Distribution Dependent Reflecting Stochastic Differential Equations*

Feng-Yu Wang

Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

December 20, 2022

Abstract

To characterize the Neumann problem for nonlinear Fokker-Planck equations, we investigate distribution dependent reflecting SDEs (DDRSDEs) in a domain. We first prove the well-posedness and establish functional inequalities for reflecting SDEs with singular drifts, then extend these results to DDRSDEs with singular or monotone coefficients, for which a general criterion deducing the well-posedness of DDRSDEs from that of reflecting SDEs is established.

AMS subject Classification: 60H10, 60G65.
Keywords: Distribution dependent reflecting SDEs, well-posedness, log-Harnack inequality.

1 Introduction

Because of intrinsic links to nonlinear Fokker-Planck equations/mean-field particle systems and many other applications, distribution dependent (McKean-Vlasov) SDEs have been intensively investigated, see for instances the monograph/surveys [7, 11, 28] among many other references. To characterize the Neumann problem for nonlinear Fokker-Planck equations in a domain, we aim to develop a counterpart theory for distribution dependent reflecting SDEs (DDRSDEs for short).

The only reference we know on this topic is [1], where DDRSDEs are studied in a convex domain for coefficients satisfying the $\mathbb{W}_2$-Lipschitz condition in the distribution variable and the semi-Lipschitz condition in the space variable. We will work on a general framework where D may be non-convex and the coefficients could be singular in both space and distribution variables.

*Supported in part by NNSFC (11831014, 11921001) and the National Key R&D Program of China (No. 2020YFA0712900)

We first state the fundamental assumption on the domain in the study of reflecting SDEs, then introduce the link of DDRSDEs and nonlinear Neumann problems, and finally summarize the main results derived in the paper with an example of (singular) granular media equation with Neumann boundary.

1.1 Assumption on the domain

Let $D \subset \mathbb{R}^d$ be a connected open domain with boundary ∂D. For any $x \in \partial D$ and $r > 0$, let

$$\mathscr{N}_{x,r} := \big\{\mathbf{n} \in \mathbb{R}^d : |\mathbf{n}| = 1, B(x - r\mathbf{n}, r) \cap D = \emptyset\big\},$$

where $B(x, r) := \{y \in \mathbb{R}^d : |x - y| < r\}$. Since $\mathscr{N}_{x,r}$ is decreasing in $r > 0$, we have

$$\mathscr{N}_x := \cup_{r>0}\mathscr{N}_{x,r} = \lim_{r\downarrow 0} \mathscr{N}_{x,r}, \quad x \in \partial D.$$

We call $\mathscr{N}_x$ the set of inward unit normal vectors of ∂D at point x. When ∂D is differentiable at x, $\mathscr{N}_x$ is a singleton set. Otherwise $\mathscr{N}_x$ may be empty or contain more than one vectors. For instance, letting D be the interior of a triangle in $\mathbb{R}^2$, at each vertex x the set $\mathscr{N}_x$ contains infinite many vectors, whereas for D being the exterior of the triangle $\mathscr{N}_x$ is empty at each vertex point x.

Following [19, 24], throughout the paper we make the following assumption on D, which automatically holds for $D = \mathbb{R}^d$ where $\partial D = \emptyset$.

(D) Either D is convex, or there exists a constant $r_0 > 0$ such that $\mathscr{N}_x = \mathscr{N}_{x,r_0} \neq \emptyset$ and

(1.1) $$\sup_{v\in\mathbb{R}^d,|v|=1} \inf\big\{\langle v, \mathbf{n}(y)\rangle : y \in B(x, r_0) \cap \partial D, \mathbf{n}(y) \in \mathscr{N}_y\big\} \geq r_0, \quad x \in \partial D.$$

Remark 2.1. We present below some facts on assumption **(D)**.

(1) According to [24, Remark 1.1], for any $x \in \partial D$ and $r > 0$, $\mathbf{n} \in \mathscr{N}_{x,r}$ if and only if $\langle y - x, \mathbf{n}\rangle \geq -\frac{|y-x|^2}{2r}$ for $\in \bar{D}$, so that the condition $\mathscr{N}_x = \mathscr{N}_{x,r_0}$ in **(D)** implies

(1.2) $$\langle y - x, \mathbf{n}(x)\rangle \geq -\frac{|y - x|^2}{2r_0}, \quad y \in \bar{D}, x \in \partial D, \mathbf{n}(x) \in \mathscr{N}_x.$$

When D is convex, **(D)** holds for any $r_0 > 0$ so that

(1.3) $$\langle y - x, \mathbf{n}(x)\rangle \geq 0, \quad y \in \bar{D}, x \in \partial D, \mathbf{n}(x) \in \mathscr{N}_x,$$

and (1.1) holds if $d = 2$ or D is bounded, see [29].

(2) When ∂D is C^1-smooth, for each $x \in \partial D$ the set $\mathscr{N}_x$ is singleton. If $\mathbf{n}(x) \in \mathscr{N}_x$ is uniformly continuous in $x \in \partial D$, then (1.1) holds for small $r_0 > 0$. In particular, **(D)** holds when $\partial D \in C_b^2$ in the following sense.

Definition 1.1. For any $r > 0$, let

$$\partial_r D := \big\{x \in \bar{D} : \mathrm{dist}(x, \partial D) \le r\big\}, \quad \partial_{-r} D := \big\{x \in D^c : \mathrm{dist}(x, \partial D) \le r\big\},$$
$$\partial_{\pm r} D := (\partial_r D) \cup \partial_{-r} D, \quad D_r := D \cup (\partial_{-r} D).$$

For any $k \in \mathbb{N}$, we write $\partial D \in C_b^k$ if there exists a constant $r_0 > 0$ such that the polar coordinate map

$$I : \partial D \times [-r_0, r_0] \ni (\theta, \rho) \mapsto \theta + \rho \mathbf{n}(\theta) \in \partial_{\pm r_0} D$$

is a C^k-diffeomorphism, such that $(\theta(x), \rho(x)) := I^{-1}(x)$ having bounded and continuous derivatives in $x \in \partial_{\pm r_0} D$ up to the k-th order, where $\theta(x)$ is the projection of x to ∂D and

$$\rho(x) = \mathrm{dist}(x, \partial D) 1_{\{\partial_{r_0} D\}}(x) - \mathrm{dist}(x, \partial D) 1_{\{\partial_{-r_0} D\}}(x), \quad x \in \partial_{\pm r_0} D. \tag{1.4}$$

Moreover, for $\varepsilon \in (0, 1)$, we denote $\partial D \in C_b^{k+\varepsilon}$ if it is in C_b^k with $\nabla^k \rho$ and $\nabla^k \theta$ being ε-Hölder continuous on $\partial_{r_0} D$. Finally, we write $\partial D \in C_b^{k,L}$ if it is C_b^k with $\nabla^k \rho$ being Lipschitiz continuous on $\partial_{r_0} D$.

Note that $\partial D \in C_b^k$ does not imply the boundedness of D or ∂D, but any bounded C^k domain satisfies $\partial D \in C_b^k$.

1.2 DDRSDE and nonlinear Neumann problem

Let $\mathscr{P}(\bar{D})$ be the space of all probability measures on the closure $\bar{D}$ of D, equipped with the weak topology. Consider the following DDRSDE on $\bar{D} \subset \mathbb{R}^d$:

$$\mathrm{d}X_t = b_t(X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma_t(X_t, \mathscr{L}_{X_t})\mathrm{d}W_t + \mathbf{n}(X_t)\mathrm{d}l_t, \quad t \ge 0, \tag{1.5}$$

where $(W_t)_{t\ge0}$ is an m-dimensional Brownian motion on a complete filtration probability space $(\Omega, \{\mathscr{F}_t\}_{t\ge0}, \mathbb{P})$, $\mathscr{L}_{X_t}$ is the distribution of X_t, $\mathbf{n}(x) \in \mathscr{N}_x$ for $x \in \partial D$, l_t is an adapted continuous increasing process which increases only when $X_t \in \partial D$, and

$$b : [0, \infty) \times D \times \mathscr{P}(\bar{D}) \to \mathbb{R}^d, \quad \sigma : [0, \infty) \times D \times \mathscr{P}(\bar{D}) \to \mathbb{R}^d \otimes \mathbb{R}^m$$

are measurable. When different probability measures are considered, we denote by $\mathscr{L}_{X|\mathbb{P}}$ the distribution of a random variable X under the probability $\mathbb{P}$.

Definition 1.2. (1) A pair $(X_t, l_t)_{t\ge0}$ is called a solution of (1.5), if X_t is an adapted continuous process on $\bar{D}$, l_t is an adapted continuous increasing process with $\mathrm{d}l_t$ supported on $\{t \ge 0 : X_t \in \partial D\}$, such that $\mathbb{P}$-a.s.

$$\int_0^t \big\{|b_r(X_r, \mathscr{L}_{X_r})| + \|\sigma_r(X_r, \mathscr{L}_{X_r})\|^2\big\}\mathrm{d}r < \infty, \quad t \ge 0,$$

and for some measurable map $\partial D \ni x \mapsto \mathbf{n}(x) \in \mathscr{N}_x$, $\mathbb{P}$-a.s.

$$X_t = X_0 + \int_0^t b_r(X_r, \mathscr{L}_{X_r})\mathrm{d}r + \int_0^t \sigma_r(X_r, \mathscr{L}_{X_r})\mathrm{d}W_r + \int_0^t \mathbf{n}(X_r)\mathrm{d}l_r, \quad t \ge 0.$$

In this case, l_t is called the local time of X_t on ∂D. We call (1.5) strongly well-posed for distributions in a subspace $\hat{\mathscr{P}} \subset \mathscr{P}(\bar{D})$, if for any $\mathscr{F}_0$-measurable variable X_0 with $\mathscr{L}_{X_0} \in \hat{\mathscr{P}}$, the equation has a unique solution with $\mathscr{L}_{X_t} \in \hat{\mathscr{P}}$ for $t \geq 0$; if this is true for $\hat{\mathscr{P}} = \mathscr{P}(\bar{D})$, we called it strongly well-posed.

(2) A triple $(X_t, l_t, W_t)_{t\geq 0}$ is called a weak solution of (1.5), if W_t is an m-dimensional Brownian motion under a probability space and $(X_t, l_t)_{t\geq 0}$ solves (1.5). (1.5) is called weakly unique (resp. jointly weakly unique), if for any two weak solutions $(X_t, l_t, W_t)_{t\geq 0}$ under probability $\mathbb{P}$ and $(\tilde{X}_t, \tilde{l}_t, \tilde{W}_t)_{t\geq 0}$ under probability $\tilde{\mathbb{P}}$, $\mathscr{L}_{X_0|\mathbb{P}} = \mathscr{L}_{\tilde{X}_0|\tilde{\mathbb{P}}}$ implies $\mathscr{L}_{(X_t,l_t)_{t\geq 0}|\mathbb{P}} = \mathscr{L}_{(\tilde{X}_t,\tilde{l}_t)_{t\geq 0}|\tilde{\mathbb{P}}}$ (resp. $\mathscr{L}_{(X_t,l_t,W_t)_{t\geq 0}|\mathbb{P}} = \mathscr{L}_{(\tilde{X}_t,\tilde{l}_t,\tilde{W}_t)_{t\geq 0}|\tilde{\mathbb{P}}}$). We call (1.5) weakly well-posed for distributions in $\hat{\mathscr{P}} \subset \mathscr{P}(\bar{D})$, if it has a unique weak solution for initial distributions in $\hat{\mathscr{P}}$ and the distribution of the solution at any time is in $\hat{\mathscr{P}}$; it is called weakly well-posed if moreover $\hat{\mathscr{P}} = \mathscr{P}(\bar{D})$.

(3) We call (1.5) well-posed (for distributions in $\hat{\mathscr{P}}$), if it is both strongly and weakly well-posed (for distributions in $\hat{\mathscr{P}}$).

To characterize the nonlinear Fokker-Planck equation associated with (1.5), consider the following time-distribution dependent second order differential operator:

$$L_{t,\mu} := \frac{1}{2}\mathrm{tr}\big\{(\sigma_t\sigma_t^*)(\cdot,\mu)\nabla^2\big\} + \nabla_{b_t(\cdot,\mu)}, \quad t \geq 0, \mu \in \mathscr{P}(\bar{D}), \tag{1.6}$$

where ∇ and ∇^2 are the gradient and Hessian operators in $\mathbb{R}^d$ respectively, and ∇_v is the directional derivative along $v \in \mathbb{R}^d$. Assume that for any $\mu \in C([0,\infty); \mathscr{P}(\bar{D}))$,

$$\sigma_t^\mu(x) := \sigma_t(x,\mu_t), \quad b_t^\mu(x) := b_t(x,\mu_t) \tag{1.7}$$

satisfy $\|\sigma^\mu\|^2 + |b^\mu| \in L^1_{loc}([0,\infty)\times\bar{D}; \mathrm{d}t\,\mu_t(\mathrm{d}x))$.

Let $C_N^2(\bar{D})$ be the class of C^2-functions on $\bar{D}$ with compact support satisfying the Neumann boundary condition $\nabla_{\mathbf{n}} f|_{\partial D} = 0$. By Itô's formula, for any (weak) solution X_t to (1.5), $\mu_t := \mathscr{L}_{X_t}$ solves the nonlinear Fokker-Planck equation

$$\partial_t \mu_t = L^*_{t,\mu_t}\mu_t \ \text{ with respect to } C_N^2(\bar{D}), \quad t \geq 0 \tag{1.8}$$

for probability measures on $\bar{D}$, in the sense that $\mu_\cdot \in C([0,\infty); \mathscr{P}(\bar{D}))$ and

$$\mu_t(f) := \int_{\bar{D}} f\mathrm{d}\mu_t = \mu_0(f) + \int_0^t \mu_s(L_{s,\mu_s}f)\mathrm{d}s, \quad t \geq 0, f \in C_N^2(\bar{D}). \tag{1.9}$$

On the other hand, by establishing the "superposition principle" as in [3, 4] based on [31], under reasonable conditions we may prove that a solution to (1.8) also provides a weak solution to (1.5). We leave this to a future study.

To understand (1.8) as a nonlinear Neumann problem on D, let L^*_{t,μ_t} be the adjoint operator of L_{t,μ_t}: for any $g \in L^1_{loc}(D, (\|\sigma_t(x,\mu_t)\|^2 + |b_t(x,\mu_t)|)\mathrm{d}x)$, $L^*_{t,\mu_t}g$ is the linear functional on $C_0^2(D)$ (the class of C^2-functions on D with compact support) given by

$$C_0^2(D) \ni f \mapsto \int_D \{fL^*_{t,\mu_t}g\}(x)\mathrm{d}x := \int_D \{gL_{t,\mu_t}f\}(x)\mathrm{d}x. \tag{1.10}$$

Assume that $\mathscr{L}_{X_t}$ has a density function ρ_t, i.e. $\mu_t := \mathscr{L}_{X_t} = \rho_t(x)\mathrm{d}x$. It is the case under a general non-degenerate or Hörmander condition (see for instance [6]), and Krylov's estimate (2.20) or (2.59) below implies the existence of ρ_t for a.e. $t > 0$. When $\partial D \in C^2$, (1.8) implies that ρ_t solves the following nonlinear Neumann problem on $\bar{D}$:

$$\partial_t \rho_t = L^*_{t,\rho_t}\rho_t, \quad \nabla_{t,\mathbf{n}}\rho_t|_{\partial D} = 0, \ t \geq 0 \tag{1.11}$$

in the weak sense, where $L_{t,\rho_t} := L_{t,\rho_t(x)\mathrm{d}x}$, and for a function g on ∂D

$$\nabla_{t,\mathbf{n}} g := \nabla_{\sigma_t\sigma_t^*\mathbf{n}} g + \mathrm{div}_{\partial D}(g\pi\sigma_t\sigma_t^*\mathbf{n})$$

for the divergence $\mathrm{div}_{\partial D}$ on ∂D and the projection π to the tangent space of ∂D:

$$\pi_x v := v - \langle v, \mathbf{n}(x)\rangle \mathbf{n}(x), \quad v \in \mathbb{R}^d, x \in \partial D.$$

If in particular $\sigma\sigma^*\mathbf{n} = \lambda\mathbf{n}$ holds on $[0,\infty) \times \partial D$ for a function $\lambda \neq 0$ a.e., $\nabla_{t,\mathbf{n}}\rho_t|_{\partial D} = 0$ is equivalent to the standard Neumann boundary condition $\nabla_{\mathbf{n}}\rho_t|_{\partial D} = 0$.

We now deduce (1.11) from (1.9). Firstly, by (1.10), (1.9) implies

$$\int_D (f\rho_t)(x)\mathrm{d}x = \int_D (f\rho_0)(x)\mathrm{d}x + \int_0^t \mathrm{d}s \int_D (fL^*_{s,\rho_s}\rho_s)(x)\mathrm{d}x, \quad f \in C_0^2(D), t \geq 0,$$

so that $\partial_t\rho_t = L^*_{t,\rho_t}\rho_t$. Next, by the integration by parts formula, (1.9) implies

$$\begin{aligned}
&\int_D (f\rho_t)(x)\mathrm{d}x = \int_D (f\rho_0)(x)\mathrm{d}x + \int_0^t \mathrm{d}s \int_D (\rho_s L_{s,\rho_s} f)(x)\mathrm{d}x \\
&= \int_D (f\rho_0)(x)\mathrm{d}x + \int_0^t \bigg(\int_D (fL^*_{s,\rho_s}\rho_s)(x)\mathrm{d}x + \int_{\partial D} \big\{ f\nabla_{\sigma_s\sigma_s^*\mathbf{n}}\rho_s - \rho_s\nabla_{\sigma_s\sigma_s^*\mathbf{n}} f \big\}(x)\mathrm{d}x \bigg)\mathrm{d}s \\
&= \int_D (f\rho_0)(x)\mathrm{d}x + \int_0^t \bigg(\int_D (f\partial_s\rho_s)(x)\mathrm{d}x + \int_{\partial D} \big\{ f\nabla_{\sigma_s\sigma_s^*\mathbf{n}}\rho_s + f\mathrm{div}_{\partial D}(\rho_s\pi\sigma_s\sigma_s^*\mathbf{n}) \big\}(x)\mathrm{d}x \bigg)\mathrm{d}s \\
&= \int_D (f\rho_t)(x)\mathrm{d}x + \int_0^t \mathrm{d}s \int_{\partial D} \big\{ f(\nabla_{t,\mathbf{n}}\rho_t) \big\}(x)\mathrm{d}x, \quad f \in C_N^2(\bar{D}), t \geq 0.
\end{aligned}$$

Thus, $\nabla_{t,\mathbf{n}}\rho_t|_{\partial D} = 0$.

1.3 Summary of main results

Theorems 2.1-2.3 provide sufficient conditions for the well-posedness and functional inequalities of reflecting SDEs with singular drifts. These results generalize the corresponding ones derived in recent years for singular SDEs without reflection, and improve some existing results for reflecting SDEs. The essential difficulty in the study of singular reflecting SDEs is explained in the beginning of Section 2.

Theorems 3.1-3.4 present the weak and strong well-posedness of the DDRSDE (1.5) under d-ifferent conditions, where the first result applies to locally integrable drifts with the distribution dependence bounded by $\|\cdot\|_{k,var} + \mathbb{W}_k$ (see Section 2 for definitions of probability distances), the second result includes a general criterion deducing the well-posedness of (1.5) from that of reflecting SDEs, and the last two results work for the monotone case with the dependence on distribution given by $\mathbb{W}_k (k > 1)$ or more general $\mathbb{W}_\psi$ induced by a cost function ψ.

Theorems 4.1 and 4.2 establish the log-Harnack inequality for solutions to (1.5) with respect to the initial distributions, which in particular implies the gradient estimate and entropy-cost inequality for the distributions of the solutions. The first result applies to the singular case and the other works for the monotone case.

To conclude this section, we consider an example of (1.11) arising from kinetic mechanics. For simplicity, we only consider bounded domain, but our general results also work for unbounded domains. See [35] for the study of exponential ergodicity.

Example 1.1 (Granular media equation with Neumann boundary). Let D be a bounded domain with $\partial D \in C_b^{2,L}$. For a potential $V : \bar{D} \to \mathbb{R}$ and an interaction functional $W : \mathbb{R}^d \to \mathbb{R}$, consider the following nonlinear PDE for probability density functions on $\bar{D}$:

$$\partial_t \varrho_t = \Delta \varrho_t + \mathrm{div}\big\{\varrho_t \nabla V + \varrho_t \nabla (W * \varrho_t)\big\}, \quad \nabla_{\mathbf{n}} \varrho_t|_{\partial D} = 0,$$

where $(W * \varrho_t)(x) := \int_{\mathbb{R}^d} W(x-z)\varrho_t(z)\mathrm{d}z$. It is easy to see that this equation is covered by (1.11) with

$$b(x,\mu) = -\nabla V(x) - \nabla (W * \mu)(x), \quad \sigma(x,\mu) = \sqrt{2}\mathbf{I}_d,$$

where $\mathbf{I}_d$ is the $d \times d$ identity matrix, and $(W * \mu)(x) := \int_{\mathbb{R}^d} W(x-z)\mu(\mathrm{d}z)$.

If V and W are weakly differentiable with $\|\nabla W\|_\infty < \infty$ and $|\nabla V| \in L^p(\bar{D})$ for some $p > d \vee 2$, then Theorem 3.1 with $k = 0$ implies that the associated SDE (1.5) is well-posed, and Theorem 4.2 provides some functional inequalities for the solution. These results apply to $W(x) := |x|^3$ which is of special interest from physics [5].

2 Reflecting SDE with singular drift

Let $\sigma_t(x,\mu) = \sigma_t(x)$ and $b_t(x,\mu) = b_t(x)$ do not depend on μ, so that (1.5) reduces to the following reflecting SDE on $\bar{D}$:

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t + \mathbf{n}(X_t)\mathrm{d}l_t, \quad t \in [0,T], \tag{2.1}$$

where $T > 0$ is a fixed time. The associated time dependent generator reads

$$L_t := \frac{1}{2}\mathrm{tr}\big\{\sigma_t \sigma_t^* \nabla^2\big\} + \nabla_{b_t}, \quad t \in [0,T]. \tag{2.2}$$

The problem of confining a stochastic process to a domain goes back to Skorokhod [26, 27], and has been well developed under monotone (or locally semi-Lipschitz) conditions, see the recent work [10] and references within. In this section, we solve (2.1) with a singular (unbounded on bounded sets) drift.

SDEs with singular coefficients have already been well investigated by using Zvokin's transform, see for instances [17, 37, 38, 41] and references within. However, the corresponding study for singular reflecting SDEs is very limited. With great effort overcoming difficulty induced by the local time, in the recent work [39] Yang and Zhang were able to prove the well-posedness of (2.1) for bounded C^3 domain, bounded b and $\sigma = \mathbf{I}_d$. So, the general setup we discussed here is new in the literature.

Before moving on, let us explain the main difficulty of the study by considering the following simple reflecting SDE on $\bar{D}$:

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t + \mathbf{n}(X_t)\mathrm{d}l_t, \quad t \in [0, T], \tag{2.3}$$

where W_t is the d-dimensional Brownian motion and $\int_0^T \|b_t\|_{L^p(\mathbb{R}^d)}^q \mathrm{d}t < \infty$ for some $p, q > 2$ with $\frac{d}{p} + \frac{2}{q} < 1$. When $\lambda > 0$ is large enough, the unique solution of the PDE

$$(\partial_t + \Delta + \nabla_{b_t})u_t = \lambda u_t - b_t, \quad t \in [0, T], u_T = 0$$

satisfies

$$\|u\|_\infty + \|\nabla u\|_\infty \le \frac{1}{2}, \quad \|\nabla^2 u\|_{L^p_q} := \left(\int_0^T \|\nabla^2 u_t\|_{L^p(\mathbb{R}^d)}^q \mathrm{d}t \right)^{\frac{1}{q}} < \infty,$$

see [17, 41]. Thus, for any $t \in [0, T]$, $\Theta_t := id + u_t$ (id is the identity map) is a homeomorphism on $\mathbb{R}^d$, and by Itô's formula, $Y_t := \Theta_t(X_t)$ solves

$$\mathrm{d}Y_t = \lambda\{u_t \circ \Theta_t^{-1}\}(Y_t)\mathrm{d}t + \mathrm{d}W_t + \{(\nabla u_t) \circ \Theta_t^{-1}\}(Y_t)\mathrm{d}W_t + \{\mathbf{n}(X_t) + \nabla_{\mathbf{n}} u_t(X_t)\}\mathrm{d}l_t.$$

When $D = \mathbb{R}^d$, we have $l_t = 0$ so that this SDE is regular enough to have well-posedness, which implies the same property of (2.3) since Θ_t is a homeomorphism, see [17]. When $D \neq \mathbb{R}^d$, to prove the pathwise uniqueness of Y_t by applying Itô's formula to $|Y_t - \tilde{Y}_t|^2$, where $\tilde{Y}_t := \Theta_t(\tilde{X}_t)$ for another solution $\tilde{X}_t$ of (2.3) with local time $\tilde{l}_t$, one needs to find a constant $c > 0$ such that

$$\begin{aligned} &\langle \Theta_t(X_t) - \Theta_t(\tilde{X}_t), (\mathbf{n} + \nabla_{\mathbf{n}} u_t)(X_t)\rangle \mathrm{d}l_t + \langle \Theta_t(\tilde{X}_t) - \Theta_t(X_t), (\mathbf{n} + \nabla_{\mathbf{n}} u_t)(\tilde{X}_t))\rangle \mathrm{d}\tilde{l}_t \\ &\le c|X_t - \tilde{X}_t|^2(\mathrm{d}l_t + \mathrm{d}\tilde{l}_t). \end{aligned} \tag{2.4}$$

This is not implied by (1.2) except for $d = 1$, since only in this case the vectors $\Theta_t(x) - \Theta_t(y)$ and $(\mathbf{n} + \nabla_{\mathbf{n}} u_t)(x)$ are in the same directions of $x - y$ and $\mathbf{n}(x)$ respectively for large $\lambda > 0$.

To overcome this difficulty, we will construct a Zvokin's transform by solving the associated Neumann problem on $\bar{D}$, for which $\nabla_{\mathbf{n}} u_t|_{\partial D} = 0$. Even in this case, Θ_t may also map a point from $\bar{D}$ to $\bar{D}^c$ such that (1.2) does not apply. To this end, we will construct a modified process of $|X_t - \tilde{X}_t|^2$ by using a function from [9]. Our construction simplifies that in [39] and enables us to work in a more general framework.

2.1 Conditions and main results

We first recall some functional spaces used in the study of singular SDEs, see for instance [37]. For any $p \ge 1$, $L^p(\mathbb{R}^d)$ is the class of measurable functions f on $\mathbb{R}^d$ such that

$$\|f\|_{L^p(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} |f(x)|^p \mathrm{d}x \right)^{\frac{1}{p}} < \infty.$$

For any $\epsilon > 0$ and $p \geq 1$, let $H^{\epsilon,p}(\mathbb{R}^d) := (1-\Delta)^{-\frac{\epsilon}{2}} L^p(\mathbb{R}^d)$ with

$$\|f\|_{H^{\epsilon,p}(\mathbb{R}^d)} := \|(1-\Delta)^{\frac{\epsilon}{2}} f\|_{L^p(\mathbb{R}^d)} < \infty, \quad f \in H^{\epsilon,p}(\mathbb{R}^d).$$

For any $z \in \mathbb{R}^d$ and $r > 0$, let $B(z,r) := \{x \in \mathbb{R}^d : |x-z| < r\}$ be the open ball centered at z with radius r. For any $p, q > 1$ and $t_0 < t_1$, let $\tilde{L}_q^p(t_0,t_1)$ denote the class of measurable functions f on $[t_0,t_1] \times \mathbb{R}^d$ such that

$$\|f\|_{\tilde{L}_q^p(t_0,t_1)} := \sup_{z\in\mathbb{R}^d} \left(\int_{t_0}^{t_1} \|1_{B(z,1)} f_t\|_{L^p(\mathbb{R}^d)}^q \mathrm{d}t \right)^{\frac{1}{q}} < \infty.$$

For any $\epsilon > 0$, let $\tilde{H}_q^{\epsilon,p}(t_0,t_1)$ be the space of $f \in \tilde{L}_q^p(t_0,t_1)$ with

$$\|f\|_{\tilde{H}_q^{\epsilon,p}(t_0,t_1)} := \sup_{z\in\mathbb{R}^d} \left(\int_{t_0}^{t_1} \|g(z+\cdot) f_t\|_{\mathbb{H}^{\epsilon,p}(\mathbb{R}^d)}^q \mathrm{d}t \right)^{\frac{1}{q}} < \infty$$

for some $g \in C_0^\infty(\mathbb{R}^d)$ satisfying $g|_{B(0,1)} = 1$, where $C_0^\infty(\mathbb{R}^d)$ is the class of C^∞ functions on $\mathbb{R}^d$ with compact support. We remark that the space $\tilde{H}_q^{\epsilon,p}(t_0,t_1)$ does not depend on the choice of g. When $t_0 = 0$, we simply denote

$$\tilde{L}_q^p(t_1) := \tilde{L}_q^p(0,t_1), \quad \tilde{H}_q^{\epsilon,p}(t_1) := \tilde{H}_q^{\epsilon,p}(0,t_1), \quad t_1 > 0.$$

For a domain $D \subset \mathbb{R}^d$, we denote $f \in \tilde{L}_q^p(t_0,t_1,D) (=: \tilde{L}_q^p(t_1,D)$ for $t_0 = 0)$, if f is a measurable function on $[t_0,t_1] \times \bar{D}$ such that

$$\|f\|_{\tilde{L}_q^p(t_0,t_1,D)} := \|1_D f\|_{\tilde{L}_q^p(t_0,t_1)} < \infty.$$

A vector or matrix valued function is said in one of the above introduced spaces, if so are its components.

We will take (p,q) from the class

$$\mathscr{K} := \Big\{ (p,q) : p,q \in (1,\infty),\ \frac{d}{p} + \frac{2}{q} < 1 \Big\},$$

and use the following assumptions on the coefficients b and σ. Let $\|\cdot\|_\infty$ denote the uniform norm for real (or vector/matrix) valued functions.

$(A_0^{\sigma,b})$ **(D)** *holds,* $a := \sigma\sigma^*$ *and* b *are extended to measurable functions on* $[0,T] \times \mathbb{R}^d$*,* b *has decomposition* $b = b^{(0)} + b^{(1)}$ *with* $b_t^{(0)}|_{\bar{D}^c} = 0$*, such that the following conditions hold:*

(1) a_t *is invertible with* $\|a\|_\infty + \|a^{-1}\|_\infty < \infty$*, and*

$$\lim_{\varepsilon\to 0} \sup_{|x-y|\leq\varepsilon, t\in[0,T]} \|a_t(x) - a_t(y)\| = 0. \tag{2.5}$$

(2) *There exists* $(p_0,q_0)\in\mathscr{K}$ *such that* $|b^{(0)}|\in\tilde L^{p_0}_{q_0}(T)$. *Moreover,* $b^{(1)}$ *is locally bounded on* $[0,T]\times\mathbb{R}^d$, *and there exist a constant* $L>1$ *and a function* $\tilde\rho\in C_b^2(\bar D)$ *such that*

$$\|\nabla b^{(1)}\|_\infty:=\sup_{t\in[0,T],x\neq y}\frac{|b_t^{(1)}(x)-b_t^{(1)}(y)|}{|x-y|}\le L, \tag{2.6}$$

$$\langle b_t^{(1)},\nabla\tilde\rho\rangle|_{\bar D}\ge -L,\quad \langle\nabla\tilde\rho,\mathbf{n}\rangle|_{\partial D}\ge 1,\quad t\in[0,T]. \tag{2.7}$$

$(A_1^{\sigma,b})$ $(A_0^{\sigma,b})$ *holds, and there exist* $\{(p_i,q_i)\}_{0\le i\le l}\subset\mathscr{K}$ *and* $0\le f_i\in\tilde L^{p_i}_{q_i}(T), 1\le i\le l$, *such that*

$$|b^{(0)}|^2\in\tilde L^{p_0}_{q_0}(T),\quad \|\nabla\sigma\|^2\le\sum_{i=1}^l f_i.$$

Remark 2.1. Each of the following two conditions implies the existence of $\tilde\rho$ in (2.7):

(a) $\partial D\in C_b^2$ and there exists a constant $K>0$ such that $\langle b_t^{(1)},\mathbf{n}\rangle|_{\partial D}\ge -K$ for $t\in[0,T]$;

(b) D is bounded and there exist $\varepsilon\in(0,1)$ and $x_0\in D$ such that

$$\langle x_0-x,\mathbf{n}(x)\rangle\ge\varepsilon|x-x_0|,\quad x\in\partial D. \tag{2.8}$$

Indeed, if (a) holds then there exists $r_0>0$ such that $\rho\in C_b^2(\partial_{r_0}D)$. Let $h\in C^\infty([0,\infty))$ with $h(r)=r$ for $r\in[0,r_0/4]$ and $h(r)=r_0/2$ for $r\ge r_0/2$. By taking $\tilde\rho=h\circ\rho$ we have $\tilde\rho\in C_b^2(\bar D)$, $\langle\nabla\tilde\rho,\mathbf{n}\rangle|_{\partial D}=1$, and for any $x\in D$ letting $\bar x\in\partial D$ such that $|x-\bar x|=\rho(x)$, we deduce from (2.6) and $h'(\rho(x))=0$ for $\rho(x)\ge r_0/2$ that

$$\langle b_t^{(1)}(x),\nabla\tilde\rho(x)\rangle=h'(\rho(x))\big\{\langle b_t^{(1)}(\bar x),\mathbf{n}(\bar x)\rangle+\langle b_t^{(1)}(x)-b_t^{(1)}(\bar x),\mathbf{n}(\bar x)\rangle\big\}\ge-(1+r_0)L\|h'\|_\infty.$$

Therefore, (2.7) holds for some (different) constant L. Next, if (b) holds, by (2.8) we may take $\tilde\rho(x)=N\sqrt{1+|x-x_0|^2}$ for large enough $N\ge1$ such that $\langle\nabla\tilde\rho,\mathbf{n}\rangle|_{\partial D}\ge1$. So, by the boundedness of D and $b^{(1)}\in C([0,T]\times\mathbb{R}^d)$, (2.7) holds for some constant $L>0$.

