Killed Distribution Dependent SDE for Nonlinear Dirichlet Problem *

Feng-Yu Wang

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

Department of Mathematics, Swansea University, Bay Campus, SA1 8EN, United Kingdom

wangfy@tju.edu.cn

December 4, 2022

Abstract

To characterize nonlinear Dirichlet problems in an open domain, we investigate killed distribution dependent SDEs. By constructing the coupling by projection and using the Zvonkin/Girsanov transforms, the well-posedness is proved for three different situations:

1) monotone case with distribution dependent noise (possibly degenerate);

2) singular case with non-degenerate distribution dependent noise;

3) singular case with non-degenerate distribution independent noise.

In the first two cases the domain is C^2 smooth such that the Lipschitz continuity in initial distributions is also derived, and in the last case the domain is arbitrary.

AMS subject Classification: 60B05, 60B10.

Keywords: Nonlinear Dirichlet problem, killed distribution dependent SDE, coupling by projection, well-posedness.

1 Introduction

The distribution dependent stochastic differential equation (DDSDE) is a crucial probability model characterizing the nonlinear Fokker-Planck equation. It is known as McKean-Vlasov SDE due to [6], and mean field SDE for its link to mean field particle systems, see for instance the lecture notes [9] and the survey [4] for the background and recent progress on the study of DDSDEs and applications. To characterize the nonlinear Neumann problem, the reflecting DDSDE has been investigated in [10], see also [1] for the convex domain case, and see the early

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

work [8] for the characterization on propagations of chaos. In this paper, we consider the killed DDSDE which in turn to describe the nonlinear Dirichlet problem. In this case, the distribution is restricted to the open domain and thus might be a sub-probability measure, i.e. the total mass may be less than 1.

Let $O \subset \mathbb{R}^d$ be a connected open domain with closure $\bar{O}$, and let

$$\mathscr{P}_O := \big\{\mu \text{ is a measure on } O,\ \mu(O) \le 1\big\}$$

be the space of sub-probability measures on O equipped with the weak topology. Consider the following time-distribution dependent second order differential operator on O:

$$L_{t,\mu} := \mathrm{tr}\{(\sigma_t\sigma_t^*)(\cdot,\mu)\nabla^2\} + \nabla_{b_t(\cdot,\mu)}, \quad t \in [0,T], \mu \in \mathscr{P}_O,$$

where $T > 0$ is a fixed constant, σ^* is the transposition of σ, ∇^2 is the Hessian operator, $\nabla_b := b \cdot \nabla$ is the derivative along b, and for some $m \in \mathbb{N}$,

$$b : [0,T] \times O \times \mathscr{P}_O \to \mathbb{R}^d, \quad \sigma : [0,T] \times O \times \mathscr{P}_O \to \mathbb{R}^d \otimes \mathbb{R}^m$$

are measurable such that

$$\int_0^T \mathrm{d}t \int_O \big\{|b_t(x,\mu_t)| + \|\sigma_t(x,\mu_t)\|^2\big\}\mu_t(\mathrm{d}x) < \infty, \quad \mu = (\mu_t)_{t\in[0,T]} \in C([0,T];\mathscr{P}_O). \tag{1.1}$$

To introduce the nonlinear Dirichlet problem for $L_{t,\mu}$ on $\mathscr{P}_O$, let $C_D^2(O)$ be the class of $f \in C_b^2(\bar{O})$ with Dirichlet condition $f|_{\partial O} = 0$, where $f \in C_b^2(\bar{O})$ means that f is a bounded C^2 function on $\bar{O}$ with bounded first and second order derivatives. For any $t \in [0,T]$ and $\mu, \nu \in \mathscr{P}_O$ such that

$$\int_O \big\{|b_t(x,\mu)| + \|\sigma_t(x,\mu)\|^2\big\}\nu(\mathrm{d}x) < \infty,$$

define the linear functional on $C_D^2(O)$:

$$L_{t,\mu}^{D*}\nu : C_D^2(O) \ni f \mapsto (L_{t,\mu}^{D*}\nu)(f) := \int_O L_{t,\mu} f \mathrm{d}\nu \in \mathbb{R}.$$

The corresponding nonlinear Dirichlet problem for $L_{t,\mu}$ is the equation

$$\partial_t \mu_t = L_{t,\mu_t}^{D*}\mu_t, \quad t \in [0,T] \tag{1.2}$$

for $\mu : [0,T] \to \mathscr{P}_O$. We call $\mu_\cdot \in C([0,T];\mathscr{P}_O)$ a solution to (1.2), if

$$\mu_t(f) = \mu_0(f) + \int_0^t \mu_s(L_{s,\mu_s} f)\mathrm{d}s, \quad t \in [0,T], f \in C_D^2(O),$$

where $\mu(f) := \int f \mathrm{d}\mu$ for a measure μ and $f \in L^1(\mu)$.

When $\mu_t(\mathrm{d}x) = \rho_t(x)\mathrm{d}x$, (1.2) reduces to the nonlinear Dirichlet problem

$$\partial_t \rho_t = L_{t,\rho_t}^{D*}\rho_t, \quad t \in [0,T],$$

where $L_{t,\rho_t} := L_{t,\rho_t(x)\mathrm{d}x}$, in the sense that

$$\int_O (f\rho_t)(x)\mathrm{d}x = \int_O (f\rho_0)(x)\mathrm{d}x + \int_0^t \mathrm{d}s \int_O (\rho_s L_{s,\rho_s} f)(x)\mathrm{d}x, \quad t \in [0,T], f \in C_D^2(O).$$

To characterize (1.2), we consider the following killed distribution dependent SDE on $\bar{O}$:

$$\mathrm{d}X_t = \mathbf{1}_{\{t<\tau(X)\}}\big\{b_t(X_t, \mathscr{L}_{X_t}^O)\mathrm{d}t + \sigma_t(X_t, \mathscr{L}_{X_t}^O)\mathrm{d}W_t\big\}, \quad t \in [0,T], \tag{1.3}$$

where $\mathbf{1}$ is the indicated function, W_t is the m-dimensional Brownian motion on a complete filtration probability space $(\Omega, \{\mathscr{F}_t\}_{t\ge 0}, \mathbb{P})$,

$$\tau(X) := \inf\{t \in [0,T] : X_t \in \partial O\}$$

with $\inf \emptyset = \infty$ by convention, and for an $\bar{O}$-valued random variable ξ,

$$\mathscr{L}_\xi^O := \mathbb{P}(\xi \in O \cap \cdot)$$

is the distribution of ξ restricted to O, which we call the O-distribution of ξ. When different probability spaces are concerned, we denote $\mathscr{L}_\xi^O$ by $\mathscr{L}_{\xi|\mathbb{P}}^O$ to emphasize the reference probability measure.

