions $\mathcal{P}_r$ and $\mathcal{Q}_r$ may depend on its value but this is omitted to simplify notation.

In Theorem 3.3, it is assumed that the function φ is of class $\mathcal{C}^\infty$ and that φ and its derivatives have at most polynomial growth (or bounded if $r = 0$). However, the right-hand side of (3.3) only depends on $\mathcal{N}_{3,r}(\varphi)$. Higher order regularity and polynomial growth of derivatives are required to prove regularity properties of solutions of Kolmogorov equations (see Sect. 5.1).

Imposing Assumption 3.2 instead of Assumption 2.4 is needed in the proof of uniform bounds (with respect to time) for derivatives of the solutions of Kolmogorov equations. In the additive noise case, *i.e.* when $\sigma_1, \ldots, \sigma_K$ are constant mappings, one has $\mathbf{C}_\sigma = 0$, therefore Assumptions 2.4 and 3.2 are both satisfied as straightforward consequences of Assumption 2.2.

To simplify the presentation, the degrees of the polynomial functions $\mathcal{P}_m$ in Theorem 3.1 and $\mathcal{P}_r$ and $\mathcal{Q}_r$ in Theorem 3.3 are not indicated. Note that their values could be tracked in the proofs. In order to insist on the polynomial dependence with respect to time T or $N\Delta t$, the exact value of the degrees does not matter.

Let us now state some consequences of Theorem 3.3 in terms of computational cost for the approximation of $\int \varphi \mathrm{d}\mu_\star$. In practice, the expectation $\mathbb{E}[\varphi(\tilde{X}_n)]$ needs to be approximated using Monte-Carlo averages. The total computational cost may be reduced for instance using the multilevel Monte-Carlo method. Since those aspects are not specific to the situation studied in this article, we only consider the cost per realization. In this setting, we use the following definition: the cost to sample a realization of $\tilde{X}_N$ is equal to the number of time-steps N. In other words, it means that the cost for the computation of $\tilde{X}_{n+1}$ when $\tilde{X}_n$ is given is set equal to 1. The values of the dimensions d of the solution and K of the noise are assumed to play no role in this study. This makes sense if one compares how different numerical schemes perform when applied to a given class of SDEs. One has the following result.

Corollary 3.4. *Let $x_0 \in \mathbb{R}^d$ and $\varphi : \mathbb{R}^d \to \mathbb{R}$ of class $\mathcal{C}^2$ such that $\mathcal{N}_r(\varphi) < \infty$. There exists a constant $C = C(x_0, \varphi) \in (0, \infty)$ and $R \in \mathbb{N}$, such that for all $\varepsilon \in (0, 1)$, the error satisfies*

$$\left|\mathbb{E}[\varphi(\tilde{X}_N)] - \int \varphi \mathrm{d}\mu_\star\right| \leq \varepsilon$$

where the required number of time-steps N is chosen such that

$$N(\varepsilon) \leq C\varepsilon^{-1}|\log(\varepsilon)|^{1+R}.$$

Comparing the behaviors of powers and logarithms near 0, for all $\alpha \in (0, 1)$ there exists $C_\alpha \in (0, \infty)$ such that one obtains the upper bound

$$N(\varepsilon) \leq C_\alpha \varepsilon^{-\frac{1}{\alpha}}.$$

The exact same upper bound would be obtained if in Theorems 3.1 and 3.3 the upper bounds were uniform in time. Indeed, in that case one could choose $R = 0$ in Corollary 3.4. One also obtains similar results when using some implicit versions of the Euler–Maruyama scheme to deal with the growth of the drift coefficient. It is worth mentioning that the polynomial growth with respect to T in moment bounds and weak error estimates only introduces a polynomial factor in $|\log(\varepsilon)|$ in the computational cost, which is negligible in our analysis. Note that if the dependence with respect to T in Theorems 3.1 and 3.3 had been exponential, an additional factor which would have been a power of ε^{-1} would have appeared in Corollary 3.4, and in turn the computational cost would have significantly increased.

Proof of Corollary 3.4. The error estimate of Theorem 3.3 is rewritten as follows: there exists $R \in \mathbb{N}$ such that

$$\left|\mathbb{E}[\varphi(\tilde{X}_N)] - \int \varphi \mathrm{d}\mu_\star\right| \leq C\Big(\exp(-\gamma N\Delta t) + (1 + (N\Delta t)^R)\Delta t\Big),$$

where C is a constant (depending on x_0 and φ). Note that R may depend on φ (more precisely on the value of r such that $\mathcal{N}_r(\varphi) < \infty$).

The parameters N and Δt are chosen such that

$$\gamma N \Delta t = -\log(\varepsilon/C)$$

and

$$(N\Delta t)^R \Delta t = C\varepsilon.$$

This leads to

$$\Delta t = C\varepsilon |\log(\varepsilon)|^{-R},$$

and finally to choose $N(\varepsilon)$ as the integer part of

$$C|\log(\varepsilon)|\Delta t^{-1} = C\varepsilon^{-1}|\log(\varepsilon)|^{1+R} + 1.$$

This concludes the proof of Corollary 3.4. □

Remark 3.5. In [20], the authors propose to use the so-called rejecting exploding trajectories technique to approximate ergordic averages $\int \varphi \mathrm{d}\mu_\star$, for SDEs with non-globally Lipschitz coefficients. This technique requires to introduce an auxiliary truncation parameter. However, even if in practice it is effective, this technique does not lead to a clean analysis of the cost as in Corollary 3.4.

4. Proof of Theorem 3.1

Let Assumptions 2.1, 2.2 and 2.3 be satisfied. It is worth mentioning that the proof of Theorem 3.1 only requires the following conditions:

- Assumption 2.2 is satisfied,
- the mappings $\sigma_1, \ldots, \sigma_K$ are bounded,
- f and its first order derivative Df have at most polynomial growth, more precisely one has $\mathcal{N}_{1,q_1}(f) < \infty$ (using the notation (2.1)).

Existence and boundedness, resp. polynomial growth, of higher-order derivatives of f, resp. of $\sigma_1, \ldots, \sigma_K$, are not required in throughout this section.

Recall that $M_n = \alpha \|f(\tilde{X}_n)\|$, and that the auxiliary process $\big(\tilde{X}(t)\big)_{t\geq 0}$ is defined by (2.10). Let $\kappa \in (0, \frac{1}{2q_1})$ and $R = \Delta t^{-\kappa}$ be auxiliary parameters.

For every $n \geq 0$, let $\Omega_{R,t_n} = \{\sup_{0\leq \ell\leq n} \|\tilde{X}_\ell\| \leq R\}$, and to simplify notation let $\chi_n = \mathbb{1}_{\Omega_{R,t_n}}$ denote the indicator function of the set Ω_{R,t_n}. Let also $\chi_{-1} = 1$.

To prove Theorem 3.1, it suffices to prove the following two auxiliary results, where again the value of Δt_0 is arbitrary.

Lemma 4.1. *For every $m \in [1,\infty)$, there exists a polynomial function $\mathcal{P}_m^1 : \mathbb{R}^2 \to \mathbb{R}$, such that for all $T \in (0,\infty)$ and $x_0 \in \mathbb{R}^d$, one has*

$$\sup_{\Delta t \in (0,\Delta t_0]} \sup_{0 \le n\Delta t \le T} \mathbb{E}[\chi_{n-1} \|\tilde{X}_n\|^m] \le \mathcal{P}_m^1(T, \|x_0\|). \tag{4.1}$$

Lemma 4.2. *For every $m \in [1,\infty)$, there exists a polynomial function $\mathcal{P}_m^2 : \mathbb{R}^2 \to \mathbb{R}$, such that for all $T \in (0,\infty)$ and $x_0 \in \mathbb{R}^d$, one has*

$$\sup_{\Delta t \in (0,\Delta t_0]} \sup_{0 \le n\Delta t \le T} \mathbb{E}[(1-\chi_n) \|\tilde{X}_n\|^m] \le \mathcal{P}_m^2(T, \|x_0\|). \tag{4.2}$$

Theorem 3.1 is then a straightforward consequence of Lemma 4.1 and Lemma 4.2.

Proof of Theorem 3.1. Since $\Omega_{R,t_n} \subset \Omega_{R,t_{n-1}}$, one has $\chi_n \le \chi_{n-1}$. Writing

$$\mathbb{E}[\|\tilde{X}_n\|^m] = \mathbb{E}[\chi_n \|X_n\|^m] + \mathbb{E}[(1-\chi_n)\|\tilde{X}_n\|^m] \le \mathbb{E}[\chi_{n-1}\|\tilde{X}_n\|^m] + \mathbb{E}[(1-\chi_n)\|\tilde{X}_n\|^m],$$

combining the auxiliary moment bounds (4.1) and (4.2) then concludes the proof. □

In the proofs, values of constants $C, C_m \in (0,\infty)$ and polynomial functions $\mathcal{P}_m$ may change from line to line.

