Bernoulli **31**(3), 2025, 1709–1734
https://doi.org/10.3150/24-BEJ1745

Limit theorems for SDEs with irregular drifts

JIANHAI BAO[a] and JIAQING HAO[b]

Center for Applied Mathematics, Tianjin University, 300072 Tianjin, P.R. China, [a]*jianhaibao@tju.edu.cn,* [b]*hjq_0227@tju.edu.cn*

In this paper, concerning SDEs with Hölder continuous drifts, which are merely dissipative at infinity, and SDEs with piecewise continuous drifts, we investigate the strong law of large numbers and the central limit theorem for underlying additive functionals and reveal the corresponding rates of convergence. To establish the limit theorems under consideration, the exponentially contractive property of solution processes under the (quasi-)Wasserstein distance plays an indispensable role. In order to achieve such contractive property, which is new and interesting in its own right for SDEs with Hölder continuous drifts or piecewise continuous drifts, the reflection coupling method is employed and meanwhile a sophisticated test function is built.

Keywords: Central limit theorem; dissipativity at infinity; Hölder continuous drift; piecewise continuous drift; strong law of large numbers

1. Introduction and main results

The research on limit theorems for Markov processes has a long and rich history. As two typical candidates of limit theorems, the law of large numbers (LLN for short) and the central limit theorem (CLT for abbreviation), depicting respectively the temporal average convergence to the ergodic limit and the normalized fluctuations around the ergodic limit, have developed greatly in various settings in a century; see [6,7,9,11], to name just a few.

For the sake of the establishment on limit theorems concerned with Markov processes, one of the essential ingredients is to investigate the corresponding ergodic property. When the Markov process under consideration possesses the strong mixing properties (e.g., ergodic in the total variation distance), the LLN and the CLT associated with the additive functionals can be derived with the respective convergence rates $t^{-\frac{1}{2}+\varepsilon}$ for any $\varepsilon\in(0,\frac{1}{2})$ and $t^{-\frac{1}{2}}$; see, for instance, [19, p.217-218] and [11, Theorem 5.1.2] for more details. Whereas, in some occasions, the Markov process under investigation does not enjoy the strong mixing properties; see, for example, [11, Example 5.1.3] concerned with functional SDEs, which have the so-called reconstruction property. Concerning this setting (e.g., ergodicity under the (quasi)-Wasserstein distance in lieu of the total variation distance), the study on limit theorems has also advanced in the past few years. In particular, the weak LLN and the CLT were explored in [10] for weakly ergodic Markov processes by examining respectively the Feller property, exponential ergodicity under the 1-Wasserstein distance and uniform moment estimates with high order. Additionally, [11, Theorem 5.3.3] and [11, Theorem 5.3.4] addressed respectively the issues on the LLN and the CLT for stationary Markov processes with weakly ergodic properties. Subsequently, under the continuous-time path coupling condition (see [11, (5.3.10)] therein), [11, Proposition 5.3.5] extended the framework in [11, Theorems 5.3.3 and 5.3.4] to the non-stationary setup. In comparison with the counterparts in [10,11], Shirikyan [19] provided much more elegant conditions for the validity of the LLN and the CLT. Particularly, Shirikyan [19] formulated respectively a general criterion to establish the strong LLN and the CLT for weakly mixing Markov processes and, most importantly, the associated convergence rates were provided therein. More precisely, [19, Theorem 2.3] shows that the convergence rate of the strong LLN is $t^{-\frac{1}{2}+r_\nu}$, where $r_\nu:=q\vee((1+\nu)/(4p))$ for $\nu\in(0,2p-1)$ and $q<1/2$. So, from a quantitative point of view, the appearance of the quantity q will attenuate the convergence rate in a certain sense.

1350-7265

On the other hand, in [19, Theorem 2.8], one of the sufficient conditions for the CLT is concerned with the requirement on the uniform moment estimates of exponential type (see [19, (2.25)] for related details) associated with weakly mixing Markov processes. In most of the circumstances, such kind of exponential estimates (uniform in time) is a formidable task to be implemented; see, for example, [1, Lemma 2.1] for the functional SDEs as one of representatives with the weakly ergodic property.

As we mentioned above, [10,11,19] have established the LLN and the CLT for weakly ergodic Markov processes under different scenarios. In addition, the framework formulated in [10,11,19] has been applied to (functional) SDEs/SPDEs with *regular coefficients*. Especially, as a byproduct of [2,3] investigated the strong LLN and the CLT for a range of functional SDEs. Later, [2] was extended in [22] to treat the setting regarding functional SDEs with infinite memory.

In recent years, the theory on strong/weak well-posedness and distribution properties (e.g., gradient estimates and Harnack inequalities) of SDEs with irregular drifts has been studied systematically (see e.g. [20,23]). Yet, the study on limit theorems for SDEs with irregular drifts is still vacant so far. Inspired by the aforementioned literature [10,11,19] as well as [2,22], in the present work we make an attempt to investigate the LLN and the CLT for several class of SDEs with irregular drifts (e.g., the Hölder continuous drifts and the piecewise continuous drifts). Most importantly, besides the establishment of the LLN and the CLT, another main goal in this work is to improve the convergence rate of the LLN in [19, Theorem 2.3] and weaken the technical condition concerned with uniformly exponential estimates imposed in [19, Theorem 2.8]. The above can be viewed as some motivations of our present work.

Another motivation arises from the significant advancements of numerical limit theorems for SDEs/SPDEs with regular coefficients. Recently, as for SDEs/SPDEs with (semi-)Lipschitz continuous coefficients, there are plenty of literature on the LLN and the CLT; see e.g. [13,17] for SDEs approximated via the forward Euler-Maruyama scheme, [8] with regard to SDEs discretized by the backward Euler-Maruyama method, and [4] concerning semilinear SPDEs approximated via the spectral Galerkin method in the spatial direction and the exponential integrator in the temporal direction. To the best of our knowledge, the study on the LLN and the CLT for numerical schemes corresponding to SDEs with irregular drifts is still infrequent. So, in this work, we aim to lay the theoretical foundation on the LLN and the CLT for SDEs with Hölder continuous or piecewise continuous drifts (which are representative SDEs with irregular drifts) so that we can pave undoubtedly the way to investigating the LLN and the CLT for the numerical SDEs with irregular drifts.

Inspired by the existing literature mentioned above, in this work we intend to address the LLN and the CLT for SDEs with Hölder continuous drifts, where one part of drifts is dissipative in the long distance, and satisfies the monotone and Lyapunov conditions, respectively. Additionally, the LLN and CLT for SDEs with piecewise continuous drifts will also be explored in detail. The preceding contents will be elaborated progressively in the following three subsections.

1.1. LLN for SDEs with Hölder continuous drifts: Partial dissipativity

In this subsection, we work on the following SDE on $\mathbb{R}^d$:

$$\mathrm{d}X_t = \big(b_0(X_t) + b_1(X_t)\big)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \tag{1.1}$$

where $b_0, b_1 : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$, $(W_t)_{t\ge 0}$ is a d-dimensional Brownian motion on the complete filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\ge 0}, \mathbb{P})$.

Concerning the drifts b_0 and b_1, we shall assume that

($\mathbf{H}_b$) $b_1 : \mathbb{R}^d \to \mathbb{R}^d$ is locally Lipschitz and there exist constants $\lambda_1, \lambda_2, \ell_0 > 0$ such that

$$2\langle x - y, b_1(x) - b_1(y)\rangle \le \lambda_1|x - y|^2 \mathbb{1}_{\{|x-y|\le\ell_0\}} - \lambda_2|x - y|^2 \mathbb{1}_{\{|x-y|\ge\ell_0\}}, \qquad x, y \in \mathbb{R}^d; \quad (1.2)$$

$b_0 \in C^\alpha(\mathbb{R}^d)$ for some $\alpha \in (0, 1)$, i.e., there exists a constant $K_1 > 0$ such that

$$|b_0(x) - b_0(y)| \le K_1|x - y|^\alpha, \qquad x, y \in \mathbb{R}^d. \quad (1.3)$$

With regard to the diffusion term σ, we shall suppose that

($\mathbf{H}_\sigma$) $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ is Lipschitz continuous, that is, there is a constant $K_2 > 0$ such that

$$\|\sigma(x) - \sigma(y)\|_{\mathrm{HS}}^2 \le K_2|x - y|^2, \qquad x, y \in \mathbb{R}^d, \quad (1.4)$$

and moreover there exists a constant $\kappa \ge 1$ such that

$$\frac{1}{\kappa}|y|^2 \le \langle(\sigma\sigma^*)(x)y, y\rangle \le \kappa|y|^2, \qquad x, y \in \mathbb{R}^d. \quad (1.5)$$

Before proceeding, we make some comments on Assumptions ($\mathbf{H}_b$) and ($\mathbf{H}_\sigma$), respectively.

Remark 1.1. In literature, the Assumption (1.2) is also named as the dissipativity at infinity; see, for example, [14]. To demonstrate the condition (1.2), we provide an example below. Define for some parameters $a > 0$ and $n \ge 1$,

$$U(x) = x^2(g_n(x))^2 + a^2 - 2axg_n(x), \qquad x \in \mathbb{R},$$

where $g_n(x) := (x \wedge n) \vee (-n)$, $x \in \mathbb{R}$. Then, $b_1(x) = -U'(x)$ satisfies the condition (1.2) (see, for example, [14]) rather than the global convexity assumption: for some constant $K > 0$,

$$2\langle x - y, b_1(x) - b_1(y)\rangle \le -K|x - y|^2, \qquad x, y \in \mathbb{R}.$$

Since the drift term b_0 is singular (i.e. Hölder continuous), the uniformly elliptic condition in (1.5) is vitally important in addressing the well-posedness of (1.1). Most importantly, the condition (1.5) also plays a crucial role in exploring the exponentially contractive property of the SDE (1.1) via the reflection coupling method; see the proof of Proposition 2.1 for related details.

Under Assumptions ($\mathbf{H}_b$) and ($\mathbf{H}_\sigma$), the SDE (1.1) admits a unique strong solution $(X_t)_{t\ge0}$. Indeed, to address the strong well-posedness, we can adopt the routine as follows: first of all, we shall show that the SDE (1.1) has a unique local solution via the Zvonkin transformation and subsequently claim that the local solution is indeed a global one. In some occasions, we shall write $(X_t^x)_{t\ge0}$ instead of $(X_t)_{t\ge0}$ to highlight the dependence on the initial value $X_0 = x \in \mathbb{R}^d$. In the following part, we shall denote $C_{\mathrm{Lip}}(\mathbb{R}^d)$ by the collection of all Lipschitz continuous functions $f : \mathbb{R}^d \to \mathbb{R}$. Moreover, for a function $f : \mathbb{R}^d \to \mathbb{R}$ and $\nu \in \mathscr{P}(\mathbb{R}^d)$ (i.e., the set of probability measures on $\mathbb{R}^d$), we shall adopt the shorthand notation $\nu(f) = \int_{\mathbb{R}^d} f(x)\,\nu(\mathrm{d}x)$ in case of the integral $\nu(|f|) < \infty$.

The following LLN reveals the convergence rate of the additive functional $A_t^{f,x} := \frac{1}{t}\int_0^t f(X_s^x)\,\mathrm{d}s$ associated with a range of SDEs, which might be dissipative merely in the long distance and, in particular, allows one part of the drift terms involved to be Hölder continuous.

Theorem 1.1. *Assume* $(\mathbf{H}_b)$ *and* $(\mathbf{H}_\sigma)$*. Then, for any* $f \in C_{\mathrm{Lip}}(\mathbb{R}^d)$ *and* $\varepsilon \in (0,1/2)$*, there exist a random time* $T_\varepsilon \geq 1$ *and a constant* $C > 0$ *(dependent on the Lipschitz constant* $\|f\|_{\mathrm{Lip}}$ *and the initial value* x*) such that for all* $t \geq T_\varepsilon$*,*

$$\left|A_t^{f,x} - \mu(f)\right| \leq C t^{-\frac{1}{2}+\varepsilon}, \tag{1.6}$$

where $\mu \in \mathscr{P}(\mathbb{R}^d)$ *stands for the unique invariant probability measure of* $(X_t^x)_{t\geq 0}$ *solving* (1.1).

Before the end of this subsection, we make some remarks and comparisons with the existing literature.

Remark 1.2. With regard to the random time T_ε mentioned in Theorem 1.1, it indeed has any finite m-th moment; see, for example, [19, Corollary 2.4] for related details. In [19], a general framework was provided to establish the LLN for mixing-type Markov processes. When the weight function therein is constant (which corresponds the setup we work on in the present paper), the observable involved must be bounded. Hence, the general criterion in [19] cannot be applied (at least) directly to handle the setting we are interested in, where the observable herein is unbounded. We have to refine the proof of [19, Theorem 2.3]. In addition, in [19, Theorem 2.3], the corresponding convergent rate is $t^{-\frac{1}{2}+r_v}$ for $r_v := q \vee ((1+v)/(4p))$ with any $q < 1/2$ and $v \in (0, 2p-1)$. Whereas, in the present scenario, the associated convergence rate is $t^{-\frac{1}{2}+r_v}$, in which $r_v := (1+v)/(2p)$ for $v \in (0, p/2-1)$. Consequently, in a certain sense (in particular, q is close enough to $1/2$), we drop the redundant parameter $q < 1/2$ and improve accordingly the convergence rate derived in [19, Theorem 2.3].