Definition 1.1. A continuous adapted process $(X_t)_{t\in[0,T]}$ on $\bar{O}$ is called a solution of (1.3), if $\mathbb{P}$-a.s.

$$\int_0^{T\wedge\tau(X)} \big\{|b_t(X_t, \mathscr{L}_{X_t}^O)| + \|\sigma_t(X_t, \mathscr{L}_{X_t}^O)\|^2\big\}\mathrm{d}t < \infty$$

and

$$X_t = X_0 + \int_0^{t\wedge\tau(X)} \big\{b_s(X_s, \mathscr{L}_{X_s}^O)\mathrm{d}s + \sigma_s(X_s, \mathscr{L}_{X_s}^O)\mathrm{d}W_s\big\}, \quad t \in [0,T].$$

We call $(\tilde{X}_t, \tilde{W}_t)$ a weak solution to (1.3), if there exists a complete filtration probability space $(\tilde{\Omega}, \{\tilde{\mathscr{F}}_t\}_{t\in[0,T]}, \tilde{\mathbb{P}})$ such that $\tilde{W}_t$ is m-dimensional Brownian motion and $\tilde{X}_t$ solves (1.3) for $\tilde{W}_t$ replacing W_t.

Remark 1.1. (1) It is easy to see that for any (weak) solution X_t of (1.3), $\mu_t := \mathscr{L}_{X_t}^O$ solves the nonlinear Dirichlet problem (1.2). Indeed, since $\mathrm{d}X_t = 0$ for $t \ge \tau(X)$, we have

$$X_t = X_{\tau(X)} \in \partial O, \quad t \ge \tau(X),$$

so that

$$X_t = X_{t\wedge\tau(X)}, \quad \mathscr{L}_{X_t}^O(\mathrm{d}x) = \mathbb{P}(t < \tau(X), X_t \in \mathrm{d}x), \quad t \in [0,T].$$

By this and Itô's formula, for any $f \in C_D^2(O)$ we have

$$\begin{aligned}\mu_t(f) &= \mathbb{E}[(\mathbf{1}_O f)(X_t)] = \mathbb{E}[f(X_t)] \\ &= \mathbb{E}[f(X_0)] + \mathbb{E}\int_0^t \mathbf{1}_{\{s<\tau(X)\}} L_{s,\mu_s} f(X_s)\mathrm{d}s\end{aligned}$$

$$= \mu_0(f) + \int_0^t \mu_s(L_{s,\mu_s} f)\mathrm{d}s, \quad t \in [0,T].$$

(2) An alternative model to (1.3) is

$$\mathrm{d}X_t = \mathbf{1}_O(X_t)\big\{b_t(X_t, \mathscr{L}_{X_t}^O)\mathrm{d}t + \sigma_t(X_t, \mathscr{L}_{X_t}^O)\mathrm{d}W_t\big\}, \quad t \in [0,T]. \tag{1.4}$$

A solution of (1.3) also solves (1.4); while for a solution X_t to (1.4),

$$\tilde{X}_t := X_{t\wedge\tau(X)}$$

solves (1.3). In general, a solution of (1.4) does not have to solve (1.3). For instance, let $d = m = 1$ and $O = (0,\infty)$, consider $\sigma_t(x,\mu) = 2x, b_t(x,\mu) = 2\sqrt{x}$. Let Y_t solve the SDE

$$\mathrm{d}Y_t = Y_t\mathrm{d}W_t + \Big(1 - \frac{1}{2}Y_t\Big)\mathrm{d}t, \quad Y_0 = 0.$$

Then $X_t := (Y_t)^2$ solves (1.4) but does not solve (1.3), since $\tau(X) = 0$ and $X_t > 0$ (i.e. $X_t \notin \partial O$) for $t > 0$. See [14] for the study of (1.4) for $\sigma_t(x,\mu) = \sigma_t(x)$ independent of μ.

(3) The SDE (1.4) can be formulated as the usual DDSDE on $\mathbb{R}^d$, so that the superposition principle in [2] applies. More precisely, let $\mathscr{P}$ be the space of probability measures on $\mathbb{R}^d$, and define

$$\bar{b}_t(x,\mu) := \mathbf{1}_O(x)b_t(x,\mu(O\cap\cdot)), \quad \bar{\sigma}_t(x,\mu) := \mathbf{1}_O(x)\sigma_t(x,\mu(O\cap\cdot))$$

for $(t,x,\mu) \in [0,T]\times\mathbb{R}^d\times\mathscr{P}$. Then (1.4) becomes the following DDSDE on $\mathbb{R}^d$:

$$\mathrm{d}X_t = \bar{b}_t(X_t, \mathscr{L}_{X_t})\mathrm{d}t + \bar{\sigma}_t(X_t, \mathscr{L}_{X_t})\mathrm{d}W_t, \quad t \in [0,T].$$

We often solve (1.3) for O-distributions in a non-empty sub-space $\hat{\mathscr{P}}_O$ of $\mathscr{P}_O$, which is equipped with the weak topology as well.

Definition 1.2. (1) If for any $\mathscr{F}_0$-measurable random variable X_0 on $\bar{O}$ with $\mathscr{L}_{X_0}^O \in \hat{\mathscr{P}}_O$, (1.3) has a unique solution starting at X_0 such that $\mathscr{L}_X^O := (\mathscr{L}_{X_t}^O)_{t\in[0,T]} \in C([0,T];\hat{\mathscr{P}}_O)$, we call the SDE strongly well-posed for O-distributions in $\hat{\mathscr{P}}_O$.

(2) We call the SDE weakly unique for O-distributions in $\hat{\mathscr{P}}_O$, if for any two weak solutions (X_t^i, W_t^i) w.r.t. $(\Omega^i, \{\mathscr{F}_t^i\}_{t\in[0,T]}, \mathbb{P}^i)(i = 1,2)$ with $\mathscr{L}_{X_0^1|\mathbb{P}^1}^O = \mathscr{L}_{X_0^2|\mathbb{P}^2}^O \in \hat{\mathscr{P}}_O$, we have $\mathscr{L}_{X^1|\mathbb{P}^1}^O = \mathscr{L}_{X^2|\mathbb{P}^2}^O$. We call (1.3) weakly well-posed for O-distributions in $\hat{\mathscr{P}}_O$, if for any initial O-distribution $\mu_0 \in \hat{\mathscr{P}}_O$, it has a unique weak solution for O-distributions in $\hat{\mathscr{P}}_O$.

(3) The SDE (1.3) is called well-posed for O-distributions in $\hat{\mathscr{P}}_O$, if it is both strongly and weakly well-posed for O-distributions in $\hat{\mathscr{P}}_O$.

When (1.3) is well-posed for O-distributions in $\hat{\mathscr{P}}_O$, for any $\mu \in \hat{\mathscr{P}}_O$ and $t \in [0,T]$, let

$$P_t^{D*}\mu = \mathscr{L}_{X_t}^O, \quad t \in [0,T], \ \mathscr{L}_{X_0}^O = \mu.$$

We will study the well-posedness under the following assumption.