Proof of Lemma 4.1. Introduce two auxiliary processes $\big(Y(t)\big)_{t\ge 0}$ and $\big(Z(t)\big)_{t\ge 0}$ as follows: for all $t \ge 0$,

$$\begin{aligned} Z(t) &:= \tilde{X}(t) - \int_0^t f(\tilde{X}(s))\mathrm{d}s \\ Y(t) &:= \tilde{X}(t) - Z(t). \end{aligned}$$

One obtains the following equality: for all $t \ge 0$

$$Y(t) = \int_0^t f(\tilde{X}(s))\mathrm{d}s = \int_0^t f(Y(s) + Z(s))\mathrm{d}s,$$

from which one obtains that the process Y solves the ordinary differential equation

$$\frac{\mathrm{d}Y(t)}{\mathrm{d}t} = f(Y(t) + Z(t)),$$

with initial value $Y(0) = 0$. Using the one-sided Lipschitz condition (2.2) satisfied by the drift coefficient f (Asm. 2.2), one obtains

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}\|Y(t)\|^2}{\mathrm{d}t} &= \langle Y(t), f(Y(t) + Z(t)\rangle \\ &\le \langle Y(t), f(Y(t) + Z(t)) - f(Z(t))\rangle + \|Y(t)\|\|f(Z(t))\| \\ &\le -\gamma\|Y(t)\|^2 + \|Y(t)\|\|f(Z(t))\|. \end{aligned}$$

Since $\gamma > 0$ by Assumption 2.2, using Young's inequality and Gronwall's lemma, for all $t \ge 0$, one has

$$\|Y(t)\|^2 \le C \int_0^t \|f(Z(s))\|^2 \mathrm{d}s.$$

Multiplying both sides of the above inequality by χ_{n-1} and using Minkowski's inequality, for every $m \in [1,\infty)$, for all $n \in \mathbb{N}$, one has

$$\big(\mathbb{E}[\chi_{n-1}\|Y(t_n)\|^{2m}]\big)^{\frac{1}{m}} \le C\int_0^{t_n}\big(\mathbb{E}[\chi_{n-1}\|f(Z(s))\|^{2m}]\big)^{\frac{1}{m}}\mathrm{d}s. \tag{4.3}$$

Since f has at most polymomial growth (Asm. 2.1), to obtain (4.1), it is thus sufficient to prove that for every $m \in [1,\infty)$, one has an estimate of the type

$$\sup_{0\le t\le t_n\le T}\mathbb{E}[\chi_{n-1}\|Z(t)\|^m] \le \mathcal{P}_m(T,\|x_0\|). \tag{4.4}$$

By definition of the auxiliary process Z, one has for all $t \ge 0$

$$Z(t) = \tilde{X}(t) - \int_0^t f(\tilde{X}(s))\mathrm{d}s = Z_0(t) + Z_1(t) + Z_2(t),$$

where

$$\begin{aligned}
Z_0(t) &= X_0 + \int_0^t \sigma(\tilde{X}_{\ell(s)})\mathrm{d}B(s)\\
Z_1(t) &= -\int_0^t \frac{\Delta t M_{\ell(s)}}{1+\Delta t M_{\ell(s)}} f(\tilde{X}_{\ell(s)})\mathrm{d}s\\
Z_2(t) &= +\int_0^t \big[f(\tilde{X}_{\ell(s)}) - f(\tilde{X}(s))\big]\mathrm{d}s.
\end{aligned}$$

First, since the mappings $\sigma_1,\ldots,\sigma_K$ are assumed to be bounded, there exists $C_m \in (0,\infty)$, such that for all $0 \le t \le t_n \le T$, one has

$$\big(\mathbb{E}[\chi_{n-1}\|Z_0(t)\|^m]\big)^{\frac{1}{m}} \le \big(\mathbb{E}[\|Z_0(t)\|^m\big)^{\frac{1}{m}} \le \|x_0\| + C_m T^{\frac{1}{2}} \le\le \|x_0\| + C_m\big(1+T\big).$$

Second, recall that $M_\ell = \alpha\|f(\tilde{X}_\ell)\|$. Owing to the condition $\mathcal{N}_{1,q_1}(f) < \infty$, f has at most polynomial growth, therefore for all $0 \le t \le t_n \le T$ one has

$$\begin{aligned}
\big(\mathbb{E}[\chi_{n-1}\|Z_1(t)\|^m]\big)^{\frac{1}{m}} &\le \alpha\Delta t\int_0^t \big(\mathbb{E}[\chi_{n-1}\|f(\tilde{X}_{\ell(s)})\|^{2m}]\big)^{\frac{1}{m}}\mathrm{d}s\\
&\le Ct\Delta t(1+R^{2q_1})\\
&\le CT(\Delta t_0 + \Delta t_0^{1-2q\kappa}),
\end{aligned}$$

where we recall that $R = \Delta t^{-\kappa}$ with $2q_1\kappa < 1$, and that the time-step size Δt is bounded from above by Δt_0.

It remains to deal with the term $\mathbb{E}[\chi_{n-1}\|Z_2(t)\|^m]$. Owing to the condition $\mathcal{N}_{1,q_1}(f) < \infty$, the first order derivative of f has at most polynomial growth. Using the Minkowski and Cauchy-Schwarz inequalities, one obtains for all $t \ge 0$

$$\begin{aligned}
&\big(\mathbb{E}[\chi_{n-1}\|Z_2(t)\|^m]\big)^{\frac{1}{m}} \le \int_0^t \big(\chi_{n-1}\|f(\tilde{X}_{\ell(s)}) - f(\tilde{X}(s)\|^m\big)^{\frac{1}{m}}\mathrm{d}s\\
&\le C\int_0^t \big(\mathbb{E}[\chi_{n-1}\|\tilde{X}_{\ell(s)} - \tilde{X}(s)\|^{2m}]\big)^{\frac{1}{2m}}\big(1+\mathbb{E}[\chi_{n-1}\|\tilde{X}_{\ell(s)}\|^{2mq_1}] + \mathbb{E}[\chi_{n-1}\|\tilde{X}(s)\|^{2mq_1}]\big)^{\frac{1}{2m}}\mathrm{d}s.
\end{aligned}$$

The first factor in the integrand above is treated as follows: for $s \leq t < t_n$, one obtains

$$\begin{aligned}\chi_{n-1}\|\tilde{X}_{\ell(s)} - \tilde{X}(s)\| &\leq \chi_{n-1}\frac{|s - t_{\ell(s)}|}{1 + \Delta t M_{\ell(s)}}\|f(\tilde{X}_{\ell(s)}\| + C\|\sigma(\tilde{X}_{\ell(s)})\big(B(s) - B(t_{\ell(s)})\big)\| \\ &\leq C\Delta t(1 + R^{q_1}) + C\|B(s) - B_{t_{\ell(s)}}\|,\end{aligned}$$

using the assumption that σ is bounded. As a consequence, one obtains

$$\big(\mathbb{E}[\chi_{n-1}\|\tilde{X}_{\ell(s)} - \tilde{X}(s)\|^{2p}]\big)^{\frac{1}{2p}} \leq C_p\big(\Delta t(1 + R^{q_1}) + \Delta t^{\frac{1}{2}}\big) \leq C\Delta t^{\frac{1}{2}}(\Delta t_0^{\frac{1}{2}-\kappa q_1} + 1),$$

for all $\Delta t \in (0, \Delta t_0]$, using the definition $R = \Delta t^{-\kappa}$ with $2q_1\kappa < 1$.

To treat the second factor of the integrand above, it suffices to write

$$\mathbb{E}[\chi_{n-1}\|\tilde{X}_{\ell(s)}\|^{2mq_1}] + \mathbb{E}[\chi_{n-1}\|\tilde{X}(s)\|^{2mq_1}] \leq C\mathbb{E}[\chi_{n-1}\|\tilde{X}_{\ell(s)}\|^{2mq_1}] + C\mathbb{E}[\chi_{n-1}\|\tilde{X}(s) - X_{\ell(s)}\|^{2mq_1}],$$

and to use the estimate on the first factor above and the inequality $\mathbb{E}[\chi_{n-1}\|\tilde{X}_{\ell(s)}\|^{2mq_1}] \leq R^{2mq_1}$.

Finally, using again the inequality $\Delta t^{\frac{1}{2}}R^{q_1} \leq \Delta t_0^{\frac{1}{2}-q_1\kappa}$ for all $\Delta t \in (0, \Delta t_0]$, one obtains

$$\big(\mathbb{E}[\chi_{n-1}\|Z_2(t)\|^m]\big)^{\frac{1}{m}} \leq C(\Delta t_0)T.$$

Gathering the estimates then yields (4.4). Inserting (4.4) in the inequality (4.3) then yields

$$\mathbb{E}[\chi_{n-1}\|Y(t_n)\|^m] \leq \mathcal{P}_m(T, \|x_0\|),$$

if $t_n \leq T$ and $\Delta t \in (0, \Delta t_0]$.

Since $X_n = \tilde{X}(t_n) = Y(t_n) + Z(t_n)$, this concludes the proof of Lemma 4.1. $\square$

Proof of Lemma 4.2. Recall that $\chi_n = \mathbb{1}_{\Omega_{R,t_n}}$, with $\Omega_{R,t_n} = \{\sup_{0\leq\ell\leq n} \|\tilde{X}_\ell\| \leq R\}$ and $\chi_{-1} = 1$. As a consequence, one has

$$\begin{aligned}1 - \chi_n = \mathbb{1}_{\Omega^c_{R,t_n}} &= \mathbb{1}_{\Omega^c_{R,t_{n-1}}} + \mathbb{1}_{\Omega_{R,t_{n-1}}}\mathbb{1}_{\|\tilde{X}_n\|>R} \\ &= 1 - \chi_{n-1} + \chi_{n-1}\mathbb{1}_{\|\tilde{X}_n\|>R}.\end{aligned}$$

One thus obtains the equality

$$1 - \chi_n = \sum_{\ell=0}^{n}\chi_{\ell-1}\mathbb{1}_{\|\tilde{X}_\ell\|>R}.$$

Let $m \in \mathbb{N}$. Using Minkowksi, Cauchy-Schwarz and Markov inequalities, one obtains

$$\begin{aligned}\big(\mathbb{E}[(1 - \chi_n)\|\tilde{X}_n\|^m]\big)^{\frac{1}{m}} &\leq \sum_{\ell=0}^{n}\big(\mathbb{E}[\chi_{\ell-1}\mathbb{1}_{\|\tilde{X}_\ell\|>R}\|X_n\|^m]\big)^{\frac{1}{m}} \\ &\leq \sum_{\ell=0}^{n}\big(\mathbb{E}[\|\tilde{X}_n\|^{2m}]\big)^{\frac{1}{2m}}\big(\mathbb{E}[\chi_{\ell-1}\frac{\|\tilde{X}_\ell\|^\theta}{R^\theta}]\big)^{\frac{1}{2m}},\end{aligned}$$

where $\theta \in \mathbb{N}$ is chosen below.