1.2. CLT for SDEs with Hölder continuous drifts: Monotone and Lyapunov conditions

In this subsection, we move forward to derive the CLT associated with SDEs with Hölder continuous drifts.

In the proof of Theorem 1.1, the exponential contractivity under the 1-Wasserstein distance is one of the important factors. Nevertheless, the function φ_f, defined in (1.10) below, is merely Lipschitz continuous under the underlying quasi-metric rather than globally Lipschitz continuous. Hence, the exponential contractivity under the 1-Wasserstein distance is insufficient to establish the corresponding CLT via the martingale approach. Conversely, the exponential contractivity under the quasi-Wasserstein distance (see Proposition 3.1 below) is adequate for our purpose. In this setup, we can further weaken Assumption $(\mathbf{H}_b)$.

Throughout this subsection, we are still interested in the SDE (1.1), where Assumption $(\mathbf{H}_\sigma)$ is the same as that in Subsection 1.1 whereas Assumption $(\mathbf{H}_b)$ is substituted with the counterpart $(\mathbf{H}_b')$ below. More precisely,

$(\mathbf{H}_b')$ $b_0 \in C^\alpha(\mathbb{R}^d)$ satisfying (1.3) and there exist constants $\lambda, \lambda^* > 0$ such that for all $x, y \in \mathbb{R}^d$,

$$2\langle x-y, b_1(x) - b_1(y)\rangle \leq \lambda |x-y|^2 \tag{1.7}$$

and

$$\langle x, b_1(x)\rangle \leq -\lambda^* |x|^2 + C_{\lambda^*}. \tag{1.8}$$

The condition (1.7) shows that the drift term b_1 satisfies the classical monotone condition, which, in literature, is also called the one-sided Lipschitz condition. Let $b_1(x) = -x + f(x), x \in \mathbb{R}^d$, where the bounded function $f:\mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz with Lipschitz constant greater than 1. Obviously, both (1.7) and (1.8) are satisfied. Nevertheless, the condition (1.2) doesn't hold any more. (1.8), in addition to (1.3) and (1.5), indicates that the SDE (1.1) fulfils the Lyapunov condition. Under ($\mathbf{H}'_b$) and ($\mathbf{H}_\sigma$), the SDE (1.1) has a unique strong solution $(X_t^x)_{t\geq 0}$ with $X_0 = x \in \mathbb{R}^d$ and admits a unique invariant probability measure μ (see Proposition 3.1 below).

Before we present the second main result concerning the CLT, we further need to introduce some additional notation. For $p \geq 2$ and $\theta \in (0,1]$, denote $C_{p,\theta}(\mathbb{R}^d)$ by the family of all continuous functions $f:\mathbb{R}^d \to \mathbb{R}$ such that

$$\|f\|_{p,\theta} := \sup_{x\neq y, x,y\in\mathbb{R}^d} \frac{|f(x)-f(y)|}{\psi_{p,\theta}(x,y)} < \infty,$$

where for any $x,y \in \mathbb{R}^d$,

$$\psi_{p,\theta}(x,y) := (1\wedge |x-y|^\theta)(1+|x|^p+|y|^p). \tag{1.9}$$

As for $f \in C_{p,\theta}(\mathbb{R}^d)$, we define the corrector R_f as below

$$R_f(x) = \int_0^\infty \big((P_t f)(x) - \mu(f)\big)\,\mathrm{d}t, \quad x \in \mathbb{R}^d,$$

where $(P_t f)(x) := \mathbb{E} f(X_t^x)$ is the Markov semigroup associated with the solution process $(X_t^x)_{t\geq 0}$. Moreover, we set

$$\varphi_f(x) := \mathbb{E}\left|\int_0^1 f(X_r^x)\,\mathrm{d}r + R_f(X_1^x) - R_f(x)\right|^2, \qquad x \in \mathbb{R}^d. \tag{1.10}$$

In terms of [2, Lemma 4.1], $0 \leq \sigma_*^2 := \mu(\varphi_f) < \infty$ for $f \in C_{p,\theta}(\mathbb{R}^d)$. The quantity σ can be used to characterize the asymptotic variance of the additive functional $\bar{A}_t^{f,x} := \frac{1}{\sqrt{t}}\int_0^t f(X_s^x)\,\mathrm{d}s$ for $f \in C_{p,\theta}(\mathbb{R}^d)$.

Next, we present another main result, which is concerned with the convergence rate of the CLT corresponding the additive functional $\bar{A}_t^{f,x}$. In detail, we have the following statement.

Theorem 1.2. *Assume* ($\mathbf{H}'_b$) *and* ($\mathbf{H}_\sigma$). *Then, for any* $f \in C_{p,\theta}(\mathbb{R}^d)$ *with* $\mu(f) = 0$, $\sigma_*^2 := \mu(\varphi_f) \geq 0$ *and* $\varepsilon \in (0, \frac{1}{4})$, *there exists a constant* $C_0 = C_0(\|f\|_{p,\theta}, \sigma_*, |x|) > 0$ *such that*

$$\sup_{z\in\mathbb{R}^d} \Big(\theta_{\sigma_*}(z)\big|\mathbb{P}(\bar{A}_t^{f,x} \leq z) - \Phi_{\sigma_*}(z)\big|\Big) \leq C_0 t^{-\frac{1}{4}+\varepsilon}, \quad t \geq 1, \tag{1.11}$$

where $\theta_{\sigma_*}(z) := \mathbb{1}_{\{0<\sigma_*<\infty\}} + (1\wedge|z|)\mathbb{1}_{\{\sigma_*=0\}}$ *and* $\Phi_{\sigma_*}(z)$ *stands for the centered Gaussian distribution function with variation* σ_*^2.

Before the end of this subsection, we make some further remarks.

Remark 1.3. As Theorem 1.1, Theorem 1.3 is applicable to an SDE with the Hölder continuous drift, where another part of the drift is monotone and satisfies the Lyapunov condition. In [19, Theorem 2.8], a general criterion was provided to explore the CLT for uniformly mixing Markov families. In particular, the uniform moment estimates of exponential type (see [19, (2.25)]), as one of the sufficient conditions, was imposed therein. However, such a uniform moment estimate is, in general, hard to check; see [2] for the setting on functional SDEs and numerical SDEs. In lieu of the requirement on the uniform exponent moment, the uniform moment in the polynomial type, which is much easier to verify, is sufficient for our purpose as shown in the proof of Theorem 1.3.

1.3. LLN and CLT for SDEs with piecewise continuous drifts

In the previous two subsections, as far as two different setups are concerned, we establish respectively the strong LLN and the CLT. Whereas, no matter what which setting, the continuity of the drift terms is necessary.

For the objective in this subsection, we consider an illustrative example:

$$b(x) := -|x| - 2, \quad |x| \geq 1; \quad b(x) := 1 - x^2, \quad |x| < 1$$

and $\sigma(x) := \frac{1}{2}(1 + \frac{1}{1+x^2})$. Apparently, the drift term b above no longer satisfies Assumption ($\mathbf{H}_b$) or ($\mathbf{H}'_b$) due to the appearance of the discontinuous points. Yet, the drift b is globally dissipative in the long distance, the associated SDE should be ergodic under an appropriate probability (quasi-)distance, which is still vacant to the best of our knowledge. Therefore, intuitively speaking, the corresponding strong LLN and the CLT should be valid. So, in this subsection, our goal is to deal with the strong LLN and the CLT for SDEs, where the drifts involved might be discontinuous. So far, the topic mentioned above is still rare.

To explain the underlying essence to handle the limit theorems for SDEs with discontinuous drifts and, most importantly, avoid the cumbersome notation, we shall consider the scalar SDE

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \tag{1.12}$$

where $b : \mathbb{R} \to \mathbb{R}$ is piecewise continuous, i.e., there exist finitely many points $\xi_1 < \cdots < \xi_k$ such that b is continuous respectively on the intervals $I_i := (\xi_i, \xi_{i+1})$, $i \in \mathbb{S}_k := \{0, 1, \cdots, k\}$, where $\xi_0 = -\infty$ and $\xi_{k+1} = \infty$; $\sigma : \mathbb{R} \to \mathbb{R}$ is continuous; $(W_t)_{t\geq 0}$ is a 1-dimensional Brownian motion. So far, the SDE (1.12) with discontinuous drifts has been applied extensively in e.g. stochastic control theory and mathematical finance.

Besides Assumption ($\mathbf{H}_\sigma$) with $d = 1$, we shall assume that

($\mathbf{A}_b$) For each integer $n \geq |\xi_1| \vee |\xi_k|$, there exists an increasing function $\phi : [0, \infty) \to [0, \infty)$

$$|b(x) - b(y)| \leq \phi(n)|x - y|, \quad x, y \in I_i \cap B_n(0), \quad i \in \mathbb{S}_k, \tag{1.13}$$

where $B_n(0) := \{x \in \mathbb{R} : |x| \leq n\}$, and there exists a constant $\lambda_\star > 0$ such that

$$(x - y)(b(x) - b(y)) \leq \lambda_\star (x - y)^2, \quad x, y \in I_i, \quad i \in \mathbb{S}_k.$$

Moreover, there are constants $\lambda^\star, C_{\lambda^\star} > 0$ and $\varepsilon^\star \in (0, 1/2]$ such that

$$\varepsilon^\star |b(x)|(1 + |x|) + x b(x) \leq C_\star - \lambda^\star x^2, \qquad x \in \mathbb{R}. \tag{1.14}$$

Below, with regard to Assumptions ($\mathbf{A}_b$) and ($\mathbf{H}_\sigma$) with $d = 1$, we make the following remarks.

Remark 1.4. In contrast to (1.8), the Lyapunov condition (1.14) is a little bit unusual. This condition is imposed naturally when we handle the ergodicity of the transformed SDE; see the proof of Proposition 4.1 for more details. Most importantly, the appearance of $\varepsilon^\star \in (0, 1/2]$ allows b to be highly nonlinear.

In [16, Lemma 3], the strong well-posedness of (1.12) was treated under the global Lyapunov condition (see (A1) therein), the local monotonicity (see (A2)(i) therein), the locally polynomial growth condition (see (A2)(ii) therein) as well as the globally polynomial growth condition (see (A3) therein) concerned with the diffusion term, which is also non-degenerate at the discontinuous points of the drift term (see (A4) therein). By a close inspection of the argument of [16, Lemma 3], the much weaker condition (1.13) can indeed take the place of (A2)(ii). Therefore, by following the exact lines of [16, Lemma 3], the SDE (1.12) is strongly well-posed under Assumptions $(\mathbf{A}_b)$ and $(\mathbf{H}_\sigma)$ with $d = 1$. If we are only concerned with the well-posedness of (1.12), Assumption $(\mathbf{H}_\sigma)$ with $d = 1$ can be replaced definitely by (A3) and (A4). Whereas, the little bit strong condition $(\mathbf{H}_\sigma)$ with $d = 1$, compared with (A3) and (A4), is imposed to achieve the exponentially contractive property under the quasi-Wasserstein distance via the reflection coupling approach.

Our third main result concerned with the strong LLN and the CLT for the SDE (1.12) with piecewise continuous drifts is stated as follows.

Theorem 1.3. *Assume $(\mathbf{A}_b)$ and $(\mathbf{H}_\sigma)$ with $d = 1$. Then,*

(1) (**Strong LLN**) *For any $f \in C_{p,\theta}(\mathbb{R})$ and $\varepsilon \in (0, 1/2)$, there exist a random time $T_\varepsilon \ge 1$ and a constant $C = C(\|f\|_{p,\theta}, |x|) > 0$ such that for all $t \ge T_\varepsilon$,*

$$\left|A_t^{f,x} - \mu(f)\right| \le C t^{-\frac{1}{2}+\varepsilon},$$

where $\mu \in \mathscr{P}(\mathbb{R}^d)$ is the unique invariant probability measure of $(X_t)_{t\ge0}$ solving (1.12); see Proposition 4.1 below.

(2) (**CLT**) *For any $f \in C_{p,\theta}(\mathbb{R})$ with $\mu(f) = 0$, and $\varepsilon \in (0, 1/4)$, If $\sigma_*^2 := \mu(\varphi_f) \ge 0$, there exists a constant $C_0 = C_0(\|f\|_{p,\theta}, \sigma_*, |x|) > 0$ such that*

$$\sup_{z\in\mathbb{R}^d} \left(\theta_{\sigma_*}(z)\left|\mathbb{P}(\bar{A}_t^{f,x} \le z) - \Phi_{\sigma_*}(z)\right|\right) \le C_0 t^{-\frac{1}{4}+\varepsilon}, \quad t \ge 1,$$

where $\theta_{\sigma_*}(z) := \mathbb{1}_{\{0<\sigma_*<\infty\}} + (1 \wedge |z|)\mathbb{1}_{\{\sigma_*=0\}}$.

Before the ending of this subsection, we make some further comments.

Remark 1.5. To finish the proof of Theorem 1.3, the 1-dimensional diffeomorphism transformation (see (4.6) below) plays a crucial role. For the multidimensional transformation to handle well-posedness and numerical approximations for SDEs with piecewise continuous drifts, we refer to [12, Theorem 3.14] for more details. With the help of the multidimensional transformation initiated in [12], Theorem 1.3 can be generalized to the multidimensional SDEs with piecewise continuous drifts. Since such a generalization will only render the notation more cumbersome without bringing any new insights into the arguments, in the present work we restrict ourselves to the 1-dimensional setup.