(H) For any $\mu \in C([0,T];\hat{\mathscr{P}}_O)$, the killed SDE

$$\mathrm{d}X_t^\mu = \mathbf{1}_{\{t<\tau(X^\mu)\}}\big\{b_t(X_t^\mu,\mu_t)\mathrm{d}t + \sigma_t(X_t^\mu,\mu_t)\mathrm{d}W_t\big\},\quad t\in[0,T] \tag{1.5}$$

is well-posed for initial value X_0^μ with $\mathscr{L}^O_{X_0^\mu} = \mu_0$, and $\mathscr{L}^O_{X^\mu} \in C([0,T];\hat{\mathscr{P}}_O)$.

Under this assumption, we define a map

$$C([0,T];\hat{\mathscr{P}}_O) \ni \mu \mapsto \Phi\mu := \mathscr{L}^O_{X^\mu} := (\mathscr{L}^O_{X_t^\mu})_{t\in[0,T]} \in C([0,T];\hat{\mathscr{P}}_O). \tag{1.6}$$

It is clear that a solution of (1.5) solves (1.3) if and only if μ is a fixed point of Φ. So, we have the following result.

Theorem 1.1. *Assume* **(H)**. *If for any $\gamma \in \hat{\mathscr{P}}_O$, Φ has a unique fixed point in*

$$\mathscr{C}^\gamma := \{\mu \in C([0,T];\hat{\mathscr{P}}_O), \mu_0 = \gamma\},$$

then (1.3) *is well-posed for O-distributions in $\hat{\mathscr{P}}_O$.*

In the remainder of the paper, we apply Theorem 1.1 to following three different situations:

1. The monotone case with distribution dependent noise (possibly degenerate);
2. The singular case with non-degenerate distribution dependent noise;
3. The singular case with non-degenerate distribution independent noise.

In the first two situations, we need O to be C^2-smooth to apply the coupling by projection, and the coefficients are local Lipschitz continuous in distributions with respect to the L^1 or truncated L^1 Wasserstein distance. In the last case, the domain is arbitrary, and $b_t(x,\cdot)$ is only local Lipschitz continuous in a weighted variation distance, but the noise has to be distribution independent, i.e. $\sigma_t(x,\mu) = \sigma_t(x)$.

2 Monotone case

In this part, we solve (1.3) under monotone conditions with respect to the L^1 or truncated L^1 Wasserstein distances:

$$\begin{aligned} &\mathbb{W}_1(\mu,\nu) := \inf_{\pi\in\mathscr{C}_O(\mu,\nu)} \int_{O\times O} |x-y|\pi(\mathrm{d}x,\mathrm{d}y),\\ &\hat{\mathbb{W}}_1(\mu,\nu) := \inf_{\pi\in\mathscr{C}_O(\mu,\nu)} \int_{O\times O} (1\wedge|x-y|)\pi(\mathrm{d}x,\mathrm{d}y),\quad \mu,\nu\in\mathscr{P}_O, \end{aligned} \tag{2.1}$$

where $\pi \in \mathscr{C}_O(\mu,\nu)$ means that π is a probability measure on $\bar{O}\times\bar{O}$ such that

$$\pi(\{\cdot\cap O\}\times\bar{O}) = \mu,\quad \pi(\bar{O}\times\{\cdot\cap O\}) = \nu.$$

2.1 Monotonicity in $\hat{\mathbb{W}}_1$

(A_1) For any $\mu \in C([0,T];\mathscr{P}_O)$, $b_t(x,\mu_t)$ and $\sigma_t(x,\mu_t)$ are continuous in $x \in O$ such that for any $N \geq 1$ and $O_N := \{x \in O : |x| \leq N\}$,

$$\int_0^T \sup_{O_N} \big\{|b_t(\cdot,\mu_t)| + \|\sigma_t(\cdot,\mu_t)\|^2\big\}\mathrm{d}t < \infty.$$

Moreover, there exists $K \in L^1([0,T];(0,\infty))$ such that for any $x,y \in O$ and $\mu,\nu \in \mathscr{P}_O$,

$$\begin{aligned}&2\langle b_t(x,\mu) - b_t(y,\nu), x-y\rangle + \|\sigma_t(x,\mu) - \sigma_t(y,\nu)\|_{HS}^2 \leq K(t)\big\{|x-y|^2 + \hat{\mathbb{W}}_1(\mu,\nu)^2\big\},\\ &2\langle b_t(x,\mu), x\rangle + \|\sigma_t(x,\mu)\|_{HS}^2 \leq K(t)\big(1+|x|^2\big), \quad t \in [0,T].\end{aligned}$$

(A_2) There exists $r_0 \in (0,1]$ such that the distance function ρ_∂ to ∂O is C^2-smooth in

$$\partial_{r_0} O := \big\{x \in \bar{O} : \ \rho_\partial(x) \leq r_0\big\},$$

and there exists a constant $\alpha > 0$ such that

$$|\sigma_t(x,\mu)^* \nabla \rho_\partial(x)|^{-2} \leq \alpha, \quad L_{t,\mu}\rho_\partial(x) \leq \alpha, \quad x \in \partial_{r_0} O, t \in [0,T].$$

Theorem 2.1. *Assume (A_1) and (A_2). Then the following assertions hold.*

(1) (1.3) *is well-posed for O-distributions in $\mathscr{P}_O$. Moreover, for any $p \geq 1$ there exists a constant $c > 0$ such that for any solution X_t to* (1.3) *for O-distributions in $\mathscr{P}_O$,*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^p \Big| \mathscr{F}_0\Big] \leq c\big(1 + |X_0|^p\big). \tag{2.2}$$

(2) *There exists a constant $c > 0$ such that*

$$\sup_{t\in[0,T]} \hat{\mathbb{W}}_1(P_t^{D*}\mu, P_t^{D*}\nu) \leq c\hat{\mathbb{W}}_1(\mu,\nu), \quad \mu,\nu \in \mathscr{P}_O. \tag{2.3}$$

Under assumption (A_1), for any $\mu \in C([0,T];\mathscr{P}_O)$ the SDE (1.5) satisfies the semi-Lipschitz condition before the hitting time $\tau(X^\mu)$, hence it is well-posed and for any $p \geq 1$ there exists a constant $c > 0$ uniformly in μ such that

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X^\mu_{t\wedge\tilde{\tau}}|^p \Big| \mathscr{F}_0\Big] = \mathbb{E}\Big[\sup_{t\in[0,T]} |X^\mu_{t\wedge\tau(X^\mu)\wedge\tilde{\tau}}|^p \Big| \mathscr{F}_0\Big] \leq c(1 + |X_0^\mu|^p) \tag{2.4}$$

holds for any solution X_t^μ of (1.5) and any stopping time $\tilde{\tau}$.

