Potential Analysis (2024) 60:1031–1065
https://doi.org/10.1007/s11118-023-10078-5

RESEARCH

Rate of Convergence in the Smoluchowski-Kramers Approximation for Mean-field Stochastic Differential Equations

Ta Cong Son[1] · **Dung Quang Le**[2] · **Manh Hong Duong**[3]

Received: 6 October 2022 / Accepted: 15 May 2023 / Published online: 3 July 2023

Abstract
In this paper we study a second-order mean-field stochastic differential systems describing the movement of a particle under the influence of a time-dependent force, a friction, a mean-field interaction and a space and time-dependent stochastic noise. Using techniques from Malliavin calculus, we establish explicit rates of convergence in the zero-mass limit (Smoluchowski-Kramers approximation) in the L^p-distances and in the total variation distance for the position process, the velocity process and a re-scaled velocity process to their corresponding limiting processes.

Keywords Smoluchowski-Kramers approximation · Stochastic differential by mean-field · Total variation distance · Malliavin calculus

Mathematics Subject Classiffication (2010) 60G22 · 60H07 · 91G30

1 Introduction

In this paper, we are interested in the following second-order mean-field stochastic differential equations

$$\begin{cases} dX_t^\alpha = Y_t^\alpha\,dt, \\ \frac{1}{\alpha}dY_t^\alpha = [-\kappa Y_t^\alpha - g(t, X_t^\alpha) - \gamma(Y_t^\alpha - \mathbb{E}(Y_t^\alpha))]\,dt + \sigma(t, X_t^\alpha)\,dW_t, \\ X_0^\alpha = x_0, Y_0^\alpha = y_0. \end{cases} \tag{1.1}$$

✉ Manh Hong Duong
H.Duong@bham.ac.uk

Ta Cong Son
congson82@gmail.com

Dung Quang Le
quangdung0110@gmail.com

1 University of Science, Vietnam National University, Hanoi, Vietnam

2 École Polytechnique, Palaiseau, France

3 University of Birmingham, Birmingham, UK

Here α, γ and κ are positive constants, $g(t,x) : [0,T] \times \mathbb{R} \to \mathbb{R}$ is a given function, $x_0, y_0 \in \mathbb{R}$ are given points in the real line, and $(W_t)_{t\geq 0}$ is the standard one-dimensional Wiener process. The notation $\mathbb{E}$ denotes the expectation with respect to the probability measure of the underlying probability space in which the Wiener process is defined.

System (1.1) describes the movement of a particle at position (displacement) $X_t^\alpha \in \mathbb{R}$ and with velocity $Y_t^\alpha \in \mathbb{R}$, at time t, under the influence of four different forces: an external, possibly time-dependent and non-potential, force $-g(t, X_t^\alpha)$; a friction $-\kappa Y_t^\alpha$; a (McKean-Vlasov type) mean-field interaction force $-\gamma(Y_t^\alpha - \mathbb{E}(Y_t^\alpha))$ (noting that here the mean-field term is acting on the velocity rather than the position) and a stochastic noise $\sigma(t, X_t^\alpha)\dot{W}_t$. Physically, α is the inverse of the mass, κ is the friction coefficient and γ is the strength of the interaction. We use the superscript α in (1.1) to emphasize the dependence on α since in the subsequent analysis we are concerned with the asymptotic behaviour of (1.1) as α tends to $+\infty$.

Under Assumptions 1.1 (see below) of this paper, system (1.1) can also be obtained as the mean-field (hydrodynamic) limit of the following interacting particle system as N tends to $+\infty$

$$\begin{cases} dX_t^{\alpha,i} = Y_t^{\alpha,i}\,dt, \\ dY_t^{\alpha,i} = [-\alpha\kappa Y_t^{\alpha,i} - \alpha g(t, X_t^{\alpha,i}) - \frac{\alpha\gamma}{N}\sum_{j=1}^N (Y_t^{\alpha,i} - Y_t^{\alpha,j})]\,dt + \alpha\sigma(t, X_t^{\alpha,i})\,dW_t^i, \\ X_0^\alpha = x_0, Y_0^\alpha = y_0, \end{cases} \tag{1.2}$$

where $\{W^i\}_{i=1}^N$ are independent one-dimensional Wiener processes. In fact, under Assumptions 1.1 the above interacting system satisfies the property of propagation of chaos, that is as N tends to infinity, it behaves more and more like a system of independent particles, in which each particle evolves according to (1.1) where the interaction term in (1.2) is replaced by the expectation one. For a detailed account on the propagation of chaos phenomenon, we refer the reader to classical papers [12, 25] and more recent papers [1, 8, 11] and references therein for degenerate diffusion systems like (1.1). The interacting particle system (1.2) and its mean-field limit (1.1) and more broadly systems of these types have been used extensively in biology, chemistry and statistical physics for the modelling of molecular dynamics, chemical reactions, flockings, social interactions, just to name a few, see for instance, the monographs [22, 23].

In this paper, we are interested in the zero-mass limit (as also known as the Smoluchowski-Kramers approximation) of (1.1), that is its asymptotic behaviour as α tends to $+\infty$. By employing techniques from Malliavin calculus, we obtain explicitly rate of convergences, in L^p-distances and in total variation distances, for both the position and velocity processes.

1.1 Main Results

Before stating our main results, we make the following assumptions.

Assumption 1.1 (A) The coefficients $g, \sigma : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ have linear growth, i.e. there exists $K > 0$ such that

$$|g(t,x)| + |\sigma(t,x)| \leq K(1 + |x|) \quad \forall x \in \mathbb{R}, t \in [0,T].$$

(B) The coefficients $g, \sigma : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ are Lipschitz, i.e. there exists $L > 0$ such that

$$|g(t,x) - g(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq L|x - y| \quad \forall x, y \in \mathbb{R}, t \in [0,T].$$

Assumption 1.2 $g(t,x), \sigma(t,x)$ are twice differentiable in x and the derivatives are bounded by some constant M > 0.

Let F, G be random variables, we denote by $d_{TV}(p_F, p_G)$ the total variation distance between the laws of F and G, that is,

$$\begin{aligned} d_{TV}(p_F, p_G) &= \sup_{A\in\mathcal{B}(\mathbb{R})} |P(F\in A) - P(G\in A)| \\ &= \frac{1}{2}\sup\{|\phi(F)-\phi(G)| : \phi : \mathbb{R}\to\mathbb{R} \text{ which is bounded by } 1\}. \end{aligned}$$

Consider the following first-order stochastic differential equation, which will be the limiting system for the displacement process

$$(\kappa+\gamma)dX_t = \Big[-g(t,X_t) - \frac{\gamma}{\kappa}\mathbb{E}[g(t,X_t)]\Big]dt + \sigma(t,X_t)dW_t, \quad X_0 = x_0 \in \mathbb{R}. \quad (1.3)$$

Our first main result provides an explicit rates of convergence for the displacement process.

Theorem 1.1 *(Quantitative rates of convergence of the displacement process) Under Assumptions 1.1 and 1.2, systems* (1.1) *and* (1.3) *have unique solutions and the following statements hold.*

1. *(rate of convergence in L^p-distances) For all $p\geq 2$, $\alpha\geq 1$ and $t\in[0,T]$,*
$$\mathbb{E}\Big[\sup_{0\leq s\leq t}|X_s^\alpha - X_s|^p\Big] \leq C\Big[(\lambda(t,\alpha(\kappa+\gamma)))^{\frac{p}{2}} + (\lambda(t,\alpha\kappa))^p\Big],$$
where $\lambda(t,a) = (1/a)[1-\exp(-at)]$ for $t, a > 0$ and C is a positive constant depending on $\{x_0, y_0, \kappa, \gamma, K, L, p, T\}$ but not on α and t.
2. *(rate of convergence in the total variation distance). We further assume that $|\sigma(t,x)|\geq \sigma_0 > 0$ for all $(t,x)\in[0,T]\times\mathbb{R}$. Then, for each $\alpha\geq 1$ and $t\in(0,T]$,*
$$d_{TV}(p_{X_t^\alpha}, p_{X_t}) \leq C\sqrt{t^{-1}(\lambda(t,\alpha(\kappa+\gamma)) + (\lambda(t,\alpha\kappa))^2)},$$
where $C>0$ is a constant depending only on $\{x_0, y_0, \sigma_0, \kappa, \gamma, K, L, M, T\}$ but not on α and t. As a corollary, if $|\sigma(t,x)|\geq\sigma_0>0$ for all $(t,x)\in[0,T]\times\mathbb{R}$ then we have.
$$d_{TV}(p_{X_t^\alpha}, p_{X_t}) \leq C\min\{t^{-1/2}\alpha^{-1/2}, 1/C\}.$$
Furthermore, the above rate of convergence in terms of α is sharp.

dummy

Theorem 1.1 combines Theorem 1.2 (for the L^p-distances) and Theorem 3.1 (for the total variation distance) in Section 3.1.

We are also interested in the asymptotic behavior, when $\alpha\to\infty$, of the velocity process Y_t^α of (1.1) and of a re-scaled velocity process, $\tilde{Y}_t^\alpha$, which is defined by

$$\tilde{Y}_t^\alpha := \frac{1}{\sqrt{\alpha}}Y_{t/\alpha}^\alpha.$$

The re-scaled process $\tilde{Y}_t^\alpha$ satisfies the following stochastic differential equation

$$\begin{cases} \tilde{Y}_t^\alpha = \frac{y_0}{\sqrt{\alpha}} - (\kappa+\gamma)\int_0^t \tilde{Y}_s^\alpha ds - \frac{1}{\sqrt{\alpha}}\int_0^t g(\frac{s}{\alpha}, X_{\frac{s}{\alpha}}^\alpha)ds - \gamma\int_0^t \mathbb{E}(\tilde{Y}_s^\alpha)ds + \int_0^t \sigma(\frac{s}{\alpha}, X_{\frac{s}{\alpha}}^\alpha)d\tilde{W}_s \\ X_0^\alpha = x_0, \end{cases}$$
(1.4)

where $\tilde{W}_t := \sqrt{\alpha} W_{t/\alpha}$ is a rescaled Brownian process.

Now we consider the following stochastic differential equation, which will be the limiting process of the rescaled velocity process

$$\begin{cases} d\tilde{Y}_t = -(\kappa+\gamma)d\tilde{Y}_t + \sigma(0, x_0)d\tilde{W}_t, \\ \tilde{Y}(0) = 0. \end{cases} \tag{1.5}$$

We now describe our result for the rescaled velocity process first since for this process we also work with a general setting where both g and σ can depend on both spatial and temporal variables. We only assume additionally the following condition.

Assumption 1.3

$$|\sigma(t, x) - \sigma(s, y)| \le L(|t-s| + |x-y|) \quad \forall x, y \in \mathbb{R}, t, s \in [0, T].$$

In the next theorem, we provide explicit rates of convergence, both in L^p-distances and in the total variation distance, for the rescaled velocity process.

Theorem 1.2 *(Quantitative rates of convergence for the rescaled velocity processes) Under Assumptions 1.1 and 1.3 the following hold.*

1. *(rate of convergence in L^p-distance for the rescaled velocity process) For all $p \ge 2$ and $\alpha \ge 1$,*
$$\mathbb{E}\left[\sup_{0\le t\le T} |\tilde{Y}_t^\alpha - \tilde{Y}_t|^p\right] \le \frac{C}{\alpha^{p/2}},$$
where C is a positive constant depending on p and other parameters but not on α.
2. *(rate of convergence in the total variation distance for the rescaled velocity process) Assume that Assumptions 1.1 and 1.3 hold and $|\sigma(0, x_0)| > 0$ for all $(t, x) \in [0, T] \times \mathbb{R}$. Then, for each $\alpha \ge 1$ and $t \in (0, T]$,*
$$d_{TV}(p_{\tilde{Y}_t^\alpha}, p_{\tilde{Y}_t}) \le C \min\{(\lambda(t, 2(\kappa+\gamma)))^{-1/2}\alpha^{-1/2}, 1/C\},$$
where $C > 0$ is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, p, T, \sigma(0, x_0)\}$.

Theorem 1.2 summarizes Theorem 3.3 (for the L^p-distances) and Theorem 3.4 (for the total variation distance) in Section 3.2.

When $g(t, x) = g(x)$ and $\sigma(t, x) = \delta$, [18, Theorem 2.3] shows that the velocity process Y_t^α converges to the normal distribution as $\alpha \to \infty$. The third aim of this paper is to generalize this result to a more general setting where g depends on both x and t while σ depends only on t, i.e. $\sigma(t, x) = \sigma(t)$, obtaining rates of convergence in the total variation distance. The following theorem is the content of Theorem 3.5 in Section 3.2.

Theorem 1.3 *(Quantitative rates of convergence for the velocity processes) Under Assumptions 1.1 the following hold. Assume additionally that $\sigma(t)$ is continuously differentiable on $[0, T]$ and that $\sigma(t) \neq 0$ for each $t \in (0, T]$. Let $N(t)$ be a normal random variable with mean 0 and variance $\dfrac{\sigma^2(t)}{2(\kappa+\gamma)}$, Then, for each $\alpha \ge 1$ and $t \in (0, T]$*

$$d_{TV}\left(p_{Y_t^\alpha/\sqrt{\alpha}}, p_{N(t)}\right) \le C \min\{(\lambda(t, 2(\kappa+\gamma)))^{-1/2}\alpha^{-1/2}, 1/C\},$$

where $C > 0$ is a constant not depending on α and t.

Theorem 1.3 is Theorem 3.5 in Section 3.2. We emphasize that in the main theorems, to obtain the existence and uniqueness as well as the rate of convergence in L^p-distances we only use Assumptions 1.1. Assumptions 1.2 and 1.3 are needed to employ techniques from Malliavin calculus, in particular to derive estimates for the Malliavin derivatives.

Corollary 1.1 *(Rate of convergence in Wasserstein distance for the laws of the displacement and velocity processes) Let μ and ν be two probability measures with finite second moments, then the p-Wasserstein distance, $W_p(\mu, \nu)$, between them can be defined by*

$$W_p(\mu, \nu) = \Big(\inf\{\mathbb{E}\big[|X - Y|^p\big] : X \sim \mu, Y \sim \nu\}\Big)^{1/p}.$$

Using this formulation, as a direct consequence of our main results, we also obtain explicit rates of convergence in p-Wasserstein distances for the laws of the displacement and the rescaled velocity processes to the corresponding limiting ones

$$\sup_{t\in[0,T]} W_p^p(\mathrm{law}(X_t^\alpha), \mathrm{law}(X_t)) \leq \frac{C}{\alpha^{p/2}},$$
$$\sup_{t\in[0,T]} W_p^p(\mathrm{law}(\tilde{Y}_t^\alpha), \mathrm{law}(\tilde{Y}_t)) \leq \frac{C}{\alpha^{p/2}}.$$

1.2 Comparison with Existing Literature

The zero-mass limit of second order differential equations has been studied intensively in the literature. In the seminal work [14], Kramers formally discusses this problem, in the context of applications to chemical reactions, for the classical underdamped Langevin dynamics, which corresponds to (1.1) with $g = -\nabla V$ (a gradient potential force), $\gamma = 0$ (no interaction force) and a constant diffusion coefficient. Due to this seminal work, this limit has become known in the literature as the Smoluchowski-Kramers approximation. Nelson rigorously shows that, under suitable rescaling, the solution to the Langevin equation converges almost surely to the solution of (3.14) with $\psi = 0$ [19]. Since then various generalizations and related results have been proved using different approaches such as stochastic methods, asymptotic expansions and variational techniques, see for instance [3, 5, 6, 9, 10, 16–18, 20]. The most relevant papers to the present one include [6, 16–18, 20]. The main novelty of the present paper lies in the fact that we consider interacting (mean-field) systems allowing time-dependent external forces and diffusion coefficients, and providing explicit rates of convergence in both L^p-distances and total variation distances for both displacement and velocity processes. Existing papers lack at least one of these features. More specifically,

Papers that Consider Mean-Field (Interaction) Systems The papers [6, 16–18] consider second order mean-field stochastic differential equations establishing the zero-mass limit, but they require much more stringent conditions that $g(t, x) = g(x)$ (time-independent force) and $\sigma(t, x) = \delta$ (constant diffusivity). On top of that, they do not provide a rate of convergence. Furthermore, our approach using Malliavin calculus is also different: Narita's papers use direct arguments while [6] employs variational methods based on Gamma-convergence and large deviation principle.