Assumption $(A_0^{\sigma,b})$ will be used to establish Krylov's estimate for functions $f\in\cap_{(p,q)\in\mathscr{K}}\tilde L^p_q(T)$, which is crucial to solve singular SDEs, see Lemma 2.5 below. To improve this estimate for (p,q) satisfying $\frac{d}{p}+\frac{2}{q}<2$ as in the case without reflecting (see [37]), we introduce one more assumption.

Consider the following differential operators on $\bar D$:

$$L_t^{\sigma,b^{(1)}}:=\frac12\mathrm{tr}\big(\sigma_t\sigma_t^*\nabla^2\big)+\nabla_{b_t^{(1)}},\quad t\in[0,T]. \tag{2.9}$$

Let $\{P_{s,t}^{\sigma,b^{(1)}}\}_{T\ge t_1\ge t\ge s\ge0}$ be the Neumann semigroup on $\bar D$ generated by $L_t^{\sigma,b^{(1)}}$, that is, for any $\phi\in C_b^2(\bar D)$, and any $t\in(0,T]$, $(P_{s,t}^{\sigma,b^{(1)}}\phi)_{s\in[0,t]}$ is the unique solution of the PDE

$$\partial_s u_s=-L_s^{\sigma,b^{(1)}}u_s,\quad \nabla_{\mathbf{n}}u_s|_{\partial D}=0\text{ for }s\in[0,t),u_t=\phi. \tag{2.10}$$

For any $t>0$, let $C_b^{1,2}([0,t]\times\bar D)$ be the set of functions $f\in C_b([0,t]\times\bar D)$ with bounded and continuous derivatives $\partial_t f,\nabla f$ and $\nabla^2 f$.

$(A_2^{\sigma,b})$ $\partial D \in C_b^{2,L}$ *and the following conditions hold for* σ *and* b *on* $[0,T]\times\bar D$:

(1) $a_t := \sigma_t\sigma_t^*$ *is invertible,* (2.5) *holds for* $x,y\in\bar D$ *and there exist* $\{(p_i,q_i)\}_{0\le i\le l}\subset\mathscr{K}$ *with* $p_i>2$ *and* $0\le f_i\in\tilde L_{q_i}^{p_i}(T), 1\le i\le l$, *such that*

$$\|\nabla\sigma\|\le\sum_{i=1}^l f_i,\quad \|a\|_\infty+\|a^{-1}\|_\infty+\|\nabla\sigma\|_{\tilde L_{q_1}^{p_1}(T,D)}<\infty.$$

(2) $b=b^{(1)}+b^{(0)}$ *with* $\nabla_{\mathbf n}b_t^{(1)}|_{\partial D}=0$, $\|\nabla b^{(1)}\|_\infty+\|1_{\partial D}\langle b^{(1)},\mathbf n\rangle\|_\infty<\infty$ *and* $|b^{(0)}|\in\tilde L_{q_0}^{p_0}(T,D)$ *for some* $(p_0,q_0)\in\mathscr{K}$ *with* $p_0>2$.

(3) *For any* $\phi\in C_b^2(\bar D)$ *and* $t\in(0,T]$, *the PDE* (2.10) *has a unique solution* $P_{\cdot,t}^{\sigma,b^{(1)}}\phi\in C^{1,2}([0,t)\times\bar D)$, *such that for some constant* $c>0$ *we have* $(\nabla^0\phi:=\phi)$

$$\begin{aligned}&\|\nabla^i P_{s,t}^{\sigma,b^{(1)}}\phi\|_\infty\le c(t-s)^{-\frac12}\|\nabla^{i-1}\phi\|_\infty,\quad 0\le s<t\le T,\ i=1,2\\ &\|\partial_s P_{s,t}^{\sigma,b^{(1)}}\phi\|_\infty\le c(t-s)^{-\frac12}\|\nabla\phi\|_\infty,\quad 0\le s<t\le T.\end{aligned}\tag{2.11}$$

Remark 2.2. (1) Let $\rho\in C_b^2(\partial_{r_0}D)$ for some $r_0>0$. Since $\nabla\rho|_{\partial D}=\mathbf n$, $\|\nabla b^{(1)}\|_\infty+\|1_{\partial D}\langle b^{(1)},\mathbf n\rangle\|_\infty<\infty$ implies $\|1_{\partial_{r_0}D}\langle b^{(1)},\nabla\rho\rangle\|_\infty<\infty$, which will be used in the proof of Lemma 2.6 below.

(2) $(A_2^{\sigma,b})(3)$ holds if D is bounded with $\partial D\in C^{2+\alpha}$ for some $\alpha\in(0,1)$, and there exists $c>0$ such that

$$\big\{|b_t^{(1)}(x)-b_s^{(1)}(y)|+\|a_t(x)-a_s(y)\|\big\}\le c(|t-s|^\alpha+|x-y|^{\frac\alpha2}),\quad s,t\in[0,T],x,y\in\bar D.\tag{2.12}$$

Indeed, $\partial D\in C^{2+\alpha}$ implies $\mathbf n\in C^{1+\alpha}(\partial D)$, so that (2.12) implies estimates (3.4) and (3.6) in [8, Theorem VI.3.1] with $\varrho=\infty$ for the Neumann heat kernel $p_{s,t}^{\sigma,b^{(1)}}(x,y)$ of $P_{s,t}^{\sigma,b^{(1)}}$. We note that according to its proof, the condition (3.3) therein is assumed for some $\alpha\in(0,1)$ rather than all $\alpha\in(0,1)$. In particular, $\nabla^2 p_{s,t}^{\sigma,b^{(1)}}(\cdot,y)(x)$ and $\partial_s p_{s,t}^{\sigma,b^{(1)}}(x,y)$ are continuous in $(s,x)\in[0,t]\times\bar D$, and there exists a constant $c>1$ such that

$$\begin{aligned}&|\nabla^i p_{s,t}^{\sigma,b^{(1)}}(\cdot,y)(x)|\le c|t-s|^{-\frac{d+i}{2}}\mathrm e^{-\frac{|x-y|^2}{c(t-s)}},\quad 0\le s<t\le T,x,y\in\bar D,i=0,1,2,\\ &|\partial_s p_{s,t}^{\sigma,b^{(1)}}(x,y)|=|L_s^{\sigma,b^{(1)}}p_{s,t}^{\sigma,b^{(1)}}(\cdot,y)(x)|\le c|t-s|^{-\frac{d+2}{2}}\mathrm e^{-\frac{|x-y|^2}{c(t-s)}},\quad 0\le s<t\le T,x,y\in\bar D.\end{aligned}$$

These properties imply (2.11). For instance, by $\int_D p_{s,t}(x,y)\mathrm dy=1$, the second estimate implies that for some constant $c'>0$,

$$\begin{aligned}|\partial_s P_{s,t}^{\sigma,b^{(1)}}\phi(x)|&=\bigg|\partial_s\int_D p_{s,t}^{\sigma,b^{(1)}}(x,y)\phi(y)\mathrm dy\bigg|=\bigg|\partial_s\int_D p_{s,t}^{\sigma,b^{(1)}}(x,y)\big\{\phi(y)-\phi(x)\big\}\mathrm dy\bigg|\\ &\le c\|\nabla\phi\|_\infty\int_D|x-y|\cdot|t-s|^{-\frac{d+2}{2}}\mathrm e^{-\frac{|x-y|^2}{c(t-s)}}\mathrm dy\le c'(t-s)^{-\frac12},\quad 0\le s<t\le T,x\in\bar D.\end{aligned}$$

When $D=\mathbb R^d$, these estimates (hence (2.11)) hold for more general σ and $b^{(1)}$, see [22].

The following are main results of this section, where Theorem 2.2 improves the main result (Theorem 6.3) in [39] for bounded C^3 domain D, bounded drift b and $\sigma=\mathbf I_d$. Moreover, going back to the case without reflection (i.e. $D=\mathbb R^d$), Theorem 2.3 covers the main result (Theorem 1.1) of [18] where $b^{(1)}=0$ is considered.

Theorem 2.1 (Weak well-posedness). *If either* $(A_1^{\sigma,b})$ *or* $(A_2^{\sigma,b})$ *holds, then* (2.1) *is weakly well-posed. Moreover, for any* $k \geq 1$ *there exists a constant* $c > 0$ *such that*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t^x|^k\Big] \leq c(1+|x|^k), \quad \mathbb{E}e^{kl_T^x} \leq c, \quad x \in \bar{D}, \tag{2.13}$$

where (X_t^x, l_t^x) *is the (weak) solution of* (2.1) *with* $X_0^x = x$.

Theorem 2.2 (Well-posedness). *Assume that one of the following conditions holds:*

(i) $d = 1$ *and* $(A_1^{\sigma,b})$ *holds;*

(ii) $(A_2^{\sigma,b})$ *holds.*

Then (2.1) *is well-posed, and for any* $k \geq 1$*, there exists a constant* $c > 0$ *such that*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t^x - X_t^y|^k\Big] \leq c|x-y|^k, \quad x, y \in \bar{D}. \tag{2.14}$$

Consequently, for any $p > 1$ *there exists a constant* $c(p) > 0$ *such that* $P_t f(x) := \mathbb{E}[f(X_t^x)]$ *satisfies*

$$|\nabla P_t f(x)| := \limsup_{y\to x} \frac{|P_t f(y) - P_t f(x)|}{|x-y|} \leq c(p)(P_t|\nabla f|^p(x))^{\frac{1}{p}}, \quad f \in C_b^1(\bar{D}), \quad t \in [0,T]. \tag{2.15}$$

Theorem 2.3 (Functional inequalities). *Assume that* $(A_2^{\sigma,b})$ *holds with* $p_1 > 2$. *Then there exist a constant* $C > 0$ *and a map* $c : (1,\infty) \to (0,\infty)$ *such that*

$$|\nabla P_t f| \leq \frac{c(p)}{\sqrt{t}}(P_t|f|^p)^{\frac{1}{p}}, \quad t \in [0,T], f \in \mathscr{B}_b(\bar{D}), \; p > 1, \tag{2.16}$$

$$P_t f^2 - (P_t f)^2 \leq tCP_t|\nabla f|^2, \quad f \in C_b^1(\bar{D}), \quad t \in [0,T], \tag{2.17}$$

$$P_t \log f(x) \leq \log P_t f(y) + \frac{C|x-y|^2}{t}, \quad t \in [0,T], x, y \in \bar{D}, 0 < f \in \mathscr{B}_b(\bar{D}). \tag{2.18}$$

To prove these results, we first establish Krylov's estimates under different conditions, then prove the weak and strong well-posedness by using Girsanov's transform and Zvokin's transforms respectively.

2.2 Krylov's estimate and Itô's formula

A crucial step in the study of singular SDEs is to establish Krylov's estimate [16]. To this end, we first introduce the following lemma taken from [40, Theorem 2.1], which extends [37, Theorem 3.2] where $b^{(1)} = 0$ is considered. See [37, 41] and references within for earlier assertions.

Lemma 2.4. *Assume* $(A_0^{\sigma,b})$*. For any* $0 \le t_0 < t_1 \le T$ *and* $f \in \tilde L_q^p(t_0,t_1)$ *for some* $p,q>1$*, the PDE*

$$(\partial_t + L_t)u_t^\lambda = \lambda u_t^\lambda + f_t, \quad t\in[t_0,t_1], u_{t_1}^\lambda = 0, \tag{2.19}$$

has a unique solution in $\tilde H_q^{2,p}(t_0,t_1)$*. Moreover, for any* $\theta\in[0,2), p'\in[p,\infty]$ *and* $q'\in[q,\infty]$ *with* $\frac{d}{p}+\frac{2}{q} < 2-\theta+\frac{d}{p'}+\frac{2}{q'}$*, there exist constants* $\lambda_0, c>0$ *increasing in* $\|b^{(0)}\|_{\tilde L_{q_0}^{p_0}(T)}$ *(i.e. they do not have to be changed when* $b^{(0)}$ *is replaced by* $\tilde b^{(0)}$ *with* $\|\tilde b^{(0)}\|_{\tilde L_{q_0}^{p_0}(T)} \le \|b^{(0)}\|_{\tilde L_{q_0}^{p_0}(T)}$*), such that for any* $\lambda\ge\lambda_0$ *and* $0\le t_0<t_1\le T, \lambda\ge\lambda_0$ *and* $f\in\tilde L_q^p(t_0,t_1)$*, the solution satisfies*

$$\lambda^{\frac12(2-\theta+\frac{d}{p'}+\frac{1}{q'}-\frac{d}{p}-\frac{2}{q})}\|u^\lambda\|_{\tilde H_{q'}^{\theta,p'}(t_0,t_1)} + \|(\partial_t+\nabla_{b^{(1)}})u^\lambda\|_{\tilde L_q^p(t_0,t_1)} + \|u^\lambda\|_{\tilde H_q^{2,p}(t_0,t_1)} \le c\|f\|_{\tilde L_q^p(t_0,t_1)}.$$

By estimating the local time, this result enables us to derive the following Krylov's estimate (2.20) and Khasminskii's estimate (2.21).

Lemma 2.5. *Assume* $(A_0^{\sigma,b})$*. Let* $(p,q)\in\mathscr K$*.*

(1) *There exist a constant* $i\ge1$ *depending only on* (p,q)*, and a constant* $c\ge1$ *increasing in* $\|b^{(0)}\|_{\tilde L_{q_0}^{p_0}(T)}$*, such that for any solution* X_t *of* (2.1)*, and any* $0\le t_0\le t_1\le T$*, the following estimates hold.*

$$\mathbb E\bigg[\bigg(\int_{t_0}^{t_1}|f_s(X_s)|\mathrm ds\bigg)^j\bigg|\mathscr F_{t_0}\bigg] \le c^j j!\|f\|_{\tilde L_q^p(t_0,t_1)}^j, \quad f\in\tilde L_q^p(t_0,t_1), j\ge1, \tag{2.20}$$

$$\mathbb E\big(\mathrm e^{\int_{t_0}^{t_1}|f_t(X_t)|\mathrm dt}\big|\mathscr F_{t_0}\big) \le \exp\big[c+c\|f\|_{\tilde L_q^p(t_0,t_1)}^i\big], \quad f\in\tilde L_q^p(t_0,t_1), \tag{2.21}$$

$$\sup_{t_0\in[0,T]}\mathbb E\big(\mathrm e^{\lambda(l_T-l_{t_0})}\big|\mathscr F_{t_0}\big) < \mathrm e^{c(1+\lambda^2)}, \quad \lambda>0. \tag{2.22}$$

(2) *For any* $u\in C([0,T]\times\mathbb R^d)$ *with continuous* ∇u *and*

$$\|u\|_\infty + \|\nabla u\|_\infty + \|(\partial_t+\nabla_{b^{(1)}})u\|_{\tilde L_q^p(T)} + \|\nabla^2 u\|_{\tilde L_q^p(T)} < \infty, \tag{2.23}$$

we have the following Itô's formula for a solution X_t *to* (2.1)*:*

$$\mathrm du_t(X_t) = (\partial_t+L_t)u_t(X_t)\mathrm dt + \langle\nabla u_t(X_t), \sigma_t(X_t)\mathrm dW_t\rangle + (\nabla_{\mathbf n}u_t)(X_t)\mathrm dl_t. \tag{2.24}$$

Proof. (1) We first prove (2.20) for $j=1$. By first using $(|f|\wedge n)1_{B(0,n)}$ replacing f then letting $n\to\infty$, we may and do assume that f is bounded with compact support. Next, by a standard approximation argument, we only need to prove for $f\in C_0^\infty([t_0,t_1]\times\mathbb R^d)$.

Let $f\in C_0^\infty([t_0,t_1]\times\mathbb R^d)$. By Lemma 2.4, for any $(p',q')\in\mathscr K$, (2.19) has a unique solution satisfying

$$\begin{aligned}&\lambda^\varepsilon\big(\|u^\lambda\|_\infty+\|\nabla u^\lambda\|_\infty\big) + \|(\partial_t+\nabla_{b^{(1)}})u^\lambda\|_{\tilde L_{q'}^{p'}(t_0,t_1)} + \|u^\lambda\|_{\tilde H_{q'}^{2,p'}(t_0,t_1)}\\ &\le c_1\|f\|_{\tilde L_{q'}^{p'}(t_0,t_1)}, \quad \lambda\ge\lambda_0,\end{aligned} \tag{2.25}$$

where $\varepsilon > 0$ depends on (p', q') and $\lambda_0, c > 0$ are constants increasing in $\|b^{(0)}\|_{\tilde L^{p_0}_{q_0}(T)}$. To apply Itô's formula, we make a standard mollifying approximation of u^λ, which is extended to $\mathbb{R}^{d+1}$ by letting $u^\lambda_t := u^\lambda_{(t\vee t_0)\wedge t_1}$ for $t \in \mathbb{R}$. Let $0 \le \varrho \in C_0^\infty(\mathbb{R}^{d+1})$ such that $\int_{\mathbb{R}^{d+1}} \varrho(z)\mathrm{d}z = 1$. For any $n \ge 1$, let

$$u^{\lambda,n}_t(x) = n^{d+1} \int_{\mathbb{R}^{d+1}} u^\lambda_{t-s}(x-y)\varrho(ns, ny)\mathrm{d}s\mathrm{d}y, \quad t \in \mathbb{R}, x \in \mathbb{R}^d. \tag{2.26}$$

Then

$$\lim_{n\to\infty} \big\{ \|(\partial_t + \nabla_{b^{(1)}})(u^{\lambda,n} - u^\lambda)\|_{\tilde L^{p'}_{q'}(t_0,t_1)} + \|u^{\lambda,n} - u^\lambda\|_{\tilde H^{2,p'}_{q'}(t_0,t_1)} \big\} = 0, \quad (p', q') \in \mathscr{K},$$

so that as shown in the proof of [38, Lemma 5.4],

$$f^{\{n\}}_t := (\partial_t + L_t - \lambda) u^{\lambda,n}_t \tag{2.27}$$

satisfies

$$\lim_{n\to\infty} \|f - f^{\{n\}}\|_{\tilde L^{p'}_{q'}(t_0,t_1)} = 0, \quad (p', q') \in \mathscr{K}, \tag{2.28}$$

and (2.25) with $(p', q') = (p, q)$ implies

$$\|u^{\lambda,n}\|_\infty + \|\nabla u^{\lambda,n}\|_\infty \le c\lambda^{-\varepsilon}\|f\|_{\tilde L^p_q(t_0,t_1)}, \quad n \ge 1, \lambda > \lambda_0. \tag{2.29}$$

By Theorem 6.2.7(ii)-(iii) in [6], the conditional distribution of X_t under P_{t_0} is absolutely continuous for $t > t_0$, so that by the dominated convergence theorem, (2.28) implies $\mathbb{P}$-a.s.

$$\mathbb{E}\bigg(\int_{t_0}^{t_1\wedge\tau_k} f_s(X_s)\mathrm{d}s \bigg| \mathscr{F}_{t_0} \bigg) = \lim_{n\to\infty} \mathbb{E}\bigg(\int_{t_0}^{t_1\wedge\tau_k} f^{\{n\}}_s(X_s)\mathrm{d}s \bigg| \mathscr{F}_{t_0} \bigg), \tag{2.30}$$

where

$$\tau_k := \inf\bigg\{ t \in [t_0, T] : l_t - l_{t_0} + \int_{t_0}^t |b_s(X_s)|\mathrm{d}s \ge k \bigg\}, \quad k \ge 1.$$

Applying Itô's formula to $u^{\lambda,n}$, we deduce from (2.27) and (2.29) that

$$\begin{aligned} &2c\lambda^{-\varepsilon}\|f\|_{\tilde L^p_q(t_0,t_1)} \ge \mathbb{E}\big\{ u^{\lambda,n}_{t_1\wedge\tau_k}(X_{t_1\wedge\tau_k}) - u^{\lambda,n}_{t_0}(X_{t_0}) \big| \mathscr{F}_{t_0} \big\} \\ &= \mathbb{E}\bigg(\int_{t_0}^{t_1\wedge\tau_k} (\partial_s + L_s) u^{\lambda,n}_s(X_s)\mathrm{d}s + \int_{t_0}^{t_1\wedge\tau_k} \{\nabla_{\mathbf{n}(X_s)} u^{\lambda,n}_s\}(X_s)\mathrm{d}l_s \bigg| \mathscr{F}_{t_0} \bigg) \\ &\ge \mathbb{E}\bigg(\int_{t_0}^{t_1\wedge\tau_k} f^{\{n\}}_s(X_s)\mathrm{d}s \bigg| \mathscr{F}_{t_0} \bigg) - c\|f\|_{\tilde L^p_q(t_0,t_1)} \big\{ \lambda + \lambda^{-\varepsilon}\mathbb{E}(l_{t_1\wedge\tau_k} - l_{t_0}|\mathscr{F}_{t_0}) \big\}. \end{aligned} \tag{2.31}$$

Therefore,

$$\begin{aligned} &\mathbb{E}\bigg(\int_{t_0}^{t_1\wedge\tau_k} f^{\{n\}}_s(X_s)\mathrm{d}s \bigg| \mathscr{F}_{t_0} \bigg) \\ &\le c\|f\|_{\tilde L^p_q(t_0,t_1)} \big\{ 2\lambda^{-\varepsilon} + \lambda + \lambda^{-\varepsilon}\mathbb{E}(l_{t_1\wedge\tau_k} - l_{t_0}|\mathscr{F}_{t_0}) \big\}, \quad n, k \ge 1, \lambda > 0. \end{aligned} \tag{2.32}$$

Combining this with (2.30), we obtain

$$
\begin{aligned}
(2.33)\qquad &\mathbb{E}\bigg(\int_{t_0}^{t_1\wedge\tau_k} f_s(X_s)\mathrm{d}s\bigg|\mathscr{F}_{t_0}\bigg) = \lim_{n\to\infty}\mathbb{E}\bigg(\int_{t_0}^{t_1\wedge\tau_k} f_s^{\{n\}}(X_s)\mathrm{d}s\bigg|\mathscr{F}_{t_0}\bigg)\\
&\le c\|f\|_{\tilde L_q^p(t_0,t_1)}\big\{2+\lambda+\lambda^{-\varepsilon}\mathbb{E}(l_{t_1}-l_{t_0}|\mathscr{F}_{t_0})\big\},\quad \lambda>0, k\ge 1.
\end{aligned}
$$

On the other hand, by (2.7) and the boundedness of σ, we find a constant $c_1>0$ such that

$$
(2.34)\qquad \mathrm{d}\tilde\rho(X_t)\ge -c_1\mathrm{d}t - c_1|b_t^{(0)}(X_t)|\mathrm{d}t+\mathrm{d}l_t+\langle\nabla\tilde\rho(X_t),\sigma_t(X_t)\mathrm{d}W_t\rangle.
$$

So, (2.33) with $(p,q)=(p_0,q_0)$ implies

$$
\begin{aligned}
&\mathbb{E}(l_{t_1\wedge\tau_k}-l_{t_0}|\mathscr{F}_{t_0})\le c_1(t-t_0)+c_1\mathbb{E}\bigg(\int_{t_0}^{t_1\wedge\tau_k}|b_s^{(0)}(X_s)|\mathrm{d}s\bigg|\mathscr{F}_{t_0}\bigg)+\|\tilde\rho\|_\infty\\
&\le c_2(1+\lambda)+c_2\lambda^{-\varepsilon}\mathbb{E}(l_{t_1\wedge\tau_k}-l_{t_0}|\mathscr{F}_{t_0}),\quad t\in[t_0,T],\quad \lambda>0, k\ge 1
\end{aligned}
$$

for some constant $c_2>0$ increasing in $\|b^{(0)}\|_{\tilde L_q^p(T)}$. Taking $\lambda>0$ large enough such that $c_2\lambda^{-\varepsilon}\le\frac{1}{2}$, we arrive at

$$
\mathbb{E}(l_{t_1\wedge\tau_k}-l_{t_0}|\mathscr{F}_{t_0})\le c_3,\quad k\ge 1
$$

for some constant $c_3>0$ increasing in $\|b^{(0)}\|_{\tilde L_q^p(T)}$. Letting $k\to\infty$ gives

$$
(2.35)\qquad \mathbb{E}(l_{t_1}-l_{t_0}|\mathscr{F}_{t_0})\le c_3,\quad t_0\le t_1\le T.
$$

This and (2.33) with $k\to\infty$ imply (2.20) for $j=1$, which further yields the inequality for any $j\ge 1$ as shown in the proof of [38, Lemma 3.5]. Moreover, taking $q'\in(2,q)$ such that $(p,q')\in\mathscr{K}$, (2.20) for $j=1$ with (p,q') replacing (p,q) yields

$$
\mathbb{E}\bigg(\int_{t_0}^{t_1} f_s(X_s)\mathrm{d}s\bigg|\mathscr{F}_{t_0}\bigg)\le c\|f\|_{\tilde L_{q'}^p(t_0,t_1)}\le c(t_1-t_0)^{\frac{q-q'}{qq'}}\|f\|_{\tilde L_q^p(t_0,t_1)}.
$$

This and [38, Lemma 3.5] with $\tilde L_{q'}^p$ replacing L_q^p imply (2.21) for $i=\frac{q}{q-q'}$. Finally, combining (2.21) with (2.34), $b^{(0)}\in\tilde L_{q_0}^{p_0}(T)$ and $\|\sigma^*\nabla\tilde\rho\|_\infty<\infty$, we derive (2.22).

(2) We first extend u to $\mathbb{R}^{d+1}$ by letting $u_t=u_0$ for $t\le 0$, and consider its mollifying approximation $u^{\{n\}}$ defined above. Then $\|\sigma\|_\infty<\infty$ and (2.23) imply

$$
(2.36)\qquad \lim_{n\to\infty}\big\{\|u-u^{\{n\}}\|_\infty+\|\nabla(u-u^{\{n\}})\|_\infty+\|(\partial_t+L_t)(u-u^{\{n\}})\|_{\tilde L_q^p(T)}\big\}=0.
$$

Combining this with $\|\sigma\|_\infty<\infty$ and (2.20), we obtain

$$
(2.37)\qquad
\begin{aligned}
&\lim_{n\to\infty}\sup_{t\in[0,T]}|u_t^{\{n\}}(X_t)-u_t(X_t)|=0,\ \mathbb{P}\text{-a.s.}\\
&\lim_{n\to\infty}\int_0^t\nabla_{\mathbf{n}}u_s^{\{n\}}(X_s)\mathrm{d}l_s=\int_0^t\nabla_{\mathbf{n}}u_s(X_s)\mathrm{d}l_s,\ \mathbb{P}\text{-a.s.}\\
&\lim_{n\to\infty}\mathbb{E}\int_0^T\big|(\partial_s+L_s)(u_s^{\{n\}}-u_s)\big|(X_s)\mathrm{d}s=0,\\
&\lim_{n\to\infty}\mathbb{E}\sup_{t\in[0,T]}\bigg|\int_0^t\langle\nabla(u_s^{\{n\}}-u_s)(X_s),\sigma_s(X_s)\mathrm{d}W_s\rangle\bigg|=0.
\end{aligned}
$$

Therefore, we prove (2.24) by letting $n \to \infty$ in the following Itô's formula:

$$u_t^{\{n\}}(X_t) = u_0^{\{n\}}(X_0) + \int_0^t (\partial_s + L_s)(u_s^{\{n\}})(X_s)\mathrm{d}s$$
$$+ \int_0^t \langle \nabla u_s^{\{n\}}(X_s), \sigma_s(X_s)\mathrm{d}W_s\rangle + \int_0^t (\nabla_{\mathbf{n}} u_s^{\{n\}})(X_s)\mathrm{d}l_s, \quad t \in [0,T].$$

□

To improve Lemma 2.5 for $(p,q) \in \mathscr{K}$ with $\frac{d}{p} + \frac{2}{q} < 2$, we first extend Lemma 2.4 to the Neumann boundary case. For any $k \in \mathbb{N}$, let $C_b^{0,k}([t_0,t_1] \times \bar{D}; \mathbb{R}^d)$ be the space of $f \in C_b([t_0,t_1] \times \bar{D}; \mathbb{R}^d)$ with bounded and continuous derivatives in $x \in \bar{D}$ up to order k. Let $C_b^{1,2}([t_0,t_1] \times \bar{D}; \mathbb{R}^d)$ denote the space of $f \in C_b^{0,2}([t_0,t_1] \times \bar{D}; \mathbb{R}^d)$ with bounded and continuous $\partial_t f$.

Lemma 2.6. *Assume $(A_2^{\sigma,b})$ but without the condition on $\|\nabla\sigma\|$. Then $(A_0^{\sigma,b})$ and the following assertions hold.*

(1) *For any $\lambda \ge 0$, $0 \le t_0 < t_1 \le T$ and $\tilde{b}, f \in C_b^{0,2}([t_0,t_1] \times \bar{D}; \mathbb{R}^d)$, the PDE*

$$(\partial_t + L_t^{\sigma,b^{(1)}} + \nabla_{\tilde{b}_t} - \lambda)\tilde{u}_t^\lambda = f_t, \quad \tilde{u}_{t_1}^\lambda = \nabla_{\mathbf{n}}\tilde{u}_t^\lambda|_{\partial D} = 0, t \in [t_0,t_1] \tag{2.38}$$

has a unique solution $\tilde{u}^\lambda \in C_b^{1,2}([t_0,t_1] \times \bar{D}; \mathbb{R}^d)$.

(2) *For any $(p,q),(p',q') \in \mathscr{K}$ and $\tilde{b} \in C_b^{0,2}([0,T] \times \bar{D}; \mathbb{R}^d)$, there exist a constant $\varepsilon > 0$ depending only on (p,q) and (p',q'), and constants $\lambda_0, c > 0$ increasing in $\|\tilde{b}\|_{\tilde{L}_{q'}^{p'}(T,D)}$, such that for any $0 \le t_0 < t_1 \le T$ and $f \in C_b^{0,2}([t_0,t_1] \times \bar{D}; \mathbb{R}^d)$,*

$$\lambda^\varepsilon(\|\tilde{u}^\lambda\|_\infty + \|\nabla\tilde{u}^\lambda\|_{\tilde{L}_q^p(t_0,t_1,D)}) \le c\|f\|_{\tilde{L}_{q/2}^{p/2}(t_0,t_1,D)}, \quad \lambda \ge \lambda_0 \text{ (when } p > 2), \tag{2.39}$$

$$\lambda^\varepsilon \|\nabla\tilde{u}^\lambda\|_\infty \le c\|f\|_{\tilde{L}_q^p(t_0,t_1,D)}, \quad \lambda \ge \lambda_0, \tag{2.40}$$

and there exists decomposition $\tilde{u}^\lambda = \tilde{u}^{\lambda,1} + \tilde{u}^{\lambda,2}$ such that

$$\begin{aligned} &\|\nabla^2\tilde{u}^{\lambda,1}\|_{\tilde{L}_q^p(t_0,t_1,D)} + \|(\partial_t + \nabla_{b^{(1)}})\tilde{u}^{\lambda,1}\|_{\tilde{L}_q^p(t_0,t_1,D)} + \|\nabla^2\tilde{u}^{\lambda,2}\|_{\tilde{L}_{q'}^{p'}(t_0,t_1,D)} \\ &+ \|(\partial_t + \nabla_{b^{(1)}})\tilde{u}^{\lambda,2}\|_{\tilde{L}_{q'}^{p'}(t_0,t_1,D)} \le c\|f\|_{\tilde{L}_q^p(t_0,t_1,D)}, \quad \lambda \ge \lambda_0. \end{aligned} \tag{2.41}$$

Proof. (1) Let $\mathbb{V} := C_b^{0,2}([t_0,t_1] \times \bar{D}; \mathbb{R}^d)$, which is a Banach space under the norm

$$\|u\|_{\mathbb{V},N} := \sup_{t \in [t_0,t_1]} \mathrm{e}^{-N(t_1-t)}\big\{\|u_t\|_\infty + \|\nabla u_t\|_\infty + \|\nabla^2 u_t\|_\infty\big\}, \quad u \in \mathbb{V}$$

for $N > 0$. To solve (2.38), for any $\lambda \ge 0$ and $u \in \mathbb{V}$, let

$$\Phi_s^\lambda(u) := \int_s^{t_1} \mathrm{e}^{-\lambda(t-s)} P_{s,t}^{\sigma,b^{(1)}}\{\nabla_{\tilde{b}_t}u_t - f_t\}\mathrm{d}t, \quad s \in [t_0,t_1].$$

Then $(A_2^{\sigma,b})$ implies $\Phi^\lambda(u) \in C_b^{1,2}([t_0,t_1]\times\bar D)$ with

$$(\partial_s + L_s^{\sigma,b^{(1)}} - \lambda)\Phi_s^\lambda(u) = f_s - \nabla_{\tilde b_s} u_s, \quad s\in[t_0,t_1], \nabla_{\mathbf n}\Phi_t^\lambda(u)|_{\partial D} = 0, \Phi_{t_1}^\lambda(u) = 0. \tag{2.42}$$

So, it suffices to prove that Φ^λ has a unique fixed point $\tilde u^\lambda \in \mathbb V$:

$$\tilde u_s^\lambda = \int_s^{t_1} \mathrm e^{-\lambda(t-s)} P_{s,t}^{\sigma,b^{(1)}}\big\{\nabla_{\tilde b_t}\tilde u_t^\lambda - f_t\big\}\mathrm dt, \quad s\in[t_0,t_1], \tag{2.43}$$

which, according to (2.42), is the unique solution of (2.38) in $C_b^{1,2}([t_0,t_1]\times\bar D;\mathbb R^d)$.