On the one hand, by construction of the tamed Euler scheme, for all $0 \le t \le T$ one has

$$\|\tilde{X}(t)\| \le \|x_0\| + \frac{T}{\alpha \Delta t} + \|\int_0^t \sigma(\tilde{X}_{\ell(s)})\mathrm{d}B(s)\|.$$

Since the mapping σ is assumed to be bounded, and using the upper bound $T^{\frac{1}{2}} \le 1 + T$ and the condition $\Delta t \le \Delta t_0$, there exists $C \in (0, \infty)$ such that when $t_n \le T$ one obtains

$$\big(\mathbb{E}[\|\tilde{X}_n\|^{2m}]\big)^{\frac{1}{2m}} \le C(\|x_0\| + \frac{T}{\Delta t} + 1).$$

On the other hand, applying Lemma 4.1 yields

$$\mathbb{E}[\chi_{\ell-1}\|\tilde{X}_\ell\|^\theta] \le \mathcal{P}_\theta^1(T, \|x_0\|).$$

Gathering the estimates yields

$$\big(\mathbb{E}[(1-\chi_n)\|\tilde{X}_n\|^m]\big)^{\frac{1}{m}} \le C\mathcal{P}_\theta^1(T, \|x_0\|)^{\frac{1}{2m}}\big(1 + \frac{T}{\Delta t}\big)(\|x_0\| + \frac{T}{\Delta t} + 1\big)R^{-\frac{\theta}{2m}}.$$

Since $R = \Delta t^{-\kappa}$, it suffices to choose θ such that $\frac{\theta\kappa}{2m} > 2$ in order to obtain (4.2).

This concludes the proof of Lemma 4.2.

□

5. Proof of Theorem 3.3

The objective of this section is to prove Theorem 3.3. In Section 5.1, some auxiliary results concerning the solution of the associated Kolmogorov equation are given. In this section, f and σ are assumed to satisfy Assumptions 2.1 and 2.3, in particular they are of class $\mathcal{C}^\infty$. Below, these assumptions ensure that solutions of associated Kolmogorov equations are also of class $\mathcal{C}^\infty$ and that the derivatives can be expressed by differentiating the stochastic differential equation. The proof of Theorem 3.3 then follows from the weak error analysis of Section 5.2.

Like in Section 4, the values of constants $C \in (0, \infty)$ and of polynomial functions $\mathcal{P}_r$ or $\mathcal{P}$ may change from line to line.

5.1. Auxiliary result: Kolmogorov equation

We refer to [6], Section 1.3 for details about the regularity results stated below. The analysis in that reference encompasses the case of non-globally Lipschitz but one-sided Lipschitz nonlinear drifts f.

Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a mapping of class $\mathcal{C}^\infty$, with at most polynomial growth: assume that there exists $r \ge 0$ such that for all $j \in \mathbb{N}$ one has $\mathcal{N}_{j,r_j}(\varphi) < \infty$ (using the notation (2.1)). For all $t \ge 0$ and $x \in \mathbb{R}^d$, define

$$u(t,x) = \mathbb{E}_x[\varphi(X_t)]. \tag{5.1}$$

Since φ, f and σ are assumed to be of class $\mathcal{C}^\infty$, with at most polynomial growth, owing to the results from [6], Section 1.3, the mapping u is of class $\mathcal{C}^\infty$ from $\mathbb{R}^+ \times \mathbb{R}^d$ to $\mathbb{R}$ and is the solution of the Kolmogorov equation

$$\partial_t u(t,x) = \mathcal{L}u(t,x) = Du(t,x).f(x) + \frac{1}{2}\sum_{k=1}^K D^2u(t,x).\big(\sigma_k(x), \sigma_k(x)\big) \tag{5.2}$$

with initial value $u(0,\cdot)=\varphi$. In addition, the derivatives of u with respect to the variable x can be computed by differentiating under the expectation sign and differentiating formally the coefficients of the stochastic differential equations, as explained below. For all $t\geq 0$ and $x\in\mathbb{R}^d$, set

$$\overline{u}(t,x)=u(t,x)-\int\varphi\mathrm{d}\mu_\star. \tag{5.3}$$

The objective of this section is to prove the following lemma.

Lemma 5.1. *Let Assumptions 2.1, 2.2, 2.3 and 3.2 be satisfied.*

For any $r\geq 0$, there exists $R\in[r+1,\infty)$ and $C_r\in(0,\infty)$ such that the following holds: for all $\varphi:\mathbb{R}^d\to\mathbb{R}$ of class $\mathcal{C}^\infty$ which satisfies $\mathcal{N}_{j,r}(\varphi)<\infty$ for all $j\in\mathbb{N}$, for all $t\geq 0$ and $x\in\mathbb{R}^d$ one has

$$\left|\overline{u}(t,x)\right|\leq C_r\mathcal{N}_{1,r}(\varphi)\mathrm{e}^{-(\gamma-\mathbf{C}_\sigma)t}(1+\|x\|^{r+1}), \tag{5.4}$$

and one has

$$\sup_{t\geq 0}\mathcal{N}_{3,R}(\overline{u}(t,\cdot))\leq C_r\mathcal{N}_{3,r}(\varphi). \tag{5.5}$$

Lemma 5.1 above states that the mappings $\overline{u}(t,\cdot)$ and its first, second and third order spatial derivatives $D\overline{u}(t,\cdot)$, $D^2\overline{u}(t,\cdot)$ and $D^3\overline{u}(t,\cdot)$ have at most polynomial growth, with some bounds which are uniform with respect to $t\in[0,\infty)$. Moreover, $\overline{u}(t,x)$ goes exponentially fast to 0 when $t\to\infty$, for all $x\in\mathbb{R}^d$, owing to (5.4).

Remark 5.2. In fact, it would be possible to prove a stronger result than the bound (5.5): there exists $\lambda\in(0,\infty)$ such that for all $t\geq 0$ one has

$$\mathcal{N}_{3,R}(\overline{u}(t,\cdot))\leq C_r\mathcal{N}_{3,r}(\varphi)\mathrm{e}^{-\lambda t}. \tag{5.6}$$

One would require this type of result in order to prove weak error estimates for the numerical schemes which are uniform in time. However, since the moment bounds from Theorem 3.1 are not uniform in time, the uniform bound (5.5) given in Lemma 5.1 is sufficient in this article. Using (5.6) instead of (5.5) would only reduce the degree of the polynomial mapping $\mathcal{P}_r$ appearing in the weak error estimates (3.3). Therefore using (5.6) would not qualitatively improve the convergence result in Theorem 3.3, this justifies why we focus on the proof of (5.5).

As will be clear below, the condition $5\mathbf{C}_\sigma<\gamma$ given by Assumption 3.2 is assumed to obtain bounds for the third order derivative $D^3\overline{u}(t,\cdot)$. To prove the bounds for the second order derivative $D^2\overline{u}(t,x)$, the weaker condition $3\mathbf{C}_\sigma<\gamma$ is needed, whereas Assumption 2.4, namely $\mathbf{C}_\sigma<\gamma$ is sufficient to obtain the bounds for the first order derivative $D\overline{u}(t,\cdot)$ and the bound (5.4). The conditions above are sufficient but may not be necessary, in this work we do not look for optimal conditions. Observe that the upper bound (5.4) is a variant of the inequality (2.7), where in Lemma 5.1 the first order derivative φ may have polynomial growth instead of being bounded.

Proof of Lemma 5.1. • Let us first prove the inequality (5.4). Let $x,x_\star\in\mathbb{R}^d$, then using the inequality (2.8) with $x_0^1=x$ and $x_0^2=x_\star$ (see the proof of the inequality (2.7) in Sect. 2), one obtains the upper bound

$$\left(\mathbb{E}[\|X_x(t)-X_{x_\star}(t)\|^2]\right)^{\frac{1}{2}}\leq\mathrm{e}^{-(\gamma-\mathbf{C}_\sigma)t}\|x-x_\star\|,$$

with $\gamma - \mathbf{C}_\sigma > 0$ owing to Assumption 2.4. As a consequence, using the definition of $\mathcal{N}_{1,r}(\varphi)$, Hölder's inequality, the inequality $\int \|x_\star\|^{r+1}\mathrm{d}\mu_\star(x_\star) < \infty$ and the moment bounds (2.5), one obtains

$$\begin{aligned}
\big|\overline{u}(t,x)\big| &= \big|u(t,x) - \int \varphi \mathrm{d}\mu_\star\big| \\
&= \big|\mathbb{E}_x[\varphi(X(t))] - \int \mathbb{E}_{x_\star}[\varphi(X(t))]\mathrm{d}\mu_\star(x_\star)\big| \\
&\leq \int \mathbb{E}[\big|\varphi(X_x(t)) - \varphi(X_{x_\star}(t))\big|]\mathrm{d}\mu_\star(x_\star) \\
&\leq \mathcal{N}_r(\varphi)\int \mathbb{E}[(1+\|X_x(t)\|^r + \|X_{x_\star}(t)\|^r)\|X_x(t) - X_{x_\star}(t)\|]\mathrm{d}\mu_\star(x_\star) \\
&\leq \mathcal{N}_r(\varphi)\int \big(\mathbb{E}[(1+\|X_x(t)\|^r + \|X_{x_\star}(t)\|^r)^2]\big)^{\frac{1}{2}}\big(\mathbb{E}[\|X_x(t) - X_{x_\star}(t)\|^2]\big)^{\frac{1}{2}}\mathrm{d}\mu_\star(x_\star) \\
&\leq C_r\mathcal{N}_r(\varphi)(1+\|x\|^r)\mathrm{e}^{-(\gamma-\mathbf{C}_\sigma)t}\int \|x - x_\star\|\mathrm{d}\mu_\star(x_\star) \\
&\leq C_r\mathcal{N}_r(\varphi)(1+\|x\|^{r+1})\mathrm{e}^{-(\gamma-\mathbf{C}_\sigma)t}.
\end{aligned}$$

This concludes the proof of the bound (5.4).