Papers that Provide a Rate of Convergence The papers [5, 20] provide a rate of convergence but only consider non-interacting systems (also using different measurements). Like our paper, [20] also utilizes techniques from Malliavin calculus, but [5] uses a completely different variational method. The recent paper [4], which studies the kinetic Vlasov-Fokker-Planck equation, is particularly interesting since it considers both interacting systems and provides a rate of convergence, but this paper is different to ours in a couple of aspects. First, the interaction force is acting on the position instead of the velocity; second, it works on the Fokker Planck equations and obtains a rate of convergence in Wasserstein distance while we work on the stochastic differential equations and obtain error quantifications in both L^p-distances and total variation distances; third, as mentioned, we use Malliavin calculus while [4] applied variational techniques like in [5, 6]. We also mention the paper [26], which provides similar rate of convergence to ours but it consider non mean-field stochastic differential equations driven by fractional Brownian motions.

1.3 Outlook

We provide further discussions on possible extensions of our results in the present paper.

Multi-dimensional Systems The analysis of the present paper is carried out only for one-dimensional processes. This is because in the proof of Lemma 2.1 below, we apply [26, Lemma 2.1] which is only applicable to one-dimensional random variables. Generalizing this lemma and our results to multi-dimensional processes would be an interesting problem. We will come back to this issue in a future work.

Non-Lipschitz and Singular Interactions The Lipschitz boundedness and differentiability Assumptions 1.1-1.2-1.3 are standard, but rather restricted since they do not cover some physically interesting interacting singular, such as Coulomb or Newton, forces. It would be interesting and challenging to extend our work to non-Lipschtizian and singular coefficients. We note that initial attempts in this direction for related models exist in the literature, see [2] for non-Lipschitzian coefficients and recent papers [4, 27] for singular forces.

Interacting Particle Systems Another interesting problem for future work is to study the Kramers-Smoluchowski approximation at the level of the N-particle system (1.2), which is a linear equation, and its relation to the mean-field model (1.1), which is nonlinear and nonlocal. In this context, one aims to obtain a rate of convergence that is independent or controllable in term of the number of particles N that enables one to pass to the limit $N \to \infty$.

1.4 Overview of the Proofs

To prove the main theorems for the general setting, with time-dependent coefficients, and obtain L^p-distances and total variations distances for the position and velocity processes, several technical improvements have been carried out.

On Existence and Uniqueness Under Assumptions 1.1, the existence and uniqueness, as well as the boundedness of the moments, of the second-order system (1.1) and the limiting first-order one (1.3) are standard results following Hölder's and the Burkholder-Davis-Gundy inequalities.

On Rate of Convergence in L^p-Distances Combining the mentioned inequalities and known estimates from [17] we can directly estimate $\mathbb{E}\Big[\sup_{0\le s\le t}|X_s^\alpha - X_s|^p\Big]$ and $\mathbb{E}\Big[\sup_{0\le t\le T}|\tilde{Y}_t^\alpha - \tilde{Y}_t|^p\Big]$ and obtain the rate of convergences in L^p-distances, proving parts (1) of both theorems.

On rate of Convergence in Total-Variation Distances The Malliavin differentiablity of the processes is followed from similar arguments as in [21]. Obtaining the rate of convergence in total variation distances is the most technically challenging. Lemma 2.1, which provides an upper bound estimate for the total variation between two random variables in terms of their Malliavin derivatives, is the key in our analysis. This lemma enables us to obtain the desired rates of convergence by estimating the corresponding quantities appearing in the right-hand side of Lemma 2.1.

1.5 Organization of the Paper

The rest of of the paper is organized as follows. In Section 2, we give an overview of some elements of Malliavin calculus and mean-field stochastic differential equations. The proofs of the main theorems are given in Section 3.

2 Preliminaries

In this section, we provide some basic and directly relevant knowledge on the Malliavin calculus and mean-field stochastic differential equations.

2.1 Malliavin Calculus

Let us recall some elements of stochastic calculus of variations (for more details see [21]). We suppose that $(W_t)_{t\in[0,T]}$ is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, where $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is a natural filtration generated by the Brownian motion W. For $h \in L^2[0,T] := \mathcal{H}$, we denote by $W(h)$ the Wiener integral

$$W(h) = \int_0^T h(t)dW_t.$$

Let $\mathcal{S}$ denote the dense subset of $L^2(\Omega, \mathcal{F}, P) := L^2(\Omega)$ consisting of smooth random variables of the form

$$F = f(W(h_1), ..., W(h_n)), \tag{2.1}$$

where $n \in \mathbb{N}$, $f \in C_b^\infty(\mathbb{R}^n)$, $h_1, ..., h_n \in \mathcal{H}$. If F has the form (2.1), we define its Malliavin derivative as the process $DF := \{D_t F, t \in [0,T]\}$ given by

$$D_t F = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(W(h_1), ..., W(h_n))h_k(t).$$

More generally, for each $k \geq 1$ we can define the iterated derivative operator on a cylindrical random variable by setting

$$D^k_{t_1,...,t_k} F = D_{t_1} ... D_{t_k} F.$$

For any $p, k \geq 1$, we shall denote by $\mathbb{D}^{k,p}$ the closure of $\mathcal{S}$ with respect to the norm

$$\|F\|^p_{k,p} := \mathbb{E}\big[|F|^p\big] + \mathbb{E}\Big[\int_0^T |D^1_{t_1} F|^p dt_1\Big] + ... + \mathbb{E}\Big[\int_0^T ... \int_0^T |D^k_{t_1,...,t_k} F|^p dt_1 ... dt_k\Big].$$

A random variable F is said to be Malliavin differentiable if it belongs to $\mathbb{D}^{1,2}$.

An important operator in the Malliavin's calculus theory is the divergence operator δ, which is the adjoint of the derivative operator D. The domain of δ is the set of all functions $u \in L^2(\Omega, \mathcal{H})$ such that

$$\mathbb{E}\big[|\langle DF, u\rangle_{\mathcal{H}}|\big] \leq C(u)\|F\|_{L^2(\Omega)},$$

where $C(u)$ is some positive constant depending on u. In particular, if $u \in \mathrm{Dom}(\delta)$, then $\delta(u)$ is characterized by the following duality relationship

$$\mathbb{E}\big[\langle DF, u\rangle_{\mathcal{H}}\big] = \mathbb{E}[F\delta(u)].$$

The following lemma provides an upper bound on the total variation distance between two random variables in terms of their Malliavin derivatives. This lemma will play an important role in the analysis of the present paper.

Lemma 2.1 *Let $F_1 \in \mathbb{D}^{2,2}$ be such that $\|DF_1\|_{\mathcal{H}} > 0$ a.s. Then, for any random variable $F_2 \in \mathbb{D}^{1,2}$ we have*

$$\begin{aligned} d_{TV}(p_{F_1}, p_{F_2}) \leq \|F_1 - F_2\|_{1,2}\Big[3\left(\mathbb{E}\|D^2F_1\|^4_{\mathcal{H}\otimes\mathcal{H}}\right)^{1/4}\left(\mathbb{E}\|DF_1\|^{-8}_{\mathcal{H}}\right)^{1/4} \\ +2\left(\mathbb{E}\|DF_1\|^{-2}_{\mathcal{H}}\right)^{1/2}\Big], \end{aligned} \tag{2.2}$$

provided that the expectations exist.

Proof From [26, Lemma 2.1] we have

$$d_{TV}(p_{F_1}, p_{F_2}) \leq \|F_1 - F_2\|_{1,2}\left[\left(\mathbb{E}\delta\left(\frac{DF_1}{\|DF_1\|^2_{\mathcal{H}}}\right)^2\right)^{1/2} + \left(\mathbb{E}\|DF_1\|^{-2}_{\mathcal{H}}\right)^{1/2}\right]. \tag{2.3}$$

Now using [21, Proposition 1.3.1], we get

$$\begin{aligned} \mathbb{E}\delta\left(\frac{DF_1}{\|DF_1\|^2_{\mathcal{H}}}\right)^2 &\leq \mathbb{E}\left\|\frac{DF_1}{\|DF_1\|^2_{\mathcal{H}}}\right\|^2_{\mathcal{H}} + \mathbb{E}\left\|D\left(\frac{DF_1}{\|DF_1\|^2_{\mathcal{H}}}\right)\right\|^2_{\mathcal{H}\otimes\mathcal{H}} \\ &= \mathbb{E}\,\|DF_1\|^{-2}_{\mathcal{H}} + \mathbb{E}\left\|D\left(\frac{DF_1}{\|DF_1\|^2_{\mathcal{H}}}\right)\right\|^2_{\mathcal{H}\otimes\mathcal{H}}. \end{aligned} \tag{2.4}$$

Moreover, observing that

$$D\left(\frac{DF_1}{\|DF_1\|^2_{\mathcal{H}}}\right) = \frac{D^2F_1}{\|DF_1\|^2_{\mathcal{H}}} - 2\frac{\langle D^2F_1, DF_1\otimes DF_1\rangle_{\mathcal{H}\otimes\mathcal{H}}}{\|DF_1\|^4_{\mathcal{H}}},$$

which implies that

$$\left\|D\left(\frac{DF_1}{\|DF_1\|^2_{\mathcal{H}}}\right)\right\|_{\mathcal{H}\otimes\mathcal{H}} \leq \frac{3\|D^2F_1\|_{\mathcal{H}\otimes\mathcal{H}}}{\|DF_1\|^2_{\mathcal{H}}}. \tag{2.5}$$

Substituting the inequality (2.5) into (2.4) and using Hölder's inequality, one can derive that

$$\mathbb{E}\delta\left(\frac{DF_1}{\|DF_1\|_{\mathcal{H}}^2}\right)^2 \leq \mathbb{E}\,\|DF_1\|_{\mathcal{H}}^{-2} + 9\mathbb{E}\left(\frac{\|D^2F_1\|_{\mathcal{H}\otimes\mathcal{H}}^2}{\|DF_1\|_{\mathcal{H}}^4}\right)$$
$$\leq \mathbb{E}\,\|DF_1\|_{\mathcal{H}}^{-2} + 9\left(\mathbb{E}\|D^2F_1\|_{\mathcal{H}\otimes\mathcal{H}}^4\right)^{1/2}\left(\mathbb{E}\|DF_1\|_{\mathcal{H}}^{-8}\right)^{1/2}.$$

Finally, substituting the above estimate back into (2.3) and using the fundamental inequality $(a+b)^{1/2} \leq a^{1/2} + b^{1/2}$ for all $a, b \geq 0$, we obtain (2.2), which completes the proof of this lemma. □

2.2 Mean-field Stochastic Differential Equations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with an increasing family $\{\mathcal{F}_t; t \geq 0\}$ of sub-σ-algebras of $\mathcal{F}$ and let $\{W_t; t \geq 0\}$ be a one-dimensional Brownian motion process adapted to $\mathcal{F}_t$.

The following lemma provides equivalent formulations of (1.1) and (1.3) as stochastic integral equations.

Lemma 2.2 *Equations* (1.1) *and* (1.3) *are, respectively, equivalent to the following equations*

$$X_t^\alpha = x_0 - \frac{1}{\kappa+\gamma}\int_0^t g(s, X_s^\alpha)\,ds - \frac{\gamma}{\kappa+\gamma}\int_0^t G^\alpha(s)\,ds + \frac{1}{\kappa+\gamma}\int_0^t \sigma(s, X_s^\alpha)dW_s + I_0^\alpha(t) + I_1^\alpha(t) - I_2^\alpha(t) - I_3^\alpha(t) + I_4^\alpha(t), \tag{2.6}$$

$$X_t = x_0 - \frac{1}{\kappa+\gamma}\int_0^t g(s, X_s)\,ds - \frac{\gamma}{\kappa+\gamma}\int_0^t G(s)\,ds + \frac{1}{\kappa+\gamma}\int_0^t \sigma(s, X_s)dW_s, \tag{2.7}$$

where $G^\alpha(t) = \mathbb{E}[g(t, X_t^\alpha)]/\kappa$, $G(t) = \mathbb{E}[g(t, X_t)]/\kappa$, *and*

$$I_0^\alpha(t) = y_0\lambda(t; \alpha(\kappa+\gamma)),$$
$$I_1^\alpha(t) = \frac{1}{\kappa+\gamma}\int_0^t \exp[\alpha(\kappa+\gamma)(u-t)]g(u, X_u^\alpha)du,$$
$$I_2^\alpha(t) = \frac{\gamma}{\kappa+\gamma}\int_0^t \exp[\alpha(\kappa+\gamma)(u-t)]n^\alpha(u)\,du,$$
$$I_3^\alpha(t) = \frac{1}{\kappa+\gamma}\int_0^t \exp[\alpha(\kappa+\gamma)(u-t)]\sigma(u, X_u^\alpha)dW_u,$$
$$I_4^\alpha(t) = \frac{\gamma}{\kappa+\gamma}\left(y_0\lambda(t;\alpha\kappa) + \int_0^t \exp[\alpha\kappa(u-t)]G^\alpha(u)du\right),$$

where $\lambda(t; a) = (1/a)(1 - \exp[-at])$, $n^\alpha(t) = E[Y_t^\alpha]$.

Proof Firstly, we can rewrite the second equation of (1.1) as follows

$$dY_t^\alpha = [-\alpha\kappa Y_t^\alpha - \alpha g(t, X_t^\alpha) - \alpha\gamma(Y_t^\alpha - n^\alpha(t))]dt + \alpha\sigma(t, X_t^\alpha)dW_t \text{ with } Y_0^\alpha = y_0.$$

Using Itô formula, we have the following expression

$$d(\exp[\alpha(\kappa+\gamma)t]Y_t^\alpha) = -\alpha\exp[\alpha(\kappa+\gamma)t]g(t, X_t^\alpha)dt + \alpha\gamma\exp[\alpha(\kappa+\gamma)t]n^\alpha(t)dt + \alpha\exp[\alpha(\kappa+\gamma)t]\sigma(t, X_t^\alpha)dW_t,$$

which implies

$$
\begin{aligned}
Y_t^\alpha &= y_0 \exp[-\alpha(\kappa+\gamma)t] - \alpha \int_0^t \exp[\alpha(\kappa+\gamma)(s-t)] g(s, X_s^\alpha) ds \\
&\quad + \alpha\gamma \int_0^t \exp[\alpha(\kappa+\gamma)(s-t)] n^\alpha(s)\, ds \\
&\quad + \alpha \int_0^t \exp[\alpha(\kappa+\gamma)(s-t)] \sigma(s, X_s^\alpha) dW_s. \qquad (2.8)
\end{aligned}
$$

Secondly, substituting this equation into the first equation of (1.1) we get

$$
\begin{aligned}
X_t^\alpha &= x_0 + y_0 \int_0^t \exp[-\alpha(\kappa+\gamma)s] ds - \alpha \int_0^t \int_0^s \exp[\alpha(\kappa+\gamma)(u-s)] g(u, X_u^\alpha) du ds \\
&\quad + \alpha\gamma \int_0^t \int_0^s \exp[\alpha(\kappa+\gamma)(u-s)] n^\alpha(u)\, du ds \\
&\quad + \alpha \int_0^t \int_0^s \exp[\alpha(\kappa+\gamma)(u-s)] \sigma(u, X_u^\alpha) dW_u ds.
\end{aligned}
$$

Now, we use integration by parts for the non-stochastic integral and Ito's product rule for the stochastic integral to get

$$
\begin{aligned}
X_t^\alpha &= x_0 - \frac{1}{\kappa+\gamma} \int_0^t g(u, X_u^\alpha)\, du + \frac{\gamma}{\kappa+\gamma} \int_0^t n^\alpha(u)\, du + \frac{1}{\kappa+\gamma} \int_0^t \sigma(s, X^\alpha) dW(s) \\
&\quad + I_0^\alpha(t) + I_1^\alpha(t) - I_2^\alpha(t) - I_3^\alpha(t), \qquad (2.9)
\end{aligned}
$$

where the terms $I_i^\alpha(t)$ $(i = 0, 1, 2, 3)$ are defined in the statement of the lemma. On the other hand, from the second equation of (1.1) we have

$$
\frac{d}{dt} n^\alpha(t) = -\alpha\kappa n^\alpha(t) - \alpha\kappa G^\alpha(t) \text{ with } n^\alpha(0) = y_0.
$$

This implies that

$$
n^\alpha(t) = y_0 \exp[-\alpha\kappa t] - \alpha\kappa \int_0^t \exp[\alpha\kappa(u-t)] G^\alpha(u) du. \qquad (2.10)
$$

Integrating this equation over the interval $[0, t]$ and changing the order of integration in the double integral, we get

$$
\int_0^t n^\alpha(s)\, ds = y_0 \lambda(t, \alpha\kappa) - \int_0^t G^\alpha(s) ds + \int_0^t \exp[\alpha\kappa(s-t)] G^\alpha(s) ds. \qquad (2.11)
$$

Substituting (2.11) back into (2.9) we obtain (2.6).

The proof for Eq. (2.7) is similar. □

The existence and uniqueness of solutions to (2.6) and (2.7) under Assumptions 1.1 is stated in the [15].