For any $u,\bar u\in\mathbb V$, by $\|\tilde b\|_\infty<\infty$, we find a constant $c_1>0$ such that

$$\|\Phi_s^\lambda(u)-\Phi_s^\lambda(\bar u)\|_\infty \le \int_s^{t_1}\|\tilde b_t\|_\infty\|\nabla(u_t-\bar u_t)\|_\infty\mathrm dt \le c_1\int_s^{t_1}\|\nabla(u_t-\bar u_t)\|_\infty\mathrm dt.$$

Similarly, (2.11) with $i=1$ implies

$$\begin{aligned}&\|\nabla\{\Phi^\lambda(u)_s-\Phi^\lambda(\bar u)_s\}\|_\infty \le c\int_s^{t_1}(t-s)^{-\frac12}\|\tilde b_t\|_\infty\|\nabla(u_t-\bar u_t)\|_\infty\mathrm dt\\ &\le c_1\int_s^t (t-s)^{-\frac12}\|\nabla(u_t-\bar u_t)\|_\infty\mathrm dt,\end{aligned}$$

while (2.11) with $i=2$ and $\|\tilde b\|_\infty+\|\nabla\tilde b_t\|_\infty<\infty$ yield

$$\begin{aligned}&\|\nabla^2\{\Phi_s^\lambda(u)-\Phi_s^\lambda(\bar u)\}\|_\infty \le c\int_s^{t_1}(t-s)^{-\frac12}\big\|\nabla\{\nabla_{\tilde b_t}(u_t-\bar u_t)\}\big\|_\infty\mathrm dt\\ &\le c_1\int_s^{t_1}(t-s)^{-\frac12}\big\{\|\nabla(u_t-\bar u_t)\|_\infty+\|\nabla^2(u_t-\bar u_t)\|_\infty\big\}\mathrm dt.\end{aligned}$$

Combining these with (2.42) and the boundedness of a and $\tilde b\in C_b^{0,1}([t_0,t_1]\times\bar D;\mathbb R^d)$, we find a constant $c_2>0$ such that

$$\begin{aligned}&\|\Phi^\lambda(u)-\Phi^\lambda(\bar u)\|_{\mathbb V,N}\\ &\le c_2\sup_{s\in[t_0,t_1]}\int_s^{t_1}\mathrm e^{-N(t_1-s)}(t-s)^{-\frac12}\Big\{\|u_t-\bar u_t\|_\infty\\ &\qquad+\|\nabla(u_t-\bar u_t)\|_\infty+\|\nabla^2(u_t-\bar u_t)\|_\infty\Big\}\mathrm dt\\ &\le c_2\|u-\bar u\|_{\mathbb V,N}\sup_{s\in[t_0,t_1]}\int_s^{t_1}\mathrm e^{-N(t-s)}(t-s)^{-\frac12}\mathrm dt.\end{aligned}$$

So, Φ^λ is contractive under the norm $\|\cdot\|_{\mathbb V,N}$ for large enough $N>0$, and hence has a unique fixed point $\tilde u^\lambda$ in $\mathbb V$.

(2) To prove (2.39) and (2.41), we extend the PDE (2.38) to a global one such that estimates in Lemma 2.4 apply. By $(A_2^{\sigma,b})$, there exists $r_0>0$ such that

$$\varphi:\partial_{-r_0D}\to\partial_{r_0}D;\quad \theta-r\mathbf n(\theta)\mapsto\theta+r\mathbf n(\theta),\quad r\in[0,r_0],\theta\in\partial D$$

is a $C_b^{1,L}$-diffeomorphism (i.e. it is a homeomorphism with $\nabla\varphi$ bounded and Lipschitz continuous) and $\rho_D := \text{dist}(\cdot, D) \in C_b^2(D_{r_0}\setminus \partial D)$, recall that $D_{r_0} = \{\rho_D \le r_0\}$. For any vector field v on $\partial_{r_0}D$, $v^\star := (\varphi^{-1})^* v$ is the vector field on $\partial^0_{-r_0}D := \partial_{-r_0}D \setminus \partial D$ given by

$$\langle v^\star, \nabla g\rangle(x) := \langle v, \nabla(g\circ\varphi^{-1})\rangle(\varphi(x)),\quad x\in\partial^0_{-r_0}D,\ g\in C^1(\partial^0_{-r_0}D).$$

We then extend $b_t^{(1)}$ and $\tilde b_t$ to $\mathbb{R}^d$ by taking

$$b_t^{(1)} := 1_{\bar D} b_t^{(1)} + h(\rho_D/2) 1_{\partial^0_{-r_0}D}(b_t^{(1)})^\star,\quad \tilde b_t := 1_{\bar D}\tilde b_t + 1_{\partial^0_{-r_0}D}(\tilde b_t)^\star, \tag{2.44}$$

where $h\in C^\infty(\mathbb{R})$ such that $0\le h\le 1, h|_{(-\infty, r_0/4]} = 1$ and $h|_{[r_0/2,\infty)} = 0$. Since $(A_2^{\sigma,b})$ implies $\|1_{\bar D}\nabla b^{(1)}\|_\infty < \infty$ and $\nabla_{\mathbf{n}} b^{(1)}|_{\partial D} = 0$, we have $\|\nabla b^{(1)}\|_\infty<\infty$. Let

$$\tilde\varphi(x) := x 1_{\bar D}(x) + \varphi(x) 1_{\partial^0_{-r_0}D}(x),\quad x\in D_{r_0}. \tag{2.45}$$

We extend $\tilde u^\lambda$ to $[t_0,t_1]\times\mathbb{R}^d$ by setting

$$u_t^\lambda = h(\rho_D)(\tilde u_t^\lambda\circ\tilde\varphi),\quad t\in[t_0,t_1]. \tag{2.46}$$

We claim that

$$u_t^\lambda \in C_b^{1,L}(\mathbb{R}^d),\quad t\in[t_0,t_1], \tag{2.47}$$

where $C_b^{1,L}(D_{r_0})$ is the class of C_b^1-functions f on D_{r_0} with Lipschitz continuous ∇f. Indeed, since φ is a $C_b^{1,L}$-diffeomorphism from $\partial_{-r_0}D$ to $\partial_{r_0}D$, $\tilde\varphi\in C_b^{1,L}(D_{r_0}\setminus\partial D)$ with bounded and continuous first and second order derivatives, which together with $\tilde u_t^\lambda \in C_b^2(\bar D)$ yields $u_t^\lambda\in C_b^{1,L}(\mathbb{R}^d\setminus\partial D)$. So, we only need to verify that $\tilde u_t^\lambda\circ\tilde\varphi\in C_b^{1,L}(D_{r_0})$. To this end, for any $x\in\partial_{-r_0}D$ and $v\in\mathbb{R}^d$, let

$$\pi_x v := v - \langle v, \mathbf{n}(\theta(x))\rangle \mathbf{n}(\theta(x))$$

be the projection of $v\in T_x\mathbb{R}^d$ to the tangent space of ∂D, recall that $\theta(x)$ is the projection of x to ∂D, i.e. $x = \theta(x) - \rho_D(x)\mathbf{n}(\theta(x))$ for $\rho_D(x) := \text{dist}(x, D)$. We have

$$\begin{aligned}\nabla_v\tilde\varphi(x) &= \nabla_{\langle v,\mathbf{n}(\theta(x))\rangle\mathbf{n}(\theta(x))}\tilde\varphi(x) + \nabla_{\pi_x v}\tilde\varphi(x)\\ &= 1_{\partial D}(x)|\langle v,\mathbf{n}(\theta(x))\rangle|\mathbf{n}(\theta(x)) + \{1_D - 1_{\partial^0_{-r_0}D}\}(x)\langle v,\mathbf{n}(\theta(x))\rangle\mathbf{n}(\theta(x))\\ &\quad + \pi_x v + \rho_D(x)(\nabla_{\pi_x v}\mathbf{n})(\theta(x)).\end{aligned} \tag{2.48}$$

Since $\tilde u_t^\lambda\in C_b^2(\bar D)$ with $\nabla_{\mathbf{n}}\tilde u_t^\lambda|_{\partial D} = 0$, (2.48) yields

$$\begin{aligned}&\nabla_v(\tilde u_t^\lambda\circ\tilde\varphi)(x) = (\nabla_v\tilde u_t^\lambda)\circ\tilde\varphi(x)\\ &- 2 1_{\partial^0_{-r_0}D}(x)\langle v,\mathbf{n}(\theta(x))\rangle\cdot\langle\mathbf{n}(\theta(x)), (\nabla\tilde u_t^\lambda)\circ\tilde\varphi(x)\rangle\\ &+ \rho_D(x)\big(\nabla_{(\nabla_{\pi_x v}\mathbf{n})(\theta(x))}\tilde u_t^\lambda\big)\circ\tilde\varphi(x),\quad x\in D_{r_0}.\end{aligned} \tag{2.49}$$

Combining this with $\nabla\tilde u_t^\lambda\in C_b^1(\bar D)$, $\nabla_{\mathbf{n}}\tilde u_t^\lambda|_{\partial D} = 0$ and $\mathbf{n}, \nabla\mathbf{n}$ are Lipschitz continuous on $\partial_{-r_0}D)$ due to $\partial D\in C_b^{2,L}$, we conclude that $\nabla(\tilde u_t^\lambda\circ\tilde\varphi)$ is Lipschitz continuous on D_{r_0}.

Next, we construct the PDE satisfied by u^λ. By (2.48), we see that

$$(\nabla\tilde{\varphi})(\nabla\tilde{\varphi})^* = Q \text{ holds on } D_{r_0}\setminus\partial D, \tag{2.50}$$

where Q is a $d\times d$ symmetric matrix valued function given by

$$\begin{aligned}\langle Q(x)v_1, v_2\rangle &:= \langle v_1, v_2\rangle + \rho_D(x)^2\big\langle(\nabla_{\pi_x v_1}\mathbf{n})(\theta(x)),\ (\nabla_{\pi_x v_2}\mathbf{n})(\theta(x))\big\rangle\\ &+ \rho_D(x)\Big\{\big\langle v_1 - 21_{\partial_{-r_0}D}(x)\langle v_1,\mathbf{n}(\theta(x))\rangle\mathbf{n}(\theta(x)),\ (\nabla_{\pi_x v_2}\mathbf{n})(\theta(x))\big\rangle\\ &\quad + \langle v_2 - 21_{\partial_{-r_0}D}(x)\langle v_2,\mathbf{n}(\theta(x))\rangle\mathbf{n}(\theta(x)),\ (\nabla_{\pi_x v_1}\mathbf{n})(\theta(x))\big\rangle\Big\}\end{aligned}$$

for $x\in D_{r_0}, v_1, v_2\in\mathbb{R}^d$. Then by taking $r_0>0$ small enough, on D_{r_0} the matrix-valued functional Q is bounded, invertible, Lipchitz continuous, and symmetric with

$$Q^{-1}(x)\geq\frac{1}{2}\mathbf{I}_d,\quad x\in D_{r_0}. \tag{2.51}$$

We extend $a_t := \frac{1}{2}\sigma_t\sigma_t^*$ from $\bar{D}$ to $\mathbb{R}^d$ by letting

$$a_t := h(\rho_D/2)(a_t\circ\tilde{\varphi})Q^{-1} + (1-h(\rho_D/2))\mathbf{I}_d. \tag{2.52}$$

Since (2.5) holds for $x,y\in\bar{D}$, with this extension of a it holds for all $x,y\in\mathbb{R}^d$. Combining this with (2.44), Remark 2.1(a) for the existence of $\tilde{\rho}$, and noting that $b_t = b_t^{(1)} + 1_{\bar{D}}b_t^{(0)}$ extends b from $\bar{D}$ to $\mathbb{R}^d$, we see that $(A_0^{\sigma,b})$ holds.

Since $h(\rho_D/2), h(\rho_D)\in C_b^2(\mathbb{R}^d)$ with $h(\rho_D/2)=1$ on $\{h(\rho_D)\neq 0\}$, by (2.38), (2.44), (2.52), (2.47) and (2.50), we see that u_t^λ in (2.46) solves the PDE

$$(\partial_t + \mathrm{tr}\{a_t\nabla^2\} + \nabla_{b_t^{(1)}+\tilde{b}_t})u_t^\lambda = \lambda u_t^\lambda + f_t^{(1)} + f_t^{(2)},\ \ t\in[t_0,t_1], u_{t_1}^\lambda = 0, \tag{2.53}$$

where outside the null set ∂D,

$$\begin{aligned} f_t^{(1)} &:= (h\circ\rho_D)f_t\circ\tilde{\varphi} + 2\big\langle a_t\nabla(h\circ\rho_D), \nabla\{\tilde{u}_t^\lambda\circ\tilde{\varphi}\}\big\rangle,\\ f_t^{(2)} &:= (\tilde{u}_t^\lambda\circ\tilde{\varphi})(L_t^{\sigma,b^{(1)}} + \nabla_{\tilde{b}_t})(h\circ\rho_D).\end{aligned} \tag{2.54}$$

By (2.48), $h\in C^\infty([0,\infty))$ with support $\mathrm{supp}\, h\subset[0,r_0/2]$, $\|a\|_\infty + \|1_{\partial_{r_0}D}\nabla_{b^{(1)}}\rho\|_\infty<\infty$ according to $(A_2^{\sigma,b})$ and Remark 2.2(1), we find a constant $c>0$ such that

$$\begin{aligned} |f_t^{(1)}| &\leq 1_{\{\rho_D\leq\frac{r_0}{2}\}}(|f_t| + |\nabla\tilde{u}_t^\lambda|)\circ\tilde{\varphi},\\ |f_t^{(2)}| &\leq c1_{\{\rho_D\leq\frac{r_0}{2}\}}\big\{(1+|\tilde{b}_t|)|\tilde{u}_t^\lambda|\big\}\circ\tilde{\varphi}.\end{aligned}$$

Since $|f|+|\tilde{b}|+|\tilde{u}^\lambda|$ is bounded on $[0,T]\times\bar{D}$, so is $|f^{(1)}|+|f^{(2)}|$ on $[0,T]\times\mathbb{R}^d$. Hence, by Lemma 2.4, the PDE (2.53) has a unique solution in $\tilde{H}_q^{2,p}(t_0,t_1)$, for each $i=1,2$ and $\lambda\geq 0$, the PDE

$$(\partial_t + \mathrm{tr}\{a_t\nabla^2\} + \nabla_{b_t^{(1)}+\tilde{b}_t})u_t^{\lambda,i} = \lambda u_t^{\lambda,i} + f_t^{(i)},\ \ t\in[t_0,t_1], u_{t_1}^{\lambda,i} = 0 \tag{2.55}$$

has a unique solution in $\tilde H^{2,p}_q(t_0,t_1)$ as well, and there exist constants $c_1,c_2>0$ increasing in $\|\tilde b\|_{\tilde L^{p'}_{q'}(T,D)}$ such that

$$\begin{aligned}&\lambda^{1-\frac{d}{p}-\frac{2}{q}}\|u^{\lambda,1}\|_\infty+\lambda^{\frac12(1-\frac{d}{p}-\frac{2}{q})}\|\nabla u^{\lambda,1}\|_{\tilde L^p_q(t_0,t_1)}\\&\le c_1\|f^{(1)}\|_{\tilde L^p_q(t_0,t_1)}\le c_2\big(\|f\|_{\tilde L^{p/2}_{q/2}(t_0,t_1,D)}+\|\tilde u^\lambda_t\|_{\tilde L^p_q(t_0,t_1,D)}\big),\quad p>2,\end{aligned}\tag{2.56}$$

$$\begin{aligned}&\lambda^{\frac12(1-\frac{d}{p}-\frac{2}{q})}\|\nabla u^{\lambda,1}\|_\infty+\|\nabla^2u^{\lambda,1}\|_{\tilde L^p_q(t_0,t_1)}+\|(\partial_t+\nabla_{b^{(1)}})u^{\lambda,1}\|_{\tilde L^p_q(t_0,t_1)}\\&\le c_1\|f^{(1)}\|_{\tilde L^p_q(t_0,t_1)}\le c_2(\|f\|_{\tilde L^p_q(t_0,t_1,D)}+\|\tilde u^\lambda\|_{\tilde L^p_q(t_0,t_1,D)}),\end{aligned}\tag{2.57}$$

and

$$\begin{aligned}&\lambda^{\frac12(1-\frac{d}{p'}-\frac{2}{q'})}(\|u^{\lambda,2}\|_\infty+\|\nabla u^{\lambda,2}\|_\infty)+\|\nabla^2u^{\lambda,2}\|_{\tilde L^{p'}_{q'}(t_0,t_1)}\\&+\|(\partial_t+\nabla_{b^{(1)}})u^{\lambda,2}\|_{\tilde L^{p'}_{q'}(t_0,t_1)}\le c_1\|f^{(2)}\|_{\tilde L^{p'}_{q'}(t_0,t_1)}\le c_2(1+\|\tilde b\|_{\tilde L^{p'}_{q'}(t_0,t_1,D)})\|\tilde u^\lambda\|_\infty,\end{aligned}\tag{2.58}$$

where the last step in these estimates follows from (2.54) and the integral transform

$$\tilde\varphi: D_{r_0}\setminus\bar D\to D$$

with $\|(\nabla\tilde\varphi)^{-1}\|_\infty<\infty$ due to (2.50) and (2.51).

By taking large enough $\lambda_0>0$ increasing in $\|\tilde b\|_{\tilde L^{p'}_{q'}(T,D)}$, we derive from (2.56) and (2.58) that

$$\begin{aligned}&\|u^{\lambda,1}\|_\infty+\|\nabla u^{\lambda,1}\|_{\tilde L^p_q(t_0,t_1)}\le\frac12\big(\|f\|_{\tilde L^{p/2}_{q/2}(t_0,t_1,D)}+\|\tilde u^\lambda_t\|_{\tilde L^p_q(t_0,t_1,D)}\big),\\&\|u^{\lambda,2}\|_\infty+\|\nabla u^{\lambda,2}\|_\infty\le\frac12\|\tilde u^\lambda\|_\infty,\quad\lambda\ge\lambda_0.\end{aligned}$$

Noting that the uniqueness of (2.53) and (2.55) implies $u^\lambda_t=u^{\lambda,1}_t+u^{\lambda,2}_t$, this and the definition of u^λ_t yield

$$\begin{aligned}&\|\tilde u^\lambda\|_\infty+\|\nabla\tilde u^\lambda\|_{\tilde L^p_q(t_0,t_1,D)}\le\sum_{i=1}^2(\|u^{\lambda,i}_t\|_\infty+\|\nabla u^{\lambda,i}\|_{\tilde L^p_q(t_0,t_1)})\\&\le\frac12\big\{\|\tilde u^\lambda\|_\infty+\|f\|_{\tilde L^{p/2}_{q/2}(t_0,t_1,D)}+\|\tilde u^\lambda_t\|_{\tilde L^p_q(t_0,t_1,D)}\big\},\end{aligned}$$

so that

$$\|\tilde u^\lambda\|_\infty+\|\nabla\tilde u^\lambda\|_{\tilde L^p_q(t_0,t_1,D)}\le\|f\|_{\tilde L^{p/2}_{q/2}(t_0,t_1,D)},\quad\lambda\ge\lambda_0.$$

This together with (2.56)-(2.58) implies (2.39), (2.40) and (2.41) for some $c,\varepsilon>0$. □

Lemma 2.7. *Assume $(A_2^{\sigma,b})$ but without the condition on $\|\nabla\sigma\|$. For any $(p,q)\in\mathscr K$ with $p>2$, there exist a constant $i\ge1$ depending only on (p,q), and a constant $c\ge1$ increasing* *in* $\|b^{(0)}\|_{\tilde L^{p_0}_{q_0}(T,D)}$*, such that* (2.22) *holds for any solution* $(X_t,l_t)_{t\in[0,T]}$ *of* (2.1)*, and for any* $0\le t_0\le t_1\le T$,

$$\mathbb{E}\bigg(\int_{t_0}^{t_1}|f_s(X_s)|\mathrm{d}s\bigg|\mathscr{F}_{t_0}\bigg)^j\le c^j j!\|f\|^j_{\tilde L^{p/2}_{q/2}(t_0,t_1)},\quad f\in\tilde L^{p/2}_{q/2}(t_0,t_1), j\ge 1, \tag{2.59}$$

$$\mathbb{E}\big(\mathrm{e}^{\int_{t_0}^T|f_t(X_t)|\mathrm{d}t}\big|\mathscr{F}_{t_0}\big)\le\exp\Big[c+c\|f\|^i_{\tilde L^{p/2}_{q/2}(t_0,T)}\Big],\quad f\in\tilde L^{p/2}_{q/2}(t_0,T), t_0\in[0,T]. \tag{2.60}$$

Proof. By Remark 2.1 (a) and step (2) of the proof of Lemma 2.6, $(\sigma_t,b_t^{(1)})$ extends to $\mathbb{R}^d$ such that $(A_0^{\sigma,b})$ holds for $b^{(0)}1_{\bar D^c}=0$. So, (2.22) is ensured by Lemma 2.5.

As explained in step (1) of the proof of Lemma 2.5, for (2.59) and (2.60) it suffices to prove (2.59) for $j=1$ and $f\in C_0^\infty([t_0,t_1]\times\mathbb{R}^d)$.

Let $(b^{0,n})_{n\ge1}$ be the mollifying approximations of $b^{(0)}=1_{\bar D}b^{(0)}$. We have

$$\|b^{0,n}\|_{\tilde L^{p_0}_{q_0}(T)}\le\|b^{(0)}\|_{\tilde L^{p_0}_{q_0}(T)},\quad\lim_{n\to\infty}\|b^{0,n}-b^{(0)}\|_{\tilde L^{p_0}_{q_0}(T)}=0. \tag{2.61}$$

By Lemma 2.6 for $(f,0,\cdots,0)$ replacing f, there exist constants $c,\lambda_0>0$ such that for any $\lambda\ge\lambda_0$, the following PDE on $\bar D$

$$(\partial_t+L_t^{\sigma,b^{(1)}}+\nabla_{b_t^{0,n}}-\lambda)u_t^{\lambda,n}=f_t,\quad t\in[t_0,t_1),\ \nabla_{\mathbf{n}}u_t^{\lambda,n}|_{\partial D}=0, u_{t_1}^{\lambda,n}=0 \tag{2.62}$$

has a unique solution in $C^{1,2}([t_0,t_1]\times\bar D)$, and for some constant $c_1>0$ we have

$$\|u^{\lambda,n}\|_\infty\le c_1\|f\|_{\tilde L^{p/2}_{q/2}(t_0,t_1,D)},\quad\|\nabla u^{\lambda,n}\|_\infty\le c_1\|f\|_\infty,\quad\lambda\ge\lambda_0, n\ge1. \tag{2.63}$$

Moreover, since $(A_2^{\sigma,b})$ implies $(A_0^{\sigma,b})$ due to Lemma 2.6, by (2.20) for $f=|b^{(0)}-b^{0,n}|$, we find a constant $c_2>0$ such that

$$\mathbb{E}\bigg(\int_{t_0}^{t_1}|b^{(0)}-b^{0,n}|(X_s)\mathrm{d}s\bigg|\mathscr{F}_{t_0}\bigg)\le c_2\|b^{(0)}-b^{0,n}\|_{\tilde L^{p_0}_{q_0}(t_0,t_1)},\quad n\ge1. \tag{2.64}$$

By (2.62) and $u^{\lambda,n}\in C_b^{1,2}([t_0,t_1]\times\bar D)$, we have the following Itô's formula

$$\begin{aligned}\mathrm{d}u_t^{\lambda,n}(X_t)&=(\partial_t+L_t)u_t^{\lambda,n}(X_t)\mathrm{d}t+\mathrm{d}M_t\\&=\{f_t+\nabla_{b_t^{(0)}-b_t^{0,n}}u_t^{\lambda,n}\}(X_t)\mathrm{d}t+\mathrm{d}M_t\end{aligned}$$

for some martingale M_t. Combining this with (2.63) and (2.64), we obtain

$$\mathbb{E}\bigg(\int_{t_0}^{t_1}f_t(X_t)\mathrm{d}t\bigg|\mathscr{F}_{t_0}\bigg)\le c_1\|f\|_{\tilde L^{p/2}_{q/2}(t_0,t_1)}+c_1c_2\|f\|_\infty\|b_t^{(0)}-b_t^{0,n}\|_{\tilde L^{p_0}_{q_0}(t_0,t_1)}.$$

Therefore, by (2.61), we may let $n\to\infty$ to derive (2.59) for $j=1$. □

2.3 Weak well-posedness: proof of Theorem 2.1

We first introduce some known results for the reflecting SDE with random coefficients:

$$\mathrm{d}X_t = J_t(X_t)\mathrm{d}t + S_t(X_t)\mathrm{d}W_t + \mathbf{n}(X_t)\mathrm{d}l_t, \quad t \in [0,T], \tag{2.65}$$

where $(W_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion on a complete filtration probability space $(\Omega, \{\mathscr{F}_t\}_{t\in[0,T]}, \mathbb{P})$,

$$J : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d, \quad S : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^m$$

are progressively measurable, and l_t is the local time of X_t on ∂D. Let Λ be the set of increasing functions $h : (0,1] \to (0,\infty)$ such that $\int_0^1 \frac{\mathrm{d}s}{h(s)} = \infty$, and let Γ be the class of increasing functions $\gamma : [0,\infty) \to [1,\infty)$ such that $\int_0^\infty \frac{\mathrm{d}s}{\gamma(s)} = \infty$. When D is convex the following result goes back to [29], and in general it is mainly summarized from [10, Theorem 1, Corollary 1 and Theorem 2], where the condition in the first assertion is more general than that stated in [10, Theorem 1.1]:

$$\|S_t(x) - S_t(y)\|_{HS}^2 + 2\langle x-y, J_t(x) - J_t(y)\rangle \le g_t h(|x-y|^2), \quad t \in [0,T], x, y \in \bar{D},$$

since in the proof of this assertion, one only uses the upper bound of

$$\|S_t(X_t) - S_t(Y_t)\|_{HS}^2 + 2\langle X_t - Y_t, J_t(X_t) - J_t(Y_t)\rangle,$$

so that the present condition is enough for the pathwise uniqueness. In Theorem 2.8(3), the term $\mathrm{tr}\{S_t S_t^* \nabla^2 V_t\}$ was formulated in [10, Theorem 1.1] as $\|S_t(x)\|^2 \Delta V_t(x)$, which should be changed into the present one according to Itô's formula of $V_t(X_t)$. Moreover, when S and J are bounded and deterministic, the weak existence is given in [23, Theorem 2.1].

Theorem 2.8 ([10, 23, 29]). *Assume* **(D)**.

(1) *For any two solutions X_t and Y_t of (2.65) with $X_0 = Y_0 \in \bar{D}$, if there exist $h \in \Lambda$ and a positive $L^1([0,T])$-valued random variable g such that $\mathbb{P}$-a.s.*

$$\|S_t(X_t) - S_t(Y_t)\|_{HS}^2 + 2\langle X_t - Y_t, J_t(X_t) - J_t(Y_t)\rangle \le g_t h(|X_t - Y_t|^2), \quad t \in [0,T],$$

then $X_t = Y_t$ up to life time.

(2) *If $\mathbb{P}$-a.s. S and J are continuous and locally bounded on $[0,\infty) \times \bar{D}$, then for any initial value in $\bar{D}$, (2.65) has a weak solution up to life time. If S and J are bounded and deterministic, (2.65) has a global weak solution.*

(3) *If either D is bounded, or there exist $1 \le V \in C^{1,2}([0,T] \times \bar{D})$ with*

$$\lim_{x\in\bar{D}, |x|\to\infty} \inf_{t\in[0,T]} V_t(x) = \infty, \quad \nabla_{\mathbf{n}} V_t|_{\partial D} \le 0,$$

and a positive $L^1([0,T])$-valued random variable g such that $\mathbb{P}$-a.s.

$$\begin{aligned}&\mathrm{tr}\{S_t S_t^* \nabla^2 V_t\} + 2\langle \nabla V(x), J_t(x)\rangle + 2\partial_t V_t(x)\\ &\le g_t \gamma(V_t(x)), \quad t \in [0,T], x \in \bar{D}\end{aligned}$$

holds for some $\gamma \in \Gamma$, then any solution to (2.65) is non-explosion.

Next, we apply Theorem 2.8 to (2.1) with coefficients satisfying the following assumption, where (1_b) is known as monotone or semi-Lipschitz condition, which comparing with (1_a) allows σ to be unbounded.

(H1) *b and σ are locally bounded satisfying the following conditions.*

(1) *One of the following conditions holds:*

(1_a) *$(A_0^{\sigma,b})$ holds with $\|\nabla\sigma\|^2 \le \sum_{i=1}^l f_i$ for some $1 \le f_i \in\in \tilde{L}_{q_i}^{p_i}(T)$ with $l \in \mathbb{N}$ and $\{(p_i, q_i)\}_{1\le i\le l} \subset \mathscr{K}$, or $(A_2^{\sigma,b})$ holds. Moreover, there exists a constant $K > 0$ such that*

$$\langle x-y, b_t(x)-b_t(y)\rangle \le K|x-y|^2, \quad t\in[0,T], x,y\in\bar{D}. \tag{2.66}$$

(1_b) *There exists an increasing function $h:[0,\infty)\to[0,\infty)$ with $\int_0^1 \frac{\mathrm{d}r}{r+h(r)} = \infty$, such that*

$$2\langle x-y, b_t(x)-b_t(y)\rangle^+ + \|\sigma_t(x)-\sigma_t(y)\|_{HS}^2 \le h(|x-y|^2), \quad t\in[0,T], x,y\in\bar{D}. \tag{2.67}$$

(2) *$\|\sigma\| \le c(1+|\cdot|^2)$ holds for some constant $c>0$, there exist $x_0 \in D$ and $\tilde{\partial} D \subset \partial D$ such that*

$$\langle x-x_0, \mathbf{n}(x)\rangle \le 0, \quad x\in\partial D\setminus\tilde{\partial}D, \ \mathbf{n}(x)\in\mathscr{N}_x; \tag{2.68}$$

and when $\tilde{\partial}D \ne \emptyset$ there exists a function $\tilde{\rho} \in C_b^2(\bar{D})$ such that

$$\langle\nabla\tilde{\rho}, \mathbf{n}\rangle|_{\partial D} \ge 1_{\tilde{\partial}D}, \quad \sup_{[0,T]\times\bar{D}} \big\{\|\sigma^*\nabla\tilde{\rho}\| + \|\mathrm{tr}\{\sigma\sigma^*\nabla^2\tilde{\rho}\}\| + \langle b, \nabla\tilde{\rho}\rangle^-\big\} \le K. \tag{2.69}$$

According to (1.3) and Remark 2.1(a), **(H1)**(2) holds with $\tilde{\rho}=0$ if either D is convex, and it holds with $\tilde{\rho}=\rho$ in $\rho_{r_0/2}D$ for some $r_0>0$ when $\partial D\in C_b^2$ and $\|\sigma\| + \langle b, \nabla\rho\rangle^-$ is bounded on $[0,T]\times\partial_{r_0}D$.