• Let us now prove the bound (5.5). First, owing to (5.4) and the condition $\mathbf{C}_\sigma < \gamma$, one has

$$\sup_{t\geq 0}\sup_{x\in\mathbb{R}^d}\frac{|\overline{u}(t,x)|}{1+\|x\|^{r+1}} \leq C_r\mathcal{N}_r(\varphi). \tag{5.7}$$

It remains to prove bounds for the first, second and third order derivatives. We refer to [6], Chapter 1 for the justification of the expressions (5.8), (5.12) and (5.18) of the derivatives given below, using the processes η^h, ζ^{h_1,h_2} and ξ^{h_1,h_2,h_3} given by (5.9), (5.13) and (5.19) respectively. We directly focus on obtaining relevant upper bounds on the moments of the auxiliary processes, in order to prove (5.5). As already mentioned above, different conditions $\mathbf{C}_\sigma < \gamma$, $3\mathbf{C}_\sigma < \gamma$ and $5\mathbf{C}_\sigma < \gamma$ are used succesively.

Recall that the notation $\sigma(\cdot)\mathrm{d}B(t) = \sum_{k=1}^K \sigma_k(\cdot)\mathrm{d}\beta^k(t)$ introduced in Section 2. Similarly, the following notation is used below:

$$\begin{aligned}
D\sigma(\cdot).h\mathrm{d}B(t) &= \sum_{k=1}^K D\sigma_k(\cdot).h\mathrm{d}\beta^k(t), \\
D^2\sigma(\cdot).(h_1,h_2)\mathrm{d}B(t) &= \sum_{k=1}^K D^2\sigma_k(\cdot).(h_1,h_2)\mathrm{d}\beta^k(t), \\
D^3\sigma(\cdot).(h_1,h_2,h_3)\mathrm{d}B(t) &= \sum_{k=1}^K D^3\sigma_k(\cdot).(h_1,h_2,h_3)\mathrm{d}\beta^k(t).
\end{aligned}$$

• Let us deal with the first order derivative. For all $t \geq 0$ and $x, h \in \mathbb{R}^d$, the first order derivative is given by

$$D\overline{u}(t,x).h = Du(t,x).h = \mathbb{E}_x[D\varphi(X(t)).\eta^h(t)], \tag{5.8}$$

where the auxiliary process $\big(\eta^h(t)\big)_{t\geq 0}$ is solution of the stochastic differential equation

$$\mathrm{d}\eta^h(t) = Df(X(t)).\eta^h(t)\mathrm{d}t + D\sigma(X(t)).\eta^h(t)\mathrm{d}B(t), \tag{5.9}$$

with initial value $\eta^h(0)=h$, using the notation introduced above. Under the condition $\mathcal{N}_{1,r}(\varphi)<\infty$, using Hölder's inequality and the moment bounds (2.5), one has

$$\begin{aligned}|D\overline{u}(t,x)| &\le \mathcal{N}_{1,r}(\varphi)\mathbb{E}_x[(1+\|X(t)\|^r)\|\eta^h(t)\|]\\ &\le \mathcal{N}_{1,r}(\varphi)\big(\mathbb{E}_x[(1+\|X(t)\|^r)^2]\big)^{1/2}\big(\mathbb{E}[\|\eta^h(t)\|^2]\big)^{1/2}\\ &\le C_r\mathcal{N}_{1,r}(\varphi)(1+\|x\|^r)\big(\mathbb{E}_x[\|\eta^h(t)\|^2]\big)^{1/2}.\end{aligned}$$

Let $p\in[1,\infty)$. Using Itô's formula, the one-sided Lipschitz continuity condition (2.2) and the definition (2.3) of $\mathbf{C}_\sigma$, one obtains the upper bounds

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}_x[\|\eta^h(t)\|^{2p}]}{\mathrm{d}t} &= \frac{\mathrm{d}\mathbb{E}_x[\psi_p(\eta^h(t))]}{\mathrm{d}t}\\ &= 2p\mathbb{E}_x[\langle Df(X(t)).\eta^h(t),\eta^h(t)\rangle\|\eta^h(t)\|^{2(p-1)}]\\ &+ p\sum_{k=1}^K\mathbb{E}_x[\|D\sigma_k(X(t)).\eta^h(t)\|^2\|\eta^h(t)\|^{2(p-1)}]\\ &+ 2p(p-1)\sum_{k=1}^K\mathbb{E}_x[\langle\eta^h(t),D\sigma_k(X(t)).\eta^h(t)\rangle^2\|\eta^h(t)\|^{2(p-2)}]\\ &\le 2p\big(-\gamma+\frac{2p-1}{2}\mathbf{C}_\sigma\big)\mathbb{E}_x[\|\eta^h(t)\|^{2p}].\end{aligned}$$

Recall that the condition $\mathbf{C}_\sigma<\gamma$ is satisfied (as a consequence of Asm. 3.2), therefore one can choose $\mathbf{p}\in(1,\infty)$ arbitrarily close to 1, such that one has $-\gamma+\frac{2\mathbf{p}-1}{2}\mathbf{C}_\sigma>0$. Using Gronwall's lemma, one thus obtains the following bound: there exists $\mathbf{p}\in(1,\infty)$ and $C_{\mathbf{p}}\in(0,\infty)$ such that

$$\sup_{t\ge0}\big(\mathbb{E}_x[\|\eta^h(t)\|^{6\mathbf{p}}]\big)^{\frac{1}{6\mathbf{p}}}\le C_{\mathbf{p}}\|h\|. \tag{5.10}$$

One thus obtains the inequality

$$\sup_{t\ge0}\sup_{\|h\|\le1}|D\overline{u}(t,x).h|\le C_r\mathcal{N}_{1,r}(\varphi)(1+\|x\|^r), \tag{5.11}$$

for some real number $C_r\in(0,\infty)$. Equivalently, one has the uniform upper bound

$$\sup_{t\ge0}\mathcal{N}_{1,r}(\overline{u}(t,\cdot))\le C_r\mathcal{N}_{1,r}(\varphi).$$

This concludes the treatment of the first order derivative.

• Let us now deal with the second order derivative. For all $t\ge0$ and $x,h_1,h_2\in\mathbb{R}^d$, the second order derivative is given by

$$\begin{aligned}D^2\overline{u}(t,x).(h_1,h_2) &= D^2u(t,x).(h_1,h_2)\\ &= \mathbb{E}_x[D\varphi(X(t)).\zeta^{h_1,h_2}(t)]+\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t))],\end{aligned} \tag{5.12}$$

where the auxiliary process $\big(\zeta^{h_1,h_2}(t)\big)_{t\geq 0}$ is solution of the stochastic differential equation

$$\begin{aligned}\mathrm{d}\zeta^{h_1,h_2}(t) &= Df(X(t)).\zeta^{h_1,h_2}(t)\mathrm{d}t + D\sigma(X(t)).\zeta^{h_1,h_2}(t)\mathrm{d}B(t)\\ &\quad + D^2f(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t))\mathrm{d}t + D^2\sigma(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t))\mathrm{d}B(t),\end{aligned} \tag{5.13}$$

with initial value $\zeta^{h_1,h_2}(0)=0$, using the notation introduced above.

Let $p\in(1,\infty)$ and denote the conjugate exponent by $p'=\frac{p}{p-1}$. Under the condition $\mathcal{N}_{2,r}(\varphi)<\infty$, using Hölder's inequality and the moment bounds (2.5), one has

$$\begin{aligned}\big|\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t))]\big| &\leq \mathcal{N}_{2,r}(\varphi)\mathbb{E}_x\big[(1+\|X(t)\|^r)\|\eta^{h_1}(t)\|\|\eta^{h_2}\|\big]\\ &\leq \mathcal{N}_{2,r}(\varphi)\big(\mathbb{E}_x[(1+\|X(t)\|^r)^{p'}]\big)^{\frac{1}{p'}}\big(\mathbb{E}_x[\|\eta^{h_1}(t)\|^{2p}]\big)^{\frac{1}{2p}}\big(\mathbb{E}_x[\|\eta^{h_2}(t)\|^{2p}]\big)^{\frac{1}{2p}}\\ &\leq C_r\mathcal{N}_{2,r}(\varphi)(1+\|x\|^r)\big(\mathbb{E}_x[\|\eta^{h_1}(t)\|^{2p}]\big)^{\frac{1}{2p}}\big(\mathbb{E}_x[\|\eta^{h_2}(t)\|^{2p}]\big)^{\frac{1}{2p}}.\end{aligned}$$

Using the inequality (5.10) and choosing the exponent $p=\mathbf{p}$ where $\mathbf{p}\in(1,\infty)$ has been introduced above in the treatment of the first order derivative, one obtains

$$\big|\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t))]\big| \leq C_r\mathcal{N}_r(\varphi)(1+\|x\|^r)\|h_1\|\|h_2\|. \tag{5.14}$$

Let us now study the other term in the right-hand side of (5.12). Under the condition $\mathcal{N}_{1,r}(\varphi)<\infty$, using Hölder's inequality and the moment bounds (2.5), one has

$$\begin{aligned}\big|\mathbb{E}_x[D\varphi(X_x(t)).\zeta^{h_1,h_2}(t)]\big| &\leq \mathcal{N}_{1,r}(\varphi)\mathbb{E}_x[(1+\|X(t)\|^r)\|\zeta^{h_1,h_2}(t)\|]\\ &\leq \mathcal{N}_{1,r}(\varphi)\big(\mathbb{E}_x[(1+\|X(t)\|^r)^2]\big)^{1/2}\big(\mathbb{E}_x[\|\zeta^{h_1,h_2}(t)\|^2]\big)^{1/2}\\ &\leq C_r\mathcal{N}_{1,r}(\varphi)(1+\|x\|^r)\big(\mathbb{E}_x[\|\zeta^{h_1,h_2}(t)\|^2]\big)^{1/2}.\end{aligned}$$