Even though the original SDE under consideration is dissipative in the long distance, the corresponding transformed SDE is no longer dissipative (at infinity). Based on this point of view, the SDE (1.12) is ergodic under the quasi-Wasserstein distance rather than the genuine Wasserstein distance; see Proposition 4.1 for more details.

The remainder of this paper is organized as follows. In Section 2, via the reflection coupling, we investigate the 1-Wasserstein exponential contractivity for SDEs with Hölder continuous drifts, which also allow the drifts involved to be dissipative in the long distance. Subsequently, the proof of Theorem 1.1 is complete. Section 3 is devoted to the proof of Theorem 1.1 based on the establishment of the exponential ergodicity under the quasi-Wasserstein distance, which is interesting in its own right for SDEs with Hölder continuous drifts, where the other drift parts satisfy the monotone and Lyapunov condition. In Section 4, we aim to complete the proof of Theorem 1.3 with the aid of the exponential contractivity under the quasi-Wasserstein distance, which is new for SDEs with piecewise continuous drifts.

2. Proof of Theorem 1.1

In this section, we aim to complete the proof of Theorem 1.1, which is based on the following 1-Wasserstein contractive property.

Proposition 2.1. *Assume* $(\mathbf{H}_b)$ *and* $(\mathbf{H}_\sigma)$. *Then, there exist constants* $C^*, \lambda^* > 0$ *such that for all* $t \geq 0$ *and* $\mu, \nu \in \mathscr{P}_1(\mathbb{R}^d)$ (*the set of probability measures on* $\mathbb{R}^d$ *with finite moments of first order*),

$$\mathbb{W}_1(\mu P_t, \nu P_t) \leq C^* \mathrm{e}^{-\lambda^* t} \mathbb{W}_1(\mu, \nu), \tag{2.1}$$

where μP_t *stands for the law of* X_t, *the solution to* (1.1), *with the initial distribution* $\mathscr{L}_{X_0} = \mu$, *and* $\mathbb{W}_1$ *means the* 1-*Wasserstein distance. Moreover,* (2.1) *implies that* $(X_t)_{t\geq 0}$ *has a unique invariant probability measure* μ.

Proof. The proof, based on the reflection coupling approach, of Proposition 2.1 is inspired by the counterpart of [21, Theorem 3.1], which indeed was traced back to [18]. In [21, Theorem 3.1], an abstract framework upon exponential ergodicity of McKean-Vlasov SDEs, which are dissipative in the long distance, was presented. In the present setup, we follow essentially the line in [21, Theorem 3.1] whereas we refine the corresponding details and provide explicit conditions imposed on the coefficients so the content is much more readable.

Due to (1.5), for each $x \in \mathbb{R}^d$, the matrix $(\sigma\sigma^*)(x) - \frac{1}{2\kappa} I_{d\times d}$ is a nonnegative-definite symmetric matrix so there exists a symmetric $d \times d$-matrix $\tilde{\sigma}(x)$ such that $\tilde{\sigma}(x)^2 = (\sigma\sigma^*)(x) - \frac{1}{2\kappa} I_{d\times d}$, where $I_{d\times d}$ means the $d \times d$-identity matrix. Therefore, we readily have $(\sigma\sigma^*)(x) = \tilde{\sigma}(x)^2 + \frac{1}{2\kappa} I_{d\times d}$, $x \in \mathbb{R}^d$. Consider the SDE

$$\mathrm{d}Y_t = b(Y_t)\mathrm{d}t + \tilde{\sigma}(Y_t)\mathrm{d}\tilde{W}_t + \frac{1}{\sqrt{2\kappa}}\mathrm{d}\hat{W}_t, \tag{2.2}$$

where $(\tilde{W}_t)_{t\geq 0}$ and $(\hat{W}_t)_{t\geq 0}$ are mutually independent d-dimensional Brownian motions defined on the same probability space. To demonstrate that (2.2) has a unique strong solution under Assumptions $(\mathbf{H}_b)$ and $(\mathbf{H}_\sigma)$, we introduce the notations

$$\hat{\sigma}(x) := \Big(\tilde{\sigma}(x), \frac{1}{\sqrt{2\kappa}} I_{d\times d}\Big) \in \mathbb{R}^d \otimes \mathbb{R}^{2d}, \quad x \in \mathbb{R}^d, \quad \text{and} \quad \bar{W}_t := (\tilde{W}_t, \hat{W}_t),$$

where $(\bar{W}_t)_{t\geq 0}$ is a $2d$-dimensional Brownian motion. Whereafter, (2.2) can be reformulated as

$$\mathrm{d}Y_t = b(Y_t)\mathrm{d}t + \hat{\sigma}(Y_t)\mathrm{d}\bar{W}_t.$$

As a result, in terms of Subsection 1.1, it is sufficient to examine that $\hat{\sigma}$ satisfies Assumption $(\mathbf{H}_{\hat{\sigma}})$ so that (2.2) is strongly well-posed. Below, we aim to check the associated details, one by one. In the first place, by invoking (1.4) and (1.5), we derive that

$$\|\hat{\sigma}(x)-\hat{\sigma}(y)\|_{\mathrm{HS}} \le 2(K_2\kappa^3)^{\frac{1}{2}}|x-y|, \qquad x,y\in\mathbb{R}^d, \tag{2.3}$$

where we also used the fact that $\|\sqrt{A}-\sqrt{B}\|_{\mathrm{HS}} \le \frac{1}{2\lambda}\|A-B\|_{\mathrm{HS}}$ for $d\times d$ symmetric positive matrices A and B with all eigenvalues greater than $\lambda>0$ (see, for example, [18, (3.3)] for related details). Therefore, (2.3) enables us to conclude that the mapping $x\mapsto\hat{\sigma}(x)$ is also Lipschitz. In the next place, it is easy to see from (1.5) that for all $x,y\in\mathbb{R}^d$,

$$\frac{1}{\kappa}|y|^2 \le \langle(\hat{\sigma}\hat{\sigma}^*)(x)y,y\rangle = \langle(\sigma\sigma^*)(x)y,y\rangle \le \kappa|y|^2.$$

For $\mathbf{0}\neq x\in\mathbb{R}^d$, set the normalized vector $\mathbf{n}(x):=x/|x|$ and define the orthogonal matrix

$$\Pi_x = I_{d\times d} - 2\mathbf{n}(x)\otimes\mathbf{n}(x)\in\mathbb{R}^d\times\mathbb{R}^d.$$

To achieve the quantitative estimate (2.1), we work on the SDE

$$\begin{cases} \mathrm{d}\hat{Y}_t = b(\hat{Y}_t)\mathrm{d}t + \tilde{\sigma}(\hat{Y}_t)\mathrm{d}\tilde{W}_t + \dfrac{1}{\sqrt{2\kappa}}\Pi_{Z_t}\mathrm{d}\hat{W}_t, & t<\tau, \\ \mathrm{d}Y_t = b(Y_t)\mathrm{d}t + \tilde{\sigma}(Y_t)\mathrm{d}\tilde{W}_t + \dfrac{1}{\sqrt{2\kappa}}\mathrm{d}\hat{W}_t, & t\ge\tau, \end{cases} \tag{2.4}$$

where the coupling time $\tau:=\inf\{t\ge0: Z_t=\mathbf{0}\}$ with $Z_t:=Y_t-\hat{Y}_t$. Since $\Pi_\cdot$ is an orthogonal matrix, (2.4) is strongly well-posed before the coupling time as shown in the analysis above. Additionally, (2.4) coincides with (2.2) after the coupling time. Thereby, (2.4) is strongly well-posed.

Owing to the existence of an optimal coupling, in the following context, we can choose the initial values Y_0 and $\hat{Y}_0$ such that $\mathbb{W}_1(\mu,\nu)=\mathbb{E}|Y_0-\hat{Y}_0|$ for given $\mu,\nu\in\mathscr{P}_1(\mathbb{R}^d)$. Applying Itô's formula, we right now obtain from (2.2) and (2.4) that

$$\begin{aligned} \mathrm{d}|Z_t| \le{}& \frac{1}{2|Z_t|}\big(2\langle Z_t, b(Y_t)-b(\hat{Y}_t)\rangle + \|\tilde{\sigma}(Y_t)-\tilde{\sigma}(\hat{Y}_t)\|_{\mathrm{HS}}^2\big)\,\mathrm{d}t \\ &+ \big\langle\mathbf{n}(Z_t),(\tilde{\sigma}(Y_t)-\tilde{\sigma}(\hat{Y}_t))\mathrm{d}\tilde{W}_t\big\rangle + \frac{1}{\sqrt{\kappa/2}}\big\langle\mathbf{n}(Z_t),\mathrm{d}\hat{W}_t\big\rangle, \quad t<\tau, \end{aligned} \tag{2.5}$$

where we also utilized the fact that for any $\mathbf{0}\neq x\in\mathbb{R}^d$,

$$\langle I_{d\times d}-\mathbf{n}(x)\otimes\mathbf{n}(x), \mathbf{n}(x)\otimes\mathbf{n}(x)\rangle_{\mathrm{HS}} = 0.$$

By invoking (2.3), it follows from (1.5) that

$$\|\tilde{\sigma}(x)-\tilde{\sigma}(y)\|_{\mathrm{HS}} \le 4\kappa^{\frac{13}{8}}d^{\frac{1}{8}}K_2^{\frac{3}{8}}|x-y|^{\frac{3}{4}}, \quad x,y\in\mathbb{R}^d. \tag{2.6}$$

With the aid of (1.3) and (1.4), we find from (2.6) that

$$2\langle x-y, b(x)-b(y)\rangle + \|\tilde{\sigma}(x)-\tilde{\sigma}(y)\|_{\mathrm{HS}}^2 \le \phi(|x-y|)|x-y|, \quad x,y\in\mathbb{R}^d, \tag{2.7}$$

where for $u\ge0$,

$$\phi(u) := \big((\lambda_1+\lambda_2)u + 2K_1u^\alpha + 16\kappa^{\frac{13}{4}}d^{\frac{1}{4}}K_2^{\frac{3}{4}}u^{\frac{1}{2}}\big)\mathbb{1}_{\{u\le\hbar_0\}} - \frac{1}{2}\lambda_2 u$$

with

$$\hbar_0 := \ell_0 \vee \left(\frac{8K_1}{\lambda_2}\right)^{\frac{1}{1-\alpha}} \vee \left(\frac{64\kappa^{\frac{13}{4}} d^{\frac{1}{4}} K_2^{\frac{3}{4}}}{\lambda_2}\right)^2. \tag{2.8}$$

Consequently, (2.5) yields

$$\mathrm{d}|Z_t| \le \frac{1}{2}\phi(|Z_t|)\,\mathrm{d}t + \langle \mathbf{n}(Z_t), (\tilde{\sigma}(Y_t) - \tilde{\sigma}(\hat{Y}_t))\mathrm{d}\tilde{W}_t\rangle + \frac{1}{\sqrt{\kappa/2}}\langle \mathbf{n}(Z_t), \mathrm{d}\hat{W}_t\rangle, \qquad t < \tau.$$

Define the test function

$$f(r) = \kappa \int_0^r \mathrm{e}^{-\frac{\kappa}{2}\int_0^u \phi(v)\,\mathrm{d}v} \int_u^\infty s\mathrm{e}^{\frac{\kappa}{2}\int_0^s \phi(v)\,\mathrm{d}v}\,\mathrm{d}s\mathrm{d}u, \quad r \ge 0.$$

Straightforward calculations show that

$$f'(r) = \kappa \mathrm{e}^{-\frac{\kappa}{2}\int_0^r \phi(v)\,\mathrm{d}v} \int_r^\infty s\mathrm{e}^{\frac{\kappa}{2}\int_0^s \phi(v)\,\mathrm{d}v}\,\mathrm{d}s, \quad r \ge 0;$$

$$f''(r) = \frac{\kappa}{2}\big(-\phi(r)f'(r) - 2r\big), \quad r \in [0, \hbar_0),$$

and that

$$f'(r) = \frac{4}{\lambda_2}, \quad r \ge \hbar_0; \quad f''(r) = 0, \quad r > \hbar_0.$$

Whence, there exist constants $c_*, c^{**} > 0$ such that

$$c_* r \le f(r) \le c^{**} r, \qquad r \ge 0 \tag{2.9}$$

and

$$\frac{1}{2}f'(r)\phi(r) + \frac{1}{\kappa}f''(r) = -r, \quad r \in [0, \hbar_0) \cup (\hbar_0, \infty). \tag{2.10}$$

Note that f introduced above is a piecewise C^2-function. Thus, Tanaka's formula, together with the continuity of f', shows that

$$\mathrm{e}^{\frac{t}{c^{**}}} f(|Z_t|) \le f(|Z_0|) + \int_0^t \mathrm{e}^{\frac{s}{c^{**}}}\Big(\frac{1}{c^{**}}f(|Z_s|) + \frac{1}{2}f'(|Z_s|)\phi(|Z_s|) + \frac{1}{\kappa}f''(|Z_s|)\Big)\,\mathrm{d}s + M_t, \quad t < \tau$$

for some martingale $(M_t)_{t\ge 0}$. This, combining (2.9) with (2.10), further gives that

$$\begin{aligned}\mathbb{E}\big(\mathrm{e}^{\frac{t\wedge\tau}{c^{**}}} f(|Z_{t\wedge\tau}|)\big) &\le \mathbb{E}f(|Z_0|) + \mathbb{E}\bigg(\int_0^{t\wedge\tau} \mathrm{e}^{\frac{s}{c^{**}}}\Big(\frac{1}{c^{**}}f(|Z_s|) + \frac{1}{2}f'(|Z_s|)\phi(|Z_s|) + \frac{1}{\kappa}f''(|Z_s|)\Big)\,\mathrm{d}s\bigg)\\ &\le \mathbb{E}f(|Z_0|).\end{aligned}$$

Thanks to $f(|Z_t|) \equiv 0$ for all $t \ge \tau$, it is apparent that for all $t \ge 0$,

$$\mathrm{e}^{\frac{t}{c^{**}}}\mathbb{E}f(|Z_t|) = \mathbb{E}\big(\mathrm{e}^{\frac{t\wedge\tau}{c^{**}}} f(|Z_{t\wedge\tau}|)\big) \le \mathbb{E}f(|Z_0|).$$

Finally, (2.1) follows immediately by recalling $\mathbb{W}_1(\mu,\nu) = \mathbb{E}|Y_0 - \hat{Y}_0|$ and taking (2.9) into consideration.