Lemma 2.9. *Assume* **(D)** *and* **(H1)**(1). *Then the reflecting SDE* (2.1) *is well-posed up to life time. If* **(H1)**(2) *holds, then the solution is non-explosive, and for any $k>0$ there exists a constant $c>0$ such that*

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|X_t^x|^k\Big] \le c(1+|x|^k), \quad x\in\bar{D}, t\in[0,T], \tag{2.70}$$

$$\sup_{x\in\bar{D}} \mathbb{E}\big(\mathrm{e}^{k(\tilde{l}_{t_1}^x - \tilde{l}_{t_0}^x)}|\mathscr{F}_{t_0}\big) \le c, \quad 0\le t_0\le t_1\le T, \tag{2.71}$$

where (X_t^x, l_t^x) is the solution with $X_0^x = x$, and $\tilde{l}_t^x := \int_0^t 1_{\tilde{\partial}(D)}(X_s^x)\mathrm{d}l_s^x$.

To prove this result, we need the following lemma on the maximal functional for nonnegative functions f on $\bar{D}$:

$$\mathscr{M}_D f(x) := \sup_{r\in(0,1)} \frac{1}{|B(0,r)|}\int_{B(0,r)} (1_D f)(x+y)\mathrm{d}y, \quad x\in\bar{D}.$$

Lemma 2.10. *Let $\partial D \in C_b^2$.*

(1) *For any real function f on $\bar{D}$ with $|\nabla f| \in L^1_{loc}(\bar{D})$,*

$$|f(x) - f(y)| \le c|x-y|\big(\mathscr{M}_D|\nabla f|(x) + \mathscr{M}_D|\nabla f|(y) + \|f\|_\infty\big), \quad \text{a.e. } x, y \in \bar{D}.$$

(2) *There exists a constant $c > 0$ such that for any nonnegative measurable function f on $[0,T]\times\bar{D}$,*

$$\|\mathscr{M}_D f\|_{\tilde{L}_q^p(T,\bar{D})} \le c\|f\|_{\tilde{L}_q^p(T,\bar{D})}, \quad p, q \ge 1.$$

Proof. We only prove (1), since (2) follows from [37, Lemma 2.1(ii)] with $1_{\bar{D}}f$ replacing f. Let $\tilde{\varphi}$ be in (2.45). Take $0 \le h \in C_b^\infty(\mathbb{R})$ with $h(r) = 1$ for $r \le r_0/4$ and $h(r) = 0$ for $r \ge r_0/2$. We then extend a function f on $\bar{D}$ to $\tilde{f}$ on $\mathbb{R}^d$ by letting

$$\tilde{f}(x) := \{h \circ \rho_D\} f \circ \tilde{\varphi},$$

where ρ_D is the distance function to D. Then there exists a constant $c > 0$ such that

$$|\nabla \tilde{f}| \le 1_{\bar{D}}|\nabla f| + c1_{\partial_{-r_0/2}D}(|f\circ\tilde{\varphi}| + |\nabla f|\circ\tilde{\varphi}).$$

By [41, Lemma 5.4] and the integral transform $x \mapsto \tilde{\varphi}(x)$ with $\|(\nabla\tilde{\varphi})^{-1}\|$ bounded on $\partial_{-r_0}D$, we find constants $c_1, c_2 > 0$ such that for any $x, y \in \bar{D}$,

$$\begin{aligned}
&|f(x) - f(y)| = |\tilde{f}(x) - \tilde{f}(y)| \\
&\le c_1|x-y|\big(\mathscr{M}|\nabla\tilde{f}|(x) + \mathscr{M}|\nabla f|(y) + \|f\|_\infty\big\} \\
&\le c_2|x-y|\big\{\mathscr{M}_D|\nabla f|(x) + \mathscr{M}_D|\nabla f|(y) + \|f\|_\infty\big\},
\end{aligned}$$

where $\mathscr{M} := \mathscr{M}_D$ for $D = \mathbb{R}^d$. □

Proof of Lemma 2.9. (1) We first prove the existence and uniqueness up to life time. Since σ and b are locally bounded, by a truncation argument we may and do assume that σ and b are bounded. Indeed, let for any $n \ge 1$ we take

$$\sigma_t^{\{n\}}(x) := \sigma_t\big(\{1 \wedge (n/|x|)\}x\big), \quad b_t^{\{n\}}(x) := h(|x|/n)b_t(x), \quad t \ge 0, x \in \bar{D},$$

where $h \in C_0^\infty([0,\infty)$ with $0 \le h \le 1$ and $h|_{[0,1]} = 1$. Then $\sigma^{\{n\}}$ and $b^{\{n\}}$ are bounded on $[0,T]\times\bar{D}$ and for some constant $K_n > 0$,

$$\begin{aligned}
&\langle b_t^{\{n\}}(x) - b_t^{\{n\}}(y), x-y\rangle^+ \\
&\le h(|x|/n)\langle b_t(x) - b_t(y), x-y\rangle^+ + \big|h(|x|/n) - h(|y|/n)\big|\langle b_t(y), x-y\rangle^+ \\
&\le \langle b_t(x) - b_t(y), x-y\rangle^+ + K_n|x-y|^2, \quad t \in [0,T], x, y \in \bar{D}, |y| \le |x|.
\end{aligned}$$

So, by the symmetry of $\langle b_t^{\{n\}}(x) - b_t^{\{n\}}(y), x-y\rangle^+$ in (x,y), under (1_a), σ and $b^{\{n\}}$ are bounded on $[0,T]\times\bar{D}$ and satisfy (2.66) with $K + K_n$ replacing K; while (1_b) and

$$|\{1\wedge(n/|x|)\}x - \{1\wedge(n/|y|)\}y| \le |x-y|$$

imply that $\sigma^{\{n\}}$ and $b^{\{n\}}$ are bounded and satisfy (2.67) for $2h(r)+K_n r$ replacing $h(r)$. Therefore, if the well-posedness is proved under **(H1)** for bounded b and σ, the SDE is well-posed up to the hitting time of $\partial B(0,n)$ for any $n\ge 1$, i.e. it is well-posed up to life time.

When σ and b are bounded, the weak existence is implied by Theorem 2.8(2). By the Yamada-Watanabe principle, it suffices to verify the pathwise uniqueness. Let X_t and Y_t be two solutions starting from $x\in\bar D$. By Lemma 2.10(1) and **(H1)**(1),

$$\|\sigma_t(X_t)-\sigma_t(Y_t)\|_{HS}^2+2\langle X_t-Y_t, b_t(X_t)-b_t(Y_t)\rangle\le\begin{cases} g_t|X_t-Y_t|^2, & \text{under } (1_a),\\ h(|X_t-Y_t|^2), & \text{under } (1_b),\end{cases}$$

where for some constant $c>0$

$$g_t:=c\big\{1+\mathscr{M}_D\|\nabla\sigma_t\|^2(X_t)+\mathscr{M}_D\|\nabla\sigma_t\|^2(Y_t)\big\}.$$

So, by Theorem 2.8(1), it suffices to prove $\int_0^T g_t\mathrm{d}t<\infty$ under (1_a). By Lemma 2.10, this follows from (2.20) under condition $(A_0^{\sigma,b})$ with $\|\nabla\sigma\|^2\le\sum_{i=1}^l f_i$ for some $1\le f_i\in\in\tilde L_{q_i}^{p_i}(T)$ with $l\in\mathbb N$ and $\{(p_i,q_i)\}_{1\le i\le l}\subset\mathscr K$, or (2.59) under condition $(A_2^{\sigma,b})$.

(2) To prove the non-explosion, we simply denote $(X_t,l_t)=(X_t^x,l_t^x)$ and let

$$\tau_n:=\inf\{t\ge0:|X_t|\ge n\},\quad n\ge1.$$

By **(H1)**(2), we find a constant $c_1>0$ such that

$$\mathrm{d}\tilde\rho(X_t)\ge -K\mathrm{d}t+\mathrm{d}M_t+\mathrm{d}\tilde l_t,\quad t\in[0,T] \tag{2.72}$$

holds for $\mathrm{d}M_t:=\langle\sigma_t(X_t)^*\nabla\tilde\rho(X_t),\mathrm{d}W_t\rangle$ satisfying $\mathrm{d}\langle M\rangle_t\le K^2\mathrm{d}t$. This implies (2.71). Next, by **(H1)**, we find a constant $c_1>0$ such that

$$\begin{aligned}&2\langle b_t(x),x-x_0\rangle+\|\sigma_t(x)\|_{HS}^2\\ &=2\langle b_t(x)-b_t(x_0),x-x_0\rangle+\|\sigma_t(x)-\sigma_t(x_0)\|_{HS}^2\\ &\quad+2\langle b_t(x_0),x-x_0\rangle+\|\sigma_t(x_0)\|_{HS}^2+2\langle\sigma_t(x_0),\sigma_t(x)\rangle_{HS}\\ &\le c_1(1+|x-x_0|^2),\quad x\in\bar D.\end{aligned}$$

Then by **(H1)**(2) and Itô's formula, for any $k\ge2$ we find a constant $c_2>0$ such that

$$\mathrm{d}|X_t-x_0|^k\le c_2(1+|X_t-x_0|^k)\mathrm{d}t+\mathrm{d}\tilde M_t+k|X_t-x_0|^{k-1}\mathrm{d}\tilde l_t,$$

where $\tilde M_t$ is a local martingale with $\mathrm{d}\langle\tilde M\rangle_t\le c_2(1+|X_t-x_0|^k)^2\mathrm{d}t$. By BDG's inequality and (2.71), we find constants $c_3,c_4>0$ such that

$$\eta_t^{\{n\}}:=\sup_{s\in[0,t\wedge\tau_n]}(1+|X_s-x_0|^k),\quad n\ge1,t\in[0,T]$$

satisfies

$$\mathbb E\eta_t^{\{n\}}\le1+|x-x_0|^k+c_3\mathbb E\int_0^t\eta_s^{\{n\}}\mathrm{d}s+2c_3\mathbb E^x\bigg(\int_0^t|\eta_s^{\{n\}}|^2\mathrm{d}s\bigg)^{\frac12}+k\mathbb E\Big[|\eta_t^{\{n\}}|^{\frac{k-1}{k}}\tilde l_t\Big]$$

$$\le \frac{1}{2}\mathbb{E}\eta_t^{\{n\}} + c_4(1+|x|^k) + c_4\int_0^t \mathbb{E}\eta_s^{\{n\}}\mathrm{d}s,\quad t\in[0,T].$$

By Gronwall's lemma, we obtain

$$\mathbb{E}[\eta_t^{\{n\}}] \le 2c_4(1+|x|^k)\mathrm{e}^{2c_4 t},\quad t\in[0,T], x\in\bar{D}, n\ge 1,$$

which implies the non-explosive of X_t and (2.70) for some constant $c>0$. □

Proof of Theorem 2.1. Let $X_0 = x\in\bar{D}$. We consider the following two cases respectively.

(a) Let $(A_1^{\sigma,b})$ hold. Then **(H1)** holds for $b^{(1)}$ replacing b. By Lemma 2.9, the reflecting SDE

$$\mathrm{d}X_t = b_t^{(1)}(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t + \mathbf{n}(X_t)\mathrm{d}l_t \tag{2.73}$$

is well-posed with (2.70) holding for all $k\ge 1$ and some constant $c>0$ depending on k. By Lemmas 2.5-2.7, (2.71) and $(A_0^{\sigma,b})$ with $|b^{(0)}|^2\in\tilde{L}_{q_0}^{p_0}(T)$, we see that (2.21) holds for $f:=|b^{(0)}|^2$, so that for some map $c:[1,\infty)\to(0,\infty)$ independent of the initial value x,

$$\sup_{x\in\bar{D}}\mathbb{E}^x|R_T|^k \le c(k),\quad k\ge 1 \tag{2.74}$$

holds for

$$R_t := \mathrm{e}^{\int_0^t\langle\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}b_s^{(0)}\}(X_s),\mathrm{d}W_s\rangle - \frac{1}{2}\int_0^t|\sigma_s^*(\sigma_s\sigma_s^*)^{-1}b_s^{(0)}|^2(X_s)\mathrm{d}s},\quad t\in[0,T].$$

By Girsanov's theorem,

$$\tilde{W}_t := W_t - \int_0^t\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}b_s^{(0)}\}(X_s)\mathrm{d}s,\quad t\in[0,T]$$

is an m-dimensional Brownian motion under the probability measure $\mathbb{Q}:=R_T\mathbb{P}$. Rewriting (2.73) as

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}\tilde{W}_t + \mathbf{n}(X_t)\mathrm{d}l_t,$$

we see that $(X_t, l_t, \tilde{W}_t)_{t\in[0,T]}$ under probability $\mathbb{Q}$ is a weak solution of (2.1). Moreover, letting $\mathbb{E}_{\mathbb{Q}}$ be the expectation under $\mathbb{Q}$, by (2.70) and (2.74), for any $k\ge 1$ we find a constant $\tilde{c}(k)>0$ independent of x such that

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\Big[\sup_{t\in[0,T]}|X_t|^k\Big] &= \mathbb{E}\Big[R_T\sup_{t\in[0,T]}|X_t|^k\Big]\\ &\le \big(\mathbb{E}\big[R_T^2\big]\big)^{\frac{1}{2}}\Big(\mathbb{E}\sup_{t\in[0,T]}|X_t|^{2k}\Big]\Big)^{\frac{1}{2}} \le \tilde{c}(k)(1+|x|^k),\quad x\in\bar{D}\end{aligned}$$

for some constant $c>0$. Similarly, (2.71) and (2.74) imply $\mathbb{E}_{\mathbb{Q}}\mathrm{e}^{kl_T}\le C(k)$ for $k\ge 1$ and constant $C(k)>0$ independent of x. So, (2.13) holds for this weak solution.

To prove the weak uniqueness, let $(\bar{X}_t, \bar{l}_t, \bar{W}_t)_{t\in[0,T]}$ under probability $\bar{\mathbb{P}}$ be another weak solution of (2.1) with $\bar{X}_0 = x$, i.e.

$$\mathrm{d}\bar{X}_t = b_t(\bar{X}_t)\mathrm{d}t + \sigma_t(\bar{X}_t)\mathrm{d}\bar{W}_t + \mathbf{n}(\bar{X}_t)\mathrm{d}\bar{l}_t,\quad t\in[0,T], \bar{X}_0 = x. \tag{2.75}$$

It suffices to show

$$\mathscr{L}_{(\bar X_t, \bar l_t)_{t\in[0,T]}|\bar{\mathbb{P}}} = \mathscr{L}_{(X_t, l_t)_{t\in[0,T]}|\mathbb{Q}}. \tag{2.76}$$

By Lemma 2.5 the estimate (2.21) holds for $\bar X_t$ and $f = |b^{(0)}|^2$, so that

$$\mathbb{E}_{\bar{\mathbb{P}}} \mathrm{e}^{\lambda \int_0^T |b_t^{(0)}(\bar X_t)|^2 \mathrm{d}t} < \infty, \quad \lambda > 0. \tag{2.77}$$

By Girsanov's theorem, this and $(A_0^{\sigma,b})$ imply that

$$G_t(\bar X, \bar W) := \bar W_t + \int_0^t \{\sigma_s^*(\sigma_s\sigma_s^*)^{-1} b_s^{(0)}\}(\bar X_s)\mathrm{d}s, \quad t\in[0,T]$$

is an m-dimensional Brownian motion under the probability $\bar{\mathbb{Q}} := R(\bar X, \bar W)\bar{\mathbb{P}}$, where

$$R(\bar X, \bar W) := \mathrm{e}^{-\int_0^T \langle \{\sigma_s^*(\sigma_s\sigma_s^*)^{-1} b_s^{(0)}\}(\bar X_s), \mathrm{d}\bar W_s\rangle - \frac{1}{2}\int_0^T |\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1} b_s^{(0)}\}(\bar X_s)|^2 \mathrm{d}s}.$$

Reformulating (2.75) as

$$\mathrm{d}\bar X_t = b_t^{(1)}(\bar X_t)\mathrm{d}t + \sigma_t(\bar X_t)\mathrm{d}G_t(\bar X, \bar W) + \mathbf{n}(\bar X_t)\mathrm{d}\bar l_t, \quad t\in[0,T],$$

and applying the well-posedness of (2.73) which implies the joint weak uniqueness, we conclude that

$$\mathscr{L}_{(\bar X_t, \bar l_t, G_t(\bar X,\bar W))_{t\in[0,T]}|\bar{\mathbb{Q}}} = \mathscr{L}_{(X_t, l_t, W_t)_{t\in[0,T]}|\mathbb{P}}.$$

Noting that

$$R(\bar X, \bar W)^{-1} = \mathrm{e}^{-\int_0^T |\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1} b_s^{(0)}\}(\bar X_s)|^2 \mathrm{d}s} R(\bar X, G(\bar X, \bar W))^{-1},$$

this implies that for any bounded continuous function F on $C([0,T];\mathbb{R}^d\times[0,\infty))$,

$$\begin{aligned}
&\mathbb{E}_{\bar{\mathbb{P}}}[F(\bar X, \bar l)] = \mathbb{E}_{\bar{\mathbb{Q}}}[R(\bar X, \bar W)^{-1} F(\bar X, \bar l)]\\
&= \mathbb{E}_{\bar{\mathbb{Q}}}[R(\bar X, G(\bar X,\bar W))^{-1} \mathrm{e}^{-\int_0^T |\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1} b_s^{(0)}\}(\bar X_s)|^2 \mathrm{d}s} F(\bar X, \bar l)]\\
&= \mathbb{E}_{\mathbb{P}}[R(X, W)^{-1} \mathrm{e}^{-\int_0^T |\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1} b_s^{(0)}\}(X_s)|^2 \mathrm{d}s} F(X, l)]\\
&= \mathbb{E}_{\mathbb{P}}[R_T F(X,l)] = \mathbb{E}_{\mathbb{Q}}[F(X,l)].
\end{aligned}$$

Therefore, (2.76) holds.

(b) Let $(A_2^{\sigma,b})$ hold. By Lemma 2.7, (2.74) and (2.77) hold, so that the desired assertions follow from Girsanov's transforms as shown in step (a). □

2.4 Well-posedness: proof of Theorem 2.2

The weak existence is implied by Theorem 2.1. By the Yamada-Watanabe principle, it suffices to prove estimate (2.14) which in particular implies the pathwise uniqueness as well as estimate (2.15):

$$|\nabla P_t f|(x) := \limsup_{\bar D\ni y\to x} \frac{|P_t f(x) - P_t f(y)|}{|x-y|} \le \limsup_{\bar D\ni y\to x} \mathbb{E}\left[\frac{|f(X_t^x) - f(X_t^y)|}{|x-y|}\right]$$

$$\le \limsup_{\bar D\ni y\to x}\Big(\mathbb{E}\frac{|f(X_t^x)-f(X_t^y)|^p}{|X_t^x-X_t^y|^p}\Big)^{\frac1p}\Big(\frac{\mathbb{E}[|X_t^x-X_t^y|^{\frac{p}{p-1}}]}{|x-y|^{\frac{p}{p-1}}}\Big)^{\frac{p-1}{p}}$$
$$\le c(p)\big(P_t|\nabla f|^p\big)^{\frac1p}(x),\quad x\in\bar D, t\in[0,T], f\in C_b^1(\bar D).$$

Let $(X_t^{(i)}, l_t^{(i)})$ be two solutions of (2.1) with $X_0^{(i)}=x^{(i)}\in\bar D, i=1,2$. Below we prove (2.14) in situations (i) and (ii) respectively.

Proof of Theorem 2.2 under (i). In this case, D is an interval or a half-line. For any $\lambda>0$, let u_t^λ be the unique solution to (2.19) with $t_0=0, t_1=T$ and $f=-b^{(0)}$, that is,

$$(\partial_t+L_t)u_t^\lambda=\lambda u_t^\lambda-b_t^{(0)},\quad t\in[0,T], u_T^\lambda=0. \tag{2.78}$$

By (2.25) with $f=-b^{(0)}\in\tilde L_{2q_0}^{2p_0}(T)$, we take large enough $\lambda>0$ such that

$$\|u^\lambda\|_\infty+\|\nabla u^\lambda\|_\infty\le\frac12,\quad \|u^\lambda\|_{\tilde H_{2q_0}^{2,2p_0}(T)}<\infty. \tag{2.79}$$

Then $\Theta_t^\lambda(x):=x+u_t^\lambda(x)$ is a diffeomorphism and there exists a constant $C>0$ such that

$$\frac12|x-y|\le|\Theta_t^\lambda(x)-\Theta_t^\lambda(y)|\le 2|x-y|,\quad x,y\in\mathbb{R}, t\in[0,T]. \tag{2.80}$$

Let $(X_t^{(i)}, l_t^{(i)})$ solve (2.1) for $X_0^{(i)}=x^{(i)}\in\bar D, i=1,2$, and let

$$Y_t^{(i)}:=\Theta_t^\lambda(X_t^{(i)})=X_t^{(i)}+u_t^\lambda(X_t^{(i)}),\quad i=1,2.$$

By Itô's formula in Lemma 2.5(2),

$$\mathrm{d}Y_t^{(i)}=B_t(Y_t^{(i)})\mathrm{d}t+\Sigma_t(Y_t^{(i)})\mathrm{d}W_t+\{1+\nabla u_t^\lambda(X_t^{(i)})\}\mathbf{n}(X_t^{(i)})\mathrm{d}l_t^{(i)},\quad i=1,2 \tag{2.81}$$

holds for

$$B_t(x):=\{b_t^{(1)}+\lambda u_t^\lambda\}\big(\{\Theta_t^\lambda\}^{-1}(x)\big),\quad \Sigma_t(x):=\big\{(1+\nabla u_t^\lambda)\sigma_t\big\}\big(\{\Theta_t^\lambda\}^{-1}(x)\big). \tag{2.82}$$

By $(A_1^{\sigma,b})$, (2.79), (2.82) and $\|\nabla b^{(1)}\|_\infty<\infty$ due to $(A_0^{\sigma,b})$, we find $0\le F_i\in\tilde L_{q_i}^{p_i}(T), 0\le i\le l$, such that

$$\|\nabla B\|_\infty<\infty,\quad \|\nabla\Sigma\|^2\le\sum_{i=0}^l F_i. \tag{2.83}$$

Since $d=1$, for any $x\in\partial D$ and $y\in D$ we have $y-x=|y-x|\mathbf{n}(x)$, so that (2.79) implies

$$\big\langle\Theta_t^\lambda(y)-\Theta_t^\lambda(x),\big\{1+\nabla u_t^\lambda(x)\big\}\mathbf{n}(x)\big\rangle\ge|y-x|(1-\|\nabla u^\lambda\|_\infty)^2\ge 0. \tag{2.84}$$

Combining this with (2.81) and Itô's formula, up to a local martingale we have

$$\mathrm{d}|Y_t^{(1)}-Y_t^{(2)}|^{2k}\le 2k|Y_t^{(1)}-Y_t^{(2)}|^{2k}\bigg\{\frac{|B_t(Y_t^{(1)})-B_t(Y_t^{(2)})|}{|Y_t^{(1)}-Y_t^{(2)}|}+\frac{k\|\Sigma_t(Y_t^{(1)})-\Sigma_t(Y_t^{(2)})\|_{HS}^2}{|Y_t^{(1)}-Y_t^{(2)}|^2}\bigg\}\mathrm{d}t.$$

So, by Lemma 2.10, we find a constant $c_1 > 0$ and a local martingale M_t such hat

$$|Y_t^{(1)} - Y_t^{(2)}|^{2k} \le |Y_0^{(1)} - Y_0^2|^{2k} + c_1 \int_0^t |Y_s^{(1)} - Y_s^{(2)}|^{2k} \mathrm{d}\mathscr{L}_s + \mathrm{d}M_t,$$

where

$$\mathscr{L}_t := \int_0^t \Big\{1 + \mathscr{M}_D \|\nabla \Sigma_s\|^2 (Y_s^{(1)}) + \mathscr{M}_D \|\nabla \Sigma_s\|^2 (Y_s^{(2)})\Big\} \mathrm{d}s. \tag{2.85}$$

Combining this with (2.83), (2.21), Lemma 2.10 and the stochastic Gronwall lemma (see [25] or [38]), for any $k > 1$ and $p \in (\frac{1}{2}, 1)$, we find constants $c_2, c_3 > 0$ such that

$$\begin{aligned}&\Big(\mathbb{E}\Big[\sup_{s\in[0,t]} \Theta_s^\lambda(X_s^{(1)}) - \Theta_s^\lambda(X_s^{(2)})|^k\Big]\Big)^2 = \Big(\mathbb{E} \sup_{s\in[0,t]} |Y_s^{(1)} - Y_s^{(2)}|^k\Big)^2\\ &\le c_2 |Y_0^{(1)} - Y_0^{(2)}|^{2k} \big(\mathbb{E} \mathrm{e}^{\frac{c_1 p}{p-1}\mathscr{L}_t}\big)^{\frac{p-1}{p}} \le c_3 |\Theta_0^\lambda(x^{(1)}) - \Theta_0^\lambda(x^{(2)})|^{2k}.\end{aligned}$$

This together with (2.80) implies (2.14) for some constant $c > 0$. □

To prove (2.14) under $(A_2^{\sigma,b})$, we need the following lemma due to [39, Lemma 5.2], which is contained in the proof of [9, Lemma 4.4]. Let ∇_1 and ∇_2 be the gradient operators in the first and second variables on $\mathbb{R}^d \times \mathbb{R}^d$.

Lemma 2.11. *There exists a function $g \in C^1(\mathbb{R}^d \times \mathbb{R}^d) \cap C^2((\mathbb{R}^d \setminus \{0\}) \times \mathbb{R}^d)$ having the following properties for some constants $k_2 > 1$ and $k_1 \in (0, 1)$:*

(1) $k_1 |x|^2 \le g(x, y) \le k_2 |x|^2, \quad x, y \in \mathbb{R}^d$;

(2) $\langle \nabla_1 g(x, y), y\rangle \le 0, \quad |y| = 1, \langle x, y\rangle \le k_1 |x|$;

(3) $|\nabla_1^i \nabla_2^j g(x, y)| \le k_2 |x|^{2-i}, \quad i, j \in \{0, 1, 2\}, i + j \le 2, x, y \in \mathbb{R}^d$.

Proof of Theorem 2.2 under (ii). Let $b^{0,n}$ be the mollifying approximation of $b^{(0)} = 1_{\bar{D}} b^{(0)}$. By Lemma 2.6, there exists $\lambda_0 > 0$ such that for any $\lambda \ge \lambda_0$ and $n \ge 1$, the PDE

$$(\partial_t + L_t + \nabla_{b_t^{0,n} - b_t^{(0)}} - \lambda) u_t^{\lambda,n} = -b_t^{0,n}, \quad u_T^{\lambda,n} = \nabla_{\mathbf{n}} u_t^{\lambda,n}|_{\partial D} = 0, \tag{2.86}$$

has a unique solution in $C_b^{1,2}([0, T] \times \bar{D})$, and there exist constants $\varepsilon, c > 0$ such that

$$\begin{aligned}&\lambda^\varepsilon \big(\|u^{\lambda,n}\|_\infty + \|\nabla u^{\lambda,n}\|_\infty\big) + \|(\partial_t + \nabla_{b^{(1)}}) u^{\lambda,n}\|_{\tilde{L}_{q_0}^{p_0}(T,D)} + \|\nabla^2 u^{\lambda,n}\|_{\tilde{L}_{q_0}^{p_0}(T,D)}\\ &\le c\|b^{(0)}\|_{\tilde{L}_{q_0}^{p_0}(T,D)}, \quad \lambda \ge \lambda_0, n \ge 1.\end{aligned} \tag{2.87}$$

Then for large enough $\lambda_0 > 0$, $\Theta_t^{\lambda,n} := id + u_t^{\lambda,n}$ satisfies

$$\frac{1}{2}|x - y|^2 \le |\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y)|^2 \le 2|x - y|^2, \quad \lambda \ge \lambda_0, x, y \in \bar{D}. \tag{2.88}$$

Since $\partial D \in C_b^{2,L}$, there exists a constant $r_0 > 0$ such that $\rho \in C_b^2(\partial_{r_0} D)$ with $\nabla^2\rho$ Lipschitz continuous on $\partial_{r_0} D$. Take $h \in C^\infty([0,\infty);[0,\infty))$ such that $h' \geq 0$, $h(r) = r$ for $r \leq r_0/2$ and $h(r) = r_0$ for $r \geq r_0$.

Let $(X_t^{(i)}, l_t^{(i)})$ solve (2.1) starting at $x^{(i)} \in \bar{D}$ for $i = 1, 2$. Alternatively to $|X_t^{(1)} - X_t^{(2)}|^2$, we consider the process

$$H_t := g\big(\Theta_t^{\lambda,n}(X_t^{(1)}) - \Theta_t^{\lambda,n}(X_t^{(2)}), \nabla(h\circ\rho)(X_t^{(1)})\big), \quad t \in [0,T],$$

where g is in Lemma 2.11. By Lemma 2.11(1) and (2.88), we have

$$\frac{k_1}{2}|X_t^{(1)} - X_t^{(2)}|^2 \leq H_t \leq 2k_2|X_t^{(1)} - X_t^{(2)}|^2, \quad t \in [0,T]. \tag{2.89}$$

Simply denote

$$\xi_t := \Theta_t^{\lambda,n}(X_t^{(1)}) - \Theta_t^{\lambda,n}(X_t^{(2)}), \quad \eta_t := \nabla(h\circ\rho)(X_t^{(1)}).$$

By Itô's formula, (2.86) and $\nabla_{\mathbf{n}}\Theta_t^{\lambda,n}|_{\partial D} = \mathbf{n}$ due to $\nabla_{\mathbf{n}} u_t^{\lambda,n}|_{\partial D} = 0$, we have

$$\begin{aligned}
\mathrm{d}\xi_t =& \big\{\lambda u_t^{\lambda,n}(X_t^{(1)}) - \lambda u_t^{\lambda,n}(X_t^{(2)}) + (b_t^{(0)} - b_t^{0,n})(X_t^{(1)}) - (b_t^{(0)} - b_t^{0,n})(X_t^{(2)})\big\}\mathrm{d}t \\
&+ \big\{[(\nabla\Theta_t^{\lambda,n})\sigma_t](X_t^{(1)}) - [(\nabla\Theta_t^{\lambda,n})\sigma_t](X_t^{(2)})\big\}\mathrm{d}W_t + \mathbf{n}(X_t^{(1)})\mathrm{d}l_t^{(1)} - \mathbf{n}(X_t^{(2)})\mathrm{d}l_t^{(2)}, \\
\mathrm{d}\eta_t =& L_t\nabla(h\circ\rho)(X_t^{(1)})\mathrm{d}t + \big\{[\nabla^2(h\circ\rho)]\sigma_t\big\}(X_t^{(1)})\mathrm{d}W_t + \{\nabla_{\mathbf{n}}\nabla(h\circ\rho)\}(X_t^{(1)})\mathrm{d}l_t^{(1)}.
\end{aligned} \tag{2.90}$$

Hence, Itô's formula for H_t reads

$$\mathrm{d}H_t = A_t\mathrm{d}t + B_t^{(1)}\mathrm{d}l_t^{(1)} - B_t^{(2)}\mathrm{d}l_t^{(2)} + \mathrm{d}M_t, \tag{2.91}$$

where

$$\begin{aligned}
A_t :=& \big\langle \nabla_1 g(\xi_t,\eta_t),\ \lambda u_t^{\lambda,n}(X_t^{(1)}) - \lambda u_t^{\lambda,n}(X_t^{(2)})\big\rangle \\
&+ \big\langle \nabla_1 g(\xi_t,\eta_t),\ \nabla_{b_t^{(0)} - b_t^{0,n}}\Theta_t^{\lambda,n}(X_t^{(1)}) - \nabla_{b_t^{(0)} - b_t^{0,n}}\Theta_t^{\lambda,n}(X_t^{(2)})\big\rangle \\
&+ \big\langle \nabla_2 g(\xi_t,\eta_t),\ L_t\nabla(h\circ\rho)(X_t^{(1)})\big\rangle + \Big\langle \nabla_1^2 g(\xi_t,\eta_t), N_tN_t^*\Big\rangle_{HS} \\
&+ \big\langle \nabla_1\nabla_2 g(\xi_t,\eta_t),\ N_t\sigma_t(X_t^{(1)})^*\nabla^2(h\circ\rho)(X_t^{(1)})\big\rangle_{HS} \\
&+ \big\langle \nabla_2^2 g(\xi_t,\eta_t),\ \big\{[\nabla^2(h\circ\rho)]\sigma_t\sigma_t^*\nabla^2(h\circ\rho)\big\}(X_t^{(1)})\big\rangle_{HS}, \\
N_t :=& \big\{(\nabla\Theta_t^{\lambda,n})\sigma_t\big\}(X_t^{(1)}) - \big\{(\nabla\Theta_t^{\lambda,n})\sigma_t\big\}(X_t^{(2)}),
\end{aligned} \tag{2.92}$$

$$\begin{aligned}
B_t^{(1)} :=& \big\langle \nabla_1 g(\xi_t,\eta_t), \mathbf{n}(X_t^{(1)})\big\rangle + \big\langle \nabla_2 g(\xi_t,\eta_t), \nabla_{\mathbf{n}}\{\nabla(h\circ\rho)\}(X_t^{(1)})\big\rangle, \\
B_t^{(2)} :=& \big\langle \nabla_1 g(\xi_t,\eta_t), \mathbf{n}(X_t^{(2)})\big\rangle,
\end{aligned} \tag{2.93}$$

$$\begin{aligned}
\mathrm{d}M_t =& \big\langle \nabla_1 g(\xi_t,\eta_t), \big[\{(\nabla\Theta_t^{\lambda,n})\sigma_t\}(X_t^{(1)}) - \{(\nabla\Theta_t^{\lambda,n})\sigma_t\}(X_t^{(2)})\big]\mathrm{d}W_t\big\rangle \\
&+ \big\langle \nabla_2 g(\xi_t,\eta_t), \big[\{\nabla^2(h\circ\rho)\}\sigma_t\big](X_t^{(1)})\mathrm{d}W_t\big\rangle.
\end{aligned} \tag{2.94}$$

In the following we estimate these terms respectively.