Let $p,\tilde{p}\in[1,\infty)$ be two exponents and introduce the conjugate exponents $p'=\frac{p}{p-1}$ and $\tilde{p}'=\frac{\tilde{p}}{\tilde{p}-1}$ respectively. Let $\epsilon\in(0,1)$ be an arbitrarily small positive real number. Recall that $\mathcal{N}_{2,q}(f)<\infty$ for some $q\in[1,\infty)$. Using Itô's formula, one obtains the equalities

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}_x[\|\zeta^{h_1,h_2}(t)\|^{2p}]}{\mathrm{d}t} &= \frac{\mathrm{d}\mathbb{E}[\psi_p(\zeta^{h_1,h_2}(t))]}{\mathrm{d}t}\\ &= 2p\mathbb{E}_x[\langle Df(X(t)).\zeta^{h_1,h_2}(t),\zeta^{h_1,h_2}(t)\rangle\|\zeta^{h_1,h_2}(t)\|^{2(p-1)}]\\ &+ p\sum_{k=1}^K\mathbb{E}_x[\|D\sigma_k(X(t)).\zeta^{h_1,h_2}(t)\|^2\|\zeta^{h_1,h_2}(t)\|^{2(p-1)}]\\ &+ 2p(p-1)\sum_{k=1}^K\mathbb{E}_x[\langle\zeta^{h_1,h_2}(t),D\sigma_k(X(t)).\zeta^{h_1,h_2}(t)\rangle^2\|\zeta^{h_1,h_2}(t)\|^{2(p-2)}]\\ &+ 2p\mathbb{E}_x[\langle D^2f(X(t).(\eta^{h_1}(t),\eta^{h_2}(t)),\zeta^{h_1,h_2}(t)\|\zeta^{h_1,h_2}(t)\|^{2(p-1)}]\\ &+ p\sum_{k=1}^K\mathbb{E}_x[\|D^2\sigma_k(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t)))\|^2\|\zeta^{h_1,h_2}(t)\|^{2(p-1)}]\\ &+ 2p(p-1)\sum_{k=1}^K\mathbb{E}_x[\langle\zeta^{h_1,h_2}(t),D^2\sigma_k(X(t).(\eta^{h_1}(t),\eta^{h_2}(t))\rangle^2\|\zeta^{h_1,h_2}(t)\|^{2(p-2)}]\end{aligned}$$

Using the definitions (2.2) and (2.3) of γ and $\mathbf{C}_\sigma$ respectively, one obtains the upper bounds

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}_x[\|\zeta^{h_1,h_2}(t)\|^{2p}]}{\mathrm{d}t} &\leq 2p\big(-\gamma+\frac{2p-1}{2}\mathbf{C}_\sigma\big)\mathbb{E}_x[\|\zeta^{h_1,h_2}(t)\|^{2p}]\\
&+C_p\mathbb{E}_x[(1+\|X(t)\|^q)\|\eta^{h_1}(t)\|\|\eta^{h_2}(t)\|\|\zeta^{h_1,h_2}(t)\|^{2p-1}]\\
&+C_p\mathbb{E}_x[\|\eta^{h_1}(t)\|^2\|\eta^{h_2}(t)\|^2\|\zeta^{h_1,h_2}(t)\|^{2(p-1)}]\\
&\leq 2p\big(-\gamma+\frac{2p-1}{2}\mathbf{C}_\sigma\big)\mathbb{E}_x[\|\zeta^{h_1,h_2}(t)\|^{2p}]\\
&+2p\epsilon\mathbb{E}_x[\|\zeta^{h_1,h_2}(t)\|^{2p}]+C_{\epsilon,p}\mathbb{E}_x[(1+\|X(t)\|^q)^{2p}\|\eta^{h_1}(t)\|^{2p}\|\eta^{h_2}(t)\|^{2p}]\\
&\leq 2p\big(-\gamma+\frac{2p-1}{2}\mathbf{C}_\sigma+\epsilon\big)\mathbb{E}_x[\|\zeta^{h_1,h_2}(t)\|^{2p}]\\
&+C_{\epsilon,p}\big(\mathbb{E}_x[(1+\|X(t)\|^{2qp\tilde{p}'})]\big)^{\frac{1}{\tilde{p}'}}\big(\mathbb{E}_x[\|\eta^{h_1}(t)\|^{4p\tilde{p}}]\big)^{\frac{1}{2\tilde{p}}}\big(\mathbb{E}_x[\|\eta^{h_2}(t)\|^{4p\tilde{p}}]\big)^{\frac{1}{2\tilde{p}}},
\end{aligned}$$

using Hölder and Young inequalities in the last steps.

Choosing a sufficiently small $\epsilon\in(0,1)$, one has $-\gamma+\frac{2\mathbf{p}-1}{2}\mathbf{C}_\sigma+\epsilon>0$ by the definition of $\mathbf{p}$ above. Choosing $p=\frac{3\mathbf{p}}{2}$ and $\tilde{p}\in(1,\mathbf{p})$ arbitrarily close to 1, using (2.5) and (5.10) and Gronwall's lemma, one obtains the following inequality: there exists $C_{\mathbf{p}}\in(0,\infty)$ such that

$$\sup_{t\geq 0}\mathbb{E}_x[\|\zeta^{h_1,h_2}(t)\|^{3\mathbf{p}}]\big)^{\frac{1}{3\mathbf{p}}}\leq C_{\mathbf{p}}(1+\|x\|^{2q\mathbf{p}})\|h_1\|\|h_2\|. \tag{5.15}$$

This yields the inequality

$$\sup_{t\geq 0}\big|\mathbb{E}_x[D\varphi(X(t)).\zeta^{h_1,h_2}(t)]\big|\leq C_{r,\mathbf{p}}\mathcal{N}_{1,r}(\varphi)(1+\|x\|^{r+2q\mathbf{p}})\|h_1\|\|h_2\|, \tag{5.16}$$

and combining the inequalities (5.14) and (5.16) with the expression (5.12) gives

$$\sup_{t\geq 0}\sup_{\|h_1\|,\|h_2\|\leq 1}|D^2\overline{u}(t,x).(h_1,h_2)|\leq C_{r,\mathbf{p}}\mathcal{N}_{2,r}(\varphi)(1+\|x\|^{r+2q\mathbf{p}}), \tag{5.17}$$

for some real number $C_{r,\mathbf{p}}\in(0,\infty)$. Equivalently, one has the uniform upper bound

$$\sup_{t\geq 0}\mathcal{N}_{2,r+2q\mathbf{p}}(\overline{u}(t,\cdot))\leq C_{r,\mathbf{p}}\mathcal{N}_{2,r}(\varphi).$$

This concludes the treatment of the second order derivative.

• Let us finally deal with the third order derivative. For all $t\geq 0$, and $x,h_1,h_2,h_3\in\mathbb{R}^d$, the third order derivative is given by

$$\begin{aligned}
D^3\overline{u}(t,x).(h_1,h_2,h_3)&=D^3u(t,x).(h_1,h_2,h_3)\\
&=\mathbb{E}_x[D\varphi(X(t)).\xi^{h_1,h_2,h_3}(t)]\\
&+\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_1}(t),\zeta^{h_2,h_3}(t))]\\
&+\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_2}(t),\zeta^{h_3,h_1}(t))]\\
&+\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_3}(t),\zeta^{h_1,h_2}(t))]\\
&+\mathbb{E}_x[D^3\varphi(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t))],
\end{aligned} \tag{5.18}$$

where the auxiliary process $\left(\zeta^{h_1,h_2,h_3}(t)\right)_{t\geq 0}$ is solution of the stochastic differential equation

$$\begin{aligned}
\mathrm{d}\xi^{h_1,h_2,h_3}(t) &= Df(X(t)).\xi^{h_1,h_2,h_3}(t)\mathrm{d}t + D\sigma(X(t)).\zeta^{h_1,h_2,h_3}(t)\mathrm{d}B(t)\\
&+ D^2f(X(t)).(\eta^{h_1}(t),\zeta^{h_2,h_3}(t))\mathrm{d}t + D^2\sigma(X(t)).(\eta^{h_1}(t),\zeta^{h_2,h_3}(t))\mathrm{d}B(t)\\
&+ D^2f(X(t)).(\eta^{h_2}(t),\zeta^{h_3,h_1}(t))\mathrm{d}t + D^2\sigma(X(t)).(\eta^{h_2}(t),\zeta^{h_3,h_1}(t))\mathrm{d}B(t)\\
&+ D^2f(X(t)).(\eta^{h_3}(t),\zeta^{h_1,h_2}(t))\mathrm{d}t + D^2\sigma(X(t)).(\eta^{h_3}(t),\zeta^{h_1,h_2}(t))\mathrm{d}B(t)\\
&+ D^3f(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t))\mathrm{d}t + D^3\sigma(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t))\mathrm{d}B(t),
\end{aligned} \tag{5.19}$$

with initial value $\xi^{h_1,h_2,h_3}(0) = 0$, using the notation introduced above. Several terms need to be treated using different arguments.

Under the condition $\mathcal{N}_{3,r}(\varphi) < \infty$, using Hölder's inequality with the exponent $p = \mathbf{p}$ introduced above (in the treatment of the first order derivative) and the conjugate exponent $p' = \frac{p}{p-1}$, one has

$$\begin{aligned}
\left|\mathbb{E}_x[D^3\varphi(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t))]\right| &\leq \mathcal{N}_{3,r}(\varphi)\mathbb{E}_x[(1+\|X(t)\|^r)\|\eta^{h_1}(t)\|\|\eta^{h_2}(t)\|\|\eta^{h_3}(t)\|]\\
&\leq \mathcal{N}_{3,r}(\varphi)\big(\mathbb{E}_x[(1+\|X(t)\|^r)^{p'}]\big)^{\frac{1}{p'}}\big(\mathbb{E}_x[\|\eta^{h_1}(t)\|^{3p}]\big)^{\frac{1}{3p}}\big(\mathbb{E}_x[\|\eta^{h_2}(t)\|^{3p}]\big)^{\frac{1}{3p}}\big(\mathbb{E}_x[\|\eta^{h_3}(t)\|^{3p}]\big)^{\frac{1}{3p}}.
\end{aligned}$$

Using the inequality (5.10) then yields

$$\sup_{t\geq 0}\left|\mathbb{E}_x[D^3\varphi(X(t)).(\eta^{h_1}(t),\eta^{h_2}(t),\eta^{h_3}(t))]\right| \leq C_r\mathcal{N}_{3,r}(\varphi)\|h_1\|\|h_2\|\|h_3\| \tag{5.20}$$

for some real numver $C_r \in (0,\infty)$.