3 Proof of the Main Results

In this section, we present the proofs of the main Theorems 1.1, 1.2 and 1.3. We start with the displacement process (Theorems 3.1 and 3.2 give Theorem 1.1) in Section 3.1. Then in Section 3.2 we deal with the rescaled velocity process and the velocity process (Theorems 3.3 and 3.4 give Theorems 1.2 and 1.3 is Theorem 3.5).

3.1 Approximation of the Displacement Process

In this section, we give explicit bounds on L^p-distances and the total variation distance between the solution X_t^α of (1.1) and the solution X_t of (1.3). We will repeatedly use the following fundamental inequalities.

(i) Minkowski's inequality: for $p \geq 1$ and n real numbers $a_1, \dots, a_n$, we have

$$\Big|\sum_{i=1}^n a_i\Big|^p \leq n^{p-1} \sum_{i=1}^n |a_i|^p. \tag{3.1}$$

(ii) Hölder's inequality: for $p \geq 1$, $t > 0$ and measurable functions f we have

$$\Big(\int_0^t |f(s)|\, ds\Big)^p \leq t^{p-1} \int_0^t |f(s)|^p\, ds. \tag{3.2}$$

(iii) The Burkholder-Davis-Gundy (BDG) inequality for Brownian stochastic integrals, see for instance [24, Section 17.7]: for $0 < p < \infty$ and $f \in L^2([0,t], \Omega)$ we have

$$\mathbb{E}\left[\sup_{s\in[0,t]} \Big|\int_0^s f_r\, dW_r\Big|^p\right] \leq C_p \mathbb{E}\left[\Big(\int_0^t |f_s|^2\, ds\Big)^{p/2}\right], \tag{3.3}$$

where C_p is a positive constant depending only on p.

Applying the BDG inequality (3.3) to solutions of (2.6) and (2.7) we obtain

$$\mathbb{E}\left[\sup_{s\in[0,t]} \Big|\int_0^s \sigma(r, X_r^\alpha)\, dW_r\Big|^p\right] \leq C_p \mathbb{E}\left[\Big(\int_0^t |\sigma(s, X_s^\alpha)|^2\, ds\Big)^{p/2}\right], \tag{3.4}$$

$$\mathbb{E}\left[\sup_{s\in[0,t]} \Big|\int_0^s (\sigma(r,X_r^\alpha)-\sigma(r,X_r))\, dW_r\Big|^p\right] \leq C_p \mathbb{E}\left[\Big(\int_0^t |\sigma(s, X_s^\alpha)-\sigma(s,X_s)|^2\, ds\Big)^{p/2}\right]. \tag{3.5}$$

The next lemma provides important estimates on the moments of the displacement process $\{X_t^\alpha, t \in [0,T]\}$, which will be helpful to prove the main results of this section. Hereafter, we denote by C a generic constant which may vary at each appearance.

Lemma 3.1 *Let $\{X^\alpha(t), t \in [0,T]\}$ be the solution of (2.6) under Assumptions 1.1. Then, for all $p \geq 2$,*

$$\sup_{\alpha>0} \mathbb{E}\Big[\sup_{0\leq t\leq T} |X^\alpha(t)|^p\Big] \leq C, \tag{3.6}$$

and for all $0 \leq t \leq T$,

$$\sup_{\alpha>0} \mathbb{E}\Big[|X^\alpha(t) - x_0|^p\Big] \leq C t^{p/2}, \tag{3.7}$$

where C is a positive constant depending only on $\{x_0, y_0, \kappa, \gamma, K, p, T\}$.

Proof We first prove (3.6). We shall divide the proof into two steps.

Step 1: We evaluate the upper bound of the moments of each $I_i^\alpha(t)$, $i = 1, 2, 3, 4$.

From definition of $I_1^\alpha(t)$ and Assumptions 1.1 we have

$$\begin{aligned}\sup_{0\leq s\leq t} |I_1^\alpha(s)| &\leq \frac{K}{\kappa+\gamma} \sup_{0\leq s\leq t} \int_0^s \exp[\alpha(\kappa+\gamma)(u-s)](1+|X_u^\alpha|)du \\ &\leq \frac{Kt}{\kappa+\gamma} + \frac{K}{\kappa+\gamma} \int_0^t \sup_{0\leq u\leq s} |X_u^\alpha| ds.\end{aligned}$$

Now Minkowski's inequality (3.1) with $n = 2$ and Hölder's inequality (3.2) yield

$$\mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big) \le \frac{2^{p-1}K^p t^p}{(\kappa+\gamma)^p} + \frac{2^{p-1}K^p t^{p-1}}{(\kappa+\gamma)^p}\int_0^t \mathbb{E}\Big(\sup_{0\le u\le s}|X_u^\alpha|^p\Big)ds \tag{3.8}$$

For $I_2^\alpha(t)$, by substituting (2.10) into $I_2^\alpha(t)$ and changing the order of integration in the double integral, one can derive that

$$\begin{aligned} I_2^\alpha(t) = \frac{y_0\gamma\lambda(t,\alpha\gamma)}{\kappa+\gamma}\exp[-\alpha(\kappa+\gamma)t] - \frac{1}{\kappa+\gamma}\int_0^t \exp[\alpha\kappa(s-t)]E[g(s,X_s^\alpha)]ds \\ + \frac{1}{\kappa+\gamma}\int_0^t \exp[\alpha(\kappa+\gamma)(s-t)]E[g(s,X_s^\alpha)]ds. \end{aligned} \tag{3.9}$$

Using Assumptions 1.1, Minkowski's inequality (3.1) with $n = 2$ and Hölder's inequality (3.2) with noting that $\lambda(t,a) \le t$ for all $t \ge 0, a \ge 0$, we get

$$\begin{aligned} \sup_{0\le s\le t}|I_2^\alpha(s)|^p &\le \left(\frac{y_0\gamma t}{\kappa+\gamma} + \frac{2K}{\kappa+\gamma}\int_0^t (1+\mathbb{E}|X_s^\alpha|)ds\right)^p \\ &\le \frac{2^{p-1}\,(y_0\gamma+2K)^p\,t^p}{(\kappa+\gamma)^p} + \frac{2^{2p-1}K^p t^{p-1}}{(\kappa+\gamma)^p}\int_0^t \mathbb{E}(\sup_{0\le u\le s}|X_u^\alpha|^p)ds. \end{aligned} \tag{3.10}$$

For I_3^α, using the BDG inequality (3.4), Hölder's inequality (3.2) and Assumptions 1.1, we get

$$\begin{aligned} \mathbb{E}\Big(\sup_{0\le s\le t}|I_3^\alpha(s)|^p\Big) &\le \frac{C_p}{(\kappa+\gamma)^p}\mathbb{E}\left(\int_0^t \exp[2\alpha(\kappa+\gamma)(s-t)]|\sigma(s,X_s^\alpha)|^2 ds\right)^{p/2} \\ &\le \frac{2^{p-1}K^p t^{\frac{p-2}{2}}C_p}{(\kappa+\gamma)^p}\int_0^t \exp[p\alpha(\kappa+\gamma)(s-t)](1+\mathbb{E}|X_s^\alpha|^p)ds \\ &\le \frac{2^{p-1}K^p t^{\frac{p}{2}}C_p}{(\kappa+\gamma)^p} + \frac{2^{p-1}K^p t^{\frac{p-2}{2}}C_p}{(\kappa+\gamma)^p}\int_0^t \mathbb{E}\Big(\sup_{0\le u\le s}|X_u^\alpha|^p\Big)ds. \end{aligned} \tag{3.11}$$

Next, from Assumptions 1.1, Minkowski's inequality (3.1) with $n = 2$ and Hölder's inequality (3.2) with noting that $\lambda(t,a) \le t$ for all $t \ge 0, a \ge 0$, we get

$$\begin{aligned} \sup_{0\le s\le t}|I_4^\alpha(s)|^p &\le \frac{\gamma^p}{(\kappa+\gamma)^p}\left[y_0 t + \frac{K}{\kappa}\int_0^t\left(1+\mathbb{E}\Big(\sup_{0\le u\le s}|X_u^\alpha|\Big)ds\right)ds\right]^p \\ &\le \frac{2^{p-1}\gamma^p}{(\kappa+\gamma)^p}\left[\left(y_0+\frac{K}{\kappa}\right)^p t^p + \frac{K^p t^{p-1}}{\kappa^p}\int_0^t \mathbb{E}\Big(\sup_{0\le u\le s}|X_u^\alpha|^p\Big)ds\right]. \end{aligned} \tag{3.12}$$

Step 2: We estimate the integrand in the integrals in the right hand side of the above expressions. From (2.6), applying Minkowski's inequality with $n=9$, Hölder's inequality (3.2), the BDG inequality (3.4) as well as Assumptions 1.1, we obtain

$$\begin{aligned}
\mathbb{E}\Big(\sup_{0\le s\le t}|X_s^\alpha|^p\Big) &\le 9^{p-1}\Bigg[x_0^p+\frac{t^{p-1}}{(\kappa+\gamma)^p}\int_0^t \mathbb{E}|g(s,X_s^\alpha)|^p ds+\frac{\gamma^p t^{p-1}}{(\kappa+\gamma)^p}\int_0^t |G^\alpha(s)|^p\,ds\\
&\quad+\frac{C_p}{(\kappa+\gamma)^p}\mathbb{E}\left(\int_0^t|\sigma(s,X_s^\alpha)|^2ds\right)^{p/2}\\
&\quad+\sup_{0\le s\le t}|I_0^\alpha(s)|^p+\mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big)\\
&\quad+\sup_{0\le s\le t}|I_2^\alpha(s)|^p+\mathbb{E}\Big(\sup_{0\le s\le t}|I_3^\alpha(s)|^p\Big)+\sup_{0\le s\le t}|I_4^\alpha(s)|^p\Bigg]\\
&\le 9^{p-1}\Bigg[x_0^p+2^{p-1}K^p\left(\frac{t^{p-1}}{(\kappa+\gamma)^p}+\frac{\gamma^p t^{p-1}}{\kappa^p(\kappa+\gamma)^p}\right)\int_0^t(1+\mathbb{E}|X_s^\alpha|^p)ds\\
&\quad+\frac{2^{p-1}K^p t^{\frac{p-2}{2}}C_p}{(\kappa+\gamma)^p}\int_0^t(1+\mathbb{E}|X_s^\alpha|^p)ds+\sup_{0\le s\le t}|I_0^\alpha(s)|^p\\
&\quad+\mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big)\\
&\quad+\sup_{0\le s\le t}|I_2^\alpha(s)|^p+\mathbb{E}\Big(\sup_{0\le s\le t}|I_3^\alpha(s)|^p\Big)+\sup_{0\le s\le t}|I_4^\alpha(s)|^p\Bigg].
\end{aligned}$$

From this, together with (3.8), (3.10), (3.11) and (3.12), we deduce that

$$\mathbb{E}\Big(\sup_{0\le s\le t}|X_s^\alpha|^p\Big)\le C+C\int_0^t\mathbb{E}\Big(\sup_{0\le u\le s}|X_u^\alpha|^p\Big)ds, \tag{3.13}$$

where C is a positive constant depending on $\{x_0, y_0, \kappa, \gamma, K, p, T\}$. From (3.13), by applying Gronwall's lemma, we have

$$\sup_{\alpha>0}\mathbb{E}\Big[\sup_{0\le t\le T}|X^\alpha(t)|^p\Big]\le C,$$

which completes the proof of (3.6).

Next we prove (3.7). From expression (2.6), applying Minkowski's inequality (3.1) with $n=8$, Hölder's inequality (3.2), the BDG inequality (3.4) and Assumptions 1.1 again, we get

$$\begin{aligned}
\mathbb{E}|X_s^\alpha-x_0|^p &\le 8^{p-1}\Bigg[\frac{t^{p-1}}{(\kappa+\gamma)^p}\int_0^t\mathbb{E}|g(s,X_s^\alpha)|^p ds+\frac{\gamma^p t^{p-1}}{(\kappa+\gamma)^p}\int_0^t|G^\alpha(s)|^p\,ds\\
&\quad+\frac{C_p}{(\kappa+\gamma)^p}\mathbb{E}\left(\int_0^t|\sigma(s,X_s^\alpha)|^2ds\right)^{p/2}+\sup_{0\le s\le t}|I_0^\alpha(s)|^p\\
&\quad+\mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big)
\end{aligned}$$

$$+ \sup_{0\le s\le t} |I_2^\alpha(s)|^p + \mathbb{E}\Big(\sup_{0\le s\le t} |I_3^\alpha(s)|^p\Big) + \sup_{0\le s\le t} |I_4^\alpha(s)|^p\Bigg]$$

$$\le 8^{p-1}\Bigg[2^{p-1}K^p\left(\frac{t^{p-1}}{(\kappa+\gamma)^p} + \frac{\gamma^p t^{p-1}}{\kappa^p(\kappa+\gamma)^p}\right)\int_0^t (1+\mathbb{E}|X_s^\alpha|^p)ds$$
$$+\frac{2^{p-1}K^p t^{\frac{p-2}{2}} C_p}{(\kappa+\gamma)^p}\int_0^t (1+\mathbb{E}|X_s^\alpha|^p)ds + \sup_{0\le s\le t}|I_0^\alpha(s)|^p$$
$$+\mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big)$$
$$+ \sup_{0\le s\le t} |I_2^\alpha(s)|^p + \mathbb{E}\Big(\sup_{0\le s\le t} |I_3^\alpha(s)|^p\Big) + \sup_{0\le s\le t} |I_4^\alpha(s)|^p\Bigg].$$

This, together with (3.8), (3.10), (3.11), (3.12) and (3.6), we get (3.7). □

In the following theorem, we obtain a rate of convergence in L^p-distances in the Smoluchowski-Kramers approximation for the displacement process.

Theorem 3.1 *Let $\{X_t^\alpha, t\in[0,T]\}$ and $\{X_t, t\in[0,T]\}$ be respectively the solution of (2.6) and of (2.7) under Assumptions 1.1. Then, for all $p\ge 2$, $\alpha\ge 1$ and $t\in[0,T]$,*

$$\mathbb{E}\Big[\sup_{0\le s\le t}|X_s^\alpha - X_s|^p\Big] \le C\Big[(\lambda(t,\alpha(\kappa+\gamma)))^{\frac{p}{2}} + (\lambda(t,\alpha\kappa))^p\Big],$$

where C is a positive constant depending on $\{x_0, y_0, \kappa, \gamma, K, L, p, T\}$ but not on α and t.