Under (1.2), (1.3) and (1.5), there exists a constant $c_\star > 0$ such that $\sup_{t\geq 0}\mathbb{E}|X_t|^2 \leq c_\star(1+\mathbb{E}|X_0|^2)$; see (2.12) below for further details. Whence, the Krylov-Bogoliubov theorem yields that $(X_t)_{t\geq 0}$ has an invariant probability measure. Thus, the contractive property (2.1) implies the uniqueness of invariant probability measures. □

Below, we move to finish the

Proof of Theorem 1.1. In the sequel, we shall assume that $f \in C_{\rm Lip}(\mathbb{R}^d)$ and set $\varepsilon \in (0,1/2)$. Evidently, it suffices to verify that (1.6) is valid as long as $\mu(f)=0$. Accordingly, we shall stipulate $\mu(f)=0$ in the subsequent analysis. Below, we shall write X_t instead of X_t^x, set $A_t(f) := \frac{1}{t}\int_0^t f(X_s)\,\mathrm{d}s$ for all $t\geq 0$, and use the notation $a \lesssim b$ for given $a,b\geq 0$ provided that there exists a constant $c_0>0$ such that $a\leq c_0 b$. Note that for any $t\geq 0$,

$$|A_t(f)| \leq \big|A_{\lfloor t\rfloor}(f)\big| + \frac{1}{t}\int_{\lfloor t\rfloor}^t |f(X_s)|\,\mathrm{d}s,$$

where $\lfloor t\rfloor$ denotes the integer part of $t\geq 0$. Whence, to obtain the assertion (1.6), it remains to show respectively that there exists a random time $T\geq 1$ (dependent on ε) such that

$$\big|A_{\lfloor t\rfloor}(f)\big| \lesssim t^{-\frac{1}{2}+\varepsilon} \qquad \text{and} \qquad \frac{1}{t}\int_{\lfloor t\rfloor}^t |f(X_s)|\,\mathrm{d}s \lesssim t^{-\frac{1}{2}} \qquad \text{for all } t\geq T. \tag{2.11}$$

By (1.2), (1.3) and (1.5), it follows that there exist constants $c_1,c_2>0$ such that for any $p\geq 2$,

$$\mathrm{d}|X_t|^p \leq \big(-c_1|X_t|^p + c_2\big)\,\mathrm{d}t + p|X_t|^{p-2}\langle X_t, \sigma(X_t)\mathrm{d}W_t\rangle.$$

Then, Gronwall's inequality implies that for all $t\geq 0$ and $x\in\mathbb{R}^d$,

$$\mathbb{E}|X_t|^p \leq \frac{c_2}{c_1} + |x|^p. \tag{2.12}$$

For any integer $q\geq 2$, direct calculations show that

$$\mathbb{E}\Big|\int_0^t f(X_s)\,\mathrm{d}s\Big|^q \lesssim \bigg(\int_0^t\int_0^s \big(\mathbb{E}\big(|f(X_u)|(|P_{s-u}f|)(X_u)\big)^{\frac{q}{2}}\big)^{\frac{2}{q}}\,\mathrm{d}u\mathrm{d}s\bigg)^{\frac{q}{2}},$$

see, for instance, [19, (2.22)]. Next, using the invariance of μ followed by exploiting the Kontorovich dual and applying Proposition 2.1 yields that for all $t\geq 0$ and $x\in\mathbb{R}^d$,

$$\big|(P_tf)(x)-\mu(f)\big| \leq \|f\|_{\rm Lip}\mathbb{W}_1(\delta_xP_t,\mu P_t) \leq C^*\mathrm{e}^{-\lambda^* t}(|x|+\mu(|\cdot|)), \tag{2.13}$$

where $\|f\|_{\rm Lip}$ is the Lipschitz constant of the function f. The previous estimate, in addition to the Lipschitz property of f, $\mu(|\cdot|)<\infty$ as well as (2.12), gives that

$$\mathbb{E}\Big|\int_0^t f(X_s)\,\mathrm{d}s\Big|^q \lesssim \bigg(\int_0^t\int_0^s \mathrm{e}^{-\lambda^*(s-u)}\big(1+\mathbb{E}|X_u|^q\big)^{\frac{2}{q}}\,\mathrm{d}u\mathrm{d}s\bigg)^{\frac{q}{2}} \lesssim t^{\frac{q}{2}}. \tag{2.14}$$

Subsequently, by means of Hölder's inequality, we have for any $q\geq 2$,

$$\mathbb{E}|A_t(f)|^q \lesssim t^{-\frac{q}{2}}. \tag{2.15}$$

For any integer $n \geq 1$, by the Chebyshev inequality, together with (2.15) for $q = \frac{2}{\varepsilon} > 4$, we deduce that

$$\mathbb{P}\big(|A_n(f)| > n^{-\frac{1}{2}+\varepsilon}\big) \leq n^{\frac{1}{\varepsilon}-2}\mathbb{E}|A_n(f)|^{\frac{2}{\varepsilon}} \lesssim n^{-2}.$$

As a consequence, the Borel-Cantelli lemma yields that there exists a random variable $T_1 \geq 1$ such that

$$\big|A_{\lfloor t\rfloor}(f)\big| \leq \lfloor t\rfloor^{-\frac{1}{2}+\varepsilon}, \quad \text{a.s.,} \qquad t \geq T_1.$$

This apparently ensures the first statement.

Next, we proceed to examine the second statement in (2.11). In view of the BDG inequality, we infer from (1.5) and (2.12) that there exist constants $c_3, c_4 > 0$ such that for all integer $k \geq 0$ and any $p \geq 2$,

$$\begin{aligned}\mathbb{E}\big(\sup_{k\leq s\leq k+1}|X_s|^p\big) &\leq c_3 + \mathbb{E}|X_k|^p + p\,\mathbb{E}\Big(\sup_{k\leq s\leq k+1}\int_k^s |X_u|^{p-2}\langle X_u, \sigma(X_u)\mathrm{d}W_u\rangle\Big)\\ &\leq c_4(1+|x|^p) + \frac{1}{2}\mathbb{E}\big(\sup_{k\leq s\leq k+1}|X_s|^p\big)\end{aligned}$$

so that for all integer $k \geq 0$ and any $p \geq 2$,

$$\mathbb{E}\big(\sup_{k\leq s\leq k+1}|X_s|^p\big) \lesssim 1 + |x|^p. \tag{2.16}$$

By retrospecting that $f : \mathbb{R}^d \to \mathbb{R}$ is of linear growth, there exists a constant $c^\star > 0$ such that $|f(x)| \leq c^\star(1+|x|)$ for all $x \in \mathbb{R}^d$. It is ready to see that for any integer $k \geq 16$,

$$\mathbb{P}\Big(\sup_{k\leq t\leq k+1}|f(X_t)| > c^\star k^{\frac{1}{4}}\Big) \leq \mathbb{P}\Big(\sup_{k\leq t\leq k+1}|X_t| > k^{\frac{1}{4}} - 1\Big) \leq \mathbb{P}\Big(\sup_{k\leq t\leq k+1}|X_t| > \frac{1}{2}k^{\frac{1}{4}}\Big). \tag{2.17}$$

Then, the Chebyshev inequality, besides (2.16), signifies that

$$\mathbb{P}\Big(\sup_{k\leq t\leq k+1}|f(X_t)| > c^\star k^{\frac{1}{4}}\Big) \leq \frac{32}{k^{\frac{5}{4}}}\mathbb{E}\big(\sup_{k\leq t\leq k+1}|X_t|^5\big) \lesssim \frac{1}{k^{\frac{5}{4}}}(1+|x|^5). \tag{2.18}$$

Once more, applying the Borel-Cantelli lemma enables us to derive that there exists a random variable $T_2 \geq 16$ such that $|f(X_t)| \leq c^\star t^{\frac{1}{4}}$, a.s., for all $t \geq T_2$. Therefore, we obtain that for all $t \geq T_2$,

$$\frac{1}{t}\int_{\lfloor t\rfloor}^t |f(X_s)|\,\mathrm{d}s \lesssim \frac{1}{\lfloor t\rfloor}\int_{\lfloor t\rfloor}^t s^{\frac{1}{4}}\,\mathrm{d}s \lesssim \lfloor t\rfloor^{\frac{1}{4}}\Big(\big(1+\frac{1}{\lfloor t\rfloor}\big)^{\frac{5}{4}} - 1\Big), \quad \text{a.s.}$$

This, combining with the fact that $(1+r)^\alpha - 1 \leq \alpha 2^{\alpha-1} r$ for any $\alpha > 1$ and $r \in [0,1]$, guarantees that for all $t \geq T_2$,

$$\frac{1}{t}\int_{\lfloor t\rfloor}^t |f(X_s)|\,\mathrm{d}s \lesssim \lfloor t\rfloor^{-\frac{3}{4}} \lesssim t^{-\frac{3}{4}}, \quad \text{a.s.}$$

As a result, the second statement in (4.14) is verifiable.

Based on the analysis above, we conclude that (1.6) follows for the random time $T := T_1 + T_2$. □

3. Proof of Theorem 1.2

Before we start to complete the proof of Theorem 1.2, we provide the following proposition, which establishes the contractive property of transition kernels under the quasi-Wasserstein distance.

Proposition 3.1. *Under the assumptions of Theorem 1.2, for any $p \ge 2$, $\theta \in (0,1]$ and $\mu,\nu \in \mathscr{P}_{\psi_{p,\theta}}(\mathbb{R}^d)$, there exist constant $C^* \ge 1, \lambda^* > 0$ such that*

$$\mathbb{W}_{\psi_{p,\theta}}(\mu P_t, \nu P_t) \le C^* \mathrm{e}^{-\lambda^* t} \mathbb{W}_{\psi_{p,\theta}}(\mu,\nu), \qquad t \ge 0, \tag{3.1}$$

where

$$\mathscr{P}_{\psi_{p,\theta}}(\mathbb{R}^d) := \big\{\mu \in \mathscr{P}(\mathbb{R}^d) : \mu(\psi_{p,\theta}(\cdot,\mathbf{0})) < \infty\big\},$$

and $\mathbb{W}_{\psi_{p,\theta}}$ denotes the quasi-Wasserstein distance (see e.g. [5, (4.3)]) induced by the cost function $\psi_{p,\theta}$, introduced in (1.9). Moreover, $(X_t)_{t\ge0}$ solving (1.1) has a unique invariant measure $\mu \in \mathscr{P}_{\psi_{p,\theta}}(\mathbb{R}^d)$.

Proof. Throughout the whole proof to be implemented, we still utilize the coupling constructed in the proof of Proposition 2.1; see (2.2) and (2.4) for more details. In view of (1.3), (1.5) and (1.8), for any $p \ge 2$ and $V_p(x) := 1 + |x|^p, x \in \mathbb{R}^d$, there are constants $C_1(p), C_2(p) > 0$ such that

$$(\mathscr{L}V_p)(x) \le -C_1(p)V_p(x) + C_2(p), \quad x \in \mathbb{R}^d, \tag{3.2}$$

where $\mathscr{L}$ is the infinitesimal generator of (1.5). For the parameters $C_1(p), C_2(p)$ above, the set

$$\mathcal{A}_p := \big\{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d : C_1(p)\big(V_p(x) + V_p(y)\big) \le 4C_2(p)\big\}$$

is of finite length since the mapping $\mathbb{R}^d \ni x \mapsto V_p(x)$ is compact. Therefore, the quantity

$$l_p^* := 1 + \sup\big\{|x-y| : (x,y) \in \mathcal{A}_p\big\} < \infty$$

is well defined.