Firstly, (1.2) implies

$$\langle \Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y), \mathbf{n}(x)\rangle \leq \frac{|x-y|^2}{2r_0} + \|\nabla u_t^{\lambda,n}\|_\infty |x-y|, \quad x \in \partial D, y \in \bar{D}.$$

Combining this with (2.87), we find constants $\varepsilon_0, \lambda_1 > 0$ such that for any $\lambda \geq \lambda_1$,

$$\begin{aligned}&\langle \Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y), \mathbf{n}(x)\rangle \leq k_1 |\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y)|,\\ &x \in \partial D, y \in \bar{D}, |x-y| \leq \varepsilon_0, n \geq 1, t \in [0,T].\end{aligned}$$

So, Lemma 2.11 yields

$$\begin{aligned}&\big\langle \nabla_1 g(\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y), \mathbf{n}(x)), \mathbf{n}(x)\big\rangle \leq k_2 1_{\{|x-y|>\varepsilon_0\}} |\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y)|\\ &\leq k_2 \varepsilon_0^{-1} |\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y)|^2, \quad x \in \partial D, y \in \bar{D}, n \geq 1, t \in [0,T].\end{aligned} \tag{2.95}$$

Next, by the same reason leading to (2.95), we find a constant $c_1 > 0$ such that

$$\begin{aligned}&\big\langle \nabla_1 g(\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y), \nabla(h\circ\rho)(x)), \mathbf{n}(y)\big\rangle\\ &\geq \big\langle \nabla_1 g(\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y), \mathbf{n}(y)), \mathbf{n}(y)\big\rangle\\ &- \big|\nabla_1 g(\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y), \nabla(h\circ\rho)(y)) - \nabla_1 g(\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y), \nabla(h\circ\rho)(x))\big|\\ &\geq -1_{\{|x-y|>\varepsilon_0\}} k_2 \varepsilon_0^{-1} |\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y)|^2\\ &- \|h'\|_\infty \|\nabla_1\nabla_2 g(\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y), \cdot)\|_\infty |\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y)|^2\\ &\geq -c_1 |\Theta_t^{\lambda,n}(x) - \Theta_t^{\lambda,n}(y)|^2, \quad x \in \bar{D}, y \in \partial D, n \geq 1, t \in [0,T].\end{aligned} \tag{2.96}$$

Moreover, by $(A_2^{\sigma,b})$ and $h\circ\rho \in C_b^{2,L}(\bar{D})$, there exists a constant $C > 0$ such that

$$|L_t\{\nabla(h\circ\rho)\}| \leq C(1 + |b_t^{(0)}|), \quad t \in [0,T].$$

Combining this with Lemma 2.11, Lemma 2.10, (2.89), and (2.92)-(2.96), we find a constant $K > 0$ such that

$$\begin{aligned}&|A_t| \leq K\big\{|b_t^{(0)} - b_t^{0,n}|^2(X_t^{(1)}) + |b_t^{(0)} - b_t^{0,n}|^2(X_t^{(2)})\big\}\\ &+ K|X_t^{(1)} - X_t^{(2)}|^2 \bigg\{1 + |b_t^{(0)}|(X_t^{(1)}) + \sum_{i=1}^2 \mathscr{M}_D \big\|\nabla\{(\nabla\Theta_t^{\lambda,n})\sigma_t\}\big\|^2(X_t^{(i)})\bigg\},\\ &\mathrm{d}\langle M\rangle_t \leq K|X_t^{(1)} - X_t^{(2)}|^4 \bigg\{1 + \sum_{i=1}^2 \mathscr{M}_D \big\|\nabla\{(\nabla\Theta_t^{\lambda,n})\sigma_t\}\big\|^2(X_t^{(i)})\bigg\},\\ &B_t^{(1)} \leq K|X_t^{(1)} - X_t^{(2)}|^2, \qquad -B_t^{(2)} \leq K|X_t^{(1)} - X_t^{(2)}|^2.\end{aligned}$$

Combining these with (2.89) and (2.91), for any $k \geq 1$, we find a constant $c_1 > 0$ such that

$$\begin{aligned}\mathrm{d}H_t^k \leq & c_1 |X_t^{(1)} - X_t^{(2)}|^{2(k-1)} \big\{|b_t^{(0)} - b_t^{0,n}|^2(X_t^{(1)}) + |b_t^{(0)} - b_t^{0,n}|^2(X_t^{(2)})\big\}\mathrm{d}t\\ &+ c_1 |X_t^{(1)} - X_t^{(2)}|^{2k} \mathrm{d}\mathscr{L}_t + kH_t^{k-1}\mathrm{d}M_t,\end{aligned} \tag{2.97}$$

where

$$\mathscr{L}_t := l_t^{(1)} + l_t^{(2)} + \int_0^t \bigg\{1 + |b_s^{(0)}|(X_s^{(1)}) + \sum_{i=1}^2 \mathscr{M}_D \big\|\nabla\{(\nabla\Theta_s^{\lambda,n})\sigma_s\}\big\|^2(X_s^{(i)})\bigg\}\mathrm{d}s. \tag{2.98}$$

For any $j \geq 1$, let

$$\tau_j := \inf\big\{t \geq 0 : |X_t^{(1)} - X_t^{(2)}| \geq j\big\}.$$

By (2.89) and (2.97), we find a constant $c_2 > 0$ such that

$$|X_{t\wedge\tau_j}^{(1)} - X_{t\wedge\tau_j}^{(2)}|^{2k} \leq G_j(t) + c_2 \int_0^{t\wedge\tau_j} |X_s^{(1)} - X_s^{(2)}|^{2k} \mathrm{d}\mathscr{L}_s + \tilde{M}_t \tag{2.99}$$

holds for some local martingale $\tilde{M}_t$ and

$$G_j(t) := c_2|x^{(1)} - x^{(2)}|^{2k} + c_2 j^{2(k-1)} \int_0^{t\wedge\tau_j} \big\{|b_s^{(0)} - b_s^{0,n}|^2(X_s^{(1)}) + |b_s^{(0)} - b_s^{0,n}|^2(X_s^{(2)})\big\}\mathrm{d}s.$$

Since $(A_2^{\sigma,b})$ and (2.87) imply

$$\sup_{n\geq 1} \big\|\nabla\{(\nabla\Theta^{\lambda,n})\sigma\}\big\| \leq \sum_{i=0}^{l} F_i$$

for some $0 \leq F_i \in \tilde{L}_{q_i}^{p_1}(T), 0 \leq i \leq l$, by (2.22), (2.59) and (2.60) in Lemma 2.7, the stochastic Gronwall lemma, and Lemma 2.10, for any $p \in (\frac{1}{2}, 1)$ there exist constants $c_3, c_4 > 0$ such that

$$\begin{aligned}\Big(\mathbb{E}\Big[\sup_{s\in[0,t\wedge\tau_j]} |X_s^{(1)} - X_s^{(2)}|^k\Big]\Big)^2 &\leq c_3(\mathbb{E}\mathrm{e}^{\frac{c_2 p}{1-p}\mathscr{L}_t})^{\frac{1-p}{p}} \mathbb{E}G_j(t)\\ &\leq c_4\big(|x^{(1)} - x^{(2)}|^{2k} + j^{2(k-1)}\|b^{(0)} - b^{0,n}\|_{\tilde{L}_{q_0}^{p_0}(T)}\big), \quad n, j \geq 1.\end{aligned}$$

By first letting $n \to \infty$ then $j \to \infty$ and applying (2.61), we prove (2.14) for some constant $c > 0$. □

2.5 Functional inequalities: proof of Theorem 2.3

Let $\{P_{s,t}\}_{t\geq s\geq 0}$ be the Markov semigroup associated with (2.1), i.e. $P_{s,t}f(x) := \mathbb{E}f(X_{s,t}^x)$ for $t \geq s, f \in \mathscr{B}_b(\bar{D})$, where $(X_{s,t}^x)_{t\geq s}$ is the unique solution of (2.1) starting from x at time s. We have

$$P_t f(x) = \mathbb{E}(P_{s,t}f)(X_s^x), \quad s \in [0,t], f \in C_b^1(\bar{D}), \tag{2.100}$$

where $X_s^x := X_{0,s}^x$. By (2.15) for (2.1) from time s, for any $p > 1$, we have

$$|\nabla P_{s,t}f| \leq c(p)(P_{s,t}|\nabla f|^p)^{\frac{1}{p}}, \quad 0 \leq s \leq t \leq T, f \in C_b^1(\bar{D}). \tag{2.101}$$

Recall that $C_N^2(\bar{D})$ is the set of C^2-functions f on $\bar{D}$ with compact support and Neumann boundary condition $\nabla_{\mathbf{n}} f|_{\partial D} = 0$. If $P_{\cdot,t}f \in C^{1,2}([0,t] \times \bar{D})$ for $f \in C_N^2(\bar{D})$ such that

$$(\partial_s + L_s)P_{s,t}f = 0, \quad f \in C_N^2(\bar{D}), \nabla_{\mathbf{n}} P_{s,t}f|_{\partial D} = 0, \tag{2.102}$$

then the desired inequalities follow from (2.101) by taking derivative in s to the following reference functions respectively:

$$P_s\{P_{s,t}(\varepsilon + f)\}^p, \quad P_s\{P_{s,t}(\varepsilon + f)\}^2, \quad P_s\{\log P_{s,t}(\varepsilon + f)\}(x + s(y - s)/t), \quad s \in [0,t],$$

see for instance the proof of [36, Theorem 3.1]. However, in the present singular setting it is not clear whether (2.102) holds or not. So, below we make an approximation argument.

(a) Proof of (2.16). Let $\{b^{0,n}\}_{n\geq 1}$ be the mollifying approximations of $b^{(0)}$. By $(A_2^{\sigma,b})$, for any $f\in C_N^2(\bar D)$ and $t\in(0,T]$, the equation

$$u_{s,t}^n = P_{s,t}^{\sigma,b^{(1)}}f + \int_s^t P_{s,r}^{\sigma,b^{(1)}}(\nabla_{b_r^{0,n}}u_{s,t}^n)\mathrm{d}r,\quad s\in[0,t].$$

has a unique solution in $C^{1,2}([0,t]\times\bar D)$, and $P_{s,t}^n f := u_{s,t}^n$ satisfies

$$(\partial_s + L_s^{\sigma,b^{(1)}} + \nabla_{b_s^{0,n}})P_{s,t}^n f = 0,\quad s\in[0,t], f\in C_N^2(\bar D). \tag{2.103}$$

By this and Itô's formula for the SDE

$$\mathrm{d}X_{s,t}^{x,n} = (b_t^{(1)} + b_t^{0,n})(X_{s,t}^{x,n})\mathrm{d}t + \sigma_t(X_{s,t}^{x,n})\mathrm{d}W_t + \mathbf{n}(X_t^{x,n})\mathrm{d}l_t,\quad t\geq s, X_{s,s}^{x,n}=x,$$

we obtain $P_{s,t}^n f(x) = \mathbb{E}f(X_{s,t}^{x,n})$ for $0\leq s\leq t$. Let X_t solve (2.1) from time s with $X_s=x$, and define

$$R_s := \mathrm{e}^{\int_0^s\langle\xi_r^n,\mathrm{d}W_r\rangle - \frac{1}{2}\int_0^s|\xi_r^n|^2\mathrm{d}r},\quad \xi_s^n := \big\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}(b_s^{(0)} - b_t^{0,n})\big\}(X_s),\quad s\in[0,t].$$

By Girsanov's theorem, we obtain

$$\begin{aligned}&|P_{s,t}f - P_{s,t}^n f|(x) = |\mathbb{E}[f(X_t) - R_t f(X_t)]|\\ &\leq \|f\|_\infty\big(\mathbb{E}\mathrm{e}^{c\int_0^t|b_s^{(0)} - b_s^{0,n}|^2(X_s)} - 1\big) =: \|f\|_\infty\varepsilon_n,\quad 0\leq s\leq t\leq T,\end{aligned}$$

where $c>0$ is a constant and due to (2.60), $\varepsilon_n\to 0$ as $n\to\infty$. Consequently,

$$\|P_{s,t}f - P_{s,t}^n f\|_\infty \leq \varepsilon_n\|f\|_\infty,\quad n\geq 1, 0\leq s\leq t\leq T. \tag{2.104}$$

Moreover, the proof of (2.101) implies that it holds for $P_{s,t}^n$ replacing $P_{s,t}$ uniformly in $n\geq 1$, since the constant is increasing in $\|b^{(0)}\|_{\tilde L_{q_0}^{p_0}(T)}$, which is not less that $\|b^{0,n}\|_{\tilde L_{q_0}^{p_0}}(T)$. Thus,

$$|\nabla P_{s,t}^n f| \leq c(p)(P_{s,t}^n|\nabla f|^p)^{\frac{1}{p}},\quad 0\leq s\leq t\leq T, f\in C_b^1(\bar D), n\geq 1. \tag{2.105}$$

Now, let $0\leq f\in C_N^2(\bar D)$ and $t\in(0,T]$. For any $\varepsilon>0$ and $p\in(1,2]$, by (2.105), (2.103), (2.104), $(A_2^{\sigma,b})$ and Itô's formula, we find constants $c_1,c_2>0$ such that

$$\begin{aligned}\mathrm{d}(\varepsilon + P_{s,t}^n f)^p(X_s) &= \big\{p(\varepsilon + P_{s,t}^n f)^{p-1}\langle b_t^{(0)} - b_t^{0,n}, \nabla P_{s,t}^n f\rangle\\ &\qquad + p(p-1)(\varepsilon + P_{s,t}^n f)^{p-2}|\sigma_s^*\nabla P_{s,t}^n f|^2\big\}(X_s)\mathrm{d}s + \mathrm{d}M_s\\ &\geq \big\{c_2(\varepsilon + P_{s,t}^n f)^{p-2}|\nabla P_{s,t}^n f|^2 - c_1\|\nabla f\|_\infty|b_t^{(0)} - b_t^{0,n}|\big\}(X_s)\mathrm{d}s + \mathrm{d}M_s,\quad s\in[0,t],\varepsilon>0\end{aligned}$$

holds for some martingale M_s. By (2.20), Hölder's inequality, and $\|b^{(0)} - b^{0,n}\|_{\tilde L_{q_0}^{p_0}(T)}\to 0$ as $n\to\infty$, we find a constant $c_3>0$ and sequence $\varepsilon_n\to 0$ as $n\to\infty$ such that

$$\varepsilon_n + P_t(\varepsilon + f)^p - (P_t^n f + \varepsilon)^p \geq c_2\int_0^t P_s\big\{(\varepsilon + P_{s,t}^n f)^{p-2}|\nabla P_{s,t}^n f|^2\big\}\mathrm{d}s$$

$$\geq c_2 \int_0^t \frac{(P_s|\nabla P_{s,t}^n f|^p)^{\frac{2}{p}}}{\{P_s(\varepsilon + P_{s,t}^n f)^p\}^{\frac{2-p}{p}}} \mathrm{d}s \geq c_3 \int_0^t \frac{|\nabla P_s P_{s,t}^n f|^2}{\{P_s(\varepsilon + P_{s,t}^n f)^p\}^{\frac{2-p}{p}}} \mathrm{d}s, \quad \varepsilon \in (0,1).$$

Thus, for any $x \in D$ and $x \neq y \in B(x,\delta) \subset D$ for small $\delta > 0$ such that

$$x_r := x + r(y-x) \in D, \quad r \in [0,1],$$

this implies

$$\begin{aligned}
&\frac{|\int_0^t (P_s P_{s,t}^n f(x) - P_s P_{s,t}^n(y))\mathrm{d}s|}{|x-y|} \leq \int_0^1 \mathrm{d}r \int_0^t |\nabla P_s P_{s,t}^n f|(x_r))\mathrm{d}s \\
&\leq \int_0^1 \bigg(\int_0^t \frac{|\nabla P_s P_{s,t}^n f|^2}{\{P_s(\varepsilon + P_{s,t}^n f)^p\}^{\frac{2-p}{p}}}(x_r)\mathrm{d}s \bigg)^{\frac{1}{2}} \bigg(\int_0^t \{P_s(\varepsilon + P_{s,t}^n f)^p)\}^{\frac{2-p}{p}}(x_r)\mathrm{d}s \bigg)^{\frac{1}{2}} \mathrm{d}r \\
&\leq \int_0^1 c_3^{-1/2} \big\{\varepsilon_n + P_t(\varepsilon + f)^p\big\}^{\frac{1}{2}}(x + r(y-x)) \bigg(\int_0^t (\varepsilon + P_s P_{s,t}^n f^p)^{\frac{2-p}{p}})(x_r)\mathrm{d}s \bigg)^{\frac{1}{2}} \mathrm{d}r.
\end{aligned}$$

Combining this with (2.104) and letting $n \to \infty, \varepsilon \to 0$, we obtain

$$\frac{|P_t f(x) - P_t f(y)|}{|x-y|} \leq \frac{1}{t} \int_0^1 (c_3^{-1} P_t f^p)^{\frac{1}{2}}(x_r) \bigg(\int_0^t (P_t f^p)^{\frac{2-p}{p}}(x_r)\mathrm{d}s \bigg)^{\frac{1}{2}} \mathrm{d}r.$$

Letting $y \to x$ we prove (2.16) for some constant c depending on p, for $p \in (1,2]$ and all $f \in C_N^2(\bar{D})$. By Jensen's inequality the estimate also holds for $p > 2$, and by approximation argument, it holds for all $f \in \mathscr{B}_b(\bar{D})$.

(b) Proof of (2.17). By (2.105), Itô's formula and $(A_2^{\sigma,b})$, we find a constant $c_4 > 0$ and a martingale M_s such that

$$\begin{aligned}
&\mathrm{d}(P_{s,t}^n f)^2(X_s) = 2\big\{\langle \nabla P_{s,t}^n f, b_s^{(0)} - b_s^{0,n}\rangle + |\sigma_s^* \nabla P_{s,t}^n f|^2\big\}(X_s)\mathrm{d}s + \mathrm{d}M_s \\
&\leq c_4 \big\{\|\nabla f\|_\infty |b_s^{(0)} - b_s^{0,n}| + P_{s,t}^n |\nabla f|^2\big\}(X_s)\mathrm{d}s + \mathrm{d}M_s, \quad s \in [0,t].
\end{aligned}$$

Integrating both sides over $s \in [0,t]$, taking expectations and letting $n \to \infty$, and combining with (2.20) and (2.104), we prove (2.17).

(c) Proof of (2.18). Let $0 < f \in C_N^2(\bar{D})$. By taking Itô's formula to $P_{s,t}^n(\varepsilon + f)(X_s)$ for $\varepsilon > 0$ and taking expectation, we derive

$$\frac{\mathrm{d}}{\mathrm{d}s} P_s \log P_{s,t}^n \{\varepsilon + f\} = -P_s |\sigma_s^* \nabla \log P_{s,t}^n f|^2 + P_s \langle b_s^{(0)} - b_s^{0,n}, \nabla \log P_{s,t}^n (\varepsilon + f)\rangle.$$

For any $x, y \in \bar{D}$, let $\gamma : [0,1] \to \bar{D}$ be a curve linking x and y such that $|\dot{\gamma}_r| \leq c|x-y|$ for some constant $c > 0$ independent of x, y. Combining these with $(A_2^{\sigma,b})$ and (2.15) for $p = 2$ we find a constant $c_5 > 0$ such that

$$\begin{aligned}
&P_t \log\{\varepsilon + f\}(x) - \log P_t^n \{\varepsilon + f\}(y) = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} P_s \log P_{s,t}^n f(\gamma_{s/t})\mathrm{d}s \\
&\leq \int_0^t \big\{ct^{-1}|x-y||\nabla P_s \log P_{s,t}^n f(\gamma_{s/t})| - P_s |\sigma_s^* \nabla \log P_{s,t}^n f|^2\big\}(\gamma_{s/t})\mathrm{d}s \\
&\leq c_5 \int_0^t \frac{|x-y|^2}{t^2}\mathrm{d}s = \frac{c_5 |x-y|^2}{t}, \quad t \in (0,T].
\end{aligned}$$

Therefore, (2.18) holds.

3 Well-posedness for DDRSDEs

To characterize the dependence on the distribution, we will use different probability distances. For a measurable function

$$\psi : \bar{D} \times \bar{D} \to [0, \infty) \text{ with } \psi(x, y) = 0 \text{ if and only if } x = y,$$

we introduce the associated Wasserstein "distance" (also called transportation cost)

$$\mathbb{W}_\psi(\mu, \nu) := \inf_{\pi \in \mathscr{C}(\mu,\nu)} \int_{\bar{D}\times\bar{D}} \psi(x, y)\pi(\mathrm{d}x, \mathrm{d}y), \quad \mu, \nu \in \mathscr{P}(\bar{D}), \tag{3.1}$$

where $\mathscr{C}(\mu, \nu)$ is the set of all couplings for μ and ν. In general, $\mathbb{W}_\psi$ is not necessarily a distance as it may be infinite and the triangle inequality may not hold. In particular, when $\psi(x, y) = |x - y|^k$ for some constant $k > 0$, the L^k-Wasserstein distance $\mathbb{W}_k := (\mathbb{W}_\psi)^{\frac{1}{1\vee k}}$ is a complete metric on the space

$$\mathscr{P}_k(\bar{D}) := \big\{\mu \in \mathscr{P}(\bar{D}) : \ \|\mu\|_k := \mu(|\cdot|^k)^{\frac{1}{k}} < \infty\big\},$$

where $\mu(f) := \int f\mathrm{d}\mu$ for $f \in L^1(\mu)$. When $k = 0$ we set $\|\mu\|_0 = 1$ such that $\mathscr{P}_2(\bar{D}) = \mathscr{P}(\bar{D})$ and $\mathbb{W}_0$ reduces to the total variation norm

$$\mathbb{W}_0(\mu, \nu) = \frac{1}{2}\|\mu - \nu\|_{var} := \frac{1}{2}\sup_{|f|\le1} |\mu(f) - \nu(f)| = \sup_{A\in\mathscr{B}(\bar{D})} |\mu(A) - \nu(A)|,$$

where $\mathscr{B}(\bar{D})$ is the Borel σ-algebra of $\bar{D}$. We will also use the weighted variation norm for $k > 0$:

$$\|\mu - \nu\|_{k,var} := \sup_{|f|\le1+|\cdot|^k} |\mu(f) - \nu(f)|, \quad \mu, \nu \in \mathscr{P}_k(\bar{D}).$$

According to [30, Theorem 6.15], there exists a constant $c > 0$ such that

$$\|\mu - \nu\|_{var} + \mathbb{W}_k(\mu, \nu)^{1\vee k} \le c\|\mu - \nu\|_{k,var}, \quad \mu, \nu \in \mathscr{P}_k(\bar{D}). \tag{3.2}$$

However, when $k > 1$, for any constant $c > 0$, $\mathbb{W}_k(\mu, \nu) \le c\|\mu - \nu\|_{k,var}$ does not hold. Indeed, by taking

$$\mu = \delta_0, \quad \nu = (1 - n^{-1-k})\delta_0 + n^{-1-k}\delta_{ne}, \quad n \ge 1, e \in \mathbb{R}^d \text{ with } |e| = 1,$$

we have $\mathbb{W}_k(\mu, \nu) = n^{-\frac{1}{k}}$, while

$$\|\mu - \nu\|_{k,var} = n^{-1-k}\|\delta_0 - \delta_{ne}\|_{k,var} \le n^{-1-k}\big\{\delta_0(1 + |\cdot|^k) + \delta_{ne}(1 + |\cdot|^k)\big\} \le \frac{3}{n}, \quad n \ge 1,$$

so that $\lim_{n\to\infty} \frac{\mathbb{W}_k(\mu,\nu)}{\|\mu-\nu\|_{k,var}} = \infty$ for $k > 1$.

In Theorem 3.1 below, we use the enlarged probability distance $\|\cdot\|_{k,var} + \mathbb{W}_k$ to measure the distribution dependence of the DDRSDE (1.5). For any subspace $\hat{\mathscr{P}}$ of $\mathscr{P}(\bar{D})$ and any $T \in (0, \infty]$, let $C([0, T]; \hat{\mathscr{P}})$ be the set of all continuous maps from $[0, T] \cap [0, \infty)$ to $\hat{\mathscr{P}}$ under the weak topology. For any $\mu \in C([0, \infty); \mathscr{P}(\bar{D}))$, let σ^μ and b^μ be in (1.7).

3.1 Singular case

We make the following assumption. Recall that $b_t^\mu := b_t(\cdot, \mu_t)$ for $\mu \in C([0,\infty); \mathscr{P}(\bar{D}))$.

(A1) *Let $T > 0$ and $k \geq 0$. $\sigma^\mu = \sigma$ does not depend on μ, and there exists $\hat{\mu} \in \mathscr{P}_k(\bar{D})$ such that at least* **one of the following two conditions** *holds.*

(1) $(A_2^{\sigma,\hat{b}})$ *holds for $\hat{b} := b(\cdot, \hat{\mu})$, and there exist a constant $\alpha \geq 0$ and $1 \leq f_i \in \tilde{L}_{q_i}^{p_i}(T, D), 0 \leq i \leq l$, such that for any $t \in [0,T]$, $x \in \bar{D}$, and $\mu, \nu \in \mathscr{P}_k(\bar{D})$,*

$$|b_t^\mu(x) - \hat{b}_t^{(1)}(x)| \leq f_0(t,x) + \alpha \|\mu\|_k, \tag{3.3}$$

$$|b_t^\mu(x) - b_t^\nu(x)| \leq \big\{\|\mu - \nu\|_{k,var} + \mathbb{W}_k(\mu,\nu)\big\} \sum_{i=0}^{l} f_i(t,x). \tag{3.4}$$

(2) $(A_1^{\sigma,\hat{b}})$ *holds, and* (3.3)-(3.4) *holds for $|f_i|^2 \in \sup_{(p,q)\in\mathscr{K}} \tilde{L}_q^p(T,D), 0 \leq i \leq l$.*

Since $\hat{b}_t^{(1)}$ is regular, (3.3) gives a control for the singular term of b^μ. Moreover, (3.4) is a Lipschitz condition on $b_t(x,\cdot)$ in $\|\cdot\|_{k,var} + \mathbb{W}_k$ with a singular Lipschitz coefficient.

Theorem 3.1. *Assume* **(A1)**.

(1) (1.5) *is weak well-posed up to time T for distributions in $\mathscr{P}_k(\bar{D})$. Moreover, for any $\gamma \in \mathscr{P}_k(\bar{D})$, and any $n > 0$, there exists a constant $c > 0$, such that*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^n \Big| X_0\Big] \leq c(1 + |X_0|^n), \quad \mathbb{E}\mathrm{e}^{nl_T} \leq c \tag{3.5}$$

holds for the solution with $\mathscr{L}_{X_0} = \gamma$.

(2) (1.5) *is well-posed up to time T for distributions in $\mathscr{P}_k(\bar{D})$ in each of the following situations:*

(i) $d = 1$ and **(A1)**(2) *holds.*

(ii) **(A1)**(1) *holds.*

To prove Theorem 3.1, we first present a general result on the well-posedness of the DDRSDE (1.5) by using that of the reflecting SDE (2.1).

For any $k \geq 0, \gamma \in \mathscr{P}_k, N \geq 2$, let

$$\mathscr{P}_{k,\gamma}^{T,N} = \Big\{\mu \in C([0,T]; \mathscr{P}_k(\bar{D})) : \mu_0 = \gamma, \sup_{t\in[0,T]} \mathrm{e}^{-Nt}(1 + \mu_t(|\cdot|^k)) \leq N\Big\}.$$

Then as $N \uparrow \infty$,

$$\mathscr{P}_{k,\gamma}^{T,N} \uparrow \mathscr{P}_{k,\gamma}^T := \big\{\mu \in C([0,T]; \mathscr{P}_k(\bar{D})) : \mu_0 = \gamma\big\}. \tag{3.6}$$

For any $\mu \in \mathscr{P}_{k,\gamma}^T$, we will assume that the reflecting SDE

$$\mathrm{d}X_t^{\mu,\gamma} = b_t(X_t^{\mu,\gamma}, \mu_t)\mathrm{d}t + \sigma_t(X_t^{\mu,\gamma})\mathrm{d}W_t + \mathbf{n}(X_t^{\mu,\gamma})\mathrm{d}l_t^{\mu,\gamma}, \quad t \in [0,T], \mathscr{L}_{X_0^{\mu,\gamma}} = \gamma \tag{3.7}$$

has a unique weak solution with

$$H_t^\gamma(\mu) := \mathscr{L}_{X_t^{\mu,\gamma}} \in \mathscr{P}_k(\bar{D}), \quad t \in [0,T].$$

(H2) Let $k \geq 0, T > 0$. For any $\gamma \in \mathscr{P}_k(\bar{D})$ and $\mu \in \mathscr{P}_{k,\gamma}^T$, (3.7) has a unique weak solution, and there exist constants $\{(p_i', q_i') > 1\}_{0 \leq i \leq l}$, $N_0 \geq 2$ and increasing maps $C : [N_0, \infty) \to (0, \infty)$ and $F : [N_0, \infty) \times [0, \infty) \to (0, \infty)$ such that for any $N \geq N_0$ and $\mu \in \mathscr{P}_{k,\gamma}^{T,N}$, the (weak) solution satisfies

$$H^\gamma(\mu) := \mathscr{L}_{(X_t^{\mu,\gamma})_{t\in[0,T]}} \in \mathscr{P}_{k,\gamma}^{T,N}, \tag{3.8}$$

$$\big(\mathbb{E}\big[(1 + |X_t^{\mu,\gamma}|^k)^2 \big| X_0^{\mu,\gamma}\big]\big)^{\frac{1}{2}} \leq C(N)(1 + |X_0^{\mu,\gamma}|^k), \quad t \in [0,T], \tag{3.9}$$

$$\begin{aligned} &\mathbb{E}\bigg(\int_0^t g_s(X_s^{\mu,\gamma})\mathrm{d}s\bigg)^2 \leq C(N)\|g\|^2_{\tilde{L}_{q_i'}^{p_i'}(t_0,t_1)}, \\ &\mathbb{E}\mathrm{e}^{\int_0^t g_s(X_s^{\mu,\gamma})\mathrm{d}s} \leq F(N, \|g\|_{\tilde{L}_{q_i}^{p_i}(t,D)}), \quad t \in [0,T], g \in \tilde{L}_{q_i'}^{p_i'}(t,D), 0 \leq i \leq l. \end{aligned} \tag{3.10}$$

Obviously, when $k = 0$, conditions (3.8) and (3.9) hold for $N_0 = 2$.

Theorem 3.2. *Assume* **(H2)** *and let $\sigma^\mu = \sigma$ do not depend on μ. Assume that there exist a measurable map $\Gamma : [0,T] \times \bar{D} \times \mathscr{P}(\bar{D}) \to \mathbb{R}^m$ such that*

$$b_t(x,\nu) - b_t(x,\mu) = \sigma_t(x)\Gamma_t(x,\nu,\mu), \quad x \in \bar{D}, t \in [0,T], \nu,\mu \in \mathscr{P}_k(\bar{D}). \tag{3.11}$$

Let $f := \big(\sum_{i=0}^l \tilde{f}_i\big)^{\frac{1}{2}}$ for some $1 \leq \tilde{f}_i \in \tilde{L}_{q_i'}^{p_i'}(T), 0 \leq i \leq l$.

(1) *If*

$$|\Gamma_t(x,\nu,\mu)| \leq f_t(x)\|\nu - \mu\|_{k,var}, \quad x \in \bar{D}, t \in [0,T], \nu,\mu \in \mathscr{P}_k(\bar{D}), \tag{3.12}$$

Then (1.5) is weak well-posed up to time T for distributions in $\mathscr{P}_k(\bar{D})$. If, furthermore, in **(H2)** *the SDE (3.7) is strongly well-posed for any $\gamma \in \mathscr{P}_k(\bar{D})$ and $\mu \in \mathscr{P}_{k,\gamma}^T$, so is (1.5) up to time T for distributions in $\mathscr{P}_k(\bar{D})$.*

(2) *Let $k > 1$ and for any $\mu, \nu \in \mathscr{P}_k(\bar{D})$,*

$$|\Gamma_t(x,\nu,\mu)| \leq f_t(x)\big\{\|\nu - \mu\|_{k,var} + \mathbb{W}_k(\mu,\nu)\big\}, \quad (t,x) \in [0,T] \times \bar{D}. \tag{3.13}$$

If for any $\gamma \in \mathscr{P}_k(\bar{D})$ and $N \geq N_0$, there exists a constant $C(N) > 0$ such that for any $\mu, \nu \in \mathscr{P}_{k,\gamma}^{T,N}$,

$$\mathbb{W}_k(H_t^\gamma(\mu), H_t^\gamma(\nu))^{2k} \leq C(N)\int_0^t \big\{\|\mu_s - \nu_s\|_{k,var}^{2k} + \mathbb{W}_k(\mu_s,\nu_s)^{2k}\big\}\mathrm{d}s, \quad t \in [0,T], \tag{3.14}$$

then assertions in (1) holds.