Proof From (2.6) and (2.7), we have

$$X_t^\alpha - X_t = \frac{1}{\kappa+\gamma}\int_0^t (g(s,X_s^\alpha) - g(s,X_s))\,ds$$
$$-\frac{\gamma}{\kappa(\kappa+\gamma)}\int_0^t (\mathbb{E}[g(t,X_t^\alpha)] - \mathbb{E}[g(t,X_t)])ds$$
$$+\frac{1}{\kappa+\gamma}\int_0^t (\sigma(s,X_s^\alpha) - \sigma(s,X_s))dW(s) + I_0^\alpha(t) + I_1^\alpha(t) - I_2^\alpha(t)$$
$$-I_3^\alpha(t) + I_4^\alpha(t).$$

Similar to the proof of the previous lemma, by applying Minkowski's inequality (3.1) with $n=8$, Hölder's inequality (3.2), the BDG inequality (3.5) and Assumptions 1.1, we get

$$\mathbb{E}\Big[\sup_{0\le s\le t}|X_s^\alpha - X_s|^p\Big] \le 8^{p-1}\Bigg[\frac{t^{p-1}}{(\kappa+\gamma)^p}\int_0^t E|g(s,X_s^\alpha) - g(s,X_s)|^p ds$$
$$+\frac{\gamma^p t^{p-1}}{\kappa^p(\kappa+\gamma)^p}\int_0^t \Big|\mathbb{E}[g(t,X_t^\alpha)] - \mathbb{E}[g(t,X_t)]\Big|^p ds$$
$$+\frac{C_p}{(\kappa+\gamma)^p}\mathbb{E}\left(\int_0^t |\sigma(s,X_s^\alpha) - \sigma(s,X_s)|^2 ds\right)^{p/2}$$
$$+ \sup_{0\le s\le t}|I_0^\alpha(s)|^p + \mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big)$$
$$+ \sup_{0\le s\le t} |I_2^\alpha(s)|^p + \mathbb{E}\Big(\sup_{0\le s\le t} |I_3^\alpha(s)|^p\Big) + \sup_{0\le s\le t} |I_4^\alpha(s)|^p\Bigg]$$

$$
\begin{aligned}
&\le 8^{p-1}\Bigg[\frac{L^p t^{p-1}}{(\kappa+\gamma)^p}\int_0^t \mathbb{E}|X_s^\alpha - X_s|^p ds \\
&\quad + \frac{L^p\gamma^p t^{p-1}}{\kappa^p(\kappa+\gamma)^p}\int_0^t \mathbb{E}|X_s^\alpha - X_s|^p ds \\
&\quad + \frac{C_p L^p t^{\frac{p}{2}}}{(\kappa+\gamma)^p}\int_0^t \mathbb{E}|X_s^\alpha - X_s|^p ds + \sup_{0\le s\le t}|I_0^\alpha(s)|^p \\
&\quad + \mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big) \\
&\quad + \sup_{0\le s\le t}|I_2^\alpha(s)|^p + \mathbb{E}\Big(\sup_{0\le s\le t}|I_3^\alpha(s)|^p\Big) + \sup_{0\le s\le t}|I_4^\alpha(s)|^p\Bigg].
\end{aligned}
$$

Next we estimate the terms $\mathbb{E}\Big(\sup_{0\le s\le t}|I_i^\alpha(s)|^p\Big)$, $i = 1, 2, 3, 4$. We start with $\mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big)$. From definition of $I_1^\alpha(t)$ and Assumptions 1.1 and Minkowski's inequality (3.1) with $n = 2$ we obtain that

$$
\begin{aligned}
|I_1^\alpha(t)|^p &\le \frac{K^p}{(\kappa+\gamma)^p}\left(\int_0^t \exp[\alpha(\kappa+\gamma)(s-t)](1+|X_s^\alpha|)ds\right)^p \\
&\le \frac{2^p K^p}{(\kappa+\gamma)^p}\left(1+\sup_{0\le t\le T}|X_t^\alpha|^p\right)(\lambda(t;(\kappa+\gamma)\alpha))^p.
\end{aligned}
$$

This, together with the fact that the function $t \mapsto \lambda(t, a)$ is increasing and Lemma 3.1, implies

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big) &\le \frac{2^p K^p}{(\kappa+\gamma)^p}\left[1+\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^\alpha|^p\Big)\right](\lambda(t;(\kappa+\gamma)\alpha))^p \\
&\le C(\lambda(t;(\kappa+\gamma)\alpha))^p,
\end{aligned} \tag{3.14}
$$

where C is a positive constant depending on $\{x_0, y_0, \kappa, \gamma, K, p, T\}$. Next, using (3.9) and Lemma 3.1, we get

$$
\begin{aligned}
|I_2^\alpha(t)| &\le y_0\lambda(t;\alpha(\kappa+\gamma)) + \frac{K}{\kappa+\gamma}\int_0^t \exp[\alpha\kappa(s-t)][1+\mathbb{E}(|X_s^\alpha|)]\,ds \\
&\quad + \frac{K}{\kappa+\gamma}\int_0^t \exp[\alpha(\kappa+\gamma)(s-t)][1+\mathbb{E}(|X_s^\alpha|)]\,ds \\
&\le C\left[\lambda(t;\alpha(\kappa+\gamma) + \int_0^t (\exp[\alpha\kappa(s-t)] + \exp[\alpha(\kappa+\gamma)(s-t)])\,ds\right] \\
&\le C\left[\lambda(t;\alpha(\kappa+\gamma) + \lambda(t,\alpha\kappa)\right],
\end{aligned}
$$

where C is constant depending on $\{x_0, y_0, \kappa, \gamma, K, p\}$. Thus,

$$
\sup_{0\le s\le t}|I_2^\alpha(s)|^p \le C\left[(\lambda(t,\alpha(\kappa+\gamma)))^p + (\lambda(t,\alpha\kappa))^p\right]. \tag{3.15}
$$

Applying the BDG inequality (3.3), Hölder's inequality (3.2) and Lemma 3.1, one can derive that

$$\begin{aligned}\mathbb{E}\Big(\sup_{0\le s\le t}|I_3^\alpha(s)|^p\Big) &\le \frac{C_p}{(\kappa+\gamma)^p}\mathbb{E}\left(\int_0^t \exp\left[2\alpha(\kappa+\gamma)(s-t)\right]|\sigma(s,X_s^\alpha)|^2ds\right)^{p/2}\\ &\le \frac{C_p}{(\kappa+\gamma)^p}\left(\int_0^t \exp\left[\frac{p}{p-1}\alpha(\kappa+\gamma)(s-t)\right]ds\right)^{p/2-1}\\ &\quad\times\int_0^t \exp\left[\frac{p}{2}\alpha(\kappa+\gamma)(s-t)\right]\mathbb{E}|\sigma(s,X_s^\alpha)|^p\,ds\\ &\le \frac{2^{p-1}C_p}{(\kappa+\gamma)^p}\left(\int_0^t \exp\left[\frac{p}{p-1}\alpha(\kappa+\gamma)(s-t)\right]ds\right)^{p/2-1}\\ &\quad\times\int_0^t \exp\left[\frac{p}{2}\alpha(\kappa+\gamma)(s-t)\right](1+\mathbb{E}|X_s^\alpha|^p)ds\\ &\le C\left(\lambda(t,\frac{p}{p-1}\alpha(\kappa+\gamma))\right)^{p/2-1}\lambda(t,\frac{p}{2}\alpha(\kappa+\gamma)),\end{aligned}$$

where C is constant depending on $\{x_0, y_0, \kappa, \gamma, K, p, T\}$. On the other hand, for all $t > 0$ and $a > 0$, we have

$$\frac{\partial\lambda(t,a)}{\partial a} = \frac{(1+at)e^{-at}-1}{a^2} < 0.$$

Thus the function $a \mapsto \lambda(t, a)$ is decreasing. Hence we get

$$\mathbb{E}\Big(\sup_{0\le s\le t}|I_3^\alpha(s)|^p\Big) \le C\,(\lambda(t,\alpha(\kappa+\gamma)))^{p/2}.$$

Now we consider $\sup_{0\le s\le t}|I_4^\alpha(s)|^p$. Using Lemma 3.1 we can derive that

$$\sup_{0\le s\le t}|I_4^\alpha(s)| \le \frac{\gamma}{\kappa+\gamma}\left[y_0\lambda(t;\alpha\kappa)+\frac{K}{\kappa}\int_0^t \exp\left[\alpha(\kappa+\gamma)(u-t)\right]\left[1+\mathbb{E}(|X_u^\alpha|)\right]du\right].$$

Thus,

$$\sup_{0\le s\le t}|I_4^\alpha(s)|^p \le C\left[(\lambda(t,\alpha(\kappa+\gamma)))^p+(\lambda(t,\alpha\kappa))^p\right],$$

where C is constant depending on $\{x_0, y_0, \kappa, \gamma, K, L, p\}$. From the above estimates, together with the fact that $\lambda(t, \alpha(\kappa+\gamma)) \le t \le T$, one sees that

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\le s\le t}|X_s^\alpha - X_s|^p\Big] &\le C\left[t^{2p-1}\int_0^t \mathbb{E}|X_s^\alpha - X_s|^p ds + \sum_{i=0}^4 \mathbb{E}\Big(\sup_{0\le s\le t}|I_i^\alpha(s)|^p\Big)\right]\\ &\le C\left[\int_0^t \mathbb{E}\Big[\sup_{0\le u\le s}|X_u^\alpha - X_u|^p\Big]ds + (\lambda(t,\alpha(\kappa+\gamma)))^{\frac{p}{2}} + (\lambda(t,\alpha\kappa))^p\right],\end{aligned}$$

where C is constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, p\}$. Using Gronwall's inequality, we obtain the claimed estimate and complete the proof. □

Remark 3.1 1. Since $1-\exp(-ta) \le ta$ for all $t > 0$, $a > 0$, we have

$$\lambda(t,\alpha(\kappa+\gamma)) = \frac{1-\exp(-\alpha(\kappa+\gamma)t)}{\alpha(\kappa+\gamma)} \le t.$$

Therefore,

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\le s\le t}|X_s^\alpha - X_s|^p\Big] &\le C\left[(\lambda(t,\alpha(\kappa+\gamma)))^{\frac{p}{2}} + (\lambda(t,\alpha\kappa))^p\right]\\ &\le C\left[(\lambda(t,\alpha(\kappa+\gamma)))^{\frac{p}{2}} + (\lambda(t,\alpha\kappa))^p\right]\\ &\le C(t^{p/2}+t^p) \to 0 \text{ as } t\to 0.\end{aligned}$$

2. In general, when the drift and diffusion coefficients depend on both space and time variables, the constant C in Theorem 3.1 depends on the final time T. However, we now show that when g and σ do not depend on x, i.e. $g(t,x)=g(t), \sigma(t,x)=\sigma(t)$, then C does not depend on T. In fact, by Assumptions 1.1 we have $|g(t)|\le K$ and $|\sigma(t)|\le K$. By virtue of proof of Theorem 3.1 we obtain

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\le s\le t}|X_s^\alpha - X_s|^p\Big] \le 5^{p-1}\Big[&\sup_{0\le s\le t}|I_0^\alpha(s)|^p + \mathbb{E}\Big(\sup_{0\le s\le t}|I_1^\alpha(s)|^p\Big) + \sup_{0\le s\le t}|I_2^\alpha(s)|^p\\ &+\mathbb{E}\Big(\sup_{0\le s\le t}|I_3^\alpha(s)|^p\Big) + \sup_{0\le s\le t}|I_4^\alpha(s)|^p\Big]\\ \le C&\left[(\lambda(t,\alpha(\kappa+\gamma)))^{\frac{p}{2}} + (\lambda(t,\alpha\kappa))^p\right],\end{aligned}$$

where C is a positive constant depending on $\{x_0, y_0, \kappa, \gamma, K, p\}$ but not on α, t and T.

In the following lemma, we show Malliavin differentiability of X_t^α and X_t.

Lemma 3.2 *Under Assumptions 1.1, the solutions $\{X_t^\alpha, t\in[0,T]\}$ and $\{X_t, t\in[0,T]\}$ of (2.6) and (2.7) respectively are Malliavin differentiable random variables. Moreover, the derivatives $D_rX_t^\alpha$, D_rX_t satisfy $D_rX_t^\alpha = D_rX_t = 0$ for $r\ge t$ and for $0\le r<t\le T$,*

$$\begin{aligned}D_rX_t^\alpha = &\frac{\sigma(r,X_r^\alpha)}{\kappa+\gamma}(\exp[\alpha(\kappa+\gamma)(r-t)]-1) - \frac{1}{\kappa+\gamma}\int_r^t \bar{g}^\alpha(s)D_rX_s^\alpha ds\\ &+\frac{1}{\kappa+\gamma}\int_r^t \bar{\sigma}^\alpha(s)D_rX_s^\alpha\, dW_s + \frac{1}{\kappa+\gamma}\int_r^t \exp[\alpha(\kappa+\gamma)(s-t)]\bar{g}^\alpha(s)D_rX_s^\alpha ds\\ &+\frac{1}{\kappa+\gamma}\int_r^t \exp[\alpha(\kappa+\gamma)(s-t)]\bar{\sigma}^\alpha(s)D_rX_s^\alpha\, dW_s,\end{aligned}\tag{3.16}$$

$$D_rX_t = \frac{\sigma(r,X_r)}{\kappa+\gamma} - \frac{1}{\kappa+\gamma}\int_r^t \bar{g}(s)D_rX_s ds + \frac{1}{\kappa+\gamma}\int_r^t \bar{\sigma}(s)D_rX_s\, dW_s, \tag{3.17}$$

where $\bar{g}(s), \bar{g}^\alpha(s), \bar{\sigma}(s), \bar{\sigma}^\alpha(s)$ are adapted stochastic processes and are bounded by L, where L is the constant from Assumption 1.1.

Proof From the second equation of (1.1) we have

$$Y_t^\alpha = y_0 - \alpha(\kappa+\gamma)\int_0^t Y_s^\alpha ds - \alpha\int_0^t g(s,X_s^\alpha)ds - \alpha\gamma\int_0^t \mathbb{E}(Y_s^\alpha)\, ds + \alpha\int_0^t \sigma(s,X_s^\alpha)\, dW_s.$$

Using Minkowski's inequality (3.1) with $n = 5$, Hölder's inequality (3.2) with $p = 2$, the BDG inequality (3.3), Assumptions 1.1 and Lemma 3.1 we get

$$\begin{aligned}\mathbb{E}(\sup_{0\le s\le t}|Y_s^\alpha|^2) &\le 5\Big(y_0^2 + t\alpha^2(\kappa+\gamma)^2\int_0^t \mathbb{E}|Y_s^\alpha|^2 ds + t\alpha^2\int_0^t g^2(s, X_s^\alpha)\,ds \\ &\quad - t\alpha^2\gamma^2\int_0^t \mathbb{E}|Y_s^\alpha|^2\,ds + \alpha^2\int_0^t \sigma^2(s, X_s^\alpha)\,ds\Big) \\ &\le C\left(1 + \int_0^t \mathbb{E}(\sup_{0\le u\le s}|Y_u^\alpha|^2)\,ds + \int_0^t (1+\mathbb{E}|X_s^\alpha|^2)\,ds\right) \\ &\le C\left(1 + \int_0^t \mathbb{E}(\sup_{0\le u\le s}|Y_u^\alpha|^2)\,ds\right).\end{aligned}$$

Applying Gronwall's lemma we obtain

$$\mathbb{E}\Big[\sup_{0\le t\le T}|Y_t^\alpha|^2\Big] \le C.$$

This, together with Lemma 3.1 we deduce that $\mathbb{E}[\sup_{0\le t\le T}|Y_t^\alpha|]$ and $\mathbb{E}[\sup_{0\le t\le T}|X_t^\alpha|]$ are bounded. Then, by Assumptions 1.1 and the dominated convergence theorem, the integrals $I_0^\alpha(t)$, $I_2^\alpha(t)$ and $I_4^\alpha(t)$ are continuous functions.

Let us define

$$f(t) := x_0 - \frac{\gamma}{\kappa+\gamma}\int_0^t G^\alpha(u)\,du + I_0^\alpha(t) - I_2^\alpha(t) + I_4^\alpha(t).$$

Then $f(t)$ is continuous function in $[0, T]$ and Equation (2.6) becomes

$$\begin{aligned}X_t^\alpha = f(t) &+ \frac{1}{\kappa+\gamma}\int_0^t (\exp[\alpha(\kappa+\gamma)(s-t)] - 1)\, g(s, X_s^\alpha)\,ds \\ &+ \frac{1}{\kappa+\gamma}\int_0^t (\exp[\alpha(\kappa+\gamma)(s-t)] - 1)\sigma(s, X_s)\,dW_s.\end{aligned} \tag{3.18}$$

Now, we consider the Picard approximation sequence $\{X_t^{\alpha,n}, t \in [0, T]\}_{n\ge 0}$ given by

$$\begin{cases} X_t^{\alpha,0} = f(t), \\ X_t^{\alpha,n+1} = f(t) + \dfrac{1}{\kappa+\gamma}\displaystyle\int_0^t (\exp[\alpha(\kappa+\gamma)(s-t)] - 1)\, g(s, X_s^{\alpha,n})\,ds \\ \qquad + \dfrac{\delta}{\kappa+\gamma}\displaystyle\int_0^t (\exp[\alpha(\kappa+\gamma)(s-t)] - 1)\sigma(s, X_s^{\alpha,n})\,dW_s,\ t\in[0,T],\ n \ge 0. \end{cases}$$

From this, using the same method as in the proof of [21, Theorem 2.2.1], we conclude that the solution $\{X_t^\alpha, t \in [0, T]\}$ of (3.18) (thus, of (2.6)) is Malliavin's differentable. Obviously, the solution $\{X_t^\alpha, t \in [0, T]\}$ is $\mathbb{F}$-adapted. Hence, we always have $D_\theta X_t^\alpha = 0$ for $\theta > t$. For $\theta \le t$, from [21, Proposition 1.2.4] and Lipschitz property of g and σ, there exist adapted processes $\bar{g}^\alpha(s)$, $\bar{\sigma}^\alpha(s)$ uniformly bounded by L such that $D_\theta g(s, X_s^\alpha) = \bar{g}^\alpha(s) D_\theta X_s^\alpha$ and $D_\theta\sigma(s, X_s^\alpha) = \bar{\sigma}^\alpha(s) D_\theta X_s^\alpha$. Then we obtain (3.16) by applying the operator D to the equation (3.18).

The proof for the solution X_t of (2.7) is similar. □

Remark 3.2 If g and σ are continuously differentiable, then $\bar{g}(s) = g_2'(s, X_s)$, $\bar{g}^\alpha(s) = g_2'(s, X_s^\alpha)$, $\bar{\sigma}(s) = \sigma_2'(s, X_s)$ and $\bar{\sigma}^\alpha(s) = \sigma_2'(s, X_s^\alpha)$. Here, for a function $h(t,x)$, we use the convention $h_2'(t,x) = \frac{\partial h(t,x)}{\partial x}$.

In the next lemma, we show that the moments of the Malliavin's derivative of solutions of (2.7) are bounded.