By Theorem 1.1, to prove the well-posedness of (1.3) for O-distributions in $\mathscr{P}_O$, it remains to show that for any $\gamma \in \mathscr{P}_O$, the map

$$\Phi\mu := \mathscr{L}^O_{X^\mu} = (\mathscr{L}^O_{X^\mu_t})_{t\in[0,T]}, \quad \mu \in C([0,T];\mathscr{P}_O)$$

has a unique fixed point in

$$\mathscr{C}^\gamma := \{\mu \in C([0,T];\mathscr{P}_O) : \mu_0 = \gamma\}.$$

To this end, for $i = 1, 2$, let $\mu^i \in C([0,T];\mathscr{P}_O)$, and let X_t^i solve (1.5) for μ^i replacing μ with $\mathscr{L}^O_{X_0^i} = \mu_0^i$, i.e.

$$\mathrm{d}X_t^i = \mathbf{1}_{\{t<\tau(X^i)\}}\big\{b_t(X_t^i,\mu_t^i)\mathrm{d}t + \sigma_t(X_t^i,\mu_t^i)\mathrm{d}W_t\big\},\quad t\in[0,T], \mathscr{L}^O_{X_0^i} = \mu_0^i. \tag{2.5}$$

Simply denote

$$\tau_i = \tau(X^i) \text{ for } i = 1,2,\quad \tau_{1,2} := \tau_1 \wedge \tau_2.$$

Since

$$\Gamma := \big\{(x,y):\ x\in O, y\in\partial O, |x-y| = \rho_\partial(x)\big\}$$

is a measurable subset of $O\times\partial O$ and $\Gamma_x := \{y\in\partial O : (x,y)\in\Gamma\} \neq \emptyset$ for any $x\in O$, by the measurable selection theorem (see [3, Theorem 1]), there exists a measurable map $P_\partial : O \to \partial O$ such that

$$|P_\partial x - x| = \rho_\partial(x),\quad x\in O. \tag{2.6}$$

We will use the following coupling by projection.

Definition 2.1. The *coupling by projection* $(\bar X_t^1, \bar X_t^2)$ for $(X_t^1, X_t^2) = (X^1_{t\wedge\tau_1}, X^2_{t\wedge\tau_2})$ is defined as

$$(\bar X_t^1,\bar X_t^2) := \begin{cases} (X_t^1, X_t^2), & \text{if } t\le\tau_{1,2},\\ (X_t^1, P_\partial X_t^1), & \text{if } \tau_2 < t\wedge\tau_1,\\ (P_\partial X_t^2, X_t^2), & \text{otherwise.}\end{cases} \tag{2.7}$$

It is easy to see that $\mathscr{L}^O_{\bar X_t^i} = \mathscr{L}^O_{X_t^i} = (\Phi\mu^i)_t$ for $i=1,2$; i.e. the distribution $\mathscr{L}_{(\bar X_t^1,\bar X_t^2)}$ of the coupling by projection $(\bar X_t^1,\bar X_t^2)$ satisfies

$$\mathscr{L}_{(\bar X_t^1,\bar X_t^2)} \in \mathscr{C}_O((\Phi\mu^1)_t, (\Phi\mu^2)_t).$$

Thus, by (2.1) and Definition 2.1,

$$\begin{aligned}\hat{\mathbb{W}}_1((\Phi\mu^1)_t,(\Phi\mu^2)_t) &\le \mathbb{E}\big[1\wedge|\bar X_t^1 - \bar X_t^2|\big] \le \mathbb{E}\big[1\wedge|X^1_{t\wedge\tau_{1,2}} - \bar X^2_{t\wedge\tau_{1,2}}|\big]\\ &\quad + r_0^{-1}\mathbb{E}\big[\{r_0\wedge\rho_\partial(X_t^1)\}\mathbf{1}_{\{t\wedge\tau_1\ge\tau_2\}}\big] + r_0^{-1}\mathbb{E}\big[\{r_0\wedge\rho_\partial(X_t^2)\}\mathbf{1}_{\{t\wedge\tau_2\ge\tau_1\}}\big].\end{aligned} \tag{2.8}$$

Lemma 2.2. *Assume (A_1). Then there exists a constant $c>1$ such that for any $t\in[0,T]$ and $\mu^1,\mu^2\in C([0,T];\mathscr{P}_O)$,*

$$\mathbb{E}\big[|X^1_{t\wedge\tau_{1,2}} - X^2_{t\wedge\tau_{1,2}}|^2\big|\mathscr{F}_0\big] \le c\,|X_0^1 - X_0^2|^2 + c\int_0^t K(s)\hat{\mathbb{W}}_1(\mu_s^1,\mu_s^2)^2\mathrm{d}s. \tag{2.9}$$

Consequently, for any $t\in[0,T]$,

$$\mathbb{E}\big[1\wedge|X^1_{t\wedge\tau_{1,2}} - X^2_{t\wedge\tau_{1,2}}|\big] \le \sqrt{c}\,\mathbb{E}[1\wedge|X_0^1 - X_0^2|] + \bigg(c\int_0^t K(s)\hat{\mathbb{W}}_1(\mu_s^1,\mu_s^2)^2\mathrm{d}s\bigg)^{\frac12}. \tag{2.10}$$

Proof. It suffices to prove (2.9), which implies (2.10) due to Jensen's inequality.

By (A_1) and Itô's formula, we obtain

$$\mathrm{d}|X_t^1 - X_t^2|^2 \le K(t)\big\{|X_t^1 - X_t^2|^2 + \hat{\mathbb{W}}_1(\mu_t^1, \mu_t^2)^2\big\}\mathrm{d}t + \mathrm{d}M_t, \quad t \in [0, T \wedge \tau_{1,2}]$$

for some local martingale M_t. This and (2.4) imply that

$$\beta_t := \mathbb{E}\big[|X^1_{t\wedge\tau_{1,2}} - X^2_{t\wedge\tau_{1,2}}|^2\big|\mathscr{F}_0\big]$$

is bounded in $t \in [0, T]$ and satisfies

$$\beta_t \le \beta_0 + \int_0^t K(s)\big\{\beta_s + \hat{\mathbb{W}}_1(\mu_s^1, \mu_s^2)^2\big\}\mathrm{d}s, \quad t \in [0, T].$$

By Gronwall's inequality, we prove (2.9). □

Lemma 2.3. *Assume (A_2). Then there exists a constant $c > 1$ independent of μ such that for any solution X_t^μ to* (1.5) *and any stopping time $\tilde{\tau}$,*

$$\mathbf{1}_{\{t\wedge\tau(X^\mu)\ge\tilde{\tau}\}}\mathbb{E}\big[r_0 \wedge \rho_\partial(X_t^\mu)\big|\mathscr{F}_{\tilde{\tau}}\big] \le c\mathbf{1}_{\{t\wedge\tau(X^\mu)\ge\tilde{\tau}\}}\rho_\partial(X^\mu_{t\wedge\tilde{\tau}}), \quad t \in [0, T].$$

Proof. By the strong Markov property of X_t^μ which is implied by the well-posedness of (1.5), we may and do assume that $\tilde{\tau} = 0$ and $x = X_0^\mu \in O$, such that the desired estimate becomes

$$\mathbb{E}^x\big[r_0 \wedge \rho_\partial(X_t^\mu)\big] \le c\rho_\partial(x), \quad t \in [0, T], \tag{2.11}$$

where $\mathbb{E}^x$ is the expectation under the probability $\mathbb{P}^x$ for X_t^μ starting at x. If $\rho_\partial(x) \ge \frac{r_0}{4}$, this inequality holds for $c := 4$. So, it suffices to prove for $\rho_\partial(x) < \frac{r_0}{4}$.