Under the condition $\mathcal{N}_{2,r}(\varphi) < \infty$, using Hölder's inequality with the exponent $p = \mathbf{p}$ and the conjugate exponent $p' = \frac{p}{p-1}$, one has

$$\begin{aligned}
\left|\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_1}(t),\zeta^{h_2,h_3}(t))]\right| &\leq \mathcal{N}_{2,r}(\varphi)\mathbb{E}_x[(1+\|X(t)\|^r)\|\eta^{h_1}(t)\|\|\zeta^{h_2,h_3}(t)\|]\\
&\leq \mathcal{N}_{2,r}(\varphi)\big(\mathbb{E}_x[(1+\|X(t)\|^r)^{p'}]\big)^{\frac{1}{p'}}\big(\mathbb{E}[\|\eta^{h_1}(t)\|^{2p}]\big)^{\frac{1}{2p}}\big(\mathbb{E}_x[\|\zeta^{h_2,h_3}(t)\|^{2p}]\big)^{\frac{1}{2p}}.
\end{aligned}$$

The terms $\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_2}(t),\zeta^{h_3,h_1}(t))]$ and $\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_3}(t),\zeta^{h_1,h_2}(t))]$ are treated similarly. Using the moment bounds (2.5) and the inequalities (5.10) and (5.15) then yields

$$\begin{aligned}
\sup_{t\geq 0}\left|\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_1}(t),\zeta^{h_2,h_3}(t))]\right| &\leq C_r\mathcal{N}_r(\varphi)(1+\|x\|^{r+2q\mathbf{p}})\|h_1\|\|h_2\|\|h_3\|\\
\sup_{t\geq 0}\left|\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_2}(t),\zeta^{h_3,h_1}(t))]\right| &\leq C_r\mathcal{N}_r(\varphi)(1+\|x\|^{r+2q\mathbf{p}})\|h_1\|\|h_2\|\|h_3\|\\
\sup_{t\geq 0}\left|\mathbb{E}_x[D^2\varphi(X(t)).(\eta^{h_3}(t),\zeta^{h_1,h_2}(t))]\right| &\leq C_r\mathcal{N}_r(\varphi)(1+\|x\|^{r+2q\mathbf{p}})\|h_1\|\|h_2\|\|h_3\|
\end{aligned} \tag{5.21}$$

It remains to deal with the term $\mathbb{E}_x[D\varphi(X_x(t)).\xi^{h_1,h_2,h_3}(t)]$. Under the condition $\mathcal{N}_{1,r}(\varphi) < \infty$, using Hölder's inequality and the moment bounds (2.5), one has

$$\begin{aligned}
\left|\mathbb{E}[D\varphi(X_x(t)).\xi^{h_1,h_2,h_3}(t)]\right| &\leq \mathcal{N}_{1,r}(\varphi)\mathbb{E}[(1+\|X(t)\|^r)\|\xi^{h_1,h_2,h_3}(t)\|]\\
&\leq \mathcal{N}_{1,r}(\varphi)\big(\mathbb{E}[(1+\|X(t)\|^r)^2]\big)^{\frac{1}{2}}\big(\mathbb{E}[\|\xi^{h_1,h_2,h_3}(t)\|^2]\big)^{\frac{1}{2}}\\
&\leq C_r\mathcal{N}_{1,r}(\varphi)(1+\|x\|^r)\big(\mathbb{E}[\|\xi^{h_1,h_2,h_3}(t)\|^2]\big)^{\frac{1}{2}}.
\end{aligned}$$

Let $\epsilon \in (0,1)$ be an arbitrarily small positive real number. Recall that $\mathcal{N}_{3,q}(f) < \infty$. Using Itô's formula and Young's inequality, and the definitions (2.2) and (2.3) of γ and $\mathbf{C}_\sigma$, one obtains

$$\begin{aligned}
&\frac{\mathrm{d}\mathbb{E}_x[\|\xi^{h_1,h_2,h_3}(t)\|^2]}{\mathrm{d}t} = \frac{\mathrm{d}\mathbb{E}_x[\psi_1(\xi^{h_1,h_2,h_3}(t))]}{\mathrm{d}t} \\
&= 2\mathbb{E}_x[\langle Df(X(t)).\xi^{h_1,h_2,h_3}(t), \xi^{h_1,h_2,h_3}(t)\rangle] + \sum_{k=1}^K \mathbb{E}_x[\|D\sigma_k(X(t)).\xi^{h_1,h_2,h_3}(t)\|^2] \\
&+ 2\mathbb{E}_x[\langle D^2 f(X(t)).(\eta^{h_1}(t), \zeta^{h_2,h_3}(t)), \xi^{h_1,h_2,h_3}(t)\rangle] + \sum_{k=1}^K \mathbb{E}_x[\|D^2\sigma_k(X(t)).(\eta^{h_1}(t), \zeta^{h_2,h_3}(t))\|^2] \\
&+ 2\mathbb{E}_x[\langle D^2 f(X(t)).(\eta^{h_2}(t), \zeta^{h_3,h_1}(t)), \xi^{h_1,h_2,h_3}(t)\rangle] + \sum_{k=1}^K \mathbb{E}_x[\|D^2\sigma_k(X(t)).(\eta^{h_2}(t), \zeta^{h_3,h_1}(t))\|^2] \\
&+ 2\mathbb{E}_x[\langle D^2 f(X(t)).(\eta^{h_3}(t), \zeta^{h_1,h_2}(t)), \xi^{h_1,h_2,h_3}(t)\rangle] + \sum_{k=1}^K \mathbb{E}_x[\|D^2\sigma_k(X(t)).(\eta^{h_3}(t), \zeta^{h_1,h_2}(t))\|^2] \\
&+ 2\mathbb{E}_x[\langle D^3 f(X(t)).(\eta^{h_1}(t), \eta^{h_2}(t), \eta^{h_3}(t)), \xi^{h_1,h_2,h_3}(t)\rangle] \\
&+ \sum_{k=1}^K \mathbb{E}_x[\|D^3\sigma_k(X(t)).(\eta^{h_1}(t), \zeta^{h_2}(t), \eta^{h_3}(t))\|^2] \\
&\leq 2\big(-\gamma + \frac{1}{2}\mathbf{C}_\sigma + \epsilon\big)\mathbb{E}_x[\|\xi^{h_1,h_2,h_3}(t)\|^2] \\
&+ C_\epsilon \mathbb{E}[(1+\|X(t)\|^q)^2\|\eta^{h_1}(t)\|^2\|\zeta^{h_2,h_3}(t)\|^2] \\
&+ C_\epsilon \mathbb{E}_x[(1+\|X(t)\|^q)^2\|\eta^{h_2}(t)\|^2\|\zeta^{h_3,h_1}(t)\|^2] \\
&+ C_\epsilon \mathbb{E}_x[(1+\|X(t)\|^q)^2\|\eta^{h_3}(t)\|^2\|\zeta^{h_1,h_2}(t)\|^2] \\
&+ C\mathbb{E}_x[(1+\|X(t)\|^q)^2\|\eta^{h_1}(t)\|^2\|\eta^{h_2}(t)\|^2\|\eta^{h_3}(t)\|^2].
\end{aligned}$$

Using Hölder's inequality with an auxiliary parameter $p \in (1,\infty)$ and $p' = \frac{p}{p-1}$, one has

$$\mathbb{E}[(1+\|X(t)\|^q)^2\|\eta^{h_1}(t)\|^2\|\zeta^{h_2,h_3}(t)\|^2] \leq \big(\mathbb{E}[(1+\|X(t)\|^q)^{2p'}]\big)^{\frac{1}{p'}}\big(\mathbb{E}[\|\eta^{h_1}(t)\|^{6p}]\big)^{\frac{1}{3p}}\big(\mathbb{E}[\|\zeta^{h_2,h_3}(t)\|^{3p}\big)^{\frac{2}{3p}}$$

and

$$\begin{aligned}
&\mathbb{E}[(1+\|X(t)\|^q)^2\|\eta^{h_1}(t)\|^2\|\eta^{h_2}(t)\|^2\|\eta^{h_3}(t)\|^2] \\
&\qquad \leq \big(\mathbb{E}[(1+\|X(t)\|^q)^{2p'}]\big)^{\frac{1}{p'}}\big(\mathbb{E}[\|\eta^{h_1}(t)\|^{6p}]\big)^{\frac{1}{3p}}\big(\mathbb{E}[\|\eta^{h_2}(t)\|^{6p}]\big)^{\frac{1}{3p}}\big(\mathbb{E}[\|\eta^{h_3}(t)\|^{6p}]\big)^{\frac{1}{3p}}.
\end{aligned}$$

Choosing the exponent $p = \mathbf{p}$ introduced above and a sufficiently small $\epsilon \in (0,1)$, using the moment bounds (2.5), the inequalities (5.10) and (5.15) and Gronwall's lemma, one obtains

$$\sup_{t\geq 0} \mathbb{E}_x[\|\xi^{h_1,h_2,h_3}(t)\|^2] \leq C_q(1+\|x\|^{2q})\|h_1\|\|h_2\|\|h_3\| \tag{5.22}$$

which then yields

$$\sup_{t\geq 0} \mathbb{E}_x[D\varphi(X_x(t)).\xi^{h_1,h_2,h_3}(t)] \leq C_{r,q}\mathcal{N}_{1,r}(\varphi)(1+\|x\|^{r+2q})\|h_1\|\|h_2\|\|h_3\|. \tag{5.23}$$