Lemma 3.3 *Let $\{X_t, t \in [0,T]\}$ be the solution of (2.7) with Assumptions 1.1. Then, for all $p \geq 2$, we have*

$$\sup_{0 \leq t, r \leq T} \mathbb{E}(|D_r X_t|^p) < \infty.$$

Proof Using Minkowski's inequality (3.1) with $n = 3$, Hölder's inequality (3.2), the BDG inequality (3.3), Assumptions 1.1 and Lemma 3.1, and noting that $|\bar{g}(s)| \leq L$, $|\bar{\sigma}(s)| \leq L$, it follows from (3.17) that

$$\begin{aligned}
\mathbb{E}[|D_r X_t|^p] &\leq 3^{p-1}\left(\frac{\mathbb{E}[|\sigma(r,X_r)|^p]}{(\kappa+\gamma)^p} - \frac{1}{(\kappa+\gamma)^p}\mathbb{E}\left(\int_r^t \bar{g}(s) D_r X_s \, ds\right)^p\right. \\
&\quad \left. + \frac{1}{(\kappa+\gamma)^p}\mathbb{E}\left(\int_r^t |\bar{\sigma}(s) D_r X_s|^2 \, ds\right)^{p/2}\right) \\
&\leq 3^{p-1}\left(\frac{K^p(1+\mathbb{E}[|X_t|^p])}{(\kappa+\gamma)^p} - \frac{L^p(t-r)^{p-1}}{(\kappa+\gamma)^p}\int_r^t \mathbb{E}[|D_r X_s|^p] \, ds\right. \\
&\quad \left. + \frac{L^p(t-r)^{p/2-1}}{(\kappa+\gamma)^p}\int_r^t \mathbb{E}[|D_r X_s|^p] ds\right) \\
&\leq C\left[1 + \int_r^t |D_r X^\alpha(s)|^p \, ds\right],
\end{aligned}$$

where C is a positive constant depending only on $\{\kappa, \gamma, K, L, p, T\}$.

Taking the expectation and using Gronwall's inequality, we obtain the claimed estimate. □

The following lemma provides an upper bound for the difference between the derivatives of the solutions of (2.6) and (2.7).

Lemma 3.4 *Let $\{X_t^\alpha, t \in [0,T]\}$ and $\{X_t, t \in [0,T]\}$ be respectively the solution of (2.6) and of (2.7) under Assumptions 1.1 and 1.2. Then, for all $\alpha \geq 1$,*

$$\mathbb{E}\left[\|DX_t^\alpha - DX_t\|_{\mathcal{H}}^2\right] \leq C(\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2),$$

where C is a positive constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T\}$.

Proof Under Assumptions 1.2, g and σ are twice differentiable, thus (see Remark 3.2) $\bar{g}(s) = g_2'(s, X_s)$, $\bar{g}^\alpha(s) = g_2'(s, X_s^\alpha)$, $\bar{\sigma}(s) = \sigma_2'(s, X_s)$ and $\bar{\sigma}^\alpha(s) = \sigma_2'(s, X_s^\alpha)$. Furthermore,

$$|g_2'(s, X_s^\alpha) - g_2'(s, X_s)| \leq M|X_s^\alpha - X_s|, \quad |\sigma_2'(s, X_s^\alpha) - \sigma_2'(s, X_s| \leq M|X_s^\alpha - X_s|. \quad (3.19)$$

From (3.16) and (3.17) we have

$$
\begin{aligned}
D_r X_t^\alpha - D_r X_t = &\left(\frac{\sigma(r, X_r^\alpha)}{\kappa+\gamma}(\exp[\alpha(\kappa+\gamma)(r-t)]-1) - \frac{\sigma(r, X_r)}{\kappa+\gamma}\right)\\
&- \frac{1}{\kappa+\gamma}\int_r^t \left(g_2'(s, X_s^\alpha) D_r X_s^\alpha - g_2'(s, X_s) D_r X_s\right) ds\\
&+ \frac{1}{\kappa+\gamma}\int_r^t \left(\sigma_2'(s, X_s^\alpha) D_r X_s^\alpha - \sigma_2'(s, X_s) D_r X_s\right) dW_s \qquad (3.20)\\
&+ \frac{1}{\kappa+\gamma}\int_r^t \exp[\alpha(\kappa+\gamma)(s-t)] g_2'(s, X_s^\alpha) D_r X_s^\alpha\, ds\\
&+ \frac{1}{\kappa+\gamma}\int_r^t \exp[\alpha(\kappa+\gamma)(s-t)] \sigma_2'(s, X_s) D_r X_s^\alpha\, dW_s.
\end{aligned}
$$

Now, we shall estimate each term in the right hand side of (3.20). First, using Assumptions 1.1, Lemma 3.1 and Theorem 3.1 for $p = 2$, we can derive that

$$
\begin{aligned}
&\mathbb{E}\left(\frac{\sigma(r, X_r^\alpha)}{\kappa+\gamma}(\exp[\alpha(\kappa+\gamma)(r-t)]-1) - \frac{\sigma(r, X_r)}{\kappa+\gamma}\right)^2\\
&\le 2\left[\frac{\mathbb{E}\big[|\sigma(r, X_r^\alpha) - \sigma(r, X_r)|^2\big]}{(\kappa+\gamma)^2} + \frac{\mathbb{E}\big[|\sigma(r, X_r^\alpha)|^2\big]}{(\kappa+\gamma)^2}\exp[2\alpha(\kappa+\gamma)(r-t)]\right]\\
&\le 2\left[\frac{L^2\mathbb{E}|X_r^\alpha - X_r|^2}{(\kappa+\gamma)^2} + \frac{2K^2(1+\mathbb{E}[|X_r^\alpha|^2])}{(\kappa+\gamma)^2}\exp[2\alpha(\kappa+\gamma)(r-t)]\right]\\
&\le C\left[\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2 + \exp[2\alpha(\kappa+\gamma)(r-t)]\right],
\end{aligned}
$$

where C is constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T\}$. From Hölder's inequality, Assumptions 1.1, Lemma 3.1, Lemma 3.3 and Theorem 3.1 for $p = 4$, together with (3.19) we get

$$
\begin{aligned}
&\mathbb{E}\left(\int_r^t \left(g_2'(s, X_s^\alpha) D_r X_s^\alpha - g_2'(s, X_s) D_r X_s\right) ds\right)^2\\
&\le 2\mathbb{E}\left(\int_r^t \left(g_2'(s, X_s^\alpha) - g_2'(s, X_s)\right) D_r X_s\, ds\right)^2\\
&\quad + 2\mathbb{E}\left(\int_r^t g_2'(s, X_s^\alpha)\left(D_r X_s^\alpha - D_r X_s\right) ds\right)^2\\
&\le 2(t-r)\Bigg[\int_r^t \mathbb{E}|\left(g_2'(s, X_s^\alpha) - g_2'(s, X_s)\right) D_r X_s|^2\, ds\\
&\quad + \int_r^t \mathbb{E}\left|g_2'(s, X_s^\alpha)\left(D_r X_s^\alpha - D_r X_s\right)\right|^2 ds\Bigg]\\
&\le 2\mathrm{M}^2(t-r)\int_r^t \left(\mathbb{E}\left[|X_s^\alpha - X_s|^4\right]\right)^{1/2}\left(\mathbb{E}|D_r X_s|^4\right)^{1/2} ds\\
&\quad + 2L^2(t-r)\int_r^t \mathbb{E}\left[|D_r X_s^\alpha - D_r X_s|^2\right] ds\\
&\le C\left(\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2 + \int_r^t \mathbb{E}\left[|D_r X_s^\alpha - D_r X_s|^2\right] ds\right).
\end{aligned}
$$

By Itô's isometry formula, Hölder's inequality, Assumptions 1.1, Lemma 3.1, Lemma 3.3 and Theorem 3.1 for $p = 4$, together with (3.19), we have

$$
\begin{aligned}
&\mathbb{E}\left(\int_r^t \left(\sigma_2'(s, X_s^\alpha)D_r X_s^\alpha - \sigma_2'(s, X_s)D_r X_s\right) dW_s\right)^2\\
&\quad \le 2\mathbb{E}\left(\int_r^t \left(\sigma_2'(s, X_s^\alpha) - \sigma_2'(s, X_s)\right) D_r X_s\, dW_s\right)^2\\
&\qquad + 2\mathbb{E}\left(\int_r^t \sigma_2'(s, X_s^\alpha)\left(D_r X_s^\alpha - D_r X_s\right) dW_s\right)^2\\
&\quad \le 2\int_r^t \mathbb{E}|\left(\sigma_2'(s,X_s^\alpha)-\sigma_2'(s,X_s)\right) D_r X_s|^2\, ds + 2\int_r^t \mathbb{E}|\sigma_2'(s,X_s^\alpha)\left(D_r X_s^\alpha - D_r X_s\right)|^2\, ds\\
&\quad \le 2\mathrm{M}^2\int_r^t \left(\mathbb{E}|X_s^\alpha - X_s|^4\right)^{1/2}\left(\mathbb{E}|D_r X_s|^4\right)^{1/2} ds + 2L^2\int_r^t \mathbb{E}|D_r X_s^\alpha - D_r X_s|^2 ds\\
&\quad \le C\left[\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2 + \int_r^t \mathbb{E}\left[|D_r X_s^\alpha - D_r X_s|^2\right] ds\right].
\end{aligned}
$$

Next, using Hölder's inequality, Assumptions 1.1 and Lemma 3.3 one sees that

$$
\begin{aligned}
&\mathbb{E}\left(\int_r^t \exp\left[\alpha(\kappa+\gamma)(s-t)\right] g_2'(s, X_s^\alpha)D_r X_s^\alpha ds\right)^2\\
&\quad \le (t-r)L^2\int_r^t \exp\left[2\alpha(\kappa+\gamma)(s-t)\right]\mathbb{E}\left[|D_r X_s^\alpha|^2\right] ds\\
&\quad \le C\int_r^t \exp\left[2\alpha(\kappa+\gamma)(s-t)\right] ds\\
&\quad \le C\lambda(t, \alpha(\kappa+\gamma)).
\end{aligned}
$$

By the same estimate for the last term in the right hand side of (3.20), we can obtain

$$
\begin{aligned}
&\mathbb{E}\left(\int_r^t \exp\left[\alpha(\kappa+\gamma)(s-t)\right]\sigma_2'(s, X_s)D_r X_s^\alpha dW_s\right)^2\\
&\quad \le L^2\int_r^t \exp\left[2\alpha(\kappa+\gamma)(s-t)\right] E\left[|D_r X_s^\alpha|^2\right] ds\\
&\quad \le C\int_r^t \exp\left[2\alpha(\kappa+\gamma)(s-t)\right] ds\\
&\quad \le C\lambda(t, \alpha(\kappa+\gamma)).
\end{aligned}
$$

From the above estimates, together with the fact that the function $a \mapsto \lambda(t, a)$ is decreasing one can derive that

$$
\begin{aligned}
&\int_0^t \mathbb{E}\left[|D_r X_t^\alpha - D_r X_t|^2\right] dr\\
&\le C\int_0^t \left(\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2 + \exp\left[2\alpha(\kappa+\gamma)(r-t)\right] + \int_r^t \mathbb{E}\left[|D_r X_u^\alpha - D_r X_u|^2\right] du\right) dr\\
&\le C\left[\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2 + \lambda(t, 2\alpha(\kappa+\gamma)) + \int_0^t dr\int_r^t \mathbb{E}\left[|D_r X_u - D_r X_u^\alpha|^2\right] du\right]\\
&\le C\left[\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2 + \int_0^t du\int_0^u \mathbb{E}\left[|D_r X_u - D_r X_u^\alpha|^2\right] dr\right],
\end{aligned}
$$

where C is constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, \mathrm{M}, T\}$.

Let $\phi(t) := \int_0^t \mathbb{E}\left[|D_r X_t - D_r X_t^\alpha|^2\right] dr$, then we have

$$\phi(t) \le C\left[\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2 + \int_0^t \phi(u)\, du\right].$$

Thus, applying Gronwall's inequality, we get

$$\phi(t) \le C(\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2) \exp(Ct),$$

where C is constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, \mathrm{M}, T\}$. This completes the proof of the lemma. □

Now, we give explicit bounds on the total variation distance between the solution $X^\alpha(t)$ of (2.6) and the solution X_t of (2.7).

Theorem 3.2 *Let $\{X_t^\alpha, t \in [0, T]\}$ and $\{X_t, t \in [0, T]\}$ be, respectively, the solution of (2.6)and of (2.7) with Assumptions 1.1 and 1.2. We further assume that $|\sigma(t, x)| \ge \sigma_0 > 0$ for all $(t, x) \in [0, T] \times \mathbb{R}$. Then, for each $\alpha \ge 1$ and $t \in (0, T]$,*

$$d_{TV}(p_{X_t^\alpha}, p_{X_t}) \le C\sqrt{t^{-1}(\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2)},$$

where C is a constant depending only on $\{x_0, y_0, \sigma_0, \kappa, \gamma, K, L, M, T\}$.

Proof Lemma 2.1 gives

$$d_{TV}(p_{X_t^\alpha}, p_{X_t}) \le \|X_t^\alpha - X_t\|_{1,2}\left[3\left(\mathbb{E}\|D^2 X_t\|^4_{\mathcal{H}\otimes\mathcal{H}}\right)^{1/4}\left(\mathbb{E}\|DX_t\|_{\mathcal{H}}^{-8}\right)^{1/4} + 2\left(\mathbb{E}\|DX_t\|_{\mathcal{H}}^{-2}\right)^{1/2}\right].$$

Thanks to Theorem 3.1 and Lemma 3.4, we obtain

$$\begin{aligned} d_{TV}(p_{X_t^\alpha}, p_{X_t}) \le\ & C\sqrt{(\lambda(t, \alpha(\kappa+\gamma)) + (\lambda(t, \alpha\kappa))^2)} \\ & \times\left[3\left(\mathbb{E}\|D^2 X_t\|^4_{\mathcal{H}\otimes\mathcal{H}}\right)^{1/4}\left(\mathbb{E}\|DX_t\|_{\mathcal{H}}^{-8}\right)^{1/4} + 2\left(\mathbb{E}\|DX_t\|_{\mathcal{H}}^{-2}\right)^{1/2}\right], \end{aligned} \tag{3.21}$$

where C is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T\}$. To proceed, we will estimate the last two terms in (3.21). To this end, we will estimate $\mathbb{E}\|DX_t\|_{\mathcal{H}}^{-\gamma}$ for $\gamma > 0$ adopting the techniques used in the proof of Proposition 2.1.3 in [21]. From (3.17) and the fundamental inequality $(a + b + c)^2 \ge \frac{a^2}{2} - 2(b^2 + c^2)$, we have

$$\begin{aligned} |D_\theta X_t|^2 \ge\ & \frac{\sigma(\theta, X_\theta)}{2(\kappa+\gamma)^2} - \frac{2}{(\kappa+\gamma)^2}\left(\int_\theta^t g_2'(s, X_s) D_\theta X_s\, ds\right)^2 \\ & - \frac{2}{(\kappa+\gamma)^2}\left(\int_\theta^t \sigma_2'(s, X_s) D_\theta X_s dW_s\right)^2. \end{aligned}$$

For each $z \ge z_0 := \frac{4(\kappa+\gamma)^2}{\sigma_0^2 t}$, the real number $\varepsilon = \frac{4(\kappa+\gamma)^2}{z\sigma_0^2 t}$ belongs to (0, 1]. Hence,

$$\|DX_t\|_{\mathcal{H}}^2 \ge \int_{t(1-\varepsilon)}^t |D_\theta X_t|^2 d\theta \ge \frac{\sigma_0^2 t\varepsilon}{2(\kappa+\gamma)^2} - I_z(t) = \frac{2}{y} - I_z(t),$$

where $I_z(t)$ is given by

$$I_z(t) := \frac{2}{(\kappa+\gamma)^2}\int_{t(1-\varepsilon)}^t \left(\int_\theta^t g_2'(s,X_s)D_\theta X_s\,ds\right)^2 d\theta + \frac{2}{(\kappa+\gamma)^2}\int_{t(1-\varepsilon)}^t \left(\int_\theta^t \sigma_2'(s,X_s)D_\theta X_s dW_s\right)^2 d\theta.$$

By Markov's inequality, for all $p \geq 2$ we get

$$P\left(\|DX_t\|_{\mathcal{H}}^2 \leq \frac{1}{z}\right) \leq P\left(\frac{2}{z} - I_z(t) \leq \frac{1}{z}\right) = P\left(I_z(t) \geq \frac{1}{z}\right) \leq z^{p/2}\mathbb{E}\left(|I_z(t)|^{p/2}\right). \tag{3.22}$$