In order to achieve the exponential contractivity (3.1), it is vital to define two auxiliary functions as below. Define for $\theta \in (0,1]$,

$$h(r) = \kappa \int_0^r \mathrm{e}^{-\frac{\kappa}{2}\int_0^u ((\lambda+2K_2\kappa^3)v + 2K_1 v^\alpha)\,\mathrm{d}v} \int_u^{l_p^*} v^\theta \mathrm{e}^{\frac{\kappa}{2}\int_0^v ((\lambda+2K_2\kappa^3)l + 2K_1 l^\alpha)\,\mathrm{d}l}\,\mathrm{d}v\,\mathrm{d}u, \quad r \ge 0, \tag{3.3}$$

where $K_1 > 0$, $\kappa > 0$ and $\lambda > 0$ were introduced in (1.3), (1.5) and (1.7), respectively, and $\alpha \in (0,1)$ is the Hölder index associated with b_0. Moreover, we define for $\theta \in (0,1]$,

$$f(r) = c^*(r \wedge l_p^*)^\theta + h(r \wedge l_p^*), \quad r \ge 0, \tag{3.4}$$

where

$$c^* := \frac{1}{\theta(\lambda + 4K_2\kappa^3 + 2K_1 r_0^{\alpha-1})} \mathbb{1}_{(0,1)}(\theta) \qquad \text{with} \qquad r_0 := 1 \wedge \Big(\frac{1-\theta}{2K_1\kappa}\Big)^{\frac{1}{1+\alpha}}. \tag{3.5}$$

In the sequel, we shall fix the function f defined in (3.4) and choose the tuneable parameter

$$\begin{aligned}\varepsilon := 1 \wedge \frac{1}{16C_2(p)(c^* + h'(0)(l_p^*)^{1-\theta})} \\ \wedge \frac{1}{\left(2^{4+\frac{4}{p}} c^\star((p-2)^{1-\frac{2}{p}} \vee (p-1)^{1-\frac{1}{p}})(c^*\theta + h'(0)(l_p^*)^{1-\theta})\right)^p},\end{aligned} \tag{3.6}$$

where

$$c^\star := \big(\frac{1}{\kappa} 2^{(p-2)\vee 1}(p-1)\big) \vee (2K_2^{\frac{1}{2}} \kappa^2). \tag{3.7}$$

Note that, for the case $p = 2$, the term $(p-2)^{1-\frac{2}{p}}$ should be understood in the limit sense, that is, $\lim_{p\downarrow 2}(p-2)^{1-\frac{2}{p}} = 1$. Moreover, in the following context, for the sake of convenience, we shall write

$$\Psi(t) := f(|Z_t|)\big(1 + \varepsilon V_p(Y_t) + \varepsilon V_p(\hat{Y}_t)\big), \quad t \geq 0.$$

By applying Itô's formula, we deduce from (2.2) and (2.4) that

$$\begin{aligned}\mathrm{d}\Psi(t) &= (1 + \varepsilon V_p(Y_t) + \varepsilon V_p(\hat{Y}_t))\mathbb{1}_{\{0<|Z_t|\leq l_p^*\}}\mathrm{d}f(|Z_t|) \\ &\quad + \varepsilon f(|Z_t|)\mathrm{d}(|Y_t|^p + |\hat{Y}_t|^p) + \varepsilon \mathrm{d}\langle |Y_\cdot|^p + |\hat{Y}_\cdot|^p, f(|Z_\cdot|)\rangle(t), \quad t < \tau,\end{aligned} \tag{3.8}$$

in which $Z_t := Y_t - \hat{Y}_t$ and $\langle \xi, \eta\rangle(t)$ means the quadratic variation of stochastic processes $(\xi_t)_{t\geq 0}$ and $(\eta_t)_{t\geq 0}$. On the one hand, applying the Itô-Tanaka formula to (2.5), followed by taking (1.3), (1.7), (2.3) and $f'' \leq 0$ into consideration yields

$$\begin{aligned}\mathrm{d}f(|Z_t|) &\leq \Big(\frac{1}{2}f'(|Z_t|)\big(\lambda|Z_t| + 4K_2\kappa^3|Z_t| + 2K_1|Z_t|^\alpha\big) + \frac{1}{\kappa}f''(|Z_t|)\Big)\,\mathrm{d}t \\ &\quad + f'(|Z_t|)\Big(\big\langle(\tilde{\sigma}(Y_t) - \tilde{\sigma}(\hat{Y}_t))^*\mathbf{n}(Z_t), \mathrm{d}\tilde{W}_t\big\rangle + \frac{1}{\sqrt{\kappa/2}}\big\langle \mathbf{n}(Z_t), \mathrm{d}\hat{W}_t\big\rangle\Big), \quad t < \tau.\end{aligned}$$

On the other hand, by applying Itô's formula once more and making use of the Lyapunov condition (3.2), we derive that

$$\begin{aligned}\mathrm{d}\big(|Y_t|^p + |\hat{Y}_t|^p\big) \leq \big\{&-C_1(p)(V_p(Y_t) + V_p(\hat{Y}_t)) + 2C_2(p)\big\}\,\mathrm{d}t \\ &+ p\langle |Y_t|^{p-2}\tilde{\sigma}(Y_t)^*Y_t + |\hat{Y}_t|^{p-2}\tilde{\sigma}(\hat{Y}_t)^*\hat{Y}_t, \mathrm{d}\tilde{W}_t\rangle \\ &+ \frac{p}{\sqrt{2\kappa}}\langle |Y_t|^{p-2}Y_t + |\hat{Y}_t|^{p-2}\Pi_{Z_t}\hat{Y}_t, \mathrm{d}\hat{W}_t\rangle, \quad t < \tau.\end{aligned}$$

Consequently, combining the estimates on $\mathrm{d}f(|Z_t|)$ and $\mathrm{d}\big(|Y_t|^p + |\hat{Y}_t|^p\big)$ with (3.8) and

$$\frac{1}{|x-y|}\langle x-y, |x|^{p-2}x + |y|^{p-2}\Pi_{x-y}y\rangle = \frac{1}{|x-y|}\langle x-y, |x|^{p-2}x - |y|^{p-2}y\rangle, \quad x \neq y$$

enables us to derive that

$$\mathrm{d}\Psi(t) \leq \big(\Theta_1 + \Theta_2\big)(Y_t, \hat{Y}_t)\,\mathrm{d}t + \mathrm{d}M_t, \quad t < \tau$$

for some underlying martingale (M_t), where for any $x, y \in \mathbb{R}^d$ with $x \neq y$,

$$
\begin{aligned}
\Theta_1(x,y) :=\ & (1+\varepsilon V_p(x)+\varepsilon V_p(y)) \\
& \times \Big(\frac{1}{2} f'(|x-y|)\big(\lambda|x-y|+4K_2\kappa^3|x-y|+2K_1|x-y|^{\alpha}\big)+\frac{1}{\kappa} f''(|x-y|)\Big) \\
& \times \mathbb{1}_{\{0<|x-y|\le l_p^*\}} + \varepsilon f(|x-y|)\big(-C_1(p)\big(V_p(x)+V_p(y)\big)+2C_2(p)\big), \\
\Theta_2(x,y) :=\ & p\varepsilon f'(|x-y|)\Big(\frac{1}{\kappa}\big||x|^{p-2}x-|y|^{p-2}y\big| \\
& \qquad + \big|(\tilde{\sigma}(x)-\tilde{\sigma}(y))^*\big(|x|^{p-2}\tilde{\sigma}(x)^*x+|y|^{p-2}\tilde{\sigma}(y)^*y\big)\big|\Big).
\end{aligned}
$$

In case of

$$
\Theta_1(x,y)+\Theta_2(x,y) \le -\lambda^\star f(|x-y|)\big(1+\varepsilon V_p(x)+\varepsilon V_p(y)\big), \qquad x, y \in \mathbb{R}^d, \tag{3.9}
$$

where

$$
\lambda^\star := \frac{1}{4(c^*+h'(0)(l_p^*)^{1-\theta})} \wedge \frac{C_1(p)\varepsilon}{1+2\varepsilon},
$$

we then at once arrive at

$$
\mathrm{d}\Psi(t) \le -\lambda^*\Psi(t)\,\mathrm{d}t + \mathrm{d}M_t, \quad t < \tau.
$$

Subsequently, via the Itô formula, the estimate

$$
\mathbb{E}\big(\mathrm{e}^{\lambda^*(t\wedge\tau)}\Psi(t\wedge\tau)\big) \le \Psi(0)
$$

is available so the assertion (3.1) is attainable by taking advantage of the fact that

$$
\big(c^* \wedge f(l_p^*)\big)(1\wedge r^\theta) \le f(r) \le \big(f(l_p^*) \vee \big(c^*+h'(0)(l_p^*)^{1-\theta}\big)\big)(1\wedge r^\theta),
$$

and $Z_t = 0$ for $t \ge \tau$. On the basis of the preceding analysis, it all boils down to the confirmation of (3.9) in order to achieve (3.1).

In accordance with the definition of the function h introduced in (3.3), a direct calculation reveals that

$$
\frac{1}{2}h'(r)(\lambda r+4K_2\kappa^3 r+2K_1 r^\alpha)+\frac{1}{\kappa}h''(r) = -r^\theta, \qquad r \in (0, l_p^*]
$$

so that for all $r \in (0, l_p^*]$,

$$
\begin{aligned}
& \frac{1}{2}f'(r)(\lambda r+4K_2\kappa^3 r+2K_1 r^\alpha)+\frac{1}{\kappa}f''(r) \\
& = c^*\theta\Big(\frac{1}{2}\big(\lambda r^\theta+4K_2\kappa^3 r^\theta+2K_1 r^{\theta+\alpha-1}\big)-\frac{1}{\kappa}(1-\theta)r^{\theta-2}\Big)-r^\theta.
\end{aligned} \tag{3.10}
$$

By noting that

$$
2K_1 r^{\theta+\alpha-1}-\frac{1}{\kappa}(1-\theta)r^{\theta-2} \le 0, \qquad r \in (0, r_0],
$$

where $r_0 > 0$ was introduced in (3.5), we right away have

$$\frac{1}{2}f'(r)(\lambda r + 4K_2\kappa^3 r + 2K_1 r^\alpha) + \frac{1}{\kappa}f''(r) \le -\big(1 - \frac{1}{2}(\lambda + 4K_2\kappa^3)c^*\theta\big)r^\theta, \quad r \in (0, r_0].$$

Furthermore, by invoking (3.10) again, we apparently infer from $\alpha \in (0,1)$ that for all $r \in [r_0, l_p^*]$,

$$\frac{1}{2}f'(r)(\lambda r + 4K_2\kappa^3 r + 2K_1 r^\alpha) + \frac{1}{\kappa}f''(r) \le -\big(1 - \frac{1}{2}c^*\theta\big(\lambda + 4K_2\kappa^3 + 2K_1 r_0^{\alpha-1}\big)\big)r^\theta.$$

Consequently, taking the choice of $c^* > 0$ given in (3.5) into consideration yields that

$$\frac{1}{2}f'(r)(\lambda r + 4K_2\kappa^3 r + 2K_1 r^\alpha) + \frac{1}{\kappa}f''(r) \le -\frac{1}{2}r^\theta, \qquad r \in (0, l_p^*].$$

This definitely implies that

$$\begin{aligned}\Theta_1(x,y) \le& -\frac{1}{2}|x-y|^\theta(1 + \varepsilon V_p(x) + \varepsilon V_p(y))\mathbb{1}_{\{0<|x-y|\le l_p^*\}} \\ &+ \varepsilon f(|x-y|)\big(-C_1(p)\big(V_p(x) + V_p(y)\big) + 2C_2(p)\big), \qquad x, y \in \mathbb{R}^d.\end{aligned} \tag{3.11}$$

Next, by means of (1.4), (2.3) and (1.5), in addition to $h'(r) \le h'(0)$ for any $r \in [0, l_p^*]$, it follows from Young's inequality that for any $\alpha, \beta > 0$ and $p > 2$,

$$\begin{aligned}\Theta_2(x,y) &\le p\varepsilon\phi(|x-y|)\Big(\frac{1}{\kappa}2^{(p-2)\vee 1}(p-1)\big(|x|^{p-2} + |y|^{p-2}\big) + 2K_2^{\frac{1}{2}}\kappa^2\big(|x|^{p-1} + |y|^{p-1}\big)\Big) \\ &\le pc^\star\varepsilon\phi(|x-y|)\big(|x|^{p-2} + |y|^{p-2} + |x|^{p-1} + |y|^{p-1}\big) \\ &\le c^\star\phi(|x-y|)\Big((p-2)\alpha\big(\varepsilon|x|^p + \varepsilon|y|^p + \frac{4}{p-2}\alpha^{-\frac{p}{2}}\varepsilon\big) \\ &\qquad + (p-1)\beta\big(\varepsilon|x|^p + \varepsilon|y|^p + \frac{2}{p-1}\beta^{-p}\varepsilon\big)\Big), \qquad x, y \in \mathbb{R}^d,\end{aligned}$$

where $c^\star > 0$ was defined as in (3.7), and $\phi(r) := c^*\theta r^\theta + h'(0)r, r \ge 0$. In particular, choosing $\alpha = \left(\frac{4\varepsilon}{p-2}\right)^{\frac{2}{p}}$ and $\beta = \left(\frac{4\varepsilon}{p-1}\right)^{\frac{1}{p}}$, respectively, and taking the alternative of ε given in (3.7) and $l_p^* \ge 1$ into account leads to

$$\begin{aligned}\Theta_2(x,y) &\le 2^{1+\frac{4}{p}}c^\star\Big((p-2)^{1-\frac{2}{p}} \vee (p-1)^{1-\frac{1}{p}}\Big)\varepsilon^{\frac{1}{p}}\phi(|x-y|)\big(1 + \varepsilon V_p(x) + \varepsilon V_p(y)\big) \\ &\le 2^{1+\frac{4}{p}}c^\star\Big((p-2)^{1-\frac{2}{p}} \vee (p-1)^{1-\frac{1}{p}}\Big)\big(c^*\theta + h'(0)(l_p^*)^{1-\theta}\big)\varepsilon^{\frac{1}{p}}|x-y|^\theta \\ &\quad\times\big(1 + \varepsilon V_p(x) + \varepsilon V_p(y)\big) \\ &\le \frac{1}{8}|x-y|^\theta\big(1 + \varepsilon V_p(x) + \varepsilon V_p(y)\big), \qquad x, y \in \mathbb{R}^d.\end{aligned} \tag{3.12}$$