Proof. Let $\gamma \in \mathscr{P}_k(\bar{D})$. Then the weak solution to (3.7) is a weak solution to (1.5) if and only if μ is a fixed point of the map H^γ in $\mathscr{P}^T_{k,\gamma}$. So, if H^γ on $\mathscr{P}^T_{k,\gamma}$ has a unique fixed point in $\mathscr{P}^T_{k,\gamma}$, then the (weak) well-posedness of (3.7) implies that of (1.5). Thus, by (3.6), it suffices to show that for any $N \ge N_0$, H^γ has a unique fixed point in $\mathscr{P}^{T,N}_{k,\gamma}$. By (3.8) and the fixed point theorem, we only need to prove that for any $N \ge N_0$, H^γ is contractive with respect to a complete metric on $\mathscr{P}^{T,N}_{k,\gamma}$.

(1) For any $\lambda > 0$, consider the metric

$$\mathbb{W}_{k,\lambda,var}(\mu,\nu) := \sup_{t\in[0,T]} \mathrm{e}^{-\lambda t}\|\mu_t - \nu_t\|_{k,var}, \quad \mu,\nu \in \mathscr{P}^{T,N}_{k,\gamma}.$$

Let $(X^{\mu,\gamma}_t, l^{\mu,\gamma}_t)$ solve (3.7) for some Brownian motion W_t on a complete probability filtration space $(\Omega, \{\mathscr{F}_t\}, \mathbb{P})$. By (3.10), (3.12) or (3.13), we find a constant $c_1 > 0$ depending on N such that

$$\begin{aligned}
&\sup_{\mu,\nu\in\mathscr{P}^{T,N}_{k,\gamma}} \mathbb{E}\big(\mathrm{e}^{2\int_0^T |\Gamma_s(X^{\mu,\gamma}_s,\nu_s,\mu_s)|^2\mathrm{d}s}|\mathscr{F}_0\big) \le c_1^2,\\
&\sup_{\mu\in\mathscr{P}^{T,N}_{k,\gamma}} \mathbb{E}\bigg(\bigg(\int_0^T g_s(X^{\mu,\gamma}_s)\mathrm{d}s\bigg)^2\bigg|\mathscr{F}_0\bigg) \le c_1^2\|g\|^2_{\tilde{L}^{p_i}_{q_i}(T)}, \quad g \in \tilde{L}^{p_i}_{q_i}(T),\ 0\le i\le l.
\end{aligned} \tag{3.15}$$

Then by Girsanov's theorem,

$$\tilde{W}_t := W_t - \int_0^t \Gamma_s(X^{\mu,\gamma}_s,\nu_s,\mu_s)\mathrm{d}s, \quad t\in[0,T]$$

is a Brownian motion under the probability $\mathbb{Q} := R_T\mathbb{P}$, where

$$R_t := \mathrm{e}^{\int_0^t \langle \Gamma_s(X^{\mu,\gamma}_s,\nu_s,\mu_s),\mathrm{d}W_s\rangle - \frac{1}{2}\int_0^t |\Gamma_s(X^{\mu,\gamma}_s,\nu_s,\mu_s)|^2\mathrm{d}s}, \quad t\in[0,T]$$

is a $\mathbb{P}$-martingale. By (3.11), we may formulate (3.7) as

$$\mathrm{d}X^{\mu,\gamma}_t = b_t(X^{\mu,\gamma}_t,\nu_t)\mathrm{d}t + \sigma_t(X^{\mu,\gamma}_t)\mathrm{d}\tilde{W}_t + \mathbf{n}(X^{\mu,\gamma}_t)\mathrm{d}l^{\mu,\gamma}_t, \quad t\in[0,T], \mathscr{L}_{X^{\mu,\gamma}_0} = \gamma.$$

By the weak uniqueness due to **(H2)**, the definition of $\|\cdot\|_{k,var}$, (3.9) and (3.11), we obtain

$$\begin{aligned}
&\|H^\gamma_t(\mu) - H^\gamma_t(\nu)\|_{k,var} = \sup_{|\tilde{f}|\le 1+|\cdot|^k} \big|\mathbb{E}\big[(R_t-1)\tilde{f}(X^{\mu,\gamma}_t)\big]\big|\\
&\le \mathbb{E}\big[(1+|X^{\mu,\gamma}_t|^k)|R_t-1|\big] \le \mathbb{E}\Big[\big\{\mathbb{E}\big((1+|X^{\mu,\gamma}_t|^k)^2|\mathscr{F}_0\big)\big\}^{\frac{1}{2}}\big\{\mathbb{E}\big(|R_t-1|^2|\mathscr{F}_0\big)\big\}^{\frac{1}{2}}\Big]\\
&\le C(N)\mathbb{E}\Big[(1+|X^{\mu,\gamma}_0|^k)\big\{\mathbb{E}(\mathrm{e}^{\int_0^t |\Gamma_s(X^{\mu,\gamma}_s,\nu_s,\mu_s)|^2\mathrm{d}s} - 1|\mathscr{F}_0)\big\}^{\frac{1}{2}}\Big].
\end{aligned} \tag{3.16}$$

Moreover, (3.15) implies

$$\begin{aligned}
&\mathbb{E}(\mathrm{e}^{\int_0^t |\Gamma_s(X^{\mu,\gamma}_s,\nu_s,\mu_s)|^2\mathrm{d}s} - 1|\mathscr{F}_0)\\
&\le \mathbb{E}\bigg(\mathrm{e}^{\int_0^t |\Gamma_s(X^{\mu,\gamma}_s,\nu_s,\mu_s)|^2\mathrm{d}s}\int_0^t |\Gamma_s(X^{\mu,\gamma}_s,\nu_s,\mu_s)|^2\mathrm{d}s\bigg|\mathscr{F}_0\bigg)
\end{aligned}$$

$$
\begin{aligned}
&\le c_1\bigg\{\mathbb{E}\bigg(\bigg(\int_0^t |f_s(X_s^{\mu,\gamma})|^2\|\mu_s-\nu_s\|_{k,var}^2\mathrm{d}s\bigg)^2\bigg|\mathscr{F}_0\bigg)\bigg\}^{\frac12}\\
&\le c_1\mathrm{e}^{2\lambda t}\mathbb{W}_{k,\lambda,var}(\mu,\nu)^2\bigg\{\mathbb{E}\bigg(\bigg(\int_0^t |f_s(X_s^{\mu,\gamma})|^2\mathrm{e}^{-2\lambda(t-s)}\mathrm{d}s\bigg)^2\bigg|\mathscr{F}_0\bigg)\bigg\}^{\frac12}\\
&\le c_1^2\mathrm{e}^{2\lambda t}\sum_{i=0}^l\|\tilde f_i\mathrm{e}^{-2\lambda(t-\cdot)}\|_{\tilde L_{q_i'}^{p_i'}(t)}\mathbb{W}_{k,\lambda,var}(\mu,\nu)^2,\quad t\in[0,T].
\end{aligned}
$$

Combining this with (3.16) and the definition of $\mathbb{W}_{k,\lambda,var}$, we obtain

$$
\mathbb{W}_{k,\lambda,var}(H^\gamma(\mu),H^\gamma(\nu))\le C(N)(1+\gamma(|\cdot|^k))c_1\sqrt{\varepsilon(\lambda)}\mathbb{W}_{k,\lambda,var}(\mu,\nu),\quad \lambda>0, \tag{3.17}
$$

where

$$
\varepsilon(\lambda):=\sup_{t\in[0,T]}\sum_{i=0}^l\|\tilde f_i\mathrm{e}^{-2\lambda(t-\cdot)}\|_{\tilde L_{q_i'}^{p_i'}(t)}\downarrow 0\ \text{ as }\ \lambda\uparrow\infty.
$$

So, H^γ is contractive on $(\mathscr{P}_{k,\gamma}^{T,N},\mathbb{W}_{k,\lambda,var})$ for large enough $\lambda>0$.

(2) Let $k>1$. We consider the metric $\tilde{\mathbb{W}}_{k,\lambda,var}:=\mathbb{W}_{k,\lambda,var}+\mathbb{W}_{k,\lambda}$, where

$$
\mathbb{W}_{k,\lambda}(\mu,\nu):=\sup_{t\in[0,T]}\mathrm{e}^{-\lambda t}\mathbb{W}_k(\mu_t,\nu_t),\quad \mu,\nu\in\mathscr{P}_{k,\gamma}^{T,N}.
$$

By using (3.13) replacing (3.12), instead of (3.17) we find constants $\{C(N,\lambda)>0\}_{\lambda>0}$ with $C(N,\lambda)\to 0$ as $\lambda\to\infty$ such that

$$
\mathbb{W}_{k,\lambda,var}(H^\gamma(\mu),H^\gamma(\nu))\le C(N,\lambda)\tilde{\mathbb{W}}_{k,\lambda,var}(\mu,\nu),\quad \lambda>0,\mu,\nu\in\mathscr{P}_{k,\gamma}^{T,N}. \tag{3.18}
$$

On the other hand, (3.14) yields

$$
\begin{aligned}
&\mathbb{W}_{k,\lambda}(H^\gamma(\mu),H^\gamma(\nu))\le\sup_{t\in[0,T]}\bigg(C(N)\mathrm{e}^{-\lambda kt}\int_0^t\big\{\|\mu_s-\nu_s\|_{k,var}^{2k}+\mathbb{W}_k(\mu_s,\nu_s)^{2k}\big\}\mathrm{d}s\bigg)^{\frac{1}{2k}}\\
&\le\tilde{\mathbb{W}}_{k,\lambda,var}(\mu,\nu)\sup_{t\in[0,T]}\bigg(C(N)\int_0^t\mathrm{e}^{-2\lambda k(t-s)}\mathrm{d}s\bigg)^{\frac{1}{2k}}\le\frac{C(N)^{\frac{1}{2k}}}{(2\lambda k)^{\frac{1}{2k}}}\tilde{\mathbb{W}}_{k,\lambda,var}(\mu,\nu),\quad\lambda>0.
\end{aligned}
$$

Combining this with (3.18), we concluded that H^γ is contractive in $\mathscr{P}_{k,\gamma}^{T,N}$ under the metric $\tilde{\mathbb{W}}_{k,\lambda,var}$ when λ is large enough, and hence finish the proof. □

Proof of Theorem 3.1. Let $\gamma\in\mathscr{P}_k(\bar D)$ be fixed. By (3.3), for any $i=1,2$, condition $(A_i^{\sigma,\hat b})$ implies (A_i^{σ,b^μ}) for any $\mu\in C([0,\infty);\mathscr{P}_k(\bar D))$. So, by Theorem 2.1, **(A1)** implies the weak well-posedness of (3.7) for distributions in $\mathscr{P}_k(\bar D)$ with

$$
H_t^\gamma(\mu)\in\mathscr{P}_k(\bar D),\quad \mathbb{E}\mathrm{e}^{\lambda l_T^{\mu,\gamma}}<\infty,\quad \lambda>0,\gamma\in\mathscr{P}_k(\bar D),\mu\in C([0,\infty);\mathscr{P}_k(\bar D)), \tag{3.19}
$$

and also implies the strong well-posedness of (3.7) in each situation of Theorem 3.1(2). Moreover, by Lemma 2.5 and Lemma 2.7, **(A1)** implies that (3.10) holds for any $(p,q)\in\mathscr{K}$, as well as for $(p,q)=(p_0/2,q_0/2)$ under $(A_2^{\sigma,\hat b})$, (3.11) with (3.12) holds for $k\le 1$ due to (3.2), and (3.11) with (3.13) holds for $k>1$. Therefore, by Theorem 3.2, it remains to verify (3.5), (3.8), (3.9), and (3.14) for $k>1$. Since (3.9) and (3.8) are trivial for $k=0$, we only need to prove: (3.5), (3.9) and (3.8) for $k>0$, (3.14) for $k>1$ for case (i), and (3.14) for $k>1$ for case (ii).

(a) Simply denote

$$f_t(x):=\sum_{i=0}^l f_i(t,x).$$

We first prove that under **(A1)**, there exits a constant $c>0$ and an increasing function $c:[1,\infty)\to(0,\infty)$ such that for any $j\ge 1$ and $\mu\in\mathscr{P}_{k,\gamma}^T$,

$$\begin{aligned}&\mathbb{E}\bigg(\int_0^t |f_s(X_s^{\mu,\gamma})|^2\mathrm{d}s\bigg)^j\le c(j)+c(j)\bigg(\int_0^t\|\mu_s\|_k^2\mathrm{d}s\bigg)^j,\\ &\mathbb{E}\exp\bigg[j\int_0^t |f_s(X_s^{\mu,\gamma})|^2\mathrm{d}s\bigg]\le c(j)\exp\bigg[c\int_0^t\|\mu_s\|_k^2\mathrm{d}s\bigg],\quad t\in[0,T],\end{aligned}\tag{3.20}$$

where $X_t^{\mu,\gamma}$ solves (3.7). We will prove these estimates by Lemmas 2.5 and 2.7 for the following reflecting SDE:

$$\mathrm{d}\hat X_s=\hat b_s(\hat X_s)\mathrm{d}s+\sigma_s(\hat X_s)\mathrm{d}W_s+\mathbf{n}(\hat X_s)\mathrm{d}\hat l_s,\quad \hat X_0=X_0^{\mu,\gamma},\, s\in[0,t].$$

By (2.60) under **(A1)**(1), and (2.21) under **(A1)**(2), for any $j\ge1$ we find a constant $c_1(j)>0$ such that

$$\mathbb{E}\mathrm{e}^{j\int_0^t(|\hat b_s^{(0)}|^2+|f_s|^2)(\hat X_s)\mathrm{d}s}\le c_1(j),\quad t\in[0,T].\tag{3.21}$$

Let $\gamma_s=\big\{[\sigma_s^*(\sigma_s\sigma_s^*)^{-1}](b_s^\mu-\hat b_s)\big\}(\hat X_s)$, and

$$R_t:=\mathrm{e}^{\int_0^t\langle\gamma_s,\mathrm{d}W_s\rangle-\frac12\int_0^t|\gamma_s|^2\mathrm{d}s},\quad \tilde W_s:=W_s-\int_0^s\gamma_r\mathrm{d}r,\quad s\in[0,t].$$

By Girsanov's theorem, $(\tilde W_s)_{s\in[0,t]}$ is a Brownian motion under $R_t\mathbb{P}$, and the SDE for $\hat X_s$ becomes

$$\mathrm{d}\hat X_s=b_s^\mu(\hat X_s)\mathrm{d}s+\sigma_s(\hat X_s)\mathrm{d}\tilde W_s+\mathbf{n}(\hat X_s)\mathrm{d}\hat l_s,\quad \hat X_0=X_0^{\mu,\gamma},\, s\in[0,t].$$

So, by (3.3), (3.21) and Hölder's inequality, we find constants $c_1,c,c(j)>0$ such that

$$\begin{aligned}&\mathbb{E}\mathrm{e}^{j\int_0^t|f_s(X_s^{\mu,\gamma})|^2\mathrm{d}s}=\mathbb{E}\big[R_t\mathrm{e}^{j\int_0^t|f_s(\hat X_s)|^2\mathrm{d}s}\big]\le\big(\mathbb{E}\mathrm{e}^{2j\int_0^t|f_s(\hat X_s)|^2\mathrm{d}s}\big)^{\frac12}\big(\mathbb{E}[R_t^2]\big)^{\frac12}\\ &\le\sqrt{c_1(2j)}\big(\mathbb{E}\mathrm{e}^{c_1\int_0^t\{|\hat b_s^{(0)}|^2+(f_s+\alpha\|\mu_s\|_k)^2\}(\hat X_s)\mathrm{d}s}\big)^{\frac12}\le c(j)\mathrm{e}^{c\int_0^t\|\mu_s\|_k^2\mathrm{d}s}.\end{aligned}$$

Next, taking $c_2(j)>0$ large enough such that the function $r\mapsto[\log(r+c_2(j))]^j$ is concave for $r\ge0$, so that this and Jensen's inequality imply

$$\mathbb{E}\bigg(\int_0^t|f_s(X_s^{\mu,\gamma})|^2\mathrm{d}s\bigg)^j\le\mathbb{E}\big(\big[\log(c_2(j)+\mathrm{e}^{\int_0^t|f_s(X_s^{\mu,\gamma})|^2\mathrm{d}s})\big]^j\big)$$

$$\leq \big[\log(c_2(j)+\mathbb{E}e^{\int_0^t |f_s(X_s^{\mu,\gamma})|^2 ds})\big]^j \leq c(j)+c(j)\bigg(\int_0^t \|\mu_s\|_k^2 ds\bigg)^j$$

holds for some constant $c(j)>0$. Therefore, (3.20) holds.

(b) Proof of (3.8). Simply denote $X_t = X_t^{\mu,\gamma}$. By (3.3), the boundedness of σ and the condition on $\hat{b}^{(1)}$ in $(A_0^{\sigma,\hat{b}})$ which follows from $(A_2^{\sigma,b})$ due to Lemma 2.6, we find a constant $c_1>0$ such that

$$L_{t,\mu} := \frac{1}{2}\mathrm{tr}\{\sigma_t\sigma_t^*\nabla^2\}+\nabla_{b_t^\mu},\quad L^{\sigma,\hat{b}^{(1)}} := \frac{1}{2}\mathrm{tr}\{\sigma_t\sigma_t^*\nabla^2\}+\nabla_{\hat{b}_t^{(1)}}$$

satisfy

$$L_{t,\mu}\tilde{\rho} \geq L_t^{\sigma,\hat{b}^{(1)}}\tilde{\rho}-|b_t^\mu-\hat{b}_t^{(1)}|\cdot|\nabla\tilde{\rho}| \geq -c_1(f_t+\|\mu_t\|_k).$$

Since $\langle \mathbf{n},\tilde{\rho}\rangle|_{\partial D}\geq 1$, by Itô's formula we obtain

$$\mathrm{d}\tilde{\rho}(X_t) \geq -c_1\big\{f_t(X_t)+\|\mu_t\|_k\big\}\mathrm{d}t+\mathrm{d}M_t+\mathrm{d}l_t \tag{3.22}$$

for some martingale M_t with $\langle M\rangle_t \leq ct$ for some constant $c>0$. This together with (3.20) yields that for some constant $k_0>0$,

$$\mathbb{E}l_t^k \leq k_0+k_0\mathbb{E}\bigg(\int_0^t \{f_s(X_s)+\|\mu_s\|_k\}\mathrm{d}s\bigg)^k.$$

Combining this with (2.20), (3.4), (3.20) and $\|\sigma\|_\infty<\infty$, and using the formula

$$X_t = X_0+\int_0^t b_s^\mu(X_s)\mathrm{d}s+\int_0^t \sigma_s(X_s)\mathrm{d}W_s+\int_0^t \mathbf{n}(X_ts\mathrm{d}l_s,\quad \mathscr{L}_{X_0}=\gamma,$$

we find constants $k_1,k_2>0$ such that

$$\begin{aligned}\mathbb{E}(1+|X_t|^k) &\leq k_1(1+\|\gamma\|_k^k)+k_1\mathbb{E}\bigg(\int_0^t \big\{|X_s|+|f_s(X_s)|+\|\mu_s\|_k\big\}\mathrm{d}s\bigg)^k \\ &\leq k_2+k_2\mathbb{E}\bigg(\int_0^t \big\{|X_s|^2+\|\mu_s\|_k^2\big\}\mathrm{d}s\bigg)^{\frac{k}{2}},\quad t\in[0,T].\end{aligned} \tag{3.23}$$

(b1) When $k\geq 2$, by (3.23) we find a constant $k_3>0$ such that

$$\mathbb{E}(1+|X_t|^k) \leq k_2+k_3\int_0^t \big\{\mathbb{E}|X_s|^k+\|\mu_s\|_k^k\big\}\mathrm{d}s,\quad t\in[0,T].$$

By Gronwall's lemma, and noting that $\mu\in\mathscr{P}_{k,\gamma}^{T,N}$, we find constant $k_4>0$ such that

$$\mathbb{E}(1+|X_t|^k) \leq k_4+k_4\int_0^t (1+\|\mu_s\|_k^k)\mathrm{d}s \leq k_4+k_4Ne^{Nt}\int_0^t e^{-N(t-s)}\mathrm{d}s \leq 2k_4e^{Nt},\quad t\in[0,T].$$

Taking $N_0=2k_4$ we prove

$$\sup_{t\in[0,T]} e^{-Nt}(1+\|H_t(\mu)\|_k^k) = \sup_{t\in[0,T]} e^{-Nt}\mathbb{E}(1+|X_t|^k) \leq N_0 \leq N,\quad N\geq N_0, \mu\in\mathscr{P}_{k,\gamma}^{T,N},$$

so that (3.8) holds.

(b2) When $k \in (0,2)$, by BDG's inequality, and by the same reason leading to (3.23), we find constants $k_5, k_6, k_7 > 0$ such that

$$\begin{aligned} U_t &:= \mathbb{E}\Big[\sup_{s\in[0,t]}(1+|X_s|^k)\Big] \le k_5 + k_5\mathbb{E}\bigg(\int_0^t \big\{|X_s|^2 + \|\mu_s\|_k^2\big\}\mathrm{d}s\bigg)^{\frac{k}{2}} \\ &\le k_6 + k_6\mathbb{E}\bigg\{\Big[\sup_{s\in[0,t]}|X_s|^k\Big]^{1-\frac{k}{2}}\bigg(\int_0^t |X_s|^k\mathrm{d}s\bigg)^{\frac{k}{2}}\bigg\} + k_6\bigg(\int_0^t \|\mu_s\|_k^2\mathrm{d}s\bigg)^{\frac{k}{2}} \\ &\le k_6 + \frac{1}{2}U_t + k_7\int_0^t U_s\mathrm{d}s + k_6\bigg(\int_0^t \|\mu_s\|_k^2\mathrm{d}s\bigg)^{\frac{k}{2}}, \quad t\in[0,T]. \end{aligned}$$

By Gronwall's lemma, we find constants $k_8, k_9 > 0$ such that for any $\mu \in \mathscr{P}_{k,\gamma}^{T,N}$,

$$\begin{aligned} \mathbb{E}(1+|X_t|^k) &\le U_t \le k_8 + k_8\bigg(\int_0^t \|\mu_s\|_k^2\mathrm{d}s\bigg)^{\frac{k}{2}} \\ &\le k_8 + k_8 N\mathrm{e}^{Nt}\bigg(\int_0^t \mathrm{e}^{-2N(t-s)/k}\mathrm{d}s\bigg)^{\frac{k}{2}} \le k_8 + k_9 N^{1-\frac{k}{2}}\mathrm{e}^{Nt}, \quad t\in[0,T]. \end{aligned}$$

Thus, there exists $N_0 > 0$ such that for any $N \ge N_0$,

$$\sup_{t\in[0,T]} \mathrm{e}^{-Nt}(1+\|H_t(\mu)\|_k^k) = \sup_{t\in[0,T]} \mathrm{e}^{-Nt}\mathbb{E}(1+|X_t|^k) \le k_8 + k_9 N^{1-\frac{k}{2}} \le N, \quad \mu \in \mathscr{P}_{k,\gamma}^{T,N},$$

which implies (3.8).

(c) Proofs of (3.9) and (3.5). Simply denote $(\hat{X}_t, \hat{l}_t) = (X_t^{\mu,\gamma}, l_t^{\mu,\gamma})$ in (3.7) for $\mu_t = \hat{\mu}, t \in [0,T]$; that is,

$$\mathrm{d}\hat{X}_t = \hat{b}_t(\hat{X}_t)\mathrm{d}t + \sigma(\hat{X}_t)\mathrm{d}W_t + \mathbf{n}(\hat{X}_t)\mathrm{d}\hat{l}_t, \quad \mathscr{L}_{\hat{X}_0} = \gamma. \tag{3.24}$$

By **(A1)** and Theorem 2.1, this SDE has a unique weak solution, and for any $n \ge 1$ there exists a constant $c > 0$ such that

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|\hat{X}_t|^n\Big|\hat{X}_0\Big] \le c(1+|\hat{X}_0|^n), \quad \mathbb{E}\mathrm{e}^{n\hat{l}_T} \le c. \tag{3.25}$$

So, by (3.4), Lemma 2.5, Lemma 2.7 under $(A_2^{\sigma,\hat{b}})$, and Girsanov's theorem,

$$\tilde{W}_t := W_t - \int_0^t \{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(\hat{X}_s)\big\{b_s^\mu(\hat{X}_s) - \hat{b}_s(\hat{X}_s)\big\}\mathrm{d}s, \quad t\in[0,T]$$

is a $\mathbb{Q}$-Brownian motion for $\mathbb{Q} := R_T\mathbb{P}$, where

$$R_T := \mathrm{e}^{\int_0^T \langle\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(\hat{X}_s)\{b_s^\mu(\hat{X}_s)-\hat{b}_s(\hat{X}_s)\},\mathrm{d}W_s\rangle - \frac{1}{2}\int_0^T |\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(\hat{X}_s)\{b_s^\mu(\hat{X}_s)-\hat{b}_s(\hat{X}_s)\}|^2\mathrm{d}s}.$$

By **(A1)**, (3.25), Lemma 2.5 when $|f_i|^2 \in \cup_{(p,q)\in\mathscr{K}} \tilde{L}_q^p(T)$, and Lemma 2.7 when $(A_2^{\sigma,\hat{b}})$ holds, we find an increasing function F such that

$$\mathbb{E}(|R_T|^2|\mathscr{F}_0) \le \mathbb{E}(\mathrm{e}^{\int_0^T |f_s(\hat{X}_s)|^2\{\|\mu_s-\hat{\mu}\|_{k,var}+\mathbb{W}_k(\mu_s,\hat{\mu})\}^2\mathrm{d}s}|\mathscr{F}_0) \le F(\|\mu\|_{k,T}),$$

where $\|\mu\|_{k,T} := \sup_{t\in[0,T]} \mu_t(|\cdot|^k)$. Reformulating (3.24) as

$$\mathrm{d}\hat{X}_t = b_t^\mu(\hat{X}_t)\mathrm{d}t + \sigma_t(\hat{X}_t)\mathrm{d}\tilde{W}_t + \mathbf{n}(\hat{X}_t)\mathrm{d}\hat{l}_t, \quad \mathscr{L}_{\hat{X}_0} = \gamma,$$

by the weak uniqueness we have $\mathscr{L}_{\hat{X}|\mathbb{Q}} = \mathscr{L}_{X^{\mu,\gamma}}$, so that (3.25) with $2n$ replacing n implies

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t^{\mu,\gamma}|^n\Big|\mathscr{F}_0\Big] = \mathbb{E}_{\mathbb{Q}}\Big[\sup_{t\in[0,T]} |\hat{X}_t|^n\Big|\mathscr{F}_0\Big] \\
&\le \bigg(\mathbb{E}\Big[\sup_{t\in[0,T]} |\hat{X}_t|^{2n}\Big|\mathscr{F}_0\Big]\bigg)^{\frac{1}{2}} (\mathbb{E}R_T^2|\mathscr{F}_0)^{\frac{1}{2}} \le c(1+|\hat{X}_0|^n)F(\|\mu\|_{k,T}).
\end{aligned}$$

Since $\sup_{\mu\in\mathscr{P}_{k,\gamma}^{T,N}} \|\mu\|_{k,T}$ is a finite increasing function of N, this implies (3.9).

Finally, since $X_t := X_t^{\mu,\gamma}$ solves (1.5) with initial distribution γ and $\mu_t = \mathscr{L}_{X_t}$ (i.e. μ is the fixed point of H^γ), and since H^γ has a unique fixed point in $\mathscr{P}_{k,\gamma}^{T,N}$ for some $N > 0$ depending on γ as proved in the proof of Theorem 2.1 using (3.10) and (3.8), we have $\mathscr{L}_{X_\cdot} \in \mathscr{P}_{k,\gamma}^{T,N}$, and hence (3.5) follows from (2.13).

(d) Proof of (3.14) for $k > 1$ in case (i). Let u_t^λ and Θ_t^λ be constructed for b^μ replacing b in the proof of Theorem 2.2 under $(A_1^{\sigma,b})$ for $d = 1$. Let $X_0^{(1)} = X_0^{(2)}$ be $\mathscr{F}_0$-measurable with $\mathscr{L}_{X_0^{(i)}} = \gamma, i = 1, 2$. As explained in the beginning in the present proof, the following reflecting SDEs are well-posed:

$$\begin{aligned}
&\mathrm{d}X_t^{(1)} = b_t(X_t^{(1)}, \mu_t)\mathrm{d}t + \sigma_t(X_t^{(1)})\mathrm{d}W_t + \mathbf{n}(X_t^{(1)})\mathrm{d}l_t^{(1)}, \\
&\mathrm{d}X_t^{(2)} = b_t(X_t^{(2)}, \nu_t)\mathrm{d}t + \sigma_t(X_t^{(2)})\mathrm{d}W_t + \mathbf{n}(X_t^{(2)})\mathrm{d}l_t^{(2)}, \quad t \in [0, T].
\end{aligned}$$

Then instead of (2.81), the processes

$$Y_t^{(i)} := \Theta_t^\lambda(X_t^{(i)}), \quad i = 1, 2$$

satisfy

$$\begin{aligned}
\mathrm{d}Y_t^{(1)} &= B_t(Y_t^{(1)})\mathrm{d}t + \Sigma_t(Y_t^{(1)})\mathrm{d}W_t + \{1 + \nabla u_t^\lambda(X_t^{(1)})\}\mathbf{n}(X_t^{(1)})\mathrm{d}l_t^{(1)}, \\
\mathrm{d}Y_t^{(2)} &= B_t(Y_t^{(2)})\mathrm{d}t + \Sigma_t(Y_t^{(2)})\mathrm{d}W_t + \{1 + \nabla u_t^\lambda(X_t^{(2)})\}\mathbf{n}(X_t^{(2)})\mathrm{d}l_t^{(2)} \\
&\qquad + \big\{b_t(X_t^{(2)}, \nu_t) - b_t(X_t^{(2)}, \mu_t)\big\}\mathrm{d}t.
\end{aligned}$$

By (3.4), $Y_0^{(1)} = Y_0^{(2)}$, Itô's formula to $|Y_t^{(1)} - Y_t^{(2)}|^{2k}$ with this formula replacing (2.81), the calculations in the proof of Theorem 2.2 under $(A_1^{\sigma,b})$ for $d = 1$ yield that when λ is large enough,

$$|Y_t^{(1)} - Y_t^{(2)}|^{2k} \le c_1 \int_0^t |Y_s^{(1)} - Y_s^{(2)}|^{2k}\mathrm{d}\mathscr{L}_s + M_t$$

$$+ c_1 \int_0^t |Y_s^{(1)} - Y_s^{(2)}|^{2k-1} f_s(X_s^{(2)}) \big\{ \|\mu_s - \nu_s\|_{k,var} + \mathbb{W}_k(\mu_s, \nu_s) \big\} \mathrm{d}s$$

$$\le c_1 \int_0^t |Y_s^{(1)} - Y_s^{(2)}|^{2k} \mathrm{d}\tilde{\mathscr{L}}_s + c_1 \int_0^t \big\{ \|\mu_s - \nu_s\|_{k,var} + \mathbb{W}_k(\mu_s, \nu_s) \big\}^{2k} \mathrm{d}s + M_t, \quad t \in [0, T]$$

holds for some constant $c_1 > 0$ depending on N uniformly in $\mu \in \mathscr{P}_{k,\gamma}^{T,N}$, some martingale M_t, $\mathscr{L}_t$ in (2.85), and

$$\tilde{\mathscr{L}}_t := \mathscr{L}_t + \int_0^t |f_s(X_s^{(2)})|^{\frac{2k}{2k-1}} \mathrm{d}s \le \mathscr{L}_t + \int_0^t |f_s(X_s^{(2)})|^2 \mathrm{d}s.$$

By the stochastic Gronwall lemma, Lemma 2.5, we find a constant $c_2 > 0$ depending on N such that

$$\Big(\mathbb{E}\Big[\sup_{s \in [0,t]} |Y_s^{(1)} - Y_s^{(2)}|^k \Big] \Big)^2 \le c_2 \int_0^t \big\{ \|\mu_s - \nu_s\|_{k,var} + \mathbb{W}_k(\mu_s, \nu_s) \big\}^{2k} \mathrm{d}s,$$

which implies (3.14) since by (2.80) and the definition of H^γ, there exists a constant $c > 0$ depending on N such that

$$(\mathbb{E}|Y_t^{(1)} - Y_t^{(2)}|^k)^2 \ge c(\mathbb{E}|X_t^{(1)} - X_t^{(2)}|^k)^2 \ge c\mathbb{W}_k(H_t^\gamma(\mu), H_t^\gamma(\nu))^{2k}.$$

(e) Proof of (3.14) for $k > 1$ in case (ii). Let $u_t^{\lambda,n}$ solve (2.86) for $L_t = L_{t,\nu}, b^{(0)} = b_t^{(0)}(\cdot, \nu_t)$ and the mollifying approximation $b^{0,n} = b_t^{0,n}(\cdot, \nu_t)$. Then in (2.90) the equation for ξ_t becomes

$$\begin{aligned} \mathrm{d}\xi_t = \Big\{ & \lambda u_t^{\lambda,n}(X_t^{(1)}) - \lambda u_t^{\lambda,n}(X_t^{(2)}) + (b_t^{(0)} - b_t^{0,n})(X_t^{(1)}) \\ & - (b_t^{(0)} - b_t^{0,n})(X_t^{(2)}) + b(X_t^{(2)}, \mu_t) - b_t(X_t^{(2)}, \nu_t) \Big\} \mathrm{d}t \\ & + \big\{ [(\nabla \Theta_t^{\lambda,n})\sigma_t](X_t^{(1)}) - [(\nabla \Theta_t^{\lambda,n})\sigma_t](X_t^{(2)}) \big\} \mathrm{d}W_t + \mathbf{n}(X_t^{(1)}) \mathrm{d}l_t^X - \mathbf{n}(X_t^{(2)}) \mathrm{d}l_t^{(2)}. \end{aligned}$$

So, as shown in step (d) by (3.4), instead of (2.99), we have

$$|X_{t\wedge\tau_m}^{(1)} - X_{t\wedge\tau_m}^{(2)}|^{2k} \le G_m(t) + c_2 \int_0^{t\wedge\tau_m} |X_{s\wedge\tau_m}^{(1)} - X_{s\wedge\tau_m}^2|^{2k} \mathrm{d}\tilde{\mathscr{L}}_s + \tilde{M}_t$$

for some local martingale $\tilde{M}_t$,

$$\tilde{\mathscr{L}}_t := \mathscr{L}_t + \int_0^t |f_s(X_s^{(2)})|^2 \mathrm{d}s, \quad t \in [0, T]$$

for $\mathscr{L}_t$ in (2.98), and due to $X_0^{(1)} = X_0^{(2)} = X_0$ in the present setting,

$$G_m(t) := \int_0^t \Big\{ c_2 m^{2(k-1)} \sum_{i=1}^2 |b_s^{(0)} - b_s^{0,n}|^2 (X_s^{(i)}) + \big(\|\mu_s - \nu_s\|_{k,var} + \mathbb{W}_k(\mu_s, \nu_s) \big)^{2k} \Big\} \mathrm{d}s.$$

By the stochastic Gronwall inequality, Lemma 2.7 and (3.20), we find a constant $c > 0$ such that

$$\begin{aligned} & \mathbb{W}_k(H_t^\gamma(\mu), H_t^\gamma(\nu))^{2k} \le (\mathbb{E}|X_t^{(1)} - X_t^{(2)}|^k)^2 \\ & \le c \liminf_{m\to\infty} \liminf_{n\to\infty} \mathbb{E} G_m(t) = c \int_0^t \big\{ \|\mu_s - \nu_s\|_{k,var}^{2k} + \mathbb{W}_k(\mu_s, \nu_s)^{2k} \big\} \mathrm{d}s. \end{aligned} \tag{3.26}$$

Thus, (3.14) holds. □

3.2 Monotone case

For any $k \geq 0$, $\mathscr{P}_k(\bar{D})$ is a complete metric space under the L^k-Wasserstein distance $\mathbb{W}_k$, where $\mathbb{W}_0(\mu,\nu) := \frac{1}{2}\|\mu-\nu\|_{var}$ and

$$\mathbb{W}_k(\mu,\nu) := \inf_{\pi\in\mathscr{C}(\mu,\nu)} \bigg(\int_{\bar{D}\times\bar{D}} |x-y|^k \pi(\mathrm{d}x,\mathrm{d}y)\bigg)^{\frac{1}{1\vee k}}, \quad \mu,\nu\in\mathscr{P}_k(\bar{D}), \quad k>0.$$

In the following, we first study the well-posedness of (1.5) for distributions in $\mathscr{P}_k(\bar{D})$ with $k>1$, then extend to a setting including $k=1$.