Let $h \in C^\infty([0, \infty))$ such that

$$h' \ge 0, \quad h'' \le 0, \ h(r) = r \text{ for } r \in [0, r_0/2], \quad h'(r) = 0 \text{ for } r \ge r_0.$$

By (A_2),

$$\mathrm{d}h(\rho_\partial(X_t^\mu)) \le \alpha\mathrm{d}t + \mathrm{d}M_t, \quad t \in [0, T \wedge \tau(X^\mu)], \tag{2.12}$$

where M_t is a martingale with

$$\mathrm{d}\langle M\rangle_t \ge \alpha^{-1}\mathrm{d}t, \quad t \le \hat{\tau} := \inf\{t \ge 0 : \rho_\partial(X_t^\mu) \ge r_0/2\}. \tag{2.13}$$

By (2.12) we obtain

$$\mathbb{E}^x[r_0 \wedge \rho_\partial(X_t^\mu)] \le 2\mathbb{E}^x[h(\rho_\partial(X^\mu_{t\wedge\tau(X^\mu)}))] \le 2\rho_\partial(x) + 2\alpha\mathbb{E}^x[t \wedge \tau(X^\mu)]. \tag{2.14}$$

On the other hand, let

$$\eta_t := \int_0^{\rho_\partial(X_t^\mu)} \mathrm{e}^{-2\alpha^2 s}\mathrm{d}s \int_s^{r_0} \mathrm{e}^{2\alpha^2\theta}\mathrm{d}\theta, \quad t \in [0, T \wedge \tau(X^\mu) \wedge \hat{\tau}].$$

Since $h(r) = r$ for $r \le \frac{r_0}{2}$, by (2.12), (2.13) and Itô's formula, we find a martingale $\tilde{M}_t$ such that

$$\mathrm{d}\eta_t \le -\mathrm{d}t + \mathrm{d}\tilde{M}_t, \quad t \in [0, T \wedge \tau(X^\mu) \wedge \hat{\tau}].$$

Consequently,

$$\mathbb{E}^x[t \wedge \tau(X^\mu) \wedge \hat{\tau}] \le \eta_0 \le c_1 \rho_\partial(x) \tag{2.15}$$

holds for some constant $c_1 > 0$. Therefore,

$$\begin{aligned} \mathbb{E}^x[t \wedge \tau(X^\mu)] &\le \mathbb{E}^x[t \wedge \tau(X^\mu) \wedge \hat{\tau}] + T\,\mathbb{E}^x\big[\mathbf{1}_{\{t\wedge\tau(X^\mu)>\hat{\tau}\}}\big] \\ &\le c_1\rho_\partial(x) + T\,\mathbb{P}^x\big(t \wedge \tau(X^\mu) > \hat{\tau}\big), \quad t \in [0,T]. \end{aligned} \tag{2.16}$$

To estimate the second term, let

$$\xi_t := \int_0^{\rho_\partial(X_t^\mu)} \mathrm{e}^{-2\alpha^2 s}\mathrm{d}s, \quad t \in [0, T \wedge \tau(X^\mu) \wedge \hat{\tau}].$$

By $h(r) = r$ for $r \in [0, \frac{r_0}{2}]$, (2.12), (2.13) and Itô's fomrula, we see that ξ_t is a sup-martingale, so that

$$\rho_\partial(x) \ge \xi_0 \ge \mathbb{E}^x[\xi_{t\wedge\tau(X^\mu)\wedge\hat{\tau}}] \ge \mathbb{P}^x\big(t \wedge \tau(X^\mu) \ge \hat{\tau}\ \big) \int_0^{r_0/2} \mathrm{e}^{-2\alpha^2 s}\mathrm{d}s. \tag{2.17}$$

Combining this with (2.14) and (2.16), we prove (2.11) for some constant $c > 0$. □

Proof of Theorem 2.1. (a) Well-posedness. Let $\gamma := \mathscr{L}_{X_0}^O$, and consider

$$\mathscr{C}^\gamma := \big\{\mu \in C([0,T]; \mathscr{P}_O) : \ \mu_0 = \gamma\big\}. \tag{2.18}$$

We intend to prove that Φ is contractive in $\mathscr{C}^\gamma$ under the complete metric

$$\hat{\mathbb{W}}_{1,\theta}(\mu^1, \mu^2) := \sup_{t\in[0,T]} \mathrm{e}^{-\theta t}\hat{\mathbb{W}}_1(\mu_t^1, \mu_t^2)$$

for large enough $\theta > 0$. Then Φ has a unique fixed point in $\mathscr{C}^\gamma$, so that the well-posedness follows from Theorem 1.1.

To this end, let $\mu^i \in \mathscr{C}^\gamma$ and let X_t^i solve (1.5) with $\mu = \mu^i$ and $X_0^i = X_0, i = 1, 2$. By $r_0 \le 1$, Lemma 2.3, and noting that

$$\mathbf{1}_{\{t\wedge\tau_2\ge\tau_1\}}\rho_\partial(X^2_{t\wedge\tau_{1,2}}) \le \mathbf{1}_{\{t\wedge\tau_2\ge\tau_1\}}|X^2_{t\wedge\tau_{1,2}} - X^1_{t\wedge\tau_{1,2}}|,$$

we obtain

$$\begin{aligned} \mathbb{E}\Big[\mathbf{1}_{\{t\wedge\tau_2\ge\tau_1\}}\big\{r_0 \wedge \rho_\partial(X^2_{t\wedge\tau_2})\big\}\Big] &= \mathbb{E}\Big(\mathbf{1}_{\{t\wedge\tau_2\ge\tau_1\}}\mathbb{E}\big[\big\{r_0 \wedge \rho(X^2_{t\wedge\tau_2})\big\}\big|\mathscr{F}_{\tau_1}\big]\Big) \\ &\le c\,\mathbb{E}\Big[\mathbf{1}_{\{t\wedge\tau_2\ge\tau_1\}}\big\{r_0 \wedge \rho_\partial(X^2_{t\wedge\tau_{1,2}})\big\}\Big] \le c\,\mathbb{E}\big[1 \wedge |X^1_{t\wedge\tau_{1,2}} - X^2_{t\wedge\tau_{1,2}}|\big]. \end{aligned} \tag{2.19}$$

By symmetry, the same estimate holds for $\mathbb{E}\Big[\mathbf{1}_{\{t\wedge\tau_1\geq\tau_2\}}\big\{r_0\wedge\rho_\partial(X^1_{t\wedge\tau_2})\big\}\Big]$. Combining these with $X_0^1 = X_0^2 = X_0$, (2.8) and (2.10), we find a constant $c_1 > 0$ such that

$$\hat{\mathbb{W}}_1((\Phi\mu^1)_t, (\Phi\mu^2)_t) \leq c_1\bigg(\int_0^t K(s)\hat{\mathbb{W}}_1(\mu_s^1, \mu_s^2)^2 \mathrm{d}s\bigg)^{\frac{1}{2}}, \quad t \in [0, T].$$

This implies that Φ is contractive in $\hat{\mathbb{W}}_{1,\theta}$ for large enough $\theta > 0$.