By using Minkowski's inequality (3.1) with $n = 2$, Hölder's inequality (3.2), the BDG inequality (3.3) and Assumptions 1.1, it follows from Lemma 3.3 that

$$\begin{aligned}
\mathbb{E}|I_z(t)|^{p/2} &\leq \frac{2^{p-1}}{(\kappa+\gamma)^p}\left(\mathbb{E}\left[\int_{t(1-\varepsilon)}^t \left(\int_\theta^t g_2'(s,X_s)D_\theta X_s\,ds\right)^2 d\theta\right]^{p/2}\right. \\
&\quad \left. + \mathbb{E}\left[\int_{t(1-\varepsilon)}^t \left(\int_\theta^t \sigma_2'(s,X_s)D_\theta X_s dW_s\right)^2 d\theta\right]^{p/2}\right) \\
&\leq \frac{2^{p-1}L^p(t\varepsilon)^{\frac{p-2}{2}}}{(\kappa+\gamma)^p}\left(\int_{t(1-\varepsilon)}^t \mathbb{E}\left(\int_\theta^t |D_\theta X_s|^2 ds\right)^{p/2} d\theta\right. \\
&\quad \left. + \int_{t(1-\varepsilon)}^t \mathbb{E}\left(\int_\theta^t |D_\theta X_s| dW_s\right)^p d\theta\right) \\
&\leq \frac{2^{p-1}L^p(t\varepsilon)^{\frac{p-2}{2}}}{(\kappa+\gamma)^p}\left(\int_{t(1-\varepsilon)}^t (t-\theta)^{\frac{p-2}{2}}\int_\theta^t \mathbb{E}|D_\theta X_s|^p ds d\theta\right. \\
&\quad \left. + \int_{t(1-\varepsilon)}^t \left(\int_\theta^t \mathbb{E}|D_\theta X_s|^2 ds\right)^{p/2} d\theta\right) \\
&\leq C(t\varepsilon)^{\frac{p-2}{2}}\left(\int_{t(1-\varepsilon)}^t (t-\theta)^{\frac{p}{2}} d\theta + \int_{t(1-\varepsilon)}^t (t-\theta)^{p/2}\,d\theta\right) \\
&= C(t\varepsilon)^{\frac{p-2}{2}}(t\varepsilon)^{\frac{p}{2}+1} = C\left(\frac{4(\kappa+\gamma)^2}{z\sigma_0^2}\right)^p.
\end{aligned} \tag{3.23}$$

Combining (3.22) and (3.23) we deduce that

$$P\left(\|DX_t\|_{\mathcal{H}}^2 \leq \frac{1}{z}\right) \leq Cz^{-p/2}\left(\frac{4(\kappa+\gamma)^2}{\|\sigma\|_0^2}\right)^p \quad \forall\ p \geq 2, z \geq z_0.$$

We recall here that $z_0 = \frac{4(\kappa+\gamma)^2}{t\sigma_0^2}$. For any $\gamma \geq 1$ and $p = 2\gamma - 1$, we have

$$
\begin{aligned}
\mathbb{E}\left(\|DX_t\|_{\mathcal{H}}^{-2\gamma}\right) &= \int_0^\infty \gamma z^{\gamma-1} P\left(\|DX_t\|_{\mathcal{H}}^{-2} > z\right) dz \\
&\leq \int_0^{z_0} \gamma z^{\gamma-1} dz + \int_{z_0}^\infty \gamma z^{\gamma-1} P\left(\|DX_t\|_{\mathcal{H}}^{2} < \frac{1}{z}\right) dz \\
&\leq z_0^\gamma + \gamma C \int_{z_0}^\infty z^{\gamma-1} z^{-p/2} \left(\frac{4(\kappa+\gamma)^2}{\sigma_0^2}\right)^p dz \\
&= z_0^\gamma + \gamma C \left(\frac{4(\kappa+\gamma)^2}{\sigma_0^2}\right)^p \frac{z_0^{\gamma-\frac{p}{2}}}{\gamma - \frac{p}{2}} \\
&\leq C\left(t^{-\gamma} + t^{\frac{p}{2}-\gamma}\right) = C\left(t^{-\gamma} + T^{-1/2}\left(\frac{T}{t}\right)^{1/2}\right) \\
&\leq C\left(t^{-\gamma} + T^{-1/2}\left(\frac{T}{t}\right)^{\gamma}\right) = Ct^{-\gamma}.
\end{aligned}
\tag{3.24}
$$

where C is a constant depending only on $\{x_0, y_0, \sigma_0, \kappa, \gamma, K, L, T\}$.

We continue to estimate the term $\mathbb{E}\|D^2X_t\|_{\mathcal{H}\otimes\mathcal{H}}^4$ appearing in the right-hand side of (3.21). From (3.16), for $\gamma, \theta \leq t$, under Assumptions 1.2, we get

$$
\begin{aligned}
D_\gamma D_\theta X_t = \sigma'(\theta, X_\theta) D_\gamma X_\theta + \sigma'(\gamma, X_\theta) D_\theta X_\gamma \\
&+ \int_{\theta\vee\gamma}^t \left(g''(s, X_s) D_\theta X_s D_\gamma X_s + g'(s, X_s) D_\gamma D_\theta X_s\right) ds \\
&+ \int_{\theta\vee\gamma}^t \left(\sigma''(s, X_s) D_\theta X_s D_\gamma X_s + \sigma'(s, X_s) D_\gamma D_\theta X_s\right) dW_s,
\end{aligned}
$$

Now, using Minkowski's inequality (3.1) with $n = 4$, Hölder's inequality (3.2), the BDG inequality (3.3), Assumptions 1.1 and 1.2, we can deduce

$$
\begin{aligned}
\mathbb{E}|D_\gamma D_\theta X_t|^4 \leq 64\Big[& L^4\mathbb{E}|D_\gamma X_\theta|^4 + L^4\mathbb{E}|D_\theta X_\gamma|^4 \\
&+ 8(t^3 + C_4 t) \int_{\theta\vee\gamma}^t \left(M^4(\mathbb{E}|D_\theta X_s|^2)^2(\mathbb{E}|D_\gamma X_s|^2)^2 + L^4\mathbb{E}|D_\gamma D_\theta X_s|^4\right) ds\Big].
\end{aligned}
$$

This, with together Lemma 3.3, gives us

$$
\mathbb{E}\left[|D_\gamma D_\theta X_t|^4\right] \leq C + C\int_{\theta\vee\gamma}^t \mathbb{E}\left[|D_\gamma D_\theta X_s|^4\right] ds,
$$

where C is a positive constant. By Gronwall's inequality, we can verify that

$$
\mathbb{E}\left[|D_\gamma D_\theta X_t|^4\right] \leq Ce^{C(t-\theta\vee\gamma)} \leq C \;\; \forall\, 0 \leq \theta, \gamma \leq t \leq T.
$$

Therefore,

$$
\mathbb{E}\|D^2X_t\|_{\mathcal{H}\otimes\mathcal{H}}^4 \leq t^2 \int_0^t \int_0^t \mathbb{E}\left[|D_\gamma D_\theta X_t|^4\right] d\theta\, d\gamma \leq t^2 \int_0^t \int_0^t C d\theta d\gamma = Ct^4, \tag{3.25}
$$

where C is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T\}$.

Combining (3.21), (3.24) and (3.25), we can conclude that

$$d_{TV}(p_{X_t^\alpha}, p_{X_t}) \le C\sqrt{(\lambda(t,\alpha(\kappa+\gamma)) + (\lambda(t,\alpha\kappa))^2)} \times \left[C + Ct^{-1/2}\right]$$
$$\le C\sqrt{t^{-1}(\lambda(t,\alpha(\kappa+\gamma)) + (\lambda(t,\alpha\kappa))^2)},$$

where C is a constant depending only on $\{x_0, y_0, \sigma_0, \kappa, \gamma, K, L, M, T\}$. This completes the proof. □

Similar to Remark 3.1, in the following Remark, we show that the constant C in Theorem 3.2 does not depend on the final time T when $g(t,x) = g(t),\ \sigma(t,x) = \sigma(t)$.

Remark 3.3 In the case $g(t,x) = g(t), \sigma(t,x) = \sigma(t)$ we see that

$$\mathbb{E}\left[\|DX_t^\alpha - DX_t\|_{\mathcal{H}}^2\right] = \frac{1}{(\kappa+\gamma)^2}\int_0^t \sigma^2(r)\exp\left[2\alpha(\kappa+\gamma)(r-t)\right]dr$$
$$\le \frac{K}{(\kappa+\gamma)^2}\lambda(t,\alpha(\kappa+\gamma)).$$

Moreover, if we assume that $\sigma(t) \ge \sigma_0$, then we get

$$d_{TV}(p_{X_t^\alpha}, p_{X_t}) \le C\sqrt{t^{-1}(\lambda(t,\alpha(\kappa+\gamma)) + (\lambda(t,\alpha\kappa))^2)},$$

where C is a positive constant depending on $\{x_0, y_0, \kappa, \gamma, K, p\}$ but not on α, t and T.

From Theorem 3.2, together with the fact that for all $t > 0$ and $a > 0$, $\lambda(t,a) < \frac{1}{a}$, we obtain the following Corollary, which provides an explicit estimate for $d_{TV}(p_{X_t^\alpha}, p_{X_t})$ in terms of α showing that it vanishes when α tends to $+\infty$.

Corollary 3.1 *Let $\{X_t^\alpha, t \in [0,T]\}$ and $\{X_t, t \in [0,T]\}$ be, respectively, the solution of (2.6) and of (2.7) with Assumptions 1.1 and 1.2. We further assume that $|\sigma(t,x)| \ge \sigma_0 > 0$ for all $(t,x) \in [0,T]\times\mathbb{R}$. Then, for each $\alpha \ge 1$ and $t \in (0,T]$,*

$$d_{TV}(p_{X_t^\alpha}, p_{X_t}) \le \min\{C\alpha^{-1/2}t^{-1/2}, 1\},$$

where C is a constant depending only on $\{x_0, y_0, \sigma_0, \kappa, \gamma, K, L, M, T\}$.

The following remark shows that the rate of convergence of X_t^α to X_t as α tends to $+\infty$ obtained in Theorem 3.2 and Corollary 3.1 is optimal.

Remark 3.4 We consider the equations (1.1) and (1.3) with $x_0 = y_0 = 0, b(t,x) = \sigma(t,x) = 1$. By Corollary 3.1 we have

$$d_{TV}(p_{X_t^\alpha}, p_{X_t}) \le Ct^{-1/2}\alpha^{-1/2} = O(\frac{1}{\sqrt{\alpha}}),\ \alpha\to\infty.$$

On the other hand, by solving (1.1) and (1.3) explicitly, we obtain

$$X_t^\alpha = -\frac{t}{\kappa+\gamma} - \frac{\gamma t}{(\kappa+\gamma)\kappa} + \frac{W_t}{\kappa+\gamma} - \frac{1}{\kappa+\gamma}\int_0^t \exp\left[\alpha(\kappa+\gamma)(u-t)\right]dW_u + \frac{1}{\kappa}\lambda(t,\alpha\kappa),$$

and

$$X_t = -\frac{t}{\kappa+\gamma} - \frac{\gamma t}{(\kappa+\gamma)\kappa} + \frac{W_t}{\kappa+\gamma}.$$

It follows that both X_t^α, X_t are normal random variables with means and variances given explicitly by

$$\mu := \mathbb{E}X_t = -\frac{t}{\kappa+\gamma} - \frac{\gamma t}{(\kappa+\gamma)\kappa}, \quad \sigma^2 := \mathrm{Var}X_t = \frac{t}{(\kappa+\gamma)^2}$$
$$\mu_\alpha = \mathbb{E}X_t^\alpha = -\frac{t}{\kappa+\gamma} - \frac{\gamma t}{(\kappa+\gamma)\kappa} + \frac{1}{\kappa}\lambda(t,\alpha\kappa),$$
$$\sigma_\alpha^2 := \mathrm{Var}(X_t^\alpha) = \frac{1}{(\kappa+\gamma)^2}\int_0^t \big(1-\exp[\alpha(\kappa+\gamma)(u-t)]\big)^2 du.$$

Then, applying Theorem 1.3 in [7], we get

$$\begin{aligned} d_{TV}(p_{X_t^\alpha}, p_{X_t}) &\geq \frac{1}{200}\min\left\{1, \max\left\{\frac{|\sigma_\alpha^2-\sigma^2|}{\min\{\sigma,\sigma_\alpha\}^2}, \frac{|\mu_\alpha-\mu|}{\min\{\sigma,\sigma_\alpha\}}\right\}\right\} \\ &\geq \frac{1}{200}\min\left\{1, \frac{|\mu_\alpha-\mu|}{\sigma_\alpha}\right\} \\ &= \frac{1}{200}\min\left\{1, \frac{(\kappa+\gamma)\lambda(t,\alpha\kappa)}{\kappa\left(\int_0^t\big(1-\exp[\alpha(\kappa+\gamma)(u-t)]\big)^2 du\right)^{1/2}}\right\} \\ &\geq \frac{1}{200}\min\left\{1, \frac{(\kappa+\gamma)\lambda(t,\alpha\kappa)}{\kappa\left(\int_0^t\big(1-\exp[\alpha(\kappa+\gamma)(u-t)]\big)du\right)^{1/2}}\right\} \\ &= \frac{1}{200}\min\left\{1, \frac{(\kappa+\gamma)\lambda(t,\alpha\kappa)}{\kappa\left(\lambda(t,\alpha(\kappa+\gamma))\right)^{1/2}}\right\} = O\left(\frac{1}{\sqrt{\alpha}}\right), \ \alpha\to\infty. \end{aligned}$$

Thus, in this simple example, we obtain an optimal rate of convergence of order $= O(\frac{1}{\sqrt{\alpha}})$ for $d_{TV}(p_{X_t^\alpha}, p_{X_t})$.

3.2 Approximation the Velocity and Rescaled Velocity Processes

In this section, we establish rates of convergence in L^p-distances and in the total variation distance for the velocity and rescaled velocity processes. We will discuss the re-scaled velocity process first since in this case, our results are applicable to more general settings where both external forces and diffusion coefficients can be dependent on both x and t, i.e. $g = g(t,x)$ and $\sigma = \sigma(t,x)$.

3.2.1 The Re-scaled Velocity Process

From the second equation of (1.1) we can see that the process $Y_{\frac{t}{\alpha}}^\alpha$ satisfies

$$\begin{cases} Y_{\frac{t}{\alpha}}^\alpha = y_0 - (\kappa+\gamma)\int_0^t Y_{\frac{s}{\alpha}}^\alpha ds - \int_0^t g(\frac{s}{\alpha}, X_{\frac{s}{\alpha}}^\alpha)ds - \gamma\int_0^t \mathbb{E}(Y_{\frac{s}{\alpha}}^\alpha)ds + \alpha\int_0^t \sigma(\frac{s}{\alpha}, X_{\frac{s}{\alpha}}^\alpha)dW_{\frac{s}{\alpha}} \\ X_0^\alpha = x_0. \end{cases} \tag{3.26}$$

We recall the definition of the re-scaled velocity process introduced in the Introduction

$$\tilde{Y}_t^\alpha = \frac{1}{\sqrt{\alpha}} Y_{\frac{t}{\alpha}}^\alpha.$$

Then $\tilde{Y}_t^\alpha$ satisfies (1.4), that is

$$\begin{cases} \tilde{Y}_t^\alpha = \dfrac{y_0}{\sqrt{\alpha}} - (\kappa+\gamma)\int_0^t \tilde{Y}_s^\alpha ds - \dfrac{1}{\sqrt{\alpha}}\int_0^t g(\frac{s}{\alpha}, X^\alpha_{\frac{s}{\alpha}})ds - \gamma\int_0^t \mathbb{E}(\tilde{Y}_s^\alpha)ds \\ \qquad + \sqrt{\alpha}\int_0^t \sigma(\frac{s}{\alpha}, X^\alpha_{\frac{s}{\alpha}})dW_{\frac{s}{\alpha}} \\ X^\alpha(0) = x_0. \end{cases} \tag{3.27}$$

Now, we put $\tilde{W}_t = \sqrt{\alpha}W_{t/\alpha}$, then $(\tilde{W}_t)_{t\geq 0}$ is a Brownian motion process and (3.27) can be rewritten in the form

$$\begin{cases} \tilde{Y}_t^\alpha = \dfrac{y_0}{\sqrt{\alpha}} - (\kappa+\gamma)\int_0^t \tilde{Y}_s^\alpha ds - \dfrac{1}{\sqrt{\alpha}}\int_0^t g(\frac{s}{\alpha}, X^\alpha_{\frac{s}{\alpha}})ds - \gamma\int_0^t \mathbb{E}(\tilde{Y}_s^\alpha)ds \\ \qquad + \int_0^t \sigma(\frac{s}{\alpha}, X^\alpha_{\frac{s}{\alpha}})d\tilde{W}_s \\ X_0^\alpha = x_0. \end{cases} \tag{3.28}$$

Our goal in this section is to study the rate of convergence in L^p-distance and in the total variation distance between $\tilde{Y}_t^\alpha$ and $\tilde{Y}_t$. Here, $\tilde{Y}_t$ is the solution of Ornstein-Uhlembeck process (1.5), which is

$$\begin{cases} d\tilde{Y}_t = -(\kappa+\gamma)d\tilde{Y}_t + \sigma(0, x_0)d\tilde{W}_t, \\ \tilde{Y}(0) = 0. \end{cases} \tag{3.29}$$

First, we obtain the rate of convergence in L^p-distances between $\tilde{Y}_t^\alpha$ and $\tilde{Y}_t$ in the following lemma.