Combining the inequalities (5.20), (5.21) and (5.23) with the expression (5.18) gives

$$\sup_{t\geq 0}\sup_{\|h_1\|,\|h_2\|,\|h_3\|\leq 1}|D^3\overline{u}(t,x).(h_1,h_2,h_3)|\leq C_{r,\mathbf{p}}\mathcal{N}_{3,r}(\varphi)(1+\|x\|^{r+2q\mathbf{p}}),\tag{5.24}$$

for some real number $C_{r,\mathbf{p}}\in(0,\infty)$. Equivalently, one has the uniform upper bound

$$\sup_{t\geq 0}\mathcal{N}_{3,r+2q\mathbf{p}}(\overline{u}(t,\cdot))\leq C_{r,\mathbf{p}}\mathcal{N}_{3,r}(\varphi).$$

• We are now in position to conclude the proof of (5.5): it suffices to set $R=r+\min(1,2q\mathbf{p})$ and to combine the bounds (5.7), (5.11), (5.17) and (5.24). The proof of Lemma 5.1 is thus completed. □

5.2. Weak error analysis

It remains to prove Theorem 3.3. This is performed by combining the moment bounds given by Theorem 3.1 and the regularity results from Lemma 5.1 with standard weak error analysis arguments to provide the proof of Theorem 3.3. By a linearity argument, without loss of generality one assumes that $\mathcal{N}_{3,r}(\varphi)\leq 1$ to simplify the presentation below.

Note that Theorem 3.1 gives moment bounds for $\tilde{X}_n$, with $0\leq n\Delta t\leq t_N$, and it is straightforward to deduce moment bounds of the same type for $\tilde{X}(t)$ defined by (2.10), with $0\leq t\leq T$: for all $m\in\mathbb{N}$, one has

$$\sup_{0\leq t\leq T}\mathbb{E}[\|\tilde{X}(t)\|^m]\leq\mathcal{P}_m(T,\|x_0\|)\tag{5.25}$$

where $\mathcal{P}_m:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is a polynomial function.

Proof of Theorem 3.3. Recall the definition (5.3) of the mapping $\overline{u}$. The weak error can then be written as

$$\begin{aligned}\mathbb{E}[\varphi(\tilde{X}_N)]-\int\varphi\mathrm{d}\mu_\star&=\mathbb{E}[\overline{u}(0,\tilde{X}_N)]\\&=\mathbb{E}[\overline{u}(0,\tilde{X}_N)]-\mathbb{E}[\overline{u}(N\Delta t,\tilde{X}_0)]+\overline{u}(N\Delta t,x_0).\end{aligned}$$

Using the inequality (5.4) from Lemma 5.1, the second term on the right-hand side above is bounded as follows: there exists a polynomial function $\mathcal{Q}_r$ such that for all $N\in\mathbb{N}$, $\Delta t\in(0,\Delta t_0)$ and $x_0\in\mathbb{R}^d$, one has

$$\big|\overline{u}(N\Delta t,x_0)\big|\leq\mathrm{e}^{-(\gamma-\mathbf{C}_\sigma)N\Delta t}\mathcal{Q}_r(\|x_0\|).\tag{5.26}$$

To proceed with the analysis of the error term $\mathbb{E}[\overline{u}(0,\tilde{X}_N)]-\mathbb{E}[\overline{u}(N\Delta t,\tilde{X}_0)]$, it is worth recalling a standard decomposition of the weak error, based on a telescoping sum argument (see for instance [24], [28], see also the recent article [7] for example): one has

$$\mathbb{E}[\overline{u}(0,\tilde{X}_N)]-\mathbb{E}[\overline{u}(N\Delta t,\tilde{X}_0)]=\sum_{n=0}^{N-1}\big(\mathbb{E}[\overline{u}(t_N-t_{n+1},\tilde{X}_{n+1})]-\mathbb{E}[\overline{u}(t_N-t_n,\tilde{X}_n)]\big).$$

Recall that the auxiliary process $\tilde{X}$ defined by (2.10) satisfies the property $\tilde{X}(t_n)=\tilde{X}_n$. Using that property, Itô's formula and the fact that $\overline{u}$ solves the Kolmogorov equation (5.2), one obtains

$$\mathbb{E}[\overline{u}(0,\tilde{X}_N)]-\mathbb{E}[\overline{u}(N\Delta t,\tilde{X}_0)]=\sum_{n=0}^{N-1}\big(\mathbb{E}[\overline{u}(t_N-t_{n+1},\tilde{X}_{n+1})]-\mathbb{E}[\overline{u}(t_N-t_n,\tilde{X}_n)]\big)$$

$$
\begin{aligned}
&= \sum_{n=0}^{N-1} \big(\mathbb{E}[\overline{u}(t_N - t_{n+1}, \tilde{X}(t_{n+1}))] - \mathbb{E}[\overline{u}(t_N - t_n, \tilde{X}(t_n))] \big) \\
&= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E}[D\overline{u}(t_N - t, \tilde{X}(t)).\big(\frac{f(\tilde{X}_n)}{1+\Delta t M_n} - f(\tilde{X}(t)) \big)]\mathrm{d}t \\
&+ \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{1}{2} \sum_{k=1}^{K} \mathbb{E}[D^2\overline{u}(t_N - t, \tilde{X}(t)).\big(\sigma_k(\tilde{X}_n), \sigma_k(\tilde{X}_n)\big) - D^2\overline{u}(t_N - t, \tilde{X}(t)).\big(\sigma_k(\tilde{X}(t)), \sigma_k(\tilde{X}(t))\big)]\mathrm{d}t,
\end{aligned}
$$

where we recall that $M_n = \alpha \| f(\tilde{X}_n) \|$.

As a result, one obtains the following decomposition

$$
\mathbb{E}[\overline{u}(0, \tilde{X}_N)] - \mathbb{E}[\overline{u}(N\Delta t, \tilde{X}_0)] = \epsilon_N^1 + \epsilon_N^2 + \epsilon_N^3 + \epsilon_N^4 + \epsilon_N^5,
$$

where the error terms in the right-hand side are given by

$$
\begin{aligned}
\epsilon_N^1 &= -\Delta t \int_0^{t_N} \mathbb{E}[M_{\ell(t)} D\overline{u}(t_N - t, \tilde{X}(t)).f(\tilde{X}_{\ell(t)})]\mathrm{d}t \\
\epsilon_N^2 &= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E}[D\overline{u}(t_N - t, \tilde{X}_n).\big(f(\tilde{X}_n) - f(\tilde{X}(t))\big)]\mathrm{d}t \\
\epsilon_N^3 &= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E}[\big(D\overline{u}(t_N - t, \tilde{X}(t)) - D\overline{u}(t_N - t, \tilde{X}_n)\big).\big(f(\tilde{X}_n) - f(\tilde{X}(t))\big)]\mathrm{d}t \\
\epsilon_N^4 &= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{1}{2} \sum_{k=1}^{K} \mathbb{E}[D^2\overline{u}(t_N - t, \tilde{X}_n).\big(\sigma_k(\tilde{X}_n), \sigma_k(\tilde{X}_n)\big) - D^2\overline{u}(t_N - t, \tilde{X}_n).\big(\sigma_k(\tilde{X}(t)), \sigma_k(\tilde{X}(t))\big)]\mathrm{d}t \\
\epsilon_N^5 &= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{1}{2} \sum_{k=1}^{K} \mathbb{E}[\big(D^2\overline{u}(t_N - t, \tilde{X}(t)) - D^2\overline{u}(t_N - t, \tilde{X}_n)\big).\big(\sigma_k(\tilde{X}_n), \sigma_k(X_n)\big)]\mathrm{d}t \\
&\quad - \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{1}{2} \sum_{k=1}^{K} \mathbb{E}[\big(D^2\overline{u}(t_N - t, \tilde{X}(t)) - D^2\overline{u}(t_N - t, \tilde{X}_n)\big).\big(\sigma_k(\tilde{X}(t)), \sigma_k(\tilde{X}(t))\big)]\mathrm{d}t
\end{aligned}
$$

It remains to provide upper bounds for each of the five error terms introduced above. In the sequel, the constants and polynomial functions are independent of the time-step size Δt.

• *Error term* ϵ_N^1. Using the inequality (5.5) from Lemma 5.1, the polynomial growth condition on f from Assumption 2.1 and the moment bounds (5.25) which follow from Theorem 3.1, one obtains

$$
\begin{aligned}
|\epsilon_N^1| &\leq \alpha \Delta t \int_0^{t_N} \mathbb{E}[(1 + \|\tilde{X}(t)\|^R) \|f(\tilde{X}_n)\|^2]\mathrm{d}t \\
&\leq C t_N \Delta t (1 + \sup_{0 \leq t \leq t_N} \big(\mathbb{E}[\|\tilde{X}(t)\|^{2R}]\big)^{\frac{1}{2}} \big(1 + \sup_{0 \leq n\Delta t \leq T} \mathbb{E}[\|\tilde{X}_n\|^{4q}]\big)^{\frac{1}{2}} \\
&\leq C \Delta t \mathcal{P}(t_N, \|x_0\|),
\end{aligned}
$$

where $\mathcal{P}$ is a polynomial function.