(b) Estimate (2.2). Let $\mu_t = \mathscr{L}^O_{X_t}$ for the unique solution of (1.3), we have $X_t = X_t^\mu$ since μ is a fixed point of Φ. So, (2.2) follows from (2.4).

(c) Estimate (2.3). Take X_0^1, X_0^2 such that

$$\mathscr{L}^O_{X_0^1} = \mu, \quad \mathscr{L}^O_{X_0^2} = \nu, \quad \mathbb{E}[1 \wedge |X_0^1 - X_0^2|] = \hat{\mathbb{W}}_1(\mu, \nu). \tag{2.20}$$

Let X_t^1 and X_t^2 solve (1.3). Then they solve (2.5) with

$$\mu_t^1 := \mathscr{L}^O_{X_t^1} = P_t^{D*}\mu, \quad \mu_t^2 := \mathscr{L}^O_{X_t^2} = P_t^{D*}\nu,$$

so that $\mu_t^i = (\Phi\mu^i)_t,\ t \in [0, T],\ i = 1, 2$. Thus, by (2.8), (2.9) and Lemma 2.3, we find a constant $c_2 > 0$ such that

$$\begin{aligned}
&\hat{\mathbb{W}}_1(P_t^{D*}\mu, P_t^{D*}\nu) = \hat{\mathbb{W}}_1((\Phi\mu^1)_t, (\Phi\mu^2)_t)\\
&\leq c_2\hat{\mathbb{W}}_1(\mu, \nu) + \bigg(c_2\int_0^t K(s)\hat{\mathbb{W}}_1(P_s^{D*}\mu, P_s^{D*}\nu)^2 \mathrm{d}s\bigg)^{\frac{1}{2}}, \quad t \in [0, T].
\end{aligned}$$

By Gronwall's inequality, we prove (2.3) for some constant $c > 0$. $\square$

2.2 Monotonicity in $\mathbb{W}_1$

Let $\mathscr{P}_O^1 = \{\mu \in \mathscr{P}_O, \|\mu\|_1 := \mu(|\cdot|) < \infty\}$. Define

$$\|\mu\|_{1,T} := \sup_{t\in[0,T]} \|\mu_t\|_1, \quad \mu \in C([0, T]; \mathscr{P}_O^1).$$

(B_1) For any $\mu \in C([0, T]; \mathscr{P}_O^1)$, $b_t(x, \mu_t)$ and $\sigma_t(x, \mu_t)$ are continuous in $x \in O$ such that for any $N \geq 1$ and $O_N := \{x \in O : |x| \leq N\}$,

$$\int_0^T \sup_{O_N} \big\{|b_t(\cdot, \mu_t)| + \|\sigma_t(\cdot, \mu_t)\|^2\big\}\mathrm{d}t < \infty.$$

Moreover, there exists $K \in L^1([0, T]; (0, \infty))$ such that for any $x, y \in O$ and $\mu, \nu \in \mathscr{P}_O^1$,

$$\begin{aligned}
&2\langle b_t(x, \mu) - b_t(y, \nu), x - y\rangle + \|\sigma_t(x, \mu) - \sigma_t(y, \nu)\|_{HS}^2 \leq K(t)\big\{|x - y|^2 + \mathbb{W}_1(\mu, \nu)^2\big\},\\
&2\langle b_t(x, \mu), x\rangle + \|\sigma_t(x, \mu)\|_{HS}^2 \leq K(t)\big\{1 + |x|^2 + \|\mu\|_1^2\big\}, \quad t \in [0, T].
\end{aligned}$$

(B_2) There exists $r_0 > 0$ such that $\rho_\partial \in C^2(\partial_{r_0} O)$, and there exists an increasing function $\alpha : [0,\infty) \to [1,\infty)$ such that

$$(2.21)\qquad |\sigma_t(x,\mu)^*\nabla \rho_\partial|^{-2} \le \alpha(\|\mu\|_1), \quad L_{t,\mu}\rho_\partial(x) \le \alpha(\|\mu\|_1), \quad x \in \partial_{r_0} O,$$

$$(2.22)\qquad \begin{aligned} &2\langle b_t(x,\mu), x-y\rangle + \|\sigma_t(x,\mu)\|_{HS}^2 \\ &\le K(t)\alpha(\|\mu\|_1)\big(1+|x-y|^2\big), \quad t \in [0,T], y \in \partial O, x \in O. \end{aligned}$$

Theorem 2.4. *Assume (B_1) and (B_2). Then the following assertions hold.*

(1) (1.3) *is well-posed for O-distributions in $\mathscr{P}_O^1$. Moreover, for any $p \ge 1$ there exists a constant $c > 0$ such that for any solution X_t to* (1.3) *for O-distributions in $\mathscr{P}_O^1$,*

$$(2.23)\qquad \mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^p \Big| \mathscr{F}_0\Big] \le c\big(1+|X_0| + \mathbb{E}[\mathbf{1}_O(X_0)|X_0|]\big)^p.$$

(2) *If α is bounded, then there exists a constant $c > 0$ such that*

$$(2.24)\qquad \sup_{t\in[0,T]} \mathbb{W}_1(P_t^{D*}\mu, P_t^{D*}\nu) \le c\,\mathbb{W}_1(\mu,\nu), \quad \mu,\nu \in \mathscr{P}_O^1.$$

It is standard that (B_1) and (B_2) imply the well-posedness of (1.5) for $\mu \in C([0,T];\mathscr{P}_O^1)$, and instead of (2.4), for any $p \ge 1$ there exists a constant $c > 0$ such that

$$(2.25)\quad \mathbb{E}\Big[\sup_{t\in[0,T]} |X_t^\mu|^p \Big| \mathscr{F}_0\Big] \le c\big(1+|X_0^\mu|^p\big) + c\int_0^t K(s)\|\mu_s\|_1^p \mathrm{d}s, \quad t \in [0,T], \mu \in C([0,T];\mathscr{P}_O^1).$$