Whereafter, for any $x, y \in \mathbb{R}^d$ with $|x-y| \le l_p^*$, estimates (3.11) and (3.12) yield that

$$\begin{aligned}\Theta_1(x,y)+\Theta_2(x,y) &\le -\frac{3}{8}(1+\varepsilon V_p(x)+\varepsilon V_p(y))|x-y|^\theta + 2C_2(p)\varepsilon\big(c^*|x-y|^\theta + h'(0)|x-y|\big)\\ &\le -\frac{1}{4}(1+\varepsilon V_p(x)+\varepsilon V_p(y))|x-y|^\theta \\ &\le -\frac{1}{4(c^*+h'(0)(l_p^*)^{1-\theta})} f(|x-y|)(1+\varepsilon V_p(x)+\varepsilon V_p(y)),\end{aligned} \tag{3.13}$$

where in the first inequality we also used the fact that $h(r) \le h'(0)r$ for all $r \in [0, l_p^*]$, in the second inequality we employed the choice of ε provided in (3.6), and in the last inequality we exploited the fact that

$$c^* r^\theta \le f(r) \le \big(c^* + h'(0)(l_p^*)^{1-\theta}\big) r^\theta, \qquad r \in [0, l_p^*].$$

For any $x, y \in \mathbb{R}^d$ with $|x-y| > l_p^*$ (which obviously indicates $(x,y) \notin \mathcal{A}_p$), with the aid of $f'(r) = 0$ for any $r > l_p^*$, we deduce from the notions of Θ_1 and Θ_2 that

$$\begin{aligned}\Theta_1(x,y)+\Theta_2(x,y) &\le -\frac{1}{2}C_1(p)\varepsilon f(|x-y|)\big(V_p(x)+V_p(y)\big)\\ &= -\frac{C_1(p)\varepsilon(V_p(x)+V_p(y))}{2(1+\varepsilon(V_p(x)+V_p(y)))} f(|x-y|)\big(1+\varepsilon V_p(x)+\varepsilon V_p(y)\big)\\ &\le -\frac{C_1(p)\varepsilon}{1+2\varepsilon} f(|x-y|)\big(1+\varepsilon V_p(x)+\varepsilon V_p(y)\big),\end{aligned} \tag{3.14}$$

where the last line is due to $V_p \ge 1$.

At length, (3.9) is verifiable by combining (3.13) with (3.14) concerning the cases $|x-y| \le l_p^*$ and $|x-y| > l_p^*$ for all $x, y \in \mathbb{R}^d$, respectively.

Once (3.1) is available, the existence and uniqueness of invariant probability measures in $\mathscr{P}_{\psi_{p,\theta}}(\mathbb{R}^d)$ can be derived by following exactly the line in [5, Corollary 4.11]. □

With Proposition 3.1 at hand, we are in position to finish the

Proof of Theorem 1.2. Inspired essentially by the procedure in [19], we shall decompose the additive functional $\bar{A}_t^{f,x} := \frac{1}{\sqrt{t}} \int_0^t f(X_s^x)\,\mathrm{d}s$ for $f \in C_{p,\theta}(\mathbb{R}^d)$ with $\mu(f) = 0$ into two parts, where the one part is concerned with the additive functional of a martingale under consideration, and the other part is the corresponding remainder term. In order to achieve the desired convergence rate in the CLT for the additive functional $\bar{A}_t^{f,x}$, we shall adopt the convergence rate concerning the CLT for martingales (see e.g. [6, Theorem 3.10]) to treat the martingale part involved, and meanwhile exploit the contractive property (i.e., (3.1)) to handle the remainder term.

In following proof, we shall fix $f \in C_{p,\theta}(\mathbb{R}^d)$ with $\mu(f) = 0$. For $x \in \mathbb{R}^d$ and $t \ge 1$, let

$$M_t^{f,x} = \int_0^t \{f(X_s^x) - (P_s f)(x)\}\,\mathrm{d}s + \int_t^\infty \{(P_{s-t}f)(X_t^x) - (P_s f)(x)\}\,\mathrm{d}s.$$

It can readily be noted that the additive functional $\bar{A}_t^{f,x}$ can be rewritten as below:

$$\begin{aligned}\bar{A}_t^{f,x} &= \frac{1}{\sqrt{\lfloor t\rfloor}} M_{\lfloor t\rfloor}^{f,x} + \Big(\big(\lfloor t\rfloor^{-\frac{1}{2}} - t^{-\frac{1}{2}}\big) M_{\lfloor t\rfloor}^{f,x}, \\ &\qquad + t^{-\frac{1}{2}}\Big(\int_{\lfloor t\rfloor}^t f(X_s^x)\,\mathrm{d}s + \int_0^\infty \big((P_s f)(x) - (P_s f)(X_{\lfloor t\rfloor}^x)\big)\,\mathrm{d}s\Big)\Big) \\ &=: \bar{M}_t^{f,x} + R_t^{f,x}.\end{aligned} \tag{3.15}$$

Recall from [19, Lemma 2.9] the basic fact that for any real-valued random variables ξ, η and any $\alpha > 0, \sigma \geq 0$,

$$\sup_{z\in\mathbb{R}} \big|\mathbb{P}(\xi \leq z) - \Phi_\sigma(z)\big| \leq \sup_{z\in\mathbb{R}} \big|\mathbb{P}(\eta \leq z) - \Phi_\sigma(z)\big| + \mathbb{P}(|\xi - \eta| > \alpha) + c_\sigma \alpha,$$

where $c_\sigma := \frac{1}{\sigma\sqrt{2\pi}} \mathbb{1}_{\{\sigma>0\}} + 2\mathbb{1}_{\{\sigma=0\}}$. Thus, the decomposition (3.15) enables us to derive that for any $\alpha > 0$ and $\sigma \geq 0$,

$$\sup_{z\in\mathbb{R}^d} \big|\mathbb{P}\big(\bar{A}_t^{f,x} \leq z\big) - \Phi_\sigma(z)\big| \leq \sup_{z\in\mathbb{R}^d} \big|\mathbb{P}\big(\bar{M}_{\lfloor t\rfloor}^{f,x} \leq z\big) - \Phi_\sigma(z)\big| + \mathbb{P}\big(|R_t^{f,x}| > \alpha\big) + c_\sigma \alpha.$$

Hence, the desired assertion (3.1) follows as soon as, for any $\varepsilon \in (0, 1/4)$, there exists a constant $C_\varepsilon(x) > 0$ such that

$$\mathbb{P}\big(|R_t^{f,x}| > t^{-\frac{1}{4}}\big) \leq C_\varepsilon(x) t^{-\frac{1}{4}}, \qquad \sup_{z\in\mathbb{R}^d} \big|\mathbb{P}\big(\bar{M}_{\lfloor t\rfloor}^{f,x} \leq z\big) - \Phi_\sigma(z)\big| \leq C_\varepsilon(x) t^{-\frac{1}{4}+\varepsilon}. \tag{3.16}$$

Due to $\lfloor t\rfloor^{-\frac{1}{2}} - t^{-\frac{1}{2}} < t^{-\frac{1}{2}}$ for any $t \geq 1$, it follows from Chebyshev's inequality that for $t \geq 1$,

$$\mathbb{P}\big(|R_t^{f,x}| > t^{-\frac{1}{4}}\big) \leq t^{\frac{1}{4}} \Big(\big(\lfloor t\rfloor^{-\frac{1}{2}} - t^{-\frac{1}{2}}\big)\mathbb{E}\Big|\int_0^{\lfloor t\rfloor} f(X_s^x)\,\mathrm{d}s\Big| + t^{-\frac{1}{2}} \Theta_p(t,x)\Big), \tag{3.17}$$

where

$$\Theta_p(t,x) := \int_{\lfloor t\rfloor}^t \mathbb{E}|f(X_s^x)|\,\mathrm{d}s + 2\int_0^\infty \big(|(P_s f)(x)| + \mathbb{E}\big|(P_s f)(X_{\lfloor t\rfloor}^x)\big|\big)\,\mathrm{d}s.$$

For any $q \geq 2$, applying Itô's formula followed by taking advantage of (3.2) yields that for some constant $C_1(q) > 0$,

$$\sup_{t\geq s} \mathbb{E}|X_t^x|^q \lesssim 1 + \mathbb{E}|X_s^x|^q, \quad s \geq 0 \tag{3.18}$$

so that, for $f \in C_{p,\theta}(\mathbb{R}^d)$,

$$\sup_{t\geq s} \mathbb{E} f(X_t^x) \lesssim 1 + \mathbb{E}|X_s^x|^p \lesssim 1 + |x|^p, \quad s \geq 0. \tag{3.19}$$

Accordingly, Proposition 3.1, together with $\mu(f) = 0$, implies that

$$\Theta_p(t,x) \lesssim 1 + |x|^p + \|f\|_{p,\theta} \int_0^\infty \mathrm{e}^{-\lambda^* s}\big(1 + |x|^p + \mathbb{E}|X_{\lfloor t\rfloor}^x|^p\big)\,\mathrm{d}s \lesssim 1 + |x|^p. \tag{3.20}$$

Further, owing to the Markov property of $(X_t^x)_{t\geq 0}$, Proposition 3.1 and (3.18), we deduce that

$$\begin{aligned}\mathbb{E}\Big|\int_0^{\lfloor t\rfloor} f(X_s^x)\,\mathrm{d}s\Big|^2 &= 2\int_0^{\lfloor t\rfloor}\int_s^{\lfloor t\rfloor}\mathbb{E}\big(f(X_s^x)(P_{u-s}f)(X_s^x)\big)\,\mathrm{d}u\mathrm{d}s\\ &\lesssim \|f\|_{p,\theta}^2\int_0^{\lfloor t\rfloor}\int_s^{\lfloor t\rfloor}\mathrm{e}^{-\lambda^*(u-s)}\big(1+\mathbb{E}|X_s^x|^{2p}\big)\,\mathrm{d}u\mathrm{d}s\\ &\lesssim \|f\|_{p,\theta}^2(1+|x|^{2p})\lfloor t\rfloor.\end{aligned}$$

This further gives that

$$\big(\lfloor t\rfloor^{-\frac{1}{2}}-t^{-\frac{1}{2}}\big)\mathbb{E}\Big|\int_0^{\lfloor t\rfloor} f(X_s^x)\,\mathrm{d}s\Big|\lesssim \|f\|_{p,\theta}(1+|x|^p)\big(t^{\frac{1}{2}}-\lfloor t\rfloor^{\frac{1}{2}}\big)t^{-\frac{1}{2}}\lesssim \|f\|_{p,\theta}(1+|x|^p)t^{-\frac{1}{2}}, \quad (3.21)$$

where the second inequality is valid thanks to $t^{\frac{1}{2}}-\lfloor t\rfloor^{\frac{1}{2}}\leq (t-\lfloor t\rfloor)^{\frac{1}{2}}\leq 1$. Subsequently, plugging (3.20) and (3.21) back into (3.17) guarantees the validity of the first statement in (3.16).