(A2) *Let* $k>1$. **(D)** *holds,* b *and* σ *are bounded on bounded subsets of* $[0,\infty)\times\bar{D}\times\mathscr{P}_k(\bar{D})$, *and the following two conditions hold.*

(1) *For any* $T>0$ *there exists a constant* $K>0$ *such that*

$$\|\sigma_t(x,\mu)-\sigma_t(y,\nu)\|_{HS}^2 + 2\langle x-y, b_t(x,\mu)-b_t(y,\nu)\rangle^+$$
$$\leq K\big\{|x-y|^2 + |x-y|\mathbb{W}_k(\mu,\nu) + 1_{\{k\geq 2\}}\mathbb{W}_k(\mu,\nu)^2\big\}, \quad t\in[0,T], x,y\in\bar{D}, \mu,\nu\in\mathscr{P}_k(\bar{D}).$$

(2) *There exists a subset* $\tilde{\partial}D\subset\partial D$ *such that*

$$\langle y-x, \mathbf{n}(x)\rangle \geq 0, \quad x\in\partial D\setminus\tilde{\partial}D, \ y\in\bar{D}, \tag{3.27}$$

and when $\tilde{\partial}D\neq\emptyset$, *there exists* $\tilde{\rho}\in C_b^2(\bar{D})$ *such that* $\tilde{\rho}|_{\partial D}=0$, $\langle\nabla\tilde{\rho},\mathbf{n}\rangle|_{\partial D}\geq 1_{\tilde{\partial}D}$ *and*

$$\sup_{(t,x)\in[0,T]\times\bar{D}} \big\{\|(\sigma_t^\mu)^*\nabla\tilde{\rho}\|^2(x) + \langle b_t^\mu, \nabla\tilde{\rho}\rangle^-(x)\big\} < \infty, \quad \mu\in C([0,T];\mathscr{P}_k(\bar{D})). \tag{3.28}$$

(A2)(1) is a monotone condition, when $k\geq 2$ it allows $\sigma_t(x,\mu)$ depending on μ, but when $k\in[1,2)$ it implies that $\sigma_t(x,\mu)=\sigma_t(x)$ does not depend on μ.

(A2)(2) holds for $\tilde{\partial}D=\emptyset$ when D is convex, and it holds for $\tilde{\partial}D=\partial D$ if $\partial D\in C_b^2$ and for some $r>0$

$$\sup_{(t,x)\in[0,T]\times\partial_{r_0}D} \big\{\|(\sigma_t^\mu)^*\nabla\rho\|^2(x) + \langle b_t^\mu, \nabla\rho\rangle^-(x)\big\} < \infty, \quad \mu\in C([0,T];\mathscr{P}_k(\bar{D})),$$

where in the second case we may take $\tilde{\rho}=h\circ\rho$ for $0\leq h\in C^\infty([0,\infty))$ with $h(r)=r$ for $r\leq r_0/2$ and $h(r)=r_0$ for $r\geq r_0$. In general, **(A2)**(2) includes the case where ∂D is partly convex and partly C_b^2.

Theorem 3.3. *Assume* **(A2)**. *Then* (1.5) *is well-posed for distributions in* $\mathscr{P}_k(\bar{D})$, *and for any* $T>0$, *there exist a constant* $C>0$ *and a map* $c:[1,\infty)\to(0,\infty)$ *such that for any solution* (X_t,l_t) *of* (1.5) *with* $\mathscr{L}_{X_0}\in\mathscr{P}_k(\bar{D})$,

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|X_t|^k\Big] \leq C(1+\mathbb{E}|X_0|^k), \tag{3.29}$$

$$\mathbb{E}\mathrm{e}^{n\tilde{l}_T} \leq c(n), \quad n\geq 1, \tilde{l}_T := \int_0^T 1_{\tilde{\partial}D}(X_t)\mathrm{d}l_t. \tag{3.30}$$

Proof. Let X_0 be $\mathscr{F}_0$-measurable with $\gamma := \mathscr{L}_{X_0} \in \mathscr{P}_k(\bar{D})$. Then

$$\mathscr{P}^T_{k,\gamma} := \big\{\mu \in C([0,T]; \mathscr{P}_k(\bar{D})) : \mu_0 = \gamma\big\}$$

is a complete space under the following metric for any $\lambda > 0$:

$$\mathbb{W}_k^{\lambda,T}(\mu,\nu) := \sup_{t\in[0,T]} \mathrm{e}^{-\lambda t}\mathbb{W}_k(\mu_t,\nu_t), \quad \mu,\nu \in \mathscr{P}^T_{k,\gamma}.$$

By Lemma 2.9, **(A2)** implies the well-posedness of the following reflecting SDE for any $\mu \in \mathscr{P}^T_{k,\gamma}$:

$$\mathrm{d}X_t^\mu = b_t(X_t^\mu,\mu_t)\mathrm{d}t + \sigma_t(X_t^\mu,\mu_t)\mathrm{d}W_t + \mathbf{n}(X_t^\mu)\mathrm{d}l_t^\mu, \quad X_0^\mu = X_0, \tag{3.31}$$

and the solution satisfies

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t^\mu|^k\Big] < \infty. \tag{3.32}$$

So, as explained in the proof of Theorem 3.2, for the well-posedness of (1.5), it suffices to prove the contraction of the map

$$\mathscr{P}^T_{k,\gamma} \ni \mu \mapsto H(\mu) := \mathscr{L}_{X^\mu} \in \mathscr{P}^T_{k,\gamma}$$

under the metric $\mathbb{W}_k^{\lambda,T}$ for large enough $\lambda > 0$.

Denote

$$\tilde{l}_t^\mu := \int_0^t 1_{\tilde{\partial} D}(X_s^\mu)\mathrm{d}l_s^\mu, \qquad \tilde{l}_t^\nu := \int_0^t 1_{\tilde{\partial} D}(X_s^\nu)\mathrm{d}l_s^\nu, \quad t \geq 0.$$

By (1.2), **(A2)** and Itô's formula, for any $k \geq 1$ we find a constant $c_1 > 0$ such that

$$\mathrm{d}|X_t^\mu - X_t^\nu|^k \leq c_1\big\{|X_t^\mu - X_t^\nu|^k + \mathbb{W}_k(\mu_t,\nu_t)^k\big\}\mathrm{d}t + \frac{k}{r_0}|X_t^\mu - X_t^\nu|^k(\mathrm{d}\tilde{l}_t^\mu + \mathrm{d}\tilde{l}_t^\nu) + \mathrm{d}M_t \tag{3.33}$$

for some martingale M_t with

$$\mathrm{d}\langle M\rangle_t \leq c_1\big\{|X_t^\mu - X_t^\nu|^{2k} + \mathbb{W}_k(\mu_t,\nu_t)^{2k}\big\}\mathrm{d}t.$$

To estimate $\int_0^t |X_s^\mu - X_s^\nu|^k(\mathrm{d}\tilde{l}_s^\mu + \mathrm{d}\tilde{l}_s^\nu)$, we take

$$0 \leq h \in C_b^\infty([0,\infty)) \text{ such that } h' \leq 0,\ h'(0) = -(1 + 2r_0^{-1}k),\ h(0) = 1, \tag{3.34}$$

where $r_0 > 0$ is in (1.2). Let

$$F(x,y) := |x-y|^k\big\{(h\circ\tilde{\rho})(x) + (h\circ\tilde{\rho})(y)\big\}, \quad x,y \in \bar{D}.$$

By **(A2)**(2), we have $\tilde{\rho}|_{\partial D} = 0$ and $\nabla_{\mathbf{n}}\tilde{\rho}|_{\partial D} \geq 1_{\tilde{\partial} D}$, so that (3.34) and (1.2) imply

$$\nabla_{\mathbf{n}}F(\cdot, X_t^\nu)(X_t^\mu)\mathrm{d}l_t^\mu + \nabla_{\mathbf{n}}F(X_t^\mu,\cdot)(X_t^\nu)\mathrm{d}l_t^\nu \leq -|X_t^\mu - X_t^\nu|^k(\mathrm{d}\tilde{l}_t^\mu + \mathrm{d}\tilde{l}_t^\nu).$$

Therefore, by **(A2)** and applying Itô's formula, we find a constant $c_2 > 0$ such that

$$\mathrm{d}F(X_t^\mu, X_t^\nu) \le c_2\big\{|X_t^\mu - X_t^\nu|^k + \mathbb{W}_k(\mu_t,\nu_t)^k\big\}\mathrm{d}t - |X_t^\mu - X_t^\nu|^k(\mathrm{d}\tilde{l}_t^\mu + \mathrm{d}\tilde{l}_t^\nu) + \mathrm{d}\tilde{M}_t$$

for some martingale $\tilde{M}_t$. This and $F(X_0^\mu, X_0^\nu) = F(X_0, X_0) = 0$ imply

$$\mathbb{E}\int_0^t |X_s^\mu - X_s^\nu|^k(\mathrm{d}\tilde{l}_s^\mu + \mathrm{d}\tilde{l}_s^\nu) \le c_2\int_0^t \big\{\mathbb{E}|X_s^\mu - X_s^\nu|^k + \mathbb{W}_k(\mu_s,\nu_s)^k\big\}\mathrm{d}s. \tag{3.35}$$

Substituting (3.35) into (3.33) and applying BDG's inequality, we find a constant $c_3 > 0$ such that

$$\zeta_t := \sup_{s\in[0,t]} |X_s^\mu - X_s^\nu|^k, \quad t\in[0,T]$$

satisfies

$$\mathbb{E}\zeta_t \le c_3\int_0^t \big\{\mathbb{E}\zeta_s + \mathbb{W}_k(\mu_s,\nu_s)^k\big\}\mathrm{d}s, \quad t\in[0,T], \tag{3.36}$$

so that for any $\lambda > c_3$,

$$\begin{aligned}\mathbb{E}\zeta_t &\le c_3\int_0^t \mathrm{e}^{c_3(t-s)}\mathbb{W}_k(\mu_s,\nu_s)^k\mathrm{d}s \le c_3\mathrm{e}^{k\lambda t}\mathbb{W}_k^{\lambda,T}(\mu,\nu)^k\int_0^t \mathrm{e}^{-(k\lambda-c_3)(t-s)}\mathrm{d}s\\ &\le \frac{c_3\mathrm{e}^{k\lambda t}}{k\lambda - c_3}\mathbb{W}_k^{\lambda,T}(\mu,\nu)^k, \quad t\in[0,T].\end{aligned} \tag{3.37}$$

Therefore, H is contractive in $\mathbb{W}_k^{\lambda,T}$ for large $\lambda > 0$ as desired.

It remains to prove (3.29) and (3.30). Let X_t be the unique solution to (1.5). By **(A2)**, for any $k > 1$, we find a constant $c(k) > 0$ such that

$$\mathrm{d}|X_t|^k \le c(k)\big\{1 + |X_t|^k + \mathbb{E}|X_t|^k\big\}\mathrm{d}t + k|X_t|^{k-2}\langle X_t, \sigma_t(X_t,\mathscr{L}_{X_t})\mathrm{d}W_t\rangle + k|X_t|^{k-1}\mathrm{d}\tilde{l}_t, \tag{3.38}$$

where $\mathrm{d}\tilde{l}_t := 1_{\tilde{\partial}D}(X_t)\mathrm{d}l_t$. By applying Itô's formula to $(1+|X_t|^k)(h\circ\tilde{\rho})(X_t)$, similarly to (3.35) we obtain

$$\mathbb{E}\int_0^t (1+|X_s|^k)\mathrm{d}\tilde{l}_s \le \tilde{c}(k)\int_0^t \mathbb{E}\big\{1+|X_s|^k\big\}\mathrm{d}s \tag{3.39}$$

for some constant $\tilde{c}(k) > 0$. Combining (3.39) with (3.38) and using Gronwall's lemma, we derive

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|X_t|^k\Big] \le c'(1+\mathbb{E}|X_0|^k)$$

for some constant $c' > 0$. Substituting this into (3.38) and using BDG's inequality, we prove (3.29) for some constant $c > 0$.

Finally, by **(A1)**(2) and applying Itô's formula to $\tilde{\rho}(X_t)$, we prove (3.30). □

We now solve (1.5) for distributions in

$$\mathscr{P}_\psi(\bar D) := \big\{\mu \in \mathscr{P}(\bar D) : \|\mu\|_\psi := \mu(\psi(|\cdot|)) < \infty\big\},$$

where ψ belongs to the following class for some $\kappa > 0$:

$$\begin{aligned}\Psi_\kappa := \big\{\psi \in C^2((0,\infty)) \cap C^1([0,\infty)) : \psi(0) = 0,\ \psi'|_{(0,\infty)} > 0, \|\psi'\|_\infty < \infty \\ r\psi'(r) + r^2\{\psi''\}^+(r) \le \kappa\psi(r) \text{ for } r > 0\big\}.\end{aligned} \tag{3.40}$$

Let $\mathbb{W}_\psi$ be in (3.1).

(A3) **(D)** *holds,* $\sigma_t(x,\mu) = \sigma_t(x)$ *does not depend on* μ*,* b *and* σ *are bounded on bounded subsets of* $[0,\infty) \times \bar D \times \mathscr{P}_\psi(\bar D)$ *for some* $\psi \in \Psi_\kappa$ *and* $\kappa > 0$*. Moreover, for any* $T > 0$ *there exists a constant* $K > 0$ *such that*

$$\begin{aligned}&\|\sigma_t(x) - \sigma_t(y)\|_{HS}^2 + 2\langle x - y, b_t(x,\mu) - b_t(y,\nu)\rangle^+ \\ &\le K|x-y|\big\{|x-y| + \mathbb{W}_\psi(\mu,\nu)\big\}, \quad t \in [0,T], x, y \in \bar D, \mu,\nu \in \mathscr{P}_k(\bar D).\end{aligned}$$

Theorem 3.4. *Assume* **(A3)** *and* **(A2)**(2)*. Then* (1.5) *is well-posed for distributions in* $\mathscr{P}_\psi(\bar D)$*, and*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \psi(|X_t|)\Big] < \infty, \quad T > 0, \mathscr{L}_{X_0} \in \mathscr{P}_\psi(\bar D). \tag{3.41}$$

Proof. Let X_0 be $\mathscr{F}_0$-measurable with $\mathbb{E}\psi(|X_0|) < \infty$, and consider the path space

$$\mathscr{P}_\psi^T := \big\{\mu \in C([0,T]; \mathscr{P}_\psi(\bar D)) : \mu_0 = \mathscr{L}_{X_0}\big\}.$$

For any $\lambda > 0$, the quasi-metric

$$\mathbb{W}_{\lambda,\psi}(\mu,\nu) := \sup_{t\in[0,T]} \mathrm{e}^{-\lambda t}\mathbb{W}_\psi(\mu_t,\nu_t), \quad \mu,\nu \in \mathscr{P}_\psi^T$$

is complete. By Lemma 2.9, **(A3)** implies the well-posedness of the SDE (3.31) for any $\mu \in \mathscr{P}_\psi^T$. By **(A2)**(2) and Itô's formula for $\gamma_t := \sqrt{1 + |X_t^\mu - X_0|^2}$, we find a constant $c_1 > 0$ such that

$$\mathrm{d}\gamma_t \le c_1\{\|\mu_t\|_\psi + \gamma_t\}\mathrm{d}t + \gamma_t^{-1}\langle X_t^\mu - X_0, \sigma_t(X_t^\mu)\mathrm{d}W_t\rangle + \mathrm{d}\tilde l_t^\mu,$$

where $\mathrm{d}\tilde l_t^\mu := 1_{\partial D}(X_t^\mu)\mathrm{d}l_t^\mu$. Combining this with $\psi \in \Psi_\kappa$ and the linear growth of $\|\sigma_t\|$ implied by **(A3)**, we find a constant $c_2 > 0$ such that

$$\mathrm{d}\psi(\gamma_t) \le c_2\{\|\mu_t\|_\psi + \psi(\gamma_t)\}\mathrm{d}t + \psi'(\gamma_t)\gamma_t^{-1}\langle X_t^\mu - X_0, \sigma_t(X_t^\mu)\mathrm{d}W_t\rangle + \psi'(\gamma_t)\mathrm{d}\tilde l_t^\mu. \tag{3.42}$$

Next, by **(A2)**(2), $\psi \in \Psi_\kappa$ which implies $\psi'(\gamma_t) \le \kappa\psi(\gamma_t)$ since $\gamma_t \ge 1$, and applying Itô's formula to $\psi(\gamma_t)\{\|\tilde\rho\|_\infty - \tilde\rho(X_t^\mu)\}$, we find a constant $c_3 > 0$ such that similarly to (3.35),

$$\mathbb{E}\int_0^t \psi'(\gamma_s)\mathrm{d}\tilde l_s^\mu \le \kappa\mathbb{E}\int_0^t \psi(\gamma_s)\mathrm{d}\tilde l_s^\mu \le c_3\mathbb{E}\int_0^t \big\{1 + \|\mu_s\|_\psi + \psi(|X_s^\mu|)\big\}\mathrm{d}s, \quad t \in [0,T]. \tag{3.43}$$

Combining this with (3.42), $r\psi'(r) \le \kappa\psi(r)$, the linear growth of σ_t ensured by **(A3)**, and applying BDG's inequality, we obtain

$$\mathbb{E}\Big[\sup_{t\in[0,T]}\psi(|X_t^\mu|)\Big] < \infty.$$

Consequently, (3.41) holds for solutions of (1.5) with $\mathscr{L}_{X_\cdot} \in \mathscr{P}_\psi^T$. So, as explained in the proof of Theorem 3.2, it remains to prove the contraction of the map

$$\mathscr{P}_\psi^T \ni \mu \mapsto H(\mu) := \mathscr{L}_{X^\mu} \in \mathscr{P}_\psi^T$$

under the metric $\mathbb{W}_{\lambda,\psi}$ for large enough $\lambda > 0$.

By (1.2), **(A2)**(2), $\|\psi'\|_\infty < \infty$ and $r\psi'(r) \le \kappa\psi(r)$, we obtain

$$\nabla_{\mathbf{n}}\{\psi(|\cdot - y|)\}(x) \le \frac{\kappa}{2r_0} 1_{\tilde{\partial} D}(x)\psi(|x-y|), \quad x \in \partial D, y \in \bar{D}. \tag{3.44}$$

Combining this with **(A3)** and Itô's formula, we find a constant $c_4 > 0$ such that

$$\mathrm{d}\psi(|X_t^\mu - X_t^\nu|) \le c_4\big\{\psi(|X_t^\mu - X_t^\nu|) + \mathbb{W}_\psi(\mu_t,\nu_t)\big\}\mathrm{d}t + c_4\psi(|X_t^\mu - X_t^\nu|)(\mathrm{d}\tilde{l}_t^\mu + \mathrm{d}\tilde{l}_t^\nu) + \mathrm{d}M_t \tag{3.45}$$

for some martingale M_t.

On the other hand, let $\varepsilon = \frac{r_0}{2\kappa}$ and take $h \in C^\infty([0,\infty))$ with $h' \ge 0, h(r) = r$ for $r \le \varepsilon/2$ and $h(r) = \varepsilon$ for $r \ge \varepsilon$. Consider

$$\eta_t := \psi(|X_t^\mu - X_t^\nu|)\big\{2\varepsilon - h\circ\tilde{\rho}(X_t^\mu) - h\circ\tilde{\rho}(X_t^\nu)\big\}.$$

By (3.44), **(A2)**(2), $\varepsilon = \frac{r_0}{2\kappa}$ and Itô's formula, we find a constant $c_5 > 0$ such that

$$\begin{aligned}\mathrm{d}\eta_t &\le c_5\big\{\psi(|X_t^\mu - X_t^\nu|) + \mathbb{W}_\psi(\mu_t,\nu_t)\big\}\mathrm{d}t + \Big(\frac{2\varepsilon\kappa}{2r_0} - 1\Big)\psi(|X_t^\mu - X_t^\nu|)(\mathrm{d}\tilde{l}_t^\mu + \mathrm{d}\tilde{l}_t^\nu) + \mathrm{d}\tilde{M}_t \\ &= c_5\big\{\psi(|X_t^\mu - X_t^\nu|) + \mathbb{W}_\psi(\mu_t,\nu_t)\big\}\mathrm{d}t - \frac{1}{2}\psi(|X_t^\mu - X_t^\nu|)(\mathrm{d}\tilde{l}_t^\mu + \mathrm{d}\tilde{l}_t^\nu) + \mathrm{d}\tilde{M}_t.\end{aligned}$$

Since $X_0^\mu = X_0^\nu = X_0$, this implies

$$\mathbb{E}\int_0^t \psi(|X_s^\mu - X_s^\nu|)(\mathrm{d}\tilde{l}_t^\mu + \mathrm{d}\tilde{l}_t^\nu) \le 2c_5\int_0^t \big\{\mathbb{E}\psi(|X_s^\mu - X_s^\nu|) + \mathbb{W}_\psi(\mu_s,\nu_s)\big\}\mathrm{d}s.$$

Substituting this into (3.45), we find a constant $c_6 > 0$ such that

$$\mathbb{W}_\psi(H_t(\mu), H_t(\nu)) \le \mathbb{E}\psi(|X_t^\mu - X_t^\nu|) \le c_6\int_0^t \mathbb{W}_\psi(\mu_s,\nu_s)\mathrm{d}s, \quad t \in [0,T],$$

so that H is contractive in $\mathbb{W}_{\lambda,\psi}$ for large $\lambda > 0$. Therefore, the proof is finished. □

4 Log-Harnack inequality and applications

In this section, we study the log-Harnack inequality introduced in [32] and applications for DDRSDEs with singular drift or under monotone conditions.

4.1 Singular case

(A4) *Let* $\partial D \in C_b^{2,L}$ *and* $T > 0$. $\sigma_t(x,\mu) = \sigma_t(x)$, *and there exists* $\hat\mu \in \mathscr{P}_2(\bar D)$ *such that* $(A_2^{\sigma,\hat b})$ *holds with* $p_1 > 2$, *where* $\hat b := b(\cdot,\hat\mu)$ *with regular term* $\hat b^{(1)}$. *Moreover, there exist a constant* $\alpha \ge 0$ *and* $1 \le f_i \in \tilde L_{q_i}^{p_i}(T), 0 \le i \le l$ *such that*

$$|b_t^\mu(x) - \hat b_t^{(1)}(x)| \le f_0(t,x) + \alpha\|\mu\|_2, \quad \mu \in \mathscr{P}_2(\bar D), (t,x) \in [0,T]\times \bar D, \tag{4.1}$$

$$|b_t^\mu(x) - b_t^\nu(x)| \le \mathbb{W}_2(\mu,\nu)\sum_{i=1}^l f_i(t,x), \quad \mu,\nu \in \mathscr{P}_2(\bar D), (t,x) \in [0,T]\times \bar D. \tag{4.2}$$

According to Theorem 3.1, **(A4)** implies the well-posedness of (1.5) up to time T for distributions in $\mathscr{P}_2(\bar D)$. Let

$$P_t^*\mu = \mathscr{L}_{X_t} \ \text{ for } \ X_t \text{ solving (1.5) with } \mathscr{L}_{X_0} = \mu \in \mathscr{P}_2(\bar D), \quad t \ge 0.$$

We consider

$$P_t f(\mu) := \int_{\bar D} f \mathrm{d}(P_t^*\mu), \quad t \ge 0, \mu \in \mathscr{P}_2(\bar D), f \in \mathscr{B}_b(\bar D),$$

where $\mathscr{B}_b(\bar D)$ is the class of all bounded measurable functions on $\bar D$.

Theorem 4.1. *Assume* **(A4)**. *For any* $N > 0$, *let* $\mathscr{P}_{2,N}(\bar D) := \{\mu \in \mathscr{P}_2(\bar D) : \|\mu\|_2 \le N\}$.

(1) *For any* $N > 0$, *there exists a constant* $C(N) > 0$ *such that for any* $\nu \in \mathscr{P}_{2,N}(\bar D)$ *and any* $t \in [0,T]$, *the following inequalities hold:*

$$\mathbb{W}_2(P_t^*\mu, P_t^*\nu)^2 \le C(N)\mathbb{W}_2(\mu,\nu)^2, \quad \mu \in \mathscr{P}_2(\bar D), \tag{4.3}$$

$$P_t \log f(\nu) \le \log P_t f(\mu) + \frac{C(N)}{t}\mathbb{W}_2(\mu,\nu)^2, \quad 0 < f \in \mathscr{B}_b(\bar D), \mu \in \mathscr{P}_{2,N}(\bar D), \tag{4.4}$$

$$\frac{1}{2}\|P_t^*\mu - P_t^*\nu\|_{var}^2 \le \mathrm{Ent}(P_t^*\nu|P_t^*\mu) \le \frac{C(N)}{t}\mathbb{W}_2(\mu,\nu)^2, \quad \mu \in \mathscr{P}_{2,N}(\bar D), \tag{4.5}$$

$$\|\nabla P_t f(\nu)\|_{\mathbb{W}_2} := \limsup_{\mu\to\nu \text{ in } \mathbb{W}_2} \frac{|P_t f(\nu) - P_t f(\mu)|}{\mathbb{W}_2(\mu,\nu)} \le \frac{\sqrt{2C(N)}}{\sqrt t}\|f\|_\infty, \quad f \in \mathscr{B}_b(\bar D). \tag{4.6}$$

(2) *Let* (4.1) *hold for* $\alpha = 0$. *Then there exists a constant* $C > 0$ *such that*

$$\mathbb{W}_2(P_t^*\mu, P_t^*\nu)^2 \le C\mathbb{W}_2(\mu,\nu)^2, \quad \mu,\nu \in \mathscr{P}_2(\bar D). \tag{4.7}$$

Moreover, if either $\sup_{1\le i\le l}\|f_i\|_\infty < \infty$ *or* D *is bounded, then* (4.4)-(4.6) *hold for some constant* C *replacing* $C(N)$ *and all* $\mu,\nu \in \mathscr{P}_2(\bar D)$.

Proof. (1) Since the relative entropy of μ with respect to ν is given by

$$\mathrm{Ent}(\nu|\mu) = \sup_{g\in\mathscr{B}^+(\bar D),\mu(g)=1} \nu(\log g),$$

(4.4) is equivalent to

$$\mathrm{Ent}(P_t^*\nu|P_t^*\mu) \le \frac{C(N)}{t}\mathbb{W}_2(\mu,\nu)^2, \quad t\in(0,T], \mu,\nu\in\mathscr{P}_{2,N}(\bar D). \tag{4.8}$$

By Pinsker's inequality

$$\frac{1}{2}\|\mu-\nu\|_{var}^2 \le \mathrm{Ent}(\nu|\mu),$$

we conclude that (4.8) implies (4.5), which further yields (4.6). So, we only need to prove (4.3) and (4.8).

For any $\mu,\nu\in\mathscr{P}_2(\bar D)$, let X_t solve (1.5) for $\mathscr{L}_{X_0}=\mu$, and denote

$$\mu_t := P_t^*\mu = \mathscr{L}_{X_t}, \quad \nu_t := P_t^*\nu, \quad \bar\mu_t := \mathscr{L}_{\bar X_t}, \quad t\in[0,T],$$

where $\bar X_t$ solves

$$\mathrm{d}\bar X_t = b_t(\bar X_t,\nu_t)\mathrm{d}t + \sigma_t(\bar X_t)\mathrm{d}W_t, \quad t\in[0,T], \bar X_0 = X_0.$$

Let σ and $\hat b := b(\cdot,\hat\mu) = \hat b^{(1)} + \hat b^{(0)}$ satisfy $(A_2^{\sigma,\hat b})$. Consider the decomposition

$$b_t^\nu := b_t(\cdot,\nu_t) = \hat b_t^{(1)} + b_t^{\nu,0}, \quad b_t^{\nu,0} := b_t^\nu - \hat b_t^{(1)}.$$

By (3.5) and (4.2), there exists a constant $K(N)>0$ such that

$$|b_t^{\nu,0}| \le |\hat b_t^{(0)}| + K(N)f_0(t,\cdot), \quad \|\nu\|_2\le N, \quad t\in[0,T]. \tag{4.9}$$

So, by Theorem 2.2 and Theorem 2.3, the estimate (2.14) and the log-Harnack inequality (2.18) hold for solutions of (2.1) with b^ν replacing b with a constant depending on N; that is, there exists a constant $c_1(N)>0$ such that

$$\mathbb{W}_2(\bar\mu_t,\nu_t)^2 \le c_1(N)\mathbb{W}_2(\mu,\nu)^2, \quad t\in[0,T], \mu\in\mathscr{P}_2(\bar D), \tag{4.10}$$

$$\mathrm{Ent}(\nu_t|\bar\mu_t) = \sup_{f>0,\bar\mu(f)=1}(P_tf)(\nu) \le \frac{c_1(N)}{t}\mathbb{W}_2(\mu,\nu)^2, \quad t\in(0,T], \ \mu\in\mathscr{P}_2(\bar D). \tag{4.11}$$

Moreover, repeating step (e) in the proof of Theorem 2.2 for $k=2$ and $(X_t,\bar X_t)$ replacing $(X_t^{(1)},X_t^{(2)})$, and using (4.2) replacing (3.4), instead of (3.26) where $\|\mu_s-\nu_s\|_{k,var}^2$ disappears in the present case, we derive

$$\mathbb{W}_2(\mu_t,\bar\mu_t)^4 \le (\mathbb{E}|X_t-\bar X_t|^2)^2 \le c_2(N)\int_0^t \mathbb{W}_2(\mu_s,\nu_s)^4\mathrm{d}s, \quad t\in[0,T]$$

for some constant $c_2(N)>0$. This together with (4.10) yields

$$\mathbb{W}_2(\mu_t,\nu_t)^4 \le 8\mathbb{W}_2(\mu_t,\bar\mu_t)^4 + 8\mathbb{W}_2(\bar\mu_t,\nu_t)^2$$

$$\leq 8c_1(N)^2\mathbb{W}_2(\mu,\nu)^4 + 8c_2(N)\int_0^t \mathbb{W}_2(\mu_s,\mu_s)^4\mathrm{d}s, \quad t\in[0,T].$$

Therefore, Gronwall's inequality implies (4.3) for some constant $C(N)>0$.