Let $\mu^i \in C([0,T];\mathscr{P}_O^1)$, $i = 1,2$, let X_t^i solve (1.5) for μ^i replacing μ with $\mathscr{L}_{X_0^i}^O = \mu_0^i$, and denote as before

$$\tau_i := \tau(X^i) \text{ for } i = 1,2, \quad \tau_{1,2} := \tau_1 \wedge \tau_2.$$

Using (B_1) replacing (A_1), the proof of (2.9) leads to

$$(2.26)\qquad \mathbb{E}\big[|X_{t\wedge\tau_{1,2}}^1 - X_{t\wedge\tau_{1,2}}^2|^2 \big| \mathscr{F}_0\big] \le c|X_0^1 - X_0^2|^2 + c\int_0^t K(s)\mathbb{W}_1(\mu_s^1,\mu_s^2)^2 \mathrm{d}s, \quad t \in [0,T],$$

and instead of (2.8), we have

$$(2.27)\qquad \begin{aligned} \mathbb{W}_1((\Phi\mu^1)_t, (\Phi\mu^2)_t) &\le \mathbb{E}\big[|\bar{X}_t^1 - \bar{X}_t^2|\big] \le \mathbb{E}\big[|X_{t\wedge\tau_{1,2}}^1 - \bar{X}_{t\wedge\tau_{1,2}}^2|\big] \\ &+ \mathbb{E}[\rho_\partial(X_t^1)\mathbf{1}_{\{t\wedge\tau_1 \ge \tau_2\}}\big] + \mathbb{E}[\rho_\partial(X_t^2)\mathbf{1}_{\{t\wedge\tau_2\ge\tau_1\}}\big]. \end{aligned}$$

The following lemma is analogous to Lemma 2.3.

Lemma 2.5. *Assume (B_2). Then there exists an increasing function $\psi : [0,\infty) \to (0,\infty)$ which is bounded if so is α, such that for any $\mu \in C([0,T];\mathscr{P}_O^1)$ and any solution X_t^μ to* (1.5) *and any stopping time $\tilde{\tau}$,*

$$\mathbf{1}_{\{t\wedge\tau(X^\mu)\ge\tilde{\tau}\}}\mathbb{E}\big[\rho_\partial(X_t^\mu)\big|\mathscr{F}_{\tilde{\tau}}\big] \le \mathbf{1}_{\{t\wedge\tau(X^\mu)\ge\tilde{\tau}\}}\psi(\|\mu\|_{1,T})\rho_\partial(X_{\tilde{\tau}}^\mu).$$

Proof. By the strong Markov property, we may assume that $\tilde{\tau} = 0$ and $x = X_0^\mu \in O$, so that it suffices to prove

$$\Gamma_t(x) := \mathbb{E}^x[\rho_\partial(X_t^\mu)] \le \psi(\|\mu\|_{1,T})\rho_\partial(x), \quad x \in O, t \in [0,T]. \tag{2.28}$$

(a) Let $\rho_\partial(x) \ge \frac{r_0}{2}$ and $y \in \partial O$ such that $\rho_\partial(x) = |y - x|$. By (2.22), we have

$$\mathrm{d}|X_t^\mu - y|^2 \le K(t)\alpha(\|\mu\|_{1,T})\big(1 + |X_t^\mu - y|^2\big)\mathrm{d}t + \mathrm{d}M_t, \quad t \in [0, T \wedge \tau(X^\mu)]$$

for some martingale M_t. Combining this with $|x - y| = \rho_\partial(x)$, we obtain

$$\begin{aligned}\mathbb{E}^x[|X_t^\mu - y|^2] &\le \rho_\partial(x)^2 + \alpha(\|\mu\|_1)\int_0^t K(s)\mathrm{d}s \\ &+ \int_0^t K(s)\alpha(\|\mu\|_{1,T})\mathbb{E}^x[|X_s^\mu - y|^2]\mathrm{d}s, \quad t \in [0,T].\end{aligned}$$

By Gronwall's inequality and $\rho_\partial(x) \ge \frac{r_0}{2}$, we find an increasing function $\psi_1 : [0,\infty) \to (0,\infty)$ which is bounded if so is α, such that

$$\begin{aligned}\mathbb{E}^x[|X_t^\mu - y|^2] &\le \left\{\rho_\partial(x)^2 + \alpha(\|\mu\|_{1,T})\int_0^T K(s)\mathrm{d}s\right\}\mathrm{e}^{\alpha(\|\mu\|_{1,T})\int_0^T K(s)\mathrm{d}s} \\ &\le \big\{\psi_1(\|\mu\|_{1,T})\rho_\partial(x)\big\}^2, \quad t \in [0,T].\end{aligned}$$

Combining this with Jensen's inequality, we prove (2.28) with $\psi = \psi_1$ holds for $\rho_\partial(x) \ge \frac{r_0}{2}$.

(b) Let $\rho_\partial(x) < \frac{r_0}{2}$. Simply denote $\alpha = \alpha(\|\mu\|_{1,T})$ and define

$$\hat{\tau} := \inf\{t \ge 0 : \rho_\partial(X_t^\mu) \ge r_0\}.$$

By (B_2) and Itô's formula, we obtain

$$\mathrm{d}\rho_\partial(X_t^\mu) \le \alpha\mathrm{d}t + \mathrm{d}M_t, \quad t \in [0, T \wedge \tau(X^\mu) \wedge \hat{\tau}]$$

for some martingale satisfying (2.13). So,

$$\mathbb{E}^x[\rho_\partial(X^\mu_{t\wedge\tau(X^\mu)\wedge\hat{\tau}})] \le \alpha\,\mathbb{E}^x[t \wedge \tau(X^\mu) \wedge \hat{\tau}].$$

Combining this with step (a) and the strong Markov property, we obtain

$$\begin{aligned}&\mathbb{E}^x[\rho_\partial(X_t^\mu)] = \mathbb{E}^x[\rho_\partial(X^\mu_{t\wedge\tau(X^\mu)})] \le \mathbb{E}^x[\rho_\partial(X^\mu_{t\wedge\tau(X^\mu)\wedge\hat{\tau}})] + \mathbb{E}^x\Big[\mathbf{1}_{\{t\wedge\tau(X^\mu)\ge\hat{\tau}\}}\Gamma_{t-\hat{\tau}}(X^\mu_{\hat{\tau}})\Big] \\ &\le \alpha\,\mathbb{E}^x[t \wedge \tau(X^\mu) \wedge \hat{\tau}] + \mathbb{P}^x\big(t \wedge \tau(X^\mu) \ge \hat{\tau}\big)\psi_1(\|\mu\|_{1,T})r_0.\end{aligned}$$

Combining this with (2.15) and (2.17), we prove (2.28) for some increasing function $\psi : [0,\infty) \to (0,\infty)$, which is bounded if so is α. □