We proceed to verify the second statement in (3.16). In light of Proposition 3.1 and by invoking the semigroup property of $(P_t)_{t\geq 0}$, it is easy to see that $(M_t^{f,x})_{t\geq 0}$ is a square integrable martingale with the zero mean. Note that $(M_n^{f,x})_{n\geq 1}$ can be reformulated as follows: for any integer $n\geq 1$,

$$M_n^{f,x}=\sum_{i=1}^n Z_i^{f,x}\qquad\text{with}\qquad Z_i^{f,x}:=M_i^{f,x}-M_{i-1}^{f,x}.$$

Trivially, according to the definition of $M_n^{f,x}$, we have for $1\leq i\leq n$,

$$Z_i^{f,x}=\int_{i-1}^i f(X_s^x)\mathrm{d}s+R_f(X_i^x)-R_f(X_{i-1}^x)\qquad\text{with}\qquad R_f(x):=\int_0^\infty (P_sf)(x)\,\mathrm{d}s.$$

By means of the property of conditional expectation and the flow property of $(X_t^x)_{t\geq 0}$, it follows readily that

$$\sum_{i=1}^n\mathbb{E}\big|Z_i^{f,x}\big|^2=\sum_{i=1}^n\mathbb{E}(\mathbb{E}(|Z_i^{f,x}|^2|\mathscr{F}_{i-1}))=\sum_{i=1}^n\mathbb{E}\varphi_f(X_{i-1}^x),$$

where

$$\varphi_f(x):=\mathbb{E}\Big|\int_0^1 f(X_s^x)+R_f(X_1^x)-R_f(x)\Big|^2.$$

By applying Proposition 3.1 and following exactly the routine of [1, Lemma 4.1& Lemma 4.2], we can deduce that $\varphi_f\in C_{2p,\theta}(\mathbb{R}^d)$ satisfying

$$0\leq\mu(\varphi_f)=2\mu(fR_f)<\infty,\quad \|\varphi_f\|_{2p,\theta}\lesssim\|f\|_{p,\theta}^2. \quad (3.22)$$

In addition, for any $q>\frac{1}{2}$ and $1\leq i\leq n$, we apparently have,

$$\mathbb{E}\big|Z_i^{f,x}\big|^{4q}\leq 3^{4q-1}\Big(\int_{i-1}^i\mathbb{E}|f(X_s^x)|^{4q}\,\mathrm{d}s+\mathbb{E}|R_f(X_i^x)|^{4q}+\mathbb{E}|R_f(X_{i-1}^x)|^{4q}\Big),$$

which, besides the fact that

$$|R_f(x)| \lesssim \|f\|_{p,\theta}(1+|x|^p), \quad x \in \mathbb{R}^d,$$

(3.18), and (3.19), leads to

$$\max_{1\le i\le n} \mathbb{E}|Z_i^{f,x}|^{4q} \lesssim 1+|x|^{4pq}, \quad n \ge 1.$$

As a consequence, by applying the Berry-Esseen type estimate associated with martingales (see, for instance, [2, Theorem 3.10]) we derive that

$$\begin{aligned}\sup_{z\in\mathbb{R}} &\Big|\mathbb{P}\big(M_n^{f,x}/\sqrt{\mu(\varphi_f)n} \le z\big) - \Phi_1(z)\Big| \\ &\lesssim \bigg(n^{-q} + (\mu(\varphi_f))^{-2q}\mathbb{E}\Big|\frac{1}{n}\sum_{i=1}^n \varphi_f(X_{i-1}^x) - \mu(\varphi_f)\Big|^{2q}\bigg)^{\frac{1}{4q+1}} \lesssim n^{-\frac{q}{4q+1}},\end{aligned} \tag{3.23}$$

where in the last display we also utilized the fact that

$$\mathbb{E}\Big|\frac{1}{n}\sum_{i=1}^n \varphi_f(X_{i-1}^x) - \mu(\varphi_f)\Big|^{2q} \lesssim n^{-q} \tag{3.24}$$

by tracing exactly the line to derive (2.14) and taking Proposition 3.1 into account. Concerning or the case $\sigma_*^2 := \mu(\varphi_f) > 0$, taking advantage of (3.23) gives that

$$\begin{aligned}\sup_{z\in\mathbb{R}^d} \Big|\mathbb{P}\big(\bar{M}_n^{f,x} \le z\big) - \Phi_{\sigma_*}(z)\Big| &= \sup_{z\in\mathbb{R}^d} \Big|\mathbb{P}\big(M_n^{f,x}/(\sqrt{n}\sigma_*) \le z/\sigma_*\big) - \Phi_{\sigma_*}(z)\Big| \\ &= \sup_{z\in\mathbb{R}^d} \Big|\mathbb{P}\big(M_n^{f,x}/(\sqrt{n}\sigma_*) \le z\big) - \Phi_1(z)\Big| \lesssim n^{-\frac{q}{4q+1}}.\end{aligned}$$

As a result, the second statement in (3.16) follows directly for the case $\mu(\varphi_f) > 0$.

Note from Chebyshev's inequality that for a random variable ξ and any $0 \ne z \in \mathbb{R}$,

$$(1\wedge|z|)\big|\mathbb{P}(\xi \le z) - \mathbb{1}_{[0,\infty)}(z)\big| = (1\wedge|z|)\big(\mathbb{P}(\xi > z)\mathbb{1}_{\{z>0\}} + \mathbb{P}(-\xi \ge -z)\mathbb{1}_{\{z<0\}}\big) \le \mathbb{E}|\xi|.$$

Therefore, with regard to the setting $\sigma_*^2 = \mu(\varphi_f) = 0$, for any integer $n \ge 1$, we have

$$\sup_{z\in\mathbb{R}} \big|\mathbb{P}\big(\bar{M}_n^{f,x} \le z\big) - \Phi_0(z)\big| \le \frac{1}{\sqrt{n}}\big(\mathbb{E}\big|M_n^{f,x}\big|^2\big)^{\frac{1}{2}}.$$

This, together with

$$\mathbb{E}|M_n^{f,x}|^2 = 2\sum_{i=1}^n\sum_{j=i}^n \mathbb{E}(Z_i^{f,x}Z_j^{f,x}) = \sum_{i=1}^n \mathbb{E}|Z_i^{f,x}|^2 = \sum_{i=1}^n \mathbb{E}\varphi_f(X_{i-1}^x)$$

by using the fact that $\mathbb{E}(Z_j^{f,x}|\mathscr{F}_i) = 0$ for $j > i$, and (3.24) with $\sigma_*^2 = \mu(\varphi_f) = 0$, implies

$$\sup_{z\in\mathbb{R}} \big|\mathbb{P}\big(\bar{M}_n^{f,x} \le z\big) - \Phi_0(z)\big| \lesssim n^{-\frac{1}{4}}.$$

Whence, the second statement concerned with the case $\sigma_*^2 = \mu(\varphi_f) = 0$ in (3.16) is verifiable. □

4. Proof of Theorem 1.3

By following respectively the procedures to implement the proof of Theorems 1.1 and 1.2, for the proof of Theorem 1.3, the key ingredient is to demonstrate the contractive property under the quasi-Wasserstein distance. More precisely, we shall prove the following statement.

Proposition 4.1. *Under Assumptions* $(\mathbf{A}_b)$ *and* $(\mathbf{H}_\sigma)$, *for any* $p \geq 2$, $\theta \in (0,1]$ *and* $\mu, \nu \in \mathscr{P}_{\psi_{p,\theta}}(\mathbb{R})$, *there exist constants* $C^* \geq 1, \lambda^* > 0$ *such that*

$$\mathbb{W}_{\psi_{p,\theta}}(\mu P_t, \nu P_t) \leq C^* \mathrm{e}^{-\lambda^* t} \mathbb{W}_{\psi_{p,\theta}}(\mu, \nu), \qquad t \geq 0, \tag{4.1}$$

where μP_t *denotes the law of* X_t *solving* (1.12) *with the initial distribution* $\mathscr{L}_{X_0} = \mu$. (4.1) *further implies that* $(X_t)_{t\geq 0}$ *solving* (1.12) *has a unique invariant probability measure in* $\mathscr{P}_{\psi_{p,\theta}}(\mathbb{R})$.

Compared with the setups treated in Subsections 1.1 and 1.2, the outstanding feature of the framework in Subsection 1.3 is due to the discontinuity of drifts associated with SDEs under investigation. Thus, the approaches in tackling Propositions 2.1 and 3.1 cannot be applied directly. In view of this, to handle the difficulty arising from the discontinuity of the drift term b, we adopt the following transformation (see, for instance, [15,16])

$$U(x) := \sum_{i=1}^{k} \alpha_i (x - \xi_i)|x - \xi_i| \phi((x - \xi_i)/\delta), \qquad x \in \mathbb{R}, \tag{4.2}$$

where

$$\alpha_i := \frac{b(\xi_i -) - b(\xi_i +)}{2\sigma(\xi_i)^2}, \qquad \phi(x) := (1 - x^2)^4 \mathbb{1}_{[-1,1]}(x),$$

and

$$\delta := \begin{cases} 1 \wedge \dfrac{\varepsilon^\star}{32|\alpha_1|}, & k = 1, \\ 1 \wedge \dfrac{\varepsilon^\star}{32 \max\{|\alpha_1|, \cdots, |\alpha_k|\}} \wedge \left(\dfrac{\varepsilon^\star}{2} \min\{\xi_2 - \xi_1, \cdots, \xi_k - \xi_{k-1}\} \right), & k \geq 2, \end{cases} \tag{4.3}$$

where the quantity $\varepsilon^\star > 0$ was introduced in (1.14).

The transformation U given in (4.2) enjoys nice properties. In particular, U and its derivative can be sufficiently small by choosing appropriate parameter δ involved in the definition of U. In terms of the definition of δ given in (4.3) and the prerequisite $\varepsilon^\star \in (0, 1/2]$, we have

$$|U(x)| \leq \delta^2 \sum_{i=1}^{k} |\alpha_i| 1_{[\xi_i - \delta, \xi_i + \delta]}(x) \leq \delta \max\{|\alpha_1|, \cdots, |\alpha_k|\} \leq \frac{\varepsilon^\star}{1 + \varepsilon^\star} \leq \frac{1}{3}, \quad x \in \mathbb{R}. \tag{4.4}$$

Moreover, a direct calculation shows that the function U is differentiable such that

$$U'(x) = 2 \sum_{i=1}^{k} \alpha_i |x - \xi_i| \big(1 - ((x - \xi_i)/\delta)^2\big)^3 \big(1 - 5((x - \xi_i)/\delta)^2\big)^2 1_{[\xi_i - \delta,\ \xi_i + \delta]}(x), \quad x \in \mathbb{R}.$$

This obviously implies that

$$|U'(x)| \le 32\delta \sum_{i=1}^{k} |\alpha_i| 1_{[\xi_i-\delta,\xi_i+\delta]}(x) \le 32\delta \max\{|\alpha_1|, \cdots, |\alpha_k|\} \le \varepsilon^\star \le \frac{1}{2}, \tag{4.5}$$

by taking advantage of the alternative of δ and $\varepsilon^\star \in (0, 1/2]$. Furthermore, we define the transformation

$$G(x) := x + U(x), \qquad x \in \mathbb{R}. \tag{4.6}$$

Apparently, (4.5) yields that

$$\frac{1}{2} \le 1 - |U'(x)| \le G'(x) \le 1 + |U'(x)| \le \frac{3}{2}. \tag{4.7}$$

Consequently, we conclude that the transformation $x \mapsto G(x)$ is a diffeomorphism.

Note that U' is differentiable on each interval $I_i, i \in \mathbb{S}_k$ so U' is piecewise differentiable. Then, for $Y_t := G(X_t)$, applying Itô's formula yields

$$\mathrm{d}Y_t = \tilde{b}(Y_t)\,\mathrm{d}t + \tilde{\sigma}(Y_t)\,\mathrm{d}W_t, \tag{4.8}$$

where

$$\tilde{b}(x) := (G'b)(G^{-1}(x)) + \frac{1}{2}(G''\sigma)(G^{-1}(x)), \quad \tilde{\sigma}(x) := (G'\sigma)(G^{-1}(x)), \quad x \in \mathbb{R}.$$

According to [16, Lemma 2], the SDE (4.8) is strongly well-posed via extending $U'' : \cup_{i=1}^{k+1} I_i \to \mathbb{R}$ to $U : \mathbb{R} \to \mathbb{R}$ by in particular taking

$$U''(\xi_i) = 2\Big(\alpha_i + \frac{b(\xi_i+) - b(\xi_i-)}{\sigma(\xi_i)^2}\Big), \quad i \in \mathbb{S}_k.$$

With the preceding preliminaries, we start to complete the

Proof of Proposition 4.1. Below, let $(X_t^\mu)_{t\ge 0}$ and $(Y_t^\mu)_{t\ge 0}$ be the solutions to (1.12) and (4.8) with $\mathscr{L}_{X_0} = \mu \in \mathscr{P}(\mathbb{R})$ and $\mathscr{L}_{Y_0} = \mu \in \mathscr{P}(\mathbb{R})$, respectively. Due to the Kontorovich dual, besides $Y_t^{\mu\circ G^{-1}} = G(X_t^\mu)$ and $Y_t^{\nu\circ G^{-1}} = G(X_t^\nu)$ for $\mu, \nu \in \mathscr{P}_{\psi_{p,\theta}}(\mathbb{R})$, we find that

$$\mathbb{W}_{\psi_{p,\theta}}(\mu P_t, \nu P_t) = \sup_{\|f\|_{\psi_{p,\theta}} \le 1} \Big|\mathbb{E}(f\circ G^{-1})(Y_t^{\mu\circ G^{-1}}) - \mathbb{E}(f\circ G^{-1})(Y_t^{\nu\circ G^{-1}})\Big|. \tag{4.9}$$

Next, by means of the mean value theorem and (4.7), it follows that

$$\Big|G^{-1}(x) - G^{-1}(y)\Big| \le 2|x-y|, \qquad |G(x) - G(y)| \le \frac{3}{2}|x-y|, \quad x, y \in \mathbb{R}. \tag{4.10}$$

This further implies that for any $x, y \in \mathbb{R}$,

$$|(f\circ G^{-1})(x) - (f\circ G^{-1})(y)| + |(f\circ G)(x) - (f\circ G)(y)| \le C_p^* \|f\|_{\psi_{p,\theta}} \psi_{p,\theta}(x,y), \tag{4.11}$$

where

$$C_p^* := \big(2^\theta (2^{2p-1} \vee \big(1 + 2^p |G^{-1}(0)|^p)\big)\big) \vee \big((3/2)^\theta \big((3^p/2) \vee (1 + 2^p |G^{(0)}|^p)\big)\big).$$

Subsequently, via the Kontorovich dual once more, along with (4.9) and (4.11), we deduce that

$$\mathbb{W}_{\psi_{p,\theta}}(\mu P_t, \nu P_t) \le C_p^* \mathbb{W}_{\psi_{p,\theta}}\big((\mu\circ G^{-1})\tilde{P}_t, (\nu\circ G^{-1})\tilde{P}_t\big), \quad \mu,\nu\in\mathscr{P}_{\psi_{p,\theta}}(\mathbb{R}) \tag{4.12}$$

where $(\mu\circ G^{-1})\tilde{P}_t := \mathscr{L}_{Y_t^{\mu\circ G^{-1}}}$. Provided that there exist constants $C^\star, \lambda^\star > 0$ such that

$$\mathbb{W}_{\psi_{p,\theta}}\big(\mu\tilde{P}_t, \nu\tilde{P}_t\big) \le C^\star \mathrm{e}^{-\lambda^\star t}\mathbb{W}_{\psi_{p,\theta}}(\mu,\nu), \quad \mu,\nu\in\mathscr{P}_{\psi_{p,\theta}}(\mathbb{R}) \tag{4.13}$$

then we derive from (4.12) that

$$\mathbb{W}_{\psi_{p,\theta}}(\mu P_t, \nu P_t) \le C_p^* C^\star \mathrm{e}^{-\lambda^\star t}\mathbb{W}_{\psi_{p,\theta}}\big(\mu\circ G^{-1}, \nu\circ G^{-1}\big) \le \big(C_p^*\big)^2 C^\star \mathrm{e}^{-\lambda^\star t}\mathbb{W}_{\psi_{p,\theta}}(\mu,\nu),$$

where the second inequality is valid due to the Kontorovich dual again and (4.11).