On the other hand, let $\|\mu\|_2\leq N$ and define

$$R_t := \exp\bigg[-\int_0^t \langle\gamma_s,\mathrm{d}W_s\rangle - \frac{1}{2}\int_0^t|\gamma_s|^2\mathrm{d}s\bigg],$$
$$\gamma_s := \big\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\big\}(X_s)\big[b_s^\mu(X_s)-b_s^\nu(X_s)\big].$$

By Girsanov's theorem, we obtain

$$\int_{\bar D}\Big(\frac{\mathrm{d}\bar\mu_t}{\mathrm{d}\mu_t}\Big)^2\mathrm{d}\mu_t = \mathbb{E}\Big\{\Big(\frac{\mathrm{d}\bar\mu_t}{\mathrm{d}\mu_t}(X_t)\Big)\Big\}^2 = \mathbb{E}\Big\{\Big(\mathbb{E}\big[R_t|X_t\big]\Big)\Big\}^2\leq \mathbb{E}R_t^2.$$

As shown in [12, p 14-15], by combining this with the Young inequality (see [2, Lemma 2.4])

$$\mu(fg)\leq\mu(f\log f)+\log\mu(\mathrm{e}^g),\quad f,g\geq 0,\mu(f)=1,\mu\in\mathscr{P}(\bar D), \tag{4.12}$$

we derive

$$\begin{aligned}\mathrm{Ent}(\nu_t|\mu_t) &= \int_{\bar D}\log\Big(\frac{\mathrm{d}\nu_t}{\mathrm{d}\mu_t}\Big)\mathrm{d}\nu_t = \int_{\bar D}\Big\{\log\frac{\mathrm{d}\nu_t}{\mathrm{d}\bar\mu_t}+\log\frac{\mathrm{d}\bar\mu_t}{\mathrm{d}\mu_t}\Big\}\mathrm{d}\nu_t\\ &= \mathrm{Ent}(\nu_t|\bar\mu_t)+\int_{\bar D}\Big(\frac{\mathrm{d}\nu_t}{\mathrm{d}\bar\mu_t}\Big)\log\frac{\mathrm{d}\bar\mu_t}{\mathrm{d}\mu_t}\mathrm{d}\bar\mu_t\leq 2\mathrm{Ent}(\nu_t|\bar\mu_t)+\log\int_{\bar D}\frac{\mathrm{d}\bar\mu_t}{\mathrm{d}\mu_t}\mathrm{d}\bar\mu_t\\ &= 2\mathrm{Ent}(\nu_t|\bar\mu_t)+\log\int_{\bar D}\Big(\frac{\mathrm{d}\bar\mu_t}{\mathrm{d}\mu_t}\Big)^2\mathrm{d}\mu_t\leq 2\mathrm{Ent}(\nu_t|\bar\mu_t)+\log\mathbb{E}R_t^2.\end{aligned} \tag{4.13}$$

Let $f_s(x):=\sum_{i=1}^l f_i(s,x)$. By (4.2), (4.3), $\|\sigma^*(\sigma\sigma^*)^{-1}\|_\infty<\infty$ and (2.60) due to (A_2^{σ,b^μ}), we find constants $c_3(N),c_4(N)>0$ such that

$$\begin{aligned}\mathbb{E}[R_t^2] &\leq \big(\mathbb{E}[R_t^2]\big)^2\leq\mathbb{E}\mathrm{e}^{c_3(N)\mathbb{W}_2(\mu,\nu)^2\int_0^t f_s(X_s)^2\mathrm{d}s}\\ &\leq 1+\mathbb{E}\bigg[c_3(N)\mathbb{W}_2(\mu,\nu)^2\Big(\int_0^t f_s(X_s)^2\mathrm{d}s\Big)\mathrm{e}^{c_3(N)\mathbb{W}_2(\mu,\nu)^2\int_0^t f_s(X_s)^2\mathrm{d}s}\bigg]\\ &\leq 1+c_3(N)\mathbb{W}_2(\mu,\nu)^2\bigg[\mathbb{E}\Big(\int_0^t f_s(X_s)^2\mathrm{d}s\Big)^2\bigg]^{\frac12}\Big[\mathbb{E}\mathrm{e}^{2c_3(N)\mathbb{W}_2(\mu,\nu)^2\int_0^t f_s(X_s)^2\mathrm{d}s}\Big]^{\frac12}\\ &\leq 1+c_4(N)\mathbb{W}_2(\mu,\nu)^2.\end{aligned} \tag{4.14}$$

Combining this with (4.11) and (4.13), we prove (4.8) for some constant $C(N)>0$.

(2) When $\alpha=0$, (4.9) holds for $K(N)=K$ independent of N, so that (4.10) and (4.11) hold for some constant $C_1(N)=C_1>0$ independent of N and all $\mu,\nu\in\mathscr{P}_2(\bar D)$, and in (4.14) the constant $C_3(N)=C_3$ is independent of N as well. Consequently, (4.7) holds and

$$\mathbb{E}[R_t^2]\leq\mathbb{E}\mathrm{e}^{C_3\mathbb{W}_2(\mu,\nu)^2\int_0^t f_s(X_s)^2\mathrm{d}s}\leq\mathrm{e}^{C\mathbb{W}_2(\mu,\nu)^2}$$

if $\sup_{1\leq i\leq l}\|f_i\|_\infty<\infty$, and when D is bounded we conclude that $C_4(N)=C_4$ in (4.14) is uniform in $N>0$. Therefore, (4.4) and hence its consequent inequalities hold for some constant independent of N. □

4.2 Monotone case

(A5) **(D)** *and* **(A2)**(2) *hold,* $\sigma_t(x,\mu)=\sigma_t(x)$ *does not depend on* μ *and is locally bounded on* $[0,\infty)\times\bar D$, $\sigma\sigma^*$ *is invertible,* b *is bounded on bounded subsets of* $[0,\infty)\times\mathbb{R}^d\times\mathscr{P}_2(\bar D)$, *and for any* $T>0$ *there exists a constant* $L>0$ *such that*

$$\|\sigma_t(x)-\sigma_t(y)\|_{HS}^2+2\langle x-y,b_t(x,\mu)-b_t(y,\nu)\rangle^+\le L|x-y|^2+L|x-y|\mathbb{W}_2(\mu,\nu),$$
$$\big\|\sigma_t(x)(\sigma_t\sigma_t^*)^{-1}(x)\big\|\le L,\quad t\in[0,T],x,y\in\bar D,\mu,\nu\in\mathscr{P}_2(\bar D).$$

By Theorem 3.3, **(A5)** implies that (1.5) is well-posed for distributions in $\mathscr{P}_2(\bar D)$.

Theorem 4.2. *Assume* **(A5)**. *Then for any* $T>0$, *there exists a constant* $C>0$ *such that the following inequalities hold for all* $t\in(0,T]$ *and* $\nu\in\mathscr{P}_2(\bar D)$:

$$\mathbb{W}_2(P_t^*\mu,P_t^*\nu)^2\le C\mathbb{W}_2(\mu,\nu)^2,\quad \mu\in\mathscr{P}_2(\bar D),\tag{4.15}$$

$$P_t\log f(\nu)\le\log P_tf(\mu)+\frac{C}{t}\mathbb{W}_2(\mu,\nu)^2,\quad 0<f\in\mathscr{B}_b(\bar D),\mu\in\mathscr{P}_2(\bar D),\tag{4.16}$$

$$\frac12\|P_t^*\mu-P_t^*\nu\|_{var}^2\le\mathrm{Ent}(P_t^*\nu|P_t^*\mu)\le\frac{C}{t}\mathbb{W}_2(\mu,\nu)^2,\quad \mu\in\mathscr{P}_2(\bar D)\tag{4.17}$$

$$\|\nabla P_tf(\nu)\|_{\mathbb{W}_2}:=\limsup_{\mu\to\nu\ in\ \mathbb{W}_2}\frac{|P_tf(\mu)-P_tf(\nu)|}{\mathbb{W}_2(\mu,\nu)}\le\frac{\sqrt{2C}\|f\|_\infty}{\sqrt t},\quad f\in\mathscr{B}_b(\bar D).\tag{4.18}$$

Proof. As explained in the proof of Theorem 4.1 that it suffices to prove (4.15) and (4.16). To this end, we modify the proof of [34, Theorem 4.1] as follows.

Firstly, for $\mu_0,\nu_0\in\mathscr{P}_2(\bar D)$, let (X_0,Y_0) be $\mathscr{F}_0$-measurable such that

$$\mathscr{L}_{X_0}=\mu_0,\quad\mathscr{L}_{Y_0}=\nu_0,\quad\mathbb{E}|X_0-Y_0|^2=\mathbb{W}_2(\mu_0,\nu_0)^2.\tag{4.19}$$

Denote

$$\mu_t:=P_t^*\mu_0,\quad\nu_t:=P_t^*\nu_0,\quad t\ge0.$$

Let X_t solve (1.5). We have

$$\mathrm{d}X_t=b_t(X_t,\mu_t)\mathrm{d}t+\sigma_t(X_t)\mathrm{d}W_t+\mathbf{n}(X_t)\mathrm{d}l_t^X,\quad t\in[0,T],\tag{4.20}$$

where l_t^X is the local time of X_t on ∂D. Next, for any $t_0\in(0,T]$ consider the SDE

$$\begin{aligned}\mathrm{d}Y_t=&\Big\{b_t(Y_t,\nu_t)+\frac{\sigma_t(Y_t)\{\sigma_t^*(\sigma_t\sigma_t^*)^{-1}\}(X_t)(X_t-Y_t)}{\xi_t}\Big\}\mathrm{d}t\\&+\sigma_t(Y_t)\mathrm{d}W_t+\mathbf{n}(Y_t)\mathrm{d}l_t^Y,\quad t\in[0,t_0),\end{aligned}\tag{4.21}$$

where l_t^Y is the local time of Y_t on ∂D. For the constant $L > 0$ in **(A5)**, let

$$\xi_t := \frac{1}{L}\Big(1 - \mathrm{e}^{L(t-t_0)}\Big), \quad t \in [0, t_0). \tag{4.22}$$

The construction of Y_t goes back to [33] for the classical SDEs, see also [34] for the extension to DDSDEs. According to Theorem 2.8, **(A5)** implies that (4.21) has a unique solution up to times

$$\tau_{n,m} := \frac{t_0 n}{n+1} \wedge \inf\big\{t \in [0, t_0) : |Y_t| \geq m\big\}, \quad n, m \geq 1.$$

Let h be in (3.34) for $k = 2$. By (1.2) and **(A2)**(2), we have

$$\langle \nabla\big\{(1 + h \circ \tilde{\rho})| \cdot - x_0|^2\big\}(Y_t), \mathbf{n}(Y_t)\rangle \mathrm{d}l_t^Y \leq 0, \quad x_0 \in \bar{D},$$

so that **(A5)**, for any $n \geq 1$ we find a constant $c(n) > 0$ such that

$$\mathrm{d}\big\{(1 + h \circ \tilde{\rho})(Y_t)|Y_t - x_0|^2\big\} \leq c(n)(1 + |Y_t|^2)\mathrm{d}t + \mathrm{d}M_t, \quad t \in [0, \tau_{n,m}], \ n, m \geq 1$$

holds for some martingale M_t. This implies $\lim_{m\to\infty} \tau_{n,m} = \frac{t_0 n}{n+1}$, and hence (4.21) has a unique solution up to time t_0.

Next, let $\tilde{Y}_t$ solve the SDE

$$\mathrm{d}\tilde{Y}_t = b_t(\tilde{Y}_t, \nu_t)\mathrm{d}t + \sigma_t(\tilde{Y}_t)\mathrm{d}W_t + \mathbf{n}(\tilde{Y}_t)\mathrm{d}l_t^{\tilde{Y}}, \quad \tilde{Y}_0 = Y_0, t \in [0, T], \tag{4.23}$$

where $l_t^{\tilde{Y}}$ is the local time of $\tilde{Y}_t$ on ∂D. By **(A5)**, (1.2) and Itô's formula, we find a constant $c_2 > 0$ such that

$$\begin{aligned} \mathbb{E}|X_t - \tilde{Y}_t|^2 \leq &\mathbb{W}_2(\mu_0, \nu_0)^2 + c_2 \int_0^t \big\{\mathbb{E}|X_s - \tilde{Y}_s|^2 + \mathbb{W}_2(\mu_s, \nu_s)^2\big\}\mathrm{d}s \\ &+ \frac{2}{r_0}\mathbb{E}\int_0^t |X_s - \tilde{Y}_s|^2(\mathrm{d}\tilde{l}_s^X + \mathrm{d}\tilde{l}_s^{\tilde{Y}}), \quad t \in [0, T]. \end{aligned} \tag{4.24}$$

For h in (3.34) with $k = 2$, we deduce from **(A2)**(2) that

$$\begin{aligned} &\big\langle \nabla\big\{|X_t - \cdot|^2(h \circ \rho(X_t) + h \circ \rho)\big\}(\tilde{Y}_t), \mathbf{n}(\tilde{Y}_t)\big\rangle \mathrm{d}\tilde{l}_t^{\tilde{Y}} \leq -|X_t - \tilde{Y}_t|^2 \mathrm{d}\tilde{l}_t^{\tilde{Y}}, \\ &\big\langle \nabla\big\{|\tilde{Y}_t - \cdot|^2(h \circ \rho(\tilde{Y}_t) + h \circ \rho)\big\}(X_t), \mathbf{n}(X_t)\big\rangle \mathrm{d}\tilde{l}_t^X \leq -|X_t - \tilde{Y}_t|^2 \mathrm{d}\tilde{l}_t^X. \end{aligned} \tag{4.25}$$

So, applying Itô's formula to

$$\eta_t := |X_t - \tilde{Y}_t|^2(h \circ \rho(X_t) + h \circ \rho(\tilde{Y}_t)),$$

and using **(A5)** and (1.2), we find a constant $c_3 > 0$ such that

$$\mathrm{d}\eta_t \leq c_3\big\{|X_t - \tilde{Y}_t|^2 + \mathbb{W}_2(\mu_t, \nu_t)^2\big\}\mathrm{d}t + \mathrm{d}M_t - |X_t - \tilde{Y}_t|^2(\mathrm{d}\tilde{l}_t^X + \mathrm{d}\tilde{l}_t^{\tilde{Y}})$$

holds for some martingale M_t. This together with (4.24) yields

$$\mathbb{E}|X_t - \tilde{Y}_t|^2 \leq \mathbb{W}_2(\mu_0, \nu_0)^2 + \mathbb{E}\eta_0 + (c_2 + c_3)\int_0^t \big\{\mathbb{E}|X_s - \tilde{Y}_s|^2 + \mathbb{W}_2(\mu_s, \nu_s)^2\big\}\mathrm{d}s$$

$$\leq 3\mathbb{W}_2(\mu_0,\nu_0)^2 + 2(c_2+c_3)\int_0^t \mathbb{E}|X_s-\tilde{Y}_s|^2\mathrm{d}s, \quad t\in[0,T],$$

where we have used the fact that $\mathbb{W}_2(\mu_s,\nu_s)^2 \leq \mathbb{E}|X_s-\tilde{Y}_s|^2$ by definition. By Gronwall's lemma, this and $\mathbb{W}_2(\mu_t,\nu_t)^2 \leq \mathbb{E}|X_t-\tilde{Y}_t|^2$, we find a constant $c_4>0$ such that

$$\mathbb{W}_2(\mu_t,\nu_t)^2 \leq \mathbb{E}|X_t-\tilde{Y}_t|^2 \leq c_4\mathbb{W}_2(\mu_0,\nu_0)^2, \quad t\in[0,T], \tag{4.26}$$

so that (4.15) holds.

Moreover, for any $n\geq 1$, let

$$\tau_n := \frac{t_0 n}{n+1} \wedge \inf\{t\in[0,t_0) : |X_t-Y_t|\geq n\}. \tag{4.27}$$

By Girsanov's theorem,

$$\tilde{W}_t := W_t + \int_0^t \frac{1}{\xi_s}\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(X_s)(X_s-Y_s)\mathrm{d}s, \quad t\in[0,\tau_n]$$

is an m-dimensional Brownian motion under the probability $\mathbb{Q}_n := R_n\mathbb{P}$, where

$$R_n := \mathrm{e}^{-\int_0^{\tau_n}\frac{1}{\xi_s}\langle\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(X_s)(X_s-Y_s),\mathrm{d}W_s\rangle - \frac{1}{2}\int_0^{\tau_n}\frac{|\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(X_s)(X_s-Y_s)|^2}{|\xi_s|^2}\mathrm{d}s}. \tag{4.28}$$

Then (4.20) and (4.21) imply

$$\begin{aligned} &\mathrm{d}X_t = \Big\{b_t(X_t,\mu_t) - \frac{X_t-Y_t}{\xi_t}\Big\}\mathrm{d}t + \sigma_t(X_t)\mathrm{d}\tilde{W}_t + \mathbf{n}(X_t)\mathrm{d}l_t^X,\\ &\mathrm{d}Y_t = b_t(Y_t,\nu_t)\mathrm{d}t + \sigma_t(Y_t)\mathrm{d}\tilde{W}_t + \mathbf{n}(Y_t)\mathrm{d}l_t^Y, \quad t\in[0,\tau_n], n\geq 1. \end{aligned} \tag{4.29}$$

Combining this with **(A5)**, (1.2), (4.26) and Itô's formula, we obtain

$$\begin{aligned} &\mathrm{d}\frac{|X_t-Y_t|^2}{\xi_t} - \mathrm{d}M_t\\ &\leq \Big\{\frac{L|X_t-Y_t|^2 + L|X_t-Y_t|\mathbb{W}_2(\mu_t,\nu_t))}{\xi_t} - \frac{|X_t-Y_t|^2(2+\xi_t')}{\xi_t^2}\Big\}\mathrm{d}t\\ &\quad + \frac{|X_t-Y_t|^2}{\xi_t^2}(\mathrm{d}\tilde{l}_t^X + \mathrm{d}\tilde{l}_t^Y)\\ &\leq \Big\{\frac{L^2\mathbb{W}_2(\mu_t,\nu_t)^2}{2} - \frac{|X_t-Y_t|^2(2+\xi_t'-L\xi_t-\frac{1}{2})}{\xi_t^2}\Big\}\mathrm{d}t + \frac{|X_t-Y_t|^2}{\xi_t^2}(\mathrm{d}\tilde{l}_t^X + \mathrm{d}\tilde{l}_t^Y)\\ &\leq \Big\{\frac{L^2\mathrm{e}^{2Lt}\mathbb{W}_2(\mu_0,\nu_0)^2}{2} - \frac{|X_t-Y_t|^2}{2\xi_t^2}\Big\}\mathrm{d}t + \frac{|X_t-Y_t|^2}{\xi_t^2}(\mathrm{d}\tilde{l}_t^X + \mathrm{d}\tilde{l}_t^Y), \quad t\in[0,\tau_n], \end{aligned} \tag{4.30}$$

where $\mathrm{d}M_t := \frac{2}{\xi_t}\big\langle X_t-Y_t, \{\sigma_t(X_t)-\sigma_t(Y_t)\}\mathrm{d}\tilde{W}_t\big\rangle$ is a $\mathbb{Q}_n$-martingale. By (4.25) for $(Y_t,\tilde{l}_t^Y)$ replacing $(\tilde{Y}_t,\tilde{l}_t^Y)$, and applying Itô's formula to $\gamma_t := \frac{|X_t-Y_t|^2}{\xi_t}(h\circ\rho(X_t)+h\circ\rho(Y_t))$, we find a constant $c_5>0$ such that

$$\mathrm{d}\gamma_t \leq c_5\gamma_t\mathrm{d}t + \mathrm{d}\tilde{M}_t - \frac{|X_t-Y_t|^2}{\xi_t}(\mathrm{d}\tilde{l}_t^X + \mathrm{d}\tilde{l}_t^Y), \quad t\in[0,\tau_n], n\geq 1$$

holds for some $\mathbb{Q}_n$-martingale $\tilde{M}_t$. This and (4.19) imply that for some constants $c_6, c_7 > 0$,

$$\mathbb{E}_{\mathbb{Q}_n}\gamma_{t\wedge\tau_n} \le e^{c_4 T}\mathbb{E}\gamma_0 \le \frac{c_6}{t_0}\mathbb{W}_2(\mu_0,\nu_0)^2,$$
$$\mathbb{E}_{\mathbb{Q}_n}\int_0^{\tau_n}\frac{|X_t-Y_t|^2}{\xi_t}(\mathrm{d}\tilde{l}_t^X+\mathrm{d}\tilde{l}_t^Y) \le \frac{c_7}{t_0}\mathbb{W}_2(\mu_0,\nu_0)^2,\quad n\ge 1, t\ge 0.$$

Combining this with (4.26), (4.30) and **(A5)**, we derive

$$\begin{aligned}\mathbb{E}[R_n\log R_n] &= \mathbb{E}_{\mathbb{Q}_n}[\log R_n] = \frac{1}{2}\mathbb{E}_{\mathbb{Q}_n}\int_0^{\tau_n}\frac{|\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(X_s)(X_s-Y_s)|^2}{|\xi_s|^2}\mathrm{d}s\\ &\le \frac{c}{t_0}\mathbb{W}_2(\mu_0,\nu_0)^2,\quad n\ge 1\end{aligned}\tag{4.31}$$

for some constant $c > 0$ uniformly in $t_0 \in (0,T]$. Therefore, by the martingale convergence theorem, $R_\infty := \lim_{n\to\infty} R_n$ exists, and

$$N_t := e^{-\int_0^t \frac{1}{\xi_s}\langle\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(X_s)(X_s-Y_s),\mathrm{d}W_s\rangle - \frac{1}{2}\int_0^t \frac{|\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(X_s)(X_s-Y_s)|^2}{|\xi_s|^2}\mathrm{d}s},\quad t\in[0,t_0]$$

is a $\mathbb{P}$-martingale.

Finally, let $\mathbb{Q} := N_{t_0}\mathbb{P}$. By Girsanov's theorem, $(\tilde{W}_t)_{t\in[0,t_0]}$ is an m-dimensional Brownian motion under the probability $\mathbb{Q}$, and $(X_t)_{t\in[0,t_0]}$ solves the SDE

$$\mathrm{d}X_t = \Big\{b_t(X_t,\mu_t) - \frac{X_t-Y_t}{\xi_t}\Big\}\mathrm{d}t + \sigma_t(X_t)\mathrm{d}\tilde{W}_t + \mathbf{n}(X_t)\mathrm{d}l_t^X,\quad t\in[0,t_0].\tag{4.32}$$

Let $(Y_t)_{t\in[0,t_0]}$ solve

$$\mathrm{d}Y_t = b_t(Y_t,\nu_t)\mathrm{d}t + \sigma_t(Y_t)\mathrm{d}\tilde{W}_t + \mathbf{n}(Y_t)\mathrm{d}l_t^Y,\quad t\in[0,t_0].\tag{4.33}$$

By the well-posedness of (1.5), this extends the second equation in (4.29) with $\mathscr{L}_{Y_{t_0}|\mathbb{Q}} = \nu_{t_0}$. Moreover, (4.31) and Fatou's lemma implies

$$\begin{aligned}&\frac{1}{2}\mathbb{E}_{\mathbb{Q}}\int_0^{t_0}\frac{|\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(X_s)(X_s-Y_s)|^2}{|\xi_s|^2}\mathrm{d}s\\ &= \mathbb{E}[N_{t_0}\log N_{t_0}] \le \liminf_{n\to\infty}\mathbb{E}[R_n\log R_n] \le \frac{c}{t_0}\mathbb{W}_2(\mu_0,\nu_0)^2,\end{aligned}\tag{4.34}$$

which in particular implies $\mathbb{Q}(X_{t_0} = Y_{t_0}) = 1$. Indeed, by **(A5)**, if $X_{t_0}(\omega) \neq Y_{t_0}(\omega)$ then there exists a small constant $\varepsilon > 0$ such that

$$|\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(X_s)(X_s-Y_s)|^2(\omega) \ge \varepsilon,\quad s\in[t_0-\varepsilon,t_0],$$

which implies $\int_0^{t_0}\frac{|\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\}(X_s)(X_s-Y_s)|^2}{|\xi_s|^2}(\omega)\mathrm{d}s = \infty$. So, (4.34) implies $\mathbb{Q}(X_{t_0} = Y_{t_0}) = 1$. Combining this with the Young's inequality (4.12), we arrive at

$$\begin{aligned}P_{t_0}\log f(\nu_0) &= \mathbb{E}[N_{t_0}\log f(Y_{t_0})] = \mathbb{E}[N_{t_0}\log f(X_{t_0})] \le \mathbb{E}[N_{t_0}\log N_{t_0}] + \log\mathbb{E}[f(X_{t_0})]\\ &\le \log P_{t_0}f(\mu_0) + \frac{c}{t_0}\mathbb{W}_2(\mu_0,\nu_0)^2,\quad t_0\in(0,T].\end{aligned}$$

Hence, (4.16) holds. □

Acknowledgement. The author would like to thank Bin Xie, Xing Huang, S. Wang, W. Hong, B. Wu, S. Hu and two referees for their helpful comments and corrections to earlier versions of the paper.

References

[1] D. Adams, G. dos Reis, R. Ravaille, W. Salkeld, J. Tugaut *Large Deviations and Exit-times for reflected McKean-Vlasov equations with self-stabilizing terms and superlinear drifts,* arXiv:2005.10057.

[2] M. Arnaudon, A. Thalmaier, F.-Y. Wang, *Gradient estimates and Harnack inequalities on non-compact Riemannian manifolds,* Stoch. Proc. Appl. 119(2009), 3653–3670.

[3] V. Barbu, M. Röckner, *Probabilistic representation for solutions to nonlinear Fokker-Planck equations,* SIAM J. Math. Anal. 50(2018), 4246–4260.

[4] V. Barbu, M. Röckner, *From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE,* Ann. Probab. 48(2020), 1902–1920.

[5] D. Benedetto, E. Caglioti, M. Pulvirenti, *A kinetic equation for granular media equation,* Rairo Modél. Math. Anal. Num. 31(1997), 615–641.

[6] V. I. Bogachev, N. V. Krylov, M. Röckner, S.V. Shaposhnikov, *Fokker-Planck-Kolmogorov equations,* American Math. Soc. 2015.

[7] R. Carmona, F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I,* Springer 2019.

[8] M. G. Carroni, J. L. Menaldi, *Green Functions for Second Order Parabolic Integro-Differential Problems,* Chapman & Hall/CRC, 1992.

[9] P. Dupuis, H. Ishii, *On oblique derivative problems for fully nonlinear second-order elliptic partial differential equations on nonsmooth domains,* Nonlin. Anal. Theo. Meth. Appl. 15(1990), 1123–1138.

[10] M. Hino, K. Matsuura, M. Yonezawa, *Pathwise uniqueness and non-explosion property of Skorohod SDEs with a class of non-Lipschitz coefficients and non-smooth domains,* J. Theo. Probab. (2020). https://doi.org/10.1007/s10959-020-01036-7

[11] X. Huang, P. Ren, F.-Y. Wang, *Distribution dependent stochastic differential equations,* Front. Math. China 16(2021), 257–301.

[12] X. Huang, Y. Song, *Well-posedness and regularity for distribution dependent SPDEs with singular drifts,* Nonlin. Anal. Theo. Meth. Appl. 203(2021), 112167.

[13] X. Huang, F.-Y. Wang, *Distribution dependent SDEs with singular coefficients,* Stoch. Proc. Appl. 129(2019), 4747–4770.

[14] X. Huang, F.-Y. Wang, *McKean-Vlasov SDEs with drifts discontinuous under Wasserstein distance,* Disc. Cont. Dyn. Syst. Ser. A. 4(2021), 1667–1679.

[15] X. Huang, F.-Y. Wang, *Singular McKean-Vlasov (reflecting) SDEs with distribution dependent noise,* J. Math. Anal. Appl. 514(2022), 126301 21pp.

[16] N.V. Krylov, *Controlled diffusion processes,* Translated from the Russian by A. B. Aries. Applications of Mathematics, 14. Springer-Verlag, New York-Berlin, 1980.

[17] N.V. Krylov, M. Röckner, *Strong solutions of stochastic equations with singular time dependent drift,* Probab. Theory Relat. Fields 131(2005), 154–196.

[18] H. Li, D. Luo, J. Wang, *Harnack inequalities for SDEs with multiplicative noise and non-regular drift,* Stoch. Dyn. 15(2015), 1550015.

[19] P. L. Lions, A. S. Sznitman, *Stochastic differential equations with reflecting boundary conditions,* Comm. Pure Appl. Math. 37(1984), 511–537.

[20] P. Marin-Rubio, J. Real, *Some results on stochastic differential equations with reflecting boundary conditions,* J. Theo. Probab. 17(2004), 705–716.

[21] H.P. McKean, *A class of Markov processes associated with nonlinear parabolic equations,* Proc. Nat. Acad. Sci. U.S.A. 56(1966), 1907–1911.

[22] S. Menozzi, A. Pesce, X. Zhang, *Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift,* J. Diff. Equat. 272(2021), 330–369.

[23] A. Rozkosz, L. Slominski, *On stability and existence of solutions of SDEs with reflection at the boundary,* Stoch. Process. Appl. 68(1997), 285–302.

[24] Y. Saisho, *Stochastic differential equations for multidimensional domain with reflecting boundary,* Probab. Theory Relat. Fields 74(1987), 455–477.

[25] M. Scheutzow, *A stochastic Gronwall's lemma,* Infin. Dimens. Anal. Quant. Probab. Relat. Topics 16(2013), 1350019.

[26] A. V. Skorohod, *Stochastic equations for diffusion processes with a boundary,* Teor. Verojatnost. i Primenen. 6(1961), 287–298.

[27] A. V. Skorohod, *Stochastic equations for diffusion processes with boundaries II,* Teor. Verojatnost. i Primenen. 7(1962), 5–25.

[28] A.-S. Sznitman, *Topics in propagations of chaos,* Lecture notes in Math. Vol. 1464, pp. 165–251, Springer, Berlin, 1991.

[29] H. Tanaka, *Stochastic differential equations with reflecting boundary conditions,* Hiroshima Math. J. 9(1979), 163–177.

[30] C. Villani, *Optimal Transport, Old and New,* Springer-Verlag, 2009.

[31] D. Trevisan, *Well-posedness of multidimensional diffusion processes,* Electr. J. Probab. 21(2016), 1–41.

[32] F.-Y. Wang, *Harnack inequalities on manifolds with boundary and applications,* J. Math. Pures Appl. 94(2010), 304–321.

[33] F.-Y. Wang, *Harnack inequality for SDE with multiplicative noise and extension to Neumann semigroup on nonconvex manifolds,* Ann. Probab. 39(2011), 1449–1467.

[34] F.-Y. Wang, *Distribution-dependent SDEs for Landau type equations,* Stoch. Proc. Appl. 128(2018), 595-621.

[35] F.-Y. Wang, *Exponential ergodicity for singular reflecting McKean-Vlasov SDEs,* arXiv:2108.03908.

[36] F.-Y. Wang, T. Zhang, *Log-Harnack inequalities for semi-linear SPDE with strongly multiplicative noise,* Stoch. Proc. Appl. 124(2014), 1261–1274.

[37] P. Xia, L. Xie, X. Zhang, G. Zhao, *$L^q(L^p)$-theory of stochastic differential equations,* Stoch. Proc. Appl. 130(2020), 5188–5211.

[38] L. Xie, X. Zhang, *Ergodicity of stochastic differential equations with jumps and singular coefficients,* Ann. Inst. Henri Poincaré Probab. Stat. 56(2020), 175–229.

[39] S. Yang, T. Zhang, *Strong solutions to reflecting stochastic differential equations with singular drift,* arXiv:2002.12150.

[40] C. Yuan, S.-Q. Zhang, *A study on Zvonkin's transformation for stochastic differential equations with singular drift and related applications,* to appear in J. Diff. Equat. arXiv:1910.05903.

[41] X. Zhang, *Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients,* Electr. J. Probab. 16(2011), 1096–1116.