Theorem 3.3 *Let $\{\tilde{Y}_t^\alpha, t\in[0,T]\}$ and $\{\tilde{Y}_t, t\in[0,T]\}$ be, respectively, the solution of* (3.28) *and of* (3.29) *with Assumptions 1.1 and 1.3. Then, for all $p\geq 2$ and $\alpha \geq 1$,*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |\tilde{Y}_t^\alpha - \tilde{Y}_t|^p\right] \leq \frac{C}{\alpha^{p/2}},$$

where C is a positive constant depending on p but not on α.

Proof From (3.28) and (3.29), together with the fact that $\mathbb{E}[\tilde{Y}_t] = 0$ for all $t\in[0,T]$, we get

$$\begin{aligned} \tilde{Y}_t^\alpha - \tilde{Y}_t = \frac{y_0}{\sqrt{\alpha}} - (\kappa+\gamma)\int_0^t (\tilde{Y}_s^\alpha - \tilde{Y}_s)ds - \frac{1}{\sqrt{\alpha}}\int_0^t g(\frac{s}{\alpha}, X^\alpha_{\frac{s}{\alpha}})ds \\ - \gamma\int_0^t \mathbb{E}(\tilde{Y}_s^\alpha - \tilde{Y}_s)ds + \int_0^t (\sigma(\frac{s}{\alpha}, X^\alpha_{\frac{s}{\alpha}}) - \sigma(0,x_0))d\tilde{W}_s. \end{aligned}$$

Using Minkowski's inequality (3.1) with $n=5$, Hölder's inequality, the BDG inequality and Assumptions 1.1 and 1.3, one can derive that

$$\begin{aligned} \mathbb{E}\left[\sup_{0\leq s\leq t} |\tilde{Y}_s^\alpha - \tilde{Y}_s|^p\right] \leq 5^{p-1}\Big[\frac{|y_0|^p}{\alpha^{p/2}} + t^{p-1}(\kappa+\gamma)^p\int_0^t \mathbb{E}\big[|\tilde{Y}_s^\alpha - \tilde{Y}_s|^p\big]ds \\ + \frac{K^p(2t)^{p-1}}{\alpha^{p/2}}\int_0^t (1+\mathbb{E}\big[|X^\alpha_{\frac{s}{\alpha}}|^p\big])\,ds \\ + t^{p-1}\gamma^p\int_0^t \mathbb{E}\big[|\tilde{Y}_s^\alpha - \tilde{Y}_s|^p\big]ds \\ + 2^{p-1}C_p t^{p/2-1}\int_0^t (\frac{s^p}{\alpha^p} + \mathbb{E}\big[|X^\alpha_{\frac{s}{\alpha}} - x_0|^p)\big]ds\Big]. \end{aligned}$$

By Lemma 3.1, with noting that $0 \le \frac{s}{\alpha} \le s \le t \le T$, we have

$$\mathbb{E}\left[\sup_{0\le s\le t} |\tilde{Y}^\alpha_s - \tilde{Y}_s|^p\right] \le \frac{C}{\alpha^{p/2}} + C\int_0^t \mathbb{E}\left[|\tilde{Y}^\alpha_s - \tilde{Y}_s|^p\right] ds + C\int_0^t (\frac{s^p}{\alpha^p} + \frac{s^{p/2}}{\alpha^{p/2}})ds$$
$$\le \frac{C}{\alpha^{p/2}} + C\int_0^t \mathbb{E}\left[\sup_{0\le u\le s} |\tilde{Y}^\alpha_u - \tilde{Y}_u|^p\right] ds,$$

where C is constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, p\}$. Using Growwall's inequality, we obtain the claimed inequality and complete the proof. □

From (3.27) and (3.29), under Assumptions 1.1 the Malliavin differentiability of the solutions $\tilde{Y}^\alpha_t$ and $\tilde{Y}_t$ can be proved by using the same method as in the proof of Lemma 3.2. Moreover, the Malliavin derivatives $D_\theta \tilde{Y}^\alpha_t$ and $D_\theta \tilde{Y}_t$ satisfy $D_r\tilde{Y}^\alpha_t = D_r\tilde{Y}_t = 0$ for $r \ge t/\alpha$ and $0 \le \alpha r < t \le T$,

$$D_r\tilde{Y}^\alpha_t = \sqrt{\alpha}\sigma(r, X^\alpha_r) - (\kappa+\gamma)\int_{\alpha r}^t D_r\tilde{Y}^\alpha_s ds - \frac{1}{\sqrt{\alpha}}\int_{\alpha r}^t \bar{g}(\frac{s}{\alpha})D_r X^\alpha_{\frac{s}{\alpha}} ds$$
$$+\sqrt{\alpha}\int_{\alpha r}^t \bar{\sigma}(\frac{s}{\alpha})D_r X^\alpha_{\frac{s}{\alpha}} dW_{\frac{s}{\alpha}},$$
$$D_r\tilde{Y}_t = \sqrt{\alpha}\sigma(0, x_0) - (\kappa+\gamma)\int_{\alpha r}^t D_r\tilde{Y}_s ds, \tag{3.30}$$

where $\bar{g}(s), \bar{g}^\alpha(s), \bar{\sigma}(s), \bar{\sigma}^\alpha(s)$ are adapted stochastic processes and bounded by L. Furthermore, if g and σ are continuously differentiable, then $\bar{g}(s) = g_2'(s, X_s)$, $\bar{g}^\alpha(s) = g_2'(s, X^\alpha_s)$, $\bar{\sigma}(s) = \sigma_2'(s, X_s)$ and $\bar{\sigma}^\alpha(s) = \sigma_2'(s, X^\alpha_s)$.

Lemma 3.5 *Let $\{\tilde{Y}^\alpha_t, t \in [0, T]\}$ and $\{\tilde{Y}_t, t \in [0, T]\}$ be defined as above. Assume that Assumptions 1.1 and 1.3 hold. Then, for all $\alpha \ge 1$,*

$$\mathbb{E}\left[\|D\tilde{Y}^\alpha_t - D\tilde{Y}_t\|^2_{\mathcal{H}}\right] \le C\alpha^{-1},$$

where C is constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T\}$.

Proof It follows from (3.30) that, for $0 \le \alpha r < t$,

$$D\tilde{Y}^\alpha_t - D\tilde{Y}_t = \sqrt{\alpha}\left(\sigma(r, X^\alpha_r) - \sigma(0, x_0)\right) - (\kappa+\gamma)\int_{\alpha r}^t \left[D\tilde{Y}^\alpha_s - D\tilde{Y}_s\right] ds$$
$$- \frac{1}{\sqrt{\alpha}}\int_{\alpha r}^t \bar{g}(\frac{s}{\alpha})D_r X^\alpha_{\frac{s}{\alpha}} ds + \sqrt{\alpha}\int_{\alpha r}^t \bar{\sigma}(\frac{s}{\alpha})D_r X^\alpha_{\frac{s}{\alpha}} dW_{\frac{s}{\alpha}}.$$

Using Cauchy-Schwarz inequality, the Itô isometry formula, Assumptions 1.1 and 1.3, Lemma 3.1, with noting that $0 \le \frac{s}{\alpha} \le s \le t \le T$, we get

$$\mathbb{E}\left[|D_r\tilde{Y}^\alpha_t - D_r\tilde{Y}_t|^2\right] \le 4\Big(2L^2\alpha\left(r^2 + \mathbb{E}\left[|X^\alpha_r - x_0|^2\right]\right)$$
$$+ (\kappa+\gamma)^2(t-\alpha r)\int_{\alpha r}^t \mathbb{E}\Big[\left|D_r\tilde{Y}^\alpha_s - D_r\tilde{Y}_s\right|^2\Big] ds$$
$$+ \frac{L^2T}{\alpha}\int_{\alpha r}^t \mathbb{E}\big[|D_r X^\alpha_{\frac{s}{\alpha}}|^2\big] ds + L^2\alpha\int_{\alpha r}^t \mathbb{E}\big[|D_r X^\alpha_{\frac{s}{\alpha}}|^2\big]\frac{ds}{\alpha}\Big)$$
$$\le C\Big(\alpha\left(r^2+r\right) + \int_{\alpha r}^t \mathbb{E}\Big[\left|D_r\tilde{Y}^\alpha_s - D_r\tilde{Y}_s\right|^2\Big] ds + \frac{(t-\alpha r)}{\alpha} + (t-\alpha r)\Big),$$

where C is constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T\}$. From this, we have

$$\int_0^{t/\alpha} \mathbb{E}\big[|D_r\tilde{Y}_t^\alpha - D_r\tilde{Y}_t|^2\big]dr$$
$$\leq C\Big(\alpha\left(\frac{t^3}{3\alpha^3}+\frac{t^2}{2\alpha^2}\right)+\int_0^{t/\alpha}\left(\int_{\alpha r}^t \mathbb{E}\Big[\big|D_r\tilde{Y}_s^\alpha - D_r\tilde{Y}_s\big|^2\Big]ds\right)dr+\frac{t^2}{2\alpha^2}+\frac{t^2}{2\alpha}\Big)$$
$$\leq \frac{C}{\alpha}+C\int_0^t\left(\int_0^{s/\alpha}\mathbb{E}\left|D_r\tilde{Y}_s^\alpha - D_r\tilde{Y}_s\right|^2 dr\right)ds.$$

Denote $\phi(t) = \int_0^{t/\alpha} E|D_r\tilde{Y}_t^\alpha - D_r\tilde{Y}_t|^2 dr$, using Gronwall's inequality, one sees easily that

$$\mathbb{E}\Big[\|D\tilde{Y}_t^\alpha - D\tilde{Y}_t\|_{\mathcal{H}}^2\Big] = \phi(t) \leq \frac{C}{\alpha}e^{Ct} \leq C\alpha^{-1},$$

where C is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T\}$. This completes our proof.□

Bringing the above lemmas together, we can get the following result.

Theorem 3.4 *Let $\{\tilde{Y}_t^\alpha, t \in [0,T]\}$ and $\{\tilde{Y}_t, t \in [0,T]\}$ be as before. Assume that Assumptions 1.1 and 1.3 hold and $|\sigma(0,x_0)| > 0$ for all $(t,x) \in [0,T]\times\mathbb{R}$. Then, for each $t \in (0,T]$,*

$$d_{TV}(p_{\tilde{Y}_t^\alpha}, p_{\tilde{Y}_t}) \leq C\,(\lambda(t, 2(\kappa+\gamma)))^{-1/2}\,\alpha^{-1/2},$$

where C is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T, \sigma(0,x_0)\}$.

Proof Using Lemma 2.1 we have

$$d_{TV}(p_{\tilde{Y}_t^\alpha}, p_{\tilde{Y}_t}) \leq \|\tilde{Y}_t^\alpha - \tilde{Y}_t\|_{1,2}\left[3\left(\mathbb{E}\|D^2\tilde{Y}_t\|^4_{\mathcal{H}\otimes\mathcal{H}}\right)^{1/4}\left(\mathbb{E}\|D\tilde{Y}_t\|_{\mathcal{H}}^{-8}\right)^{1/4}+2\left(\mathbb{E}\|D\tilde{Y}_t\|_{\mathcal{H}}^{-2}\right)^{1/2}\right].$$

Thanks to Theorem 3.3 and Lemma 3.5, we obtain

$$d_{TV}(p_{\tilde{Y}_t^\alpha}, p_{\tilde{Y}_t}) \leq C\alpha^{-1/2}\left[3\left(\mathbb{E}\|D^2\tilde{Y}_t\|^4_{\mathcal{H}\otimes\mathcal{H}}\right)^{1/4}\left(\mathbb{E}\|D\tilde{Y}_t\|_{\mathcal{H}}^{-8}\right)^{1/4}+2\left(\mathbb{E}\|D\tilde{Y}_t\|_{\mathcal{H}}^{-2}\right)^{1/2}\right]. \tag{3.31}$$

Moreover, we have $D_r\tilde{Y}_t = 0$ for $r \geq t/\alpha$ and $D_r\tilde{Y}_t = \sqrt{\alpha}\sigma(0,x_0) - (\kappa+\gamma)\int_{\alpha r}^t D_r\tilde{Y}_s ds$, for $0 \leq \alpha r < t$. Solving this ODE directly yields

$$D_r\tilde{Y}_t = \begin{cases} 0 & \text{if } r\alpha \geq t \\ \sqrt{\alpha}\sigma(0,x_0)\exp\left[(\kappa+\gamma)(r\alpha - t)\right] & \text{if } r\alpha < t. \end{cases} \tag{3.32}$$

Thus, one can easily show that

$$\|D\tilde{Y}_t\|_{\mathcal{H}}^2 = \alpha\sigma^2(0,x_0)\int_0^{t/\alpha}\exp\left[2(\kappa+\gamma)(r\alpha - t)\right]dr = \sigma^2(0,x_0)\lambda(t, 2(\kappa+\gamma)).$$

This implies, for all $p \geq 2$,

$$\mathbb{E}\big[\|D\tilde{Y}_t\|_{\mathcal{H}}^{-p}\big] = \sigma^{-p}(0,x_0)\,(\lambda(t, 2(\kappa+\gamma)))^{-p/2}. \tag{3.33}$$

Applying the Malliavin derivative to (3.32), we have

$$D_\theta D_r\tilde{Y}_t = 0. \tag{3.34}$$

Combining (3.31), (3.33) and (3.34), we obtain

$$d_{TV}(p_{\tilde{Y}_t^\alpha}, p_{\tilde{Y}_t}) \le C\,(\lambda(t, 2(\kappa+\gamma)))^{-1/2}\,\alpha^{-1/2},$$

where C is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T, \sigma(0, x_0)\}$. This completes our proof. □

3.2.2 The Velocity Process

As mentioned in the introduction, when $g(t, x) = g(x)$ and $\sigma(t, x) = \delta$, [18, Theorem 2.3] shows that the velocity process Y_t^α converges to the normal distribution as $\alpha \to \infty$. In the rest of this section, we generalize this result to a much more general setting where g depends on both x and t while σ depends only on t, i.e. $\sigma(t, x) = \sigma(t)$.

From (2.8) we get

$$Y_t^\alpha = y_0 \exp[-\alpha(\kappa+\gamma)t] - \alpha(\kappa+\gamma)I_1^\alpha(t) + \alpha(\kappa+\gamma)I_2^\alpha(t) + \alpha(\kappa+\gamma)I_3^\alpha(t). \tag{3.35}$$

Since X_t^α is Malliavin differentiable, Y_t^α is also Malliavin differentiable satisfying $D_r Y_t^\alpha = 0$ for $r \ge t$, and for $0 \le r < t \le T$

$$D_r Y_t^\alpha = \alpha\sigma(r)\exp[\alpha(\kappa+\gamma)(r-t)] - \alpha\int_r^t \exp[\alpha(\kappa+\gamma)(t-s)]g'(s, X_s^\alpha)D_r X_s^\alpha\,ds. \tag{3.36}$$

Define

$$W^\alpha(t) = \sqrt{\alpha}(\kappa+\gamma)I_3^\alpha(t) = \sqrt{\alpha}\int_0^t \exp[\alpha(\kappa+\gamma)(u-t)]\sigma(u)dW_u.$$

Then $W^\alpha(t)$ is also Malliavin differentiable and $D_r W_t^\alpha = 0$ for $r \ge t$ and for $0 \le r < t \le T$

$$D_r W^\alpha(t) = \sqrt{\alpha}\sigma(r)\exp[\alpha(\kappa+\gamma)(r-t)]. \tag{3.37}$$

Lemma 3.6 *Let Y_t^α be the solution of (3.26) with Assumptions 1.1. Then, for all $p \ge 2$ and $t \in (0, T]$,*

$$\mathbb{E}\left[\left|\frac{Y_t^\alpha}{\sqrt{\alpha}} - W^\alpha(t)\right|^p\right] \le C\alpha^{-p/2},$$

where C is constant depending only on $\{y_0, \kappa, \gamma, K, L, p\}$.