Based on the analysis above, to achieve (4.1), it remains to claim that (4.13) is verifiable. By applying Proposition 3.1 with $b_0 = 0$, for the validity of (4.13), it amounts to proving that

($\mathbf{H}_{\tilde\sigma}$) there exist constants $K^* > 0$ and $\kappa^* \ge 1$ such that for all $x, y \in \mathbb{R}$,

$$\frac{1}{\kappa^*} \le \tilde\sigma(x) \le \kappa^*, \qquad |\tilde\sigma(x) - \tilde\sigma(x)| \le K^*|x-y|;$$

($\mathbf{H}_{\tilde b}$) there exist constants $\lambda_0, \lambda_0^*, C_{\lambda_0^*} > 0$ such that for all $x \in \mathbb{R}$,

$$2(x-y)\big(\tilde b(x) - \tilde b(y)\big) \le \lambda_0(x-y)^2, \qquad x\,\tilde b(x) \le -\lambda_0^* x^2 + C_{\lambda_0^*}.$$

By recalling the definition of $\tilde\sigma$, we obtain from ($\mathbf{H}_\sigma$) with $d = 1$ and (4.7) that for any $x, y \in \mathbb{R}$,

$$\frac{1}{4\kappa} \le \tilde\sigma(x)^2 \le 4\kappa, \qquad |\tilde\sigma(x) - \tilde\sigma(y)| \le \sqrt{\kappa}|(G'\circ G^{-1})(x) - (G'\circ G^{-1})(y)| + 3\sqrt{K_2}|x-y|.$$

Notice that

$$G''(x) = -2\alpha_i\psi_i(x), \quad x\in(\xi_i-\varepsilon, \xi_i); \quad G''(x) = 2\alpha_i\psi_i(x), \quad x\in(\xi_i, \xi_i+\varepsilon),$$

and that, otherwise, $G''(x) = 0$. Thus, a straightforward calculation, besides the continuity of $G' : \mathbb{R} \to \mathbb{R}$, reveals that there exists a constant $c_0 > 0$ such that

$$|G'(x) - G'(y)| \le c_0|x-y|, \quad x, y\in\mathbb{R} \tag{4.14}$$

so by invoking (4.10) there is a constant $c_1 > 0$ satisfying that

$$|\tilde\sigma(x) - \tilde\sigma(y)| \le c_1|x-y|, \qquad x, y\in\mathbb{R}.$$

Therefore, we conclude that the assertion ($\mathbf{H}_{\tilde\sigma}$) follows.

By following the exact line to derive [16, (A2')(i)], there exists a constant $c_2 > 0$ such that

$$2(x-y)\big(\tilde b(x) - \tilde b(y)\big) \le c_2|x-y|^2, \quad x, y\in\mathbb{R}.$$

Next, by taking the definition of $\tilde b$ into consideration, we find readily from (4.5) and (4.4) that for some constant $c_3 > 0$,

$$x\tilde b(x) = \big(G^{-1}(x) + U(G^{-1}(x))\big)\big(b(G^{-1}(x)) + (U'b)(G^{-1}(x))\big) + \frac{1}{2}x(G''\sigma)(G^{-1}(x))$$

$$
\begin{aligned}
&\le G^{-1}(x)b(G^{-1}(x)) + |G^{-1}|(x)|U'b|(G^{-1}(x)) \\
&\quad + |U|(G^{-1}(x))\big(|b|(G^{-1}(x)) + |U'b|(G^{-1}(x))\big) + \frac{1}{2}|x| \cdot |G''\sigma|(G^{-1}(x)) \\
&\le G^{-1}(x)b(G^{-1}(x)) + \varepsilon^{\star}\big(1 + |G^{-1}(x)|\big)|b|(G^{-1}(x)) + c_3|x|, \qquad x \in \mathbb{R},
\end{aligned}
$$

where in the identity we used $G(x) = x + U(x)$ and in the last inequality we employed (1.5) and (4.14). Whereafter, (1.14) yields that

$$
x\tilde{b}(x) \le C_{\star} - \lambda^*(G^{-1}(x))^2 + c_3|x|, \qquad x \in \mathbb{R}.
$$

This, together with the fact that

$$
|(G^{-1}(x)|^2 = |x - U((G^{-1}(x))|^2 \ge \frac{1}{2}|x^2| - |U((G^{-1}(x))|^2 \ge \frac{1}{2}|x^2| - \frac{1}{9}, \quad x \in \mathbb{R},
$$

by making use of the basic inequality: $(a-b)^2 \ge \frac{1}{2}a^2 - b^2$ for $a, b \in \mathbb{R}$ and (4.4), leads to

$$
x\tilde{b}(x) \le c_4 - c_5|x|^2, \quad x \in \mathbb{R}
$$

for some constants $c_4, c_5 > 0$. As a consequence, we reach the assertion $(\mathbf{H}_{\tilde{b}})$.

Based on the contractivity (4.13), the transformed SDE (4.8) has a unique invariant probability measure $\nu \in \mathscr{P}_{\psi_{p,\theta}}(\mathbb{R})$ by following the line of [5, Corollary 4.11]. Note that the transformation G constructed above is a diffeomorphism. Thus, via integrals with respect to image measures, we conclude that $\mu := \nu \circ G \in \mathscr{P}_{\psi_{p,\theta}}(\mathbb{R})$ is the unique invariant probability measure of $(X_t)_{t\ge 0}$ solving (1.12). □

Finally, with the aid of Proposition 4.1, we complete

Proof of Theorem 1.3. Applying Proposition 4.1 yields that for any $f \in C_{p,\theta}(\mathbb{R})$.

$$
\big|(P_t f)(x) - \mu(f)\big| \le \|f\|_{p,\theta}\mathbb{W}_{\psi_{p,\theta}}(\delta_x P_t, \mu P_t) \le C^* \mathrm{e}^{-\lambda^* t}(|x|^p + \mu(|\cdot|^p)), \quad t \ge 0,\ x \in \mathbb{R}, \tag{4.15}
$$

With this estimate at hand, the strong LLN can be verifiable by tracing the line in the proof of Theorem 1.1 and, in particular, replacing $|X_t|$ in (2.14), (2.17) and (2.18) by $|X_t|^p$, respectively. Moreover, with the help of (4.15), the CLT can de derived by following exactly the procedure to tackle Theorem 1.2 so we omit the corresponding details herein. □

Acknowledgements

We would like to thank the AE and referees for careful comments and corrections.

Funding

The research of Jianhai Bao is supported by the National Key R&D Program of China (2022YFA1006 004) and NSF of China (No. 12071340).

References

[1] Bao, J., Wang, F.-Y. and Yuan, C. (2015). Hypercontractivity for functional stochastic differential equations. *Stochastic Process. Appl.* **125** 3636–3656. MR3357623 https://doi.org/10.1016/j.spa.2015.04.001

[2] Bao, J., Wang, F.-Y. and Yuan, C. (2020). Limit theorems for additive functionals of path-dependent SDEs. *Discrete Contin. Dyn. Syst.* **40** 5173–5188. MR4128304 https://doi.org/10.3934/dcds.2020224

[3] Bao, J., Wang, F.-Y. and Yuan, C. (2020). Ergodicity for neutral type SDEs with infinite length of memory. *Math. Nachr.* **293** 1675–1690. MR4148827 https://doi.org/10.1002/mana.201800539

[4] Chen, C., Dang, T., Hong, J. and Zhou, T. (2023). CLT for approximating ergodic limit of SPDEs via a full discretization. *Stochastic Process. Appl.* **157** 1–41. MR4521762 https://doi.org/10.1016/j.spa.2022.11.015

[5] Hairer, M., Mattingly, J.C. and Scheutzow, M. (2011). Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields* **149** 223–259. MR2773030 https://doi.org/10.1007/s00440-009-0250-6

[6] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. *Probability and Mathematical Statistics*. New York-London: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0624435

[7] Jacod, J. and Shiryaev, A.N. (1987). *Limit Theorems for Stochastic Processes*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **288**. Berlin: Springer. MR0959133 https://doi.org/10.1007/978-3-662-02514-7

[8] Jin, D. (2025). Central limit theorem for temporal average of backward Euler-Maruyama method. *J. Comput. Math.* **43** 588–614. MR4829500 https://doi.org/10.4208/jcm.2311-m2023-0147

[9] Komorowski, T., Landim, C. and Olla, S. (2012). *Fluctuations in Markov Processes: Time Symmetry and Martingale Approximation*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **345**. Heidelberg: Springer. MR2952852 https://doi.org/10.1007/978-3-642-29880-6

[10] Komorowski, T. and Walczuk, A. (2012). Central limit theorem for Markov processes with spectral gap in the Wasserstein metric. *Stochastic Process. Appl.* **122** 2155–2184. MR2921976 https://doi.org/10.1016/j.spa.2012.03.006

[11] Kulik, A. (2018). *Ergodic Behavior of Markov Processes with Applications to Limit Theorems*. *De Gruyter Studies in Mathematics* **67**. Berlin: de Gruyter. MR3791835

[12] Leobacher, G. and Szölgyenyi, M. (2017). A strong order 1/2 method for multidimensional SDEs with discontinuous drift. *Ann. Appl. Probab.* **27** 2383–2418. MR3693529 https://doi.org/10.1214/16-AAP1262

[13] Lu, J., Tan, Y. and Xu, L. (2022). Central limit theorem and self-normalized Cramér-type moderate deviation for Euler-Maruyama scheme. *Bernoulli* **28** 937–964. MR4388925 https://doi.org/10.3150/21-bej1372

[14] Majka, M.B., Mijatović, A. and Szpruch, Ł. (2020). Nonasymptotic bounds for sampling algorithms without log-concavity. *Ann. Appl. Probab.* **30** 1534–1581. MR4132634 https://doi.org/10.1214/19-AAP1535

[15] Müller-Gronbach, T. and Yaroslavtseva, L. (2022). A strong order 3/4 method for SDEs with discontinuous drift coefficient. *IMA J. Numer. Anal.* **42** 229–259. MR4367654 https://doi.org/10.1093/imanum/draa078

[16] Müller-Gronbach, T., Sabanis, S. and Yaroslavtseva, L. (2024). Existence, uniqueness and approximation of solutions of SDEs with superlinear coefficients in the presence of discontinuities of the drift coefficient. *Commun. Math. Sci.* **23** 389–429. MR4846712 https://doi.org/10.4310/cms.241217213438

[17] Pagès, G. and Panloup, F. (2012). Ergodic approximation of the distribution of a stationary diffusion: Rate of convergence. *Ann. Appl. Probab.* **22** 1059–1100. MR2977986 https://doi.org/10.1214/11-AAP779

[18] Priola, E. and Wang, F.-Y. (2006). Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.* **236** 244–264. MR2227134 https://doi.org/10.1016/j.jfa.2005.12.010

[19] Shirikyan, A. (2006). Law of large numbers and central limit theorem for randomly forced PDE's. *Probab. Theory Related Fields* **134** 215–247. MR2222383 https://doi.org/10.1007/s00440-005-0427-6

[20] Wang, F.-Y. (2016). Gradient estimates and applications for SDEs in Hilbert space with multiplicative noise and Dini continuous drift. *J. Differ. Equ.* **260** 2792–2829. MR3427683 https://doi.org/10.1016/j.jde.2015.10.020

[21] Wang, F.-Y. (2023). Exponential ergodicity for non-dissipative McKean-Vlasov SDEs. *Bernoulli* **29** 1035–1062. MR4550214 https://doi.org/10.3150/22-bej1489

[22] Wang, Y., Wu, F. and Zhu, C. (2022). Limit theorems for additive functionals of stochastic functional differential equations with infinite delay. *J. Differ. Equ.* **308** 421–454. MR4340783 https://doi.org/10.1016/j.jde.2021.11.011

[23] Xia, P., Xie, L., Zhang, X. and Zhao, G. (2020). $L^q(L^p)$-theory of stochastic differential equations. *Stochastic Process. Appl.* **130** 5188–5211. MR4108486 https://doi.org/10.1016/j.spa.2020.03.004

Received October 2023 and revised March 2024