• *Error term* ϵ_N^2. Applying Itô's formula to $t \in [t_n, t_{n+1}] \mapsto f(\tilde{X}(t))$ and a conditional expectation argument, for all $t \in [t_n, t_{n+1}]$ one has

$$\mathbb{E}[D\overline{u}(t_N - t, \tilde{X}_n).\big(f(\tilde{X}_n) - f(\tilde{X}(t))\big)]\mathrm{d}t = \mathbb{E}[D\overline{u}(t_N - t, \tilde{X}_n).\big(\int_{t_n}^{t} Df(\tilde{X}(s)).\frac{f(\tilde{X}_n)}{1+\Delta t M_n}\mathrm{d}s\big)]$$
$$+ \mathbb{E}[D\overline{u}(t_N - t, \tilde{X}_n).\big(\frac{1}{2}\int_{t_n}^{t}\sum_{k=1}^{K} D^2 f(\tilde{X}(s)).(\sigma_k(\tilde{X}_n), \sigma_k(\tilde{X}_n))\mathrm{d}s\big)].$$

Using the inequality (5.5) from Lemma 5.1, the polynomial growth assumption on f (Asm. 2.1), the boundedness of $\sigma_1, \ldots, \sigma_K$ (Asm. 2.3) and the moment bounds (5.25) which follow from Theorem 3.1, one obtains

$$\begin{aligned}
|\epsilon_N^2| &\leq C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}\int_{t_n}^{t_{n+1}} \mathbb{E}\big[(1+\|\tilde{X}_n\|)^R(1+\|\tilde{X}(t)\|^q)(1+\|\tilde{X}_n\|^q)\big]\mathrm{d}s\mathrm{d}t \\
&+ C\Delta t\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}\int_{t_n}^{t} \mathbb{E}[(1+\|\tilde{X}_n\|^R)(1+\|\tilde{X}(s)\|^q)]\mathrm{d}s\mathrm{d}t \\
&\leq C\Delta t\mathcal{P}(t_N, \|x_0\|),
\end{aligned}$$

where $\mathcal{P}$ is a polynomial function.

• *Error term* ϵ_N^3. Using a Taylor expansion argument, the inequality (5.5) from Lemma 5.1, the polynomial growth assumption on f (Asm. 2.1), one has

$$\begin{aligned}
|\epsilon_N^3| &\leq C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \mathbb{E}[(1+\|\tilde{X}_n\|^{R+q}+\|\tilde{X}(t)\|^{R+q})\|\tilde{X}(t)-\tilde{X}_n\|^2]\mathrm{d}t \\
&\leq C\Delta t\mathcal{P}(t_N, \|x_0\|),
\end{aligned}$$

where $\mathcal{P}$ is a polynomial function, using Hólder's inequality, the moment bounds (5.25) which follow from Theorem 3.1, and the inequality

$$\mathbb{E}[\|\tilde{X}(t)-\tilde{X}_n\|^4] \leq C\Delta t^4\mathbb{E}[\|f(\tilde{X}_n)\|^4] + \mathbb{E}[\|\sigma(\tilde{X}_n)(B(t)-B(t_n))\|^4] \leq C\Delta t^2\mathcal{P}(t_N, \|x_0\|)$$

for all $t \in [t_n, t_{n+1})$.

• *Error term* ϵ_N^4. The error term ϵ_N^4 is treated using the same arguments as in the treatment of the error term ϵ_N^2: using Itô's formula, a conditional expectation argument, the assumption that $\sigma_1, \ldots, \sigma_K$ have bounded first and second order derivatives (Asm. 2.3) and the inequality (5.5) from Lemma 5.1, one obtains

$$\begin{aligned}
|\epsilon_N^4| &\leq C\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}\int_{t_n}^{t_{n+1}} \mathbb{E}\big[(1+\|\tilde{X}_n\|)^R(1+\|\tilde{X}_n\|^q)\big]\mathrm{d}s\mathrm{d}t \\
&\leq \Delta t\mathcal{P}(t_N, \|x_0\|),
\end{aligned}$$

where $\mathcal{P}$ is a polynomial function.

• *Error term* ϵ_N^5. The error term ϵ_N^5 is treated using the same arguments as in the treatment of the error term ϵ_N^3: using a Taylor expansion argument, the assumption that $\sigma_1, \ldots, \sigma_K$ are bounded and globally Lipschitz continuous (Asm. 2.3) and the inequality (5.5) from Lemma 5.1, one obtains

$$\begin{aligned}|\epsilon_N^5| &\leq C \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E}[(1 + \|\tilde{X}_n\|^R + \|\tilde{X}(t)\|^R)\|\tilde{X}(t) - \tilde{X}_n\|^2]\mathrm{d}t \\ &\leq C\Delta t \mathcal{P}(t_N, \|x_0\|),\end{aligned}$$

where $\mathcal{P}$ is a polynomial function.

• Gathering the estimates on the error terms ϵ_N^1, ϵ_N^2, ϵ_N^3, ϵ_N^4 and ϵ_N^5, we obtain the following result: there exists a polynomial function $\mathcal{P}$ such that for all $N \in \mathbb{N}$, $\Delta t \in (0, \Delta t_0)$ and $x_0 \in \mathbb{R}^d$, one has

$$\left|\mathbb{E}[\overline{u}(0, \tilde{X}_N)] - \mathbb{E}[\overline{u}(N\Delta t, \tilde{X}_0)] \leq \Delta t \mathcal{P}(N\Delta t, \|x_0\|).\right. \tag{5.27}$$

Combining the inequalities (5.26) and (5.27) then provides the weak error estimate (3.3), which concludes the proof of Theorem 3.3. □

Acknowledgements. The author would like to thank the anonymous referees for useful suggestions which helped improved the quality of the article. This work is partially supported by the SIMALIN project ANR-19-CE40-0016 of the French National Research Agency.

References

[1] A. Abdulle, I. Almuslimani and G. Vilmart, Optimal explicit stabilized integrator of weak order 1 for stiff and ergodic stochastic differential equations. *SIAM/ASA J. Uncertain. Quantif.* **6** (2018) 937–964.

[2] A. Abdulle, G. Vilmart and K.C. Zygalakis, High order numerical approximation of the invariant measure of ergodic SDEs. *SIAM J. Numer. Anal.* **52** (2014) 1600–1622.

[3] L. Angeli, D. Crisan and M. Ottobre, Uniform in time convergence of numerical schemes for stochastic differential equations via strong exponential stability: Euler methods, split-step and tamed schemes. Preprint, 2023.

[4] N. Bou-Rabee and E. Vanden-Eijnden, Pathwise accuracy and ergodicity of metropolized integrators for SDEs. *Commun. Pure Appl. Math.* **63** (2010) 655–696.

[5] C.-E. Bréhier, Approximation of the invariant distribution for a class of ergodic SPDEs using an explicit tamed exponential Euler scheme. *ESAIM Math. Model. Numer. Anal.* **56** (2022) 151–175.

[6] S. Cerrai, Second order PDE's in finite and infinite dimension. Vol. 1762 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin (2001).

[7] D. Crisan, P. Dobson and M. Ottobre, Uniform in time estimates for the weak error of the Euler method for SDEs and a pathwise approach to derivative estimates for diffusion semigroups. *Trans. Am. Math. Soc.* **374** (2021) 3289–3330.

[8] M. Hutzenthaler and A. Jentzen, Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients. *Mem. Am. Math. Soc.* **236** (2015) v+99.

[9] M. Hutzenthaler, A. Jentzen and P.E. Kloeden, Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **467** (2011) 1563–1576.

[10] M. Hutzenthaler, A. Jentzen and P.E. Kloeden, Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab.* **22** (2012) 1611–1641.

[11] C. Kelly and G. J. Lord, Adaptive time-stepping strategies for nonlinear stochastic systems. *IMA J. Numer. Anal.* **38** (2018) 1523–1549.

[12] M. Kopec, Weak backward error analysis for overdamped Langevin processes. *IMA J. Numer. Anal.* **35** (2015) 583–614.

[13] D. Lamberton and G. Pagès, Recursive computation of the invariant distribution of a diffusion. *Bernoulli* **8** (2002) 367–405.

[14] A. Laurent and G. Vilmart, Exotic aromatic B-series for the study of long time integrators for a class of ergodic SDEs. *Math. Comp.* **89** (2020) 169–202.

[15] B. Leimkuhler, C. Matthews and G. Stoltz, The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.* **36** (2016) 13–79.

[16] B. Leimkuhler, C. Matthews and M. V. Tretyakov, On the long-time integration of stochastic gradient systems. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **470** (2014) 20140120, 16.

[17] B. Leimkuhler, C. Matthews and M. V. Tretyakov, On the long-time integration of stochastic gradient systems. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **470** (2014) 20140120, 16.

[18] J.C. Mattingly, A.M. Stuart and D.J. Higham, Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.* **101** (2002) 185–232.

[19] J.C. Mattingly, A.M. Stuart and M.V. Tretyakov, Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM J. Numer. Anal.* **48** (2010) 552–577.
[20] G.N. Milstein and M.V. Tretyakov, Computing ergodic limits for Langevin equations. *Phys. D* **229** (2007) 81–95.
[21] G. Pagès and C. Rey, Recursive computation of invariant distributions of Feller processes. *Stochastic Process. Appl.* **130** (2020) 328–365.
[22] C. Pang, X. Wang and Y. Wu, Linear implicit approximations of invariant measures of semi-linear sdes with non-globally Lipschitz coefficients. Preprint, 2023.
[23] S. Sabanis, A note on tamed Euler approximations. *Electron. Commun. Probab.* **18** (2013) 10.
[24] D. Talay, Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution. *RAIRO Modél. Math. Anal. Numér.* **20** (1986) 141–179.
[25] D. Talay, Classification of discretization schemes of diffusions according to an ergodic criterium, in Stochastic modelling and filtering (Rome, 1984). Vol. 91 of *Lect. Notes Control Inf. Sci.*. Springer, Berlin (1987) 207–218.
[26] D. Talay, Second-order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stochastics Stochastic Rep.* **29** (1990) 13–36.
[27] D. Talay, Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. Vol. 8. (2002) 163–198.
[28] D. Talay and L. Tubaro, Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.* **8** (1991) 483–509.
[29] G. Vilmart, Postprocessed integrators for the high order integration of ergodic SDEs. *SIAM J. Sci. Comput.* **37** (2015) A201–A220.