Proof From (3.35), we get

$$\frac{Y_t^\alpha}{\sqrt{\alpha}} - W^\alpha(t) = \frac{y_0}{\sqrt{\alpha}}\exp[-\alpha(\kappa+\gamma)t] - \sqrt{\alpha}(\kappa+\gamma)I_1^\alpha(t) + \sqrt{\alpha}(\kappa+\gamma)I_2^\alpha(t).$$

Then, by (3.14), (3.15) and Lemma 3.3, we have the following estimation

$$\begin{aligned}
\mathbb{E}\left[\left|\frac{Y_t^\alpha}{\sqrt{\alpha}} - W^\alpha(t)\right|^p\right] &\le 3^{p-1}\left[\frac{|y_0|^p}{\alpha^{p/2}} + \alpha^{p/2}(\kappa+\gamma)^p\mathbb{E}\big(|I_1^\alpha(t)|^p\big) + \alpha^{p/2}(\kappa+\gamma)^p|I_2^\alpha(t)|^p\right] \\
&\le C\left[\frac{1}{\alpha^{p/2}} + \alpha^{p/2}\left(2+\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^\alpha|^p\Big)\right)(\lambda(t;(\kappa+\gamma)\alpha))^p + \alpha^{p/2}(\lambda(t,\alpha\kappa))^p\right] \\
&\le \frac{C}{\alpha^{p/2}}\left[1+\mathbb{E}\Big(\sup_{0\le t\le T}|X_t^\alpha|^p\Big)\right] \\
&\le \frac{C}{\alpha^{p/2}},
\end{aligned}$$

where C is a constant depending only on $\{y_0, \kappa, \gamma, K, L\}$. This completes the proof of the lemma.

Lemma 3.7 *Let Y_t^α be the solution of (3.26) with Assumptions 1.1. Assume that $\sigma(t)$ is continuous on $[0, T]$. Then, for all $\alpha \ge 1$,*

$$\mathbb{E}\int_0^t \left|\frac{D_r Y_t^\alpha}{\sqrt{\alpha}} - D_r W^\alpha(t)\right|^p dr \le C\|\sigma\|_\infty^p \alpha^{-p/2},$$

where $\|\sigma\|_\infty = \sup_{t\in[0,T]}|\sigma(t)|$ and C is constant depending only on $\{\kappa, \gamma, K, L, p, T\}$.

Proof From (3.36) and (3.37) we have

$$\begin{aligned}
\mathbb{E}\int_0^t \left|\frac{D_r Y_t^\alpha}{\sqrt{\alpha}} - D_r W^\alpha(t)\right|^p dr &= \mathbb{E}\int_0^t \left|\sqrt{\alpha}\int_r^t \exp[\alpha(\kappa+\gamma)(t-s)]g'(s, X_s^\alpha)D_r X_s^\alpha ds\right|^p dr \\
&\le C\alpha^{p/2}\int_0^t \mathbb{E}\left|\int_r^t \exp[\alpha(\kappa+\gamma)(t-s)]|D_r X_s^\alpha|ds\right|^p dr.
\end{aligned} \tag{3.38}$$

On the other hand, from (3.16), together with Assumptions 1.1 and the fact that $\bar{\sigma}^\alpha(s) = 0$, we have

$$\begin{aligned}
|D_r X_t^\alpha| &\le \frac{|\sigma(r)|}{\kappa+\gamma}(1-\exp[\alpha(\kappa+\gamma)(r-t)]) + \frac{1}{\kappa+\gamma}\int_r^t |\bar{g}^\alpha(s)||D_r X_s^\alpha|ds \\
&\quad + \frac{1}{\kappa+\gamma}\int_r^t \exp[\alpha(\kappa+\gamma)(s-t)]|\bar{b}^\alpha(s)||D_r X_s^\alpha|ds \\
&\le \frac{\|\sigma\|_\infty}{\kappa+\gamma} + \frac{2L}{\kappa+\gamma}\int_r^t |D_r X_s^\alpha|\,ds.
\end{aligned}$$

Thus, by Gronwall's inequality one sees that

$$|D_r X_t^\alpha| \le \frac{\|\sigma\|_\infty}{\kappa+\gamma}e^{\frac{2L(t-r)}{\kappa+\gamma}} \le C\|\sigma\|_\infty,$$

where C is constant depending only on $\{\kappa, \gamma, K, L, T\}$. Substituting this into (3.38) yields

$$\begin{aligned}
\mathbb{E}\int_0^t \left|\frac{D_r Y_t^\alpha}{\sqrt{\alpha}} - D_r W^\alpha(t)\right|^p dr &\le C\|\sigma\|_\infty^p\alpha^{p/2}\int_0^t\left|\int_r^t \exp[\alpha(\kappa+\gamma)(t-s)]ds\right|^p dr \\
&\le C\|\sigma\|_\infty^p\alpha^{p/2}\int_0^t \lambda^p(t-r;\alpha(\kappa+\gamma))dr \\
&\le C\|\sigma\|_\infty^p\alpha^{-p/2},
\end{aligned}$$

which is the desired conclusion. □

Now we are ready to prove the rate of convergence in total variation distance for the velocity process Y_t^α as $\alpha \to \infty$.

Theorem 3.5 *Let Y_t^α be the solution of (3.26) with Assumptions 1.1. Assume that $\sigma(t)$ is a continuously differentiable function on $[0, T]$ and that $\sigma(t) \neq 0$ for each $t \in (0, T]$. Then, for each $\alpha \geq 1$ and $t \in (0, T]$*

$$d_{TV}\left(p_{Y_t^\alpha/\sqrt{\alpha}}, p_{N(t)}\right) \leq C\left(\lambda(t, 2(\kappa+\gamma))\right)^{-1/2}\alpha^{-1/2},$$

where N is a normal random variable with mean 0 and variance $\dfrac{\sigma^2(t)}{2(\kappa+\gamma)}$, and C is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T, \sigma\}$.

Proof Using Lemma 2.1, we have:

$$d_{TV}\left(p_{Y_t^\alpha/\sqrt{\alpha}}, p_{W^\alpha(t)}\right) \leq \left\| \frac{Y_t^\alpha}{\sqrt{\alpha}} - W^\alpha(t) \right\|_{1,2}\Bigg[3\left(\mathbb{E}\|D^2W^\alpha(t)\|^4_{\mathcal{H}\otimes\mathcal{H}}\right)^{1/4}\left(\mathbb{E}\|DW^\alpha(t)\|^{-8}_{\mathcal{H}}\right)^{1/4} + 2\left(\mathbb{E}\|DW^\alpha(t)\|^{-2}_{\mathcal{H}}\right)^{1/2}\Bigg].$$

By Lemmas 3.6 and 3.7, one can derive that

$$d_{TV}\left(p_{Y_t^\alpha/\sqrt{\alpha}}, p_{W^\alpha(t)}\right) \leq C\alpha^{-1/2}\Bigg[3\left(\mathbb{E}\|D^2W^\alpha(t)\|^4_{\mathcal{H}\otimes\mathcal{H}}\right)^{1/4}\left(\mathbb{E}\|DW^\alpha(t)\|^{-8}_{\mathcal{H}}\right)^{1/4} + 2\left(\mathbb{E}\|DW^\alpha(t)\|^{-2}_{\mathcal{H}}\right)^{1/2}\Bigg], \tag{3.39}$$

where C is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T\}$.

Now we calculate the derivatives of $W^\alpha(t)$. One can easily show that

$$D_rW^\alpha(t) = \sqrt{\alpha}\sigma(r)\exp\left[\alpha(\kappa+\gamma)(r-t)\right] \text{ and } D_\theta D_r W^\alpha(t) = 0. \tag{3.40}$$

Therefore,

$$\begin{aligned}\|DW^\alpha(t)\|^2_{\mathcal{H}} &= \int_0^t \sigma^2(r)\alpha\exp\left[2\alpha(\kappa+\gamma)(r-t)\right]dr \\ &\geq \alpha\lambda(t; 2\alpha(\kappa+\gamma))\min_{t\in[0,T]}|\sigma(t)|^2 \\ &\geq \lambda(t; 2(\kappa+\gamma))\min_{t\in[0,T]}|\sigma(t)|^2.\end{aligned}$$

Thus, for all $p \geq 2$,

$$\mathbb{E}\left[\|D\tilde{Y}_t\|^{-p}_{\mathcal{H}}\right] \leq \frac{1}{\lambda^{p/2}(t; 2(\kappa+\gamma))\min_{t\in[0,T]}|\sigma(t)|^p}. \tag{3.41}$$

From (3.39), (3.40), (3.41), we obtain

$$d_{TV}\left(p_{Y_t^\alpha/\sqrt{\alpha}}, p_{W^\alpha(t)}\right) \leq C\alpha^{-1/2}\left(\frac{2}{(\lambda(t; 2(\kappa+\gamma)))^{1/2}\min_{t\in[0,T]}|\sigma(t)|}\right),$$

where C is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T, \sigma\}$. Thus,

$$d_{TV}\left(p_{Y_t^\alpha/\sqrt{\alpha}}, p_{W^\alpha(t)}\right) \le C\,(\lambda(t, 2(\kappa+\gamma)))^{-1/2}\,\alpha^{-1/2},$$

where C is a constant depending only on $\{x_0, y_0, \kappa, \gamma, K, L, T, \sigma\}$.

Note that by Itô's isometry and using integration by parts for the non-stochastic integral, we have

$$\begin{aligned}\mathbb{E}\big[\,[W^\alpha(t)]^2\,\big] &= \int_0^t \alpha\sigma^2(r)\exp[2\alpha(\kappa+\gamma)(r-t)]dr\\ &= \frac{\sigma^2(t)}{2(\kappa+\gamma)} - \frac{\sigma^2(0)\exp[-2\alpha(\kappa+\gamma)t]}{2(\kappa+\gamma)}\\ &\quad - \int_0^t \sigma(r)\sigma'(r)\exp[2\alpha(\kappa+\gamma)(r-t)]dr.\end{aligned}$$

Thus, we can deduce that $W^\alpha(t)$ is random variable with normal distribution with mean 0 and variance

$$\frac{\sigma^2(t)}{2(\kappa+\gamma)} - \frac{\sigma^2(0)\exp[-2\alpha(\kappa+\gamma)t]}{2(\kappa+\gamma)} - \int_0^t \sigma(r)\sigma'(r)\exp[2\alpha(\kappa+\gamma)(r-t)]dr.$$

Now, applying Lemma 4.9, [13], we derive that

$$\begin{aligned}d_{TV}(p_{W^\alpha(t)}, p_{N(t)}) &\le C\left(\frac{\sigma^2(0)\exp[-2\alpha(\kappa+\gamma)t]}{2(\kappa+\gamma)} + \int_0^t |\sigma(r)\sigma'(r)|\exp[2\alpha(\kappa+\gamma)(r-t)]dr\right)\\ &\le C\left(\frac{\sigma^2(0)\exp[-2\alpha(\kappa+\gamma)t]}{2(\kappa+\gamma)} + \|\sigma\|_\infty\|\sigma'\|_\infty\lambda(t, 2\alpha(\kappa+\gamma))\right)\\ &\le C\|\sigma\|_\infty\|\sigma'\|_\infty\alpha^{-1},\end{aligned}$$

where $\|\sigma\|_\infty = \sup_{t\in[0,T]}|\sigma(t)|$, $\|\sigma'\|_\infty = \sup_{t\in[0,T]}|\sigma'(t)|$ and C is an universal constant. Thus,

$$\begin{aligned}d_{TV}\left(p_{Y_t^\alpha/\sqrt{\alpha}}, p_{N(t)}\right) &\le d_{TV}(p_{W^\alpha(t)}, p_{N(t)}) + d_{TV}\left(p_{Y_t^\alpha/\sqrt{\alpha}}, p_{W^\alpha(t)}\right)\\ &\le C\,(\lambda(t, 2(\kappa+\gamma)))^{-1/2}\,\alpha^{-1/2},\end{aligned}$$

where C is a constant depending on $\{x_0, y_0, \kappa, \gamma, K, L, T, \sigma\}$. This completes the proof. □

Acknowledgements Research of MHD was supported by EPSRC Grants EP/W008041/1 and EP/V038516/1.

Author Contributions All the authors contributed to conducting the research and writing the manuscript.

Data Availability There is no additional data and materials.

Declarations

Ethical Approval Not applicable.

Competing interests There is no competing interests.

References

1. Bolley, F., Guillin, A., Malrieu, F.: Trend to equilibrium and particle approximation for a weakly selfconsistent vlasov-fokker-planck equation. ESAIM: M2AN **44**(5):867–884 (2010)
2. Breimhorst, N.: Smoluchowski-Kramers Approximation for Stochastic Differential Equations with non-Lipschitzian coefficients. PhD thesis (2009)
3. Cerrai, S., Freidlin, M.: On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom. Probab. Theory Related Fields **135**(3), 363–394 (2006)
4. Choi, Y.P., Tse, O.: Quantified overdamped limit for kinetic vlasov-fokker-planck equations with singular interaction forces. J. Differ. Equ. **330**, 150–207 (2022)
5. Duong, M.H., Lamacz, A., Peletier, M.A., Schlichting, A., Sharma, U.: Quantification of coarse-graining error in langevin and overdamped langevin dynamics. Nonlinearity **31**(10), 4517–4566 (2018)
6. Duong, M.H., Lamacz, A., Peletier, M.A., Sharma, U.: Variational approach to coarse-graining of generalized gradient flows. Calc. Var. Partial Differential Equations **56**(4), 100 (2017)
7. Devroye, L., Mehrabian, A., Reddad, T.: The total variation distance between high-dimensional gaussians with the same mean. arXiv:1810.08693 (2022)
8. Duong, M.H.: Long time behaviour and particle approximation of a generalised vlasov dynamic. Nonlinear Anal. Theory Methods Appl. **127**, 1–16 (2015)
9. Freidlin, M.: Some remarks on the Smoluchowski-Kramers approximation. J. Stat. Phys. **117**(3–4), 617–634 (2004)
10. Hottovy, S., Volpe, G., Wehr, J.: Noise-induced drift in stochastic differential equations with arbitrary friction and diffusion in the Smoluchowski-Kramers limit. J. Stat. Phys. **146**(4), 762–773 (2012)
11. Jabin, P.E., Wang, Z.: Mean Field Limit for Stochastic Particle Systems, pages 379–402. Springer International Publishing, Cham (2017)
12. Kac, M.: Foundations of kinetic theory. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954-1955, vol. III, University of California Press, Berkeley, Los Angeles (1956)
13. Klartag, B.: A central limit theorem for convex sets. Inventiones mathematicae **168**(1), 91–131 (2007)
14. Kramers, H.A.: Brownian motion in a field of force and the diffusion model of chemical reactions. Physica **7**(4), 284–304 (1940)
15. McKean, H.P.: Propagation of chaos for a class of non-linear parabolic equations, pages 41–57. Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ. (1967)
16. Narita, K.: Asymptotic behavior of velocity process in the smoluchowski-kramers approximation for stochastic differential equations. Adv. Appl. Probab. **23**(2), 317–326 (1991)
17. Narita, K.: The smoluchowski-kramers approximation for the stochastic liénard equation by mean-field. Adv. Appl. Probab. **23**(2), 303–316 (1991)
18. Narita, K.: Asymptotic behavior of fluctuation and deviation from limit system in the smoluchowski-kramers approximation for sde. Yokohama Mathematical Journal **42**, 41–76 (1994)
19. Nelson, E.: Dynamical Theories of Brownian Motion, volume 17. Princeton University Press Princeton (1967)
20. Nguyen, V.T., Nguyen, T.D.: A berry-esseen bound in the smoluchowski-kramers approximation. J. Stat. Phys. **179**(4), 871–884 (2020)
21. Nualart, D.: The Malliavin Calculus and Related Topics. Probability and Its Applications. Springer, Berlin Heidelberg (2006)
22. Pavliotis, G.A.: Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations. Texts in Applied Mathematics. Springer, New York (2014)
23. Risken, H., Frank, T.: The Fokker-Planck Equation: Methods of Solution and Applications. Springer Series in Synergetics. Springer, Berlin Heidelberg (1996)

24. Schilling, R.L., Partzsch, L.: Brownian Motion: An Introduction to Stochastic Processes. Walter de Gruyter (2012)
25. Sznitman, A.S.: Topics in propagation of chaos. In Paul-Louis Hennequin, editor, Ecole d'Eté de Probabilités de Saint-Flour XIX – 1989, volume 1464 of Lecture Notes in Mathematics, pages 165–251. Springer Berlin Heidelberg (1991)
26. Ta, C.S.: The rate of convergence for the smoluchowski-kramers approximation for stochastic differential equations with fbm. J. Stat. Phys. **181**(5), 1730–1745 (2020)
27. Xie, L., Yang, L.: The smoluchowski-kramers limits of stochastic differential equations with irregular coefficients. Stochastic Processes and their Applications **150**, 91–115 (2022)