Proof of Theorem 2.4. Let X_t solve (1.3) for O-distributions in $\mathscr{P}_O^1$. Then $X_t = X_t^\mu$ for $\mu_t := \mathscr{L}_{X_t}^O$, so that

$$\|\mu_s\|_1 = \mathbb{E}[\mathbf{1}_O(X_s)|X_s|] = \mathbb{E}[\mathbf{1}_{\{t<\tau(X)\}}|X_s|], \quad s \in [0,T].$$

Combining this with (2.25), we obtain

$$\begin{aligned}\|\mu_t\|_1^2 &\le \Big(\mathbb{E}\sqrt{\mathbb{E}[\mathbf{1}_{\{t<\tau(X)\}}|X_t|^2|\mathscr{F}_0]}\Big)^2 \le 2\Big(\mathbb{E}\sqrt{c(1+\mathbf{1}_O(X_0)|X_0|^2)}\Big)^2 + 2c\int_0^t K(s)\|\mu_s\|_1^2 \mathrm{d}s \\ &\le 2c(1+\mathbb{E}[\mathbf{1}_O(X_0)|X_0|])^2 + 2c\int_0^t K(s)\|\mu_s\|_1^2\mathrm{d}s, \quad t \in [0,T].\end{aligned}$$

By Gronwall's inequality, we find a constant $c_1 > 0$ such that

$$\sup_{t\in[0,T]} \|\mu_t\|_1^2 \le c_1(1+\mathbb{E}[\mathbf{1}_O(X_0)|X_0|])^2.$$

This together (2.25) yields (2.23) for some different constant $c > 0$. It remains to prove the well-posedness and (2.24).

(a) Well-posedness. Let $\gamma := \mathscr{L}_{X_0}^O \in \mathscr{P}_O^1$. For any $N > 0$, let

$$\mathscr{C}_N^\gamma := \Big\{\mu \in C([0,T];\mathscr{P}_O^1):\ \mu_0 = \gamma,\ \sup_{t\in[0,T]} \mathrm{e}^{-Nt}\|\mu_t\|_1 \le N\Big\}. \tag{2.29}$$

We first observe that for some constant $N_0 > 0$,

$$\Phi\mathscr{C}_N^\gamma \subset \mathscr{C}_N^\gamma, \quad N \ge N_0. \tag{2.30}$$

Let $\mu \in \mathscr{C}_N^\gamma$ and let X_t^μ solve (1.5) for $X_0^\mu = X_0$. Then $(\Phi\mu)_t = \mathscr{L}_{X_t^\mu}^O$. By (2.25) and

$$\|(\Phi\mu)_t\|_1 \le \mathbb{E}\sqrt{\mathbb{E}[\mathbf{1}_O(X_0)|X_{t\wedge\tau(X^\mu)}|^2|\mathscr{F}_0]},$$

we find a constant $c_1 > 0$ such that

$$\|(\Phi\mu)_t\|_1 \le c_1(1+\|\gamma\|_1) + c_1\Big(\int_0^t \|\mu_s\|_1^2\mathrm{d}s\Big)^{\frac12}, \quad t \in [0,T].$$

Then for any $N \ge N_0 := c_1 + 2c_1(1+\|\gamma\|_1)$, we have

$$\begin{aligned}\sup_{t\in[0,T]} \mathrm{e}^{-Nt}\|(\Phi\mu)_t\|_1 &\le c_1(1+\|\gamma\|_1) + c_1 \sup_{t\in[0,T]}\Big(\int_0^t \mathrm{e}^{-2Ns}\|\mu_s\|_1^2 \mathrm{e}^{-2N(t-s)}\mathrm{d}s\Big)^{\frac12} \\ &\le c_1(1+\|\gamma\|_1) + c_1 N \sup_{t\in[0,T]}\Big(\int_0^t \mathrm{e}^{-2N(t-s)}\mathrm{d}s\Big)^{\frac12} \\ &\le c_1(1+\|\gamma\|_1) + c_1\sqrt{N} \le N.\end{aligned}$$

Next, for any $N \ge N_0$, we intend to prove that Φ is contractive in $\mathscr{C}_N^\gamma$ under the complete metric

$$\mathbb{W}_{1,\theta}(\mu^1,\mu^2) := \sup_{t\in[0,T]} \mathrm{e}^{-\theta t}\mathbb{W}_1(\mu_t^1,\mu_t^2)$$

for large enough $\theta > 0$, so that Φ has a unique fixed point in $\mathscr{C}^\gamma = \cup_{N\ge N_0}\mathscr{C}_N^\gamma$, hence the well-posedness follows from Theorem 1.1.

To this end, let $\mu^i \in \mathscr{C}_N^\gamma$ and X_t^i solve (1.5) for $\mu = \mu^i$ and $X_0^i = X_0, i = 1,2$. By Lemma 2.5 and noting that $\rho_\partial(x) \le |x-y|$ for $x \in O$ and $y \in \partial O$, we find a constant $c_2 > 0$ depending on N such that for any $\mu^1, \mu^2 \in \mathscr{C}_N^\gamma$,

$$\begin{aligned}&\mathbb{E}\big[\rho_\partial(X_t^1)\mathbf{1}_{\{t\wedge\tau_1\ge\tau_2\}} + \rho_\partial(X_t^2)\mathbf{1}_{\{t\wedge\tau_2\ge\tau_1\}}\big]\\ &\le c_2\mathbb{E}\big[\rho_\partial(X^1_{t\wedge\tau_{1,2}})\mathbf{1}_{\{t\wedge\tau_1\ge\tau_2\}} + \rho_\partial(X^2_{t\wedge\tau_{1,2}})\mathbf{1}_{\{t\wedge\tau_2\ge\tau_1\}}\big] \le 2c_2\mathbb{E}[|X^1_{t\wedge\tau_{1,2}} - X^2_{t\wedge\tau_{1,2}}|].\end{aligned}$$

Combining this with (2.26) and (2.27), we find a constant $c_3 > 0$ depending on N such that

$$\mathbb{W}_1((\Phi\mu^1)_t, (\Phi\mu^2)_t) \le c_3\mathbb{E}[|X_0^1 - X_0^2|] + c_3\bigg(\int_0^t K(s)\mathbb{W}_1(\mu_s^1,\mu_s^2)^2\mathrm{d}s\bigg)^{\frac12}, \quad \mu^1,\mu^2 \in \mathscr{C}_N^\gamma. \tag{2.31}$$

Since $X_0^1 = X_0^2 = X_0$, this implies the contraction of Φ in $\mathbb{W}_{1,\theta}$ for large enough $\theta > 0$.

(b) Estimate (2.24). Now, for $\mu_0^1, \mu_0^2 \in \mathscr{P}_O^1$, let X_0^1, X_0^2 be $\mathscr{F}_0$-measurable random variables on $\bar O$ such that

$$\mathscr{L}^O_{X_0^1} = \mu_0^1, \quad \mathscr{L}^O_{X_0^2} = \mu_0^2, \quad \mathbb{E}[|X_0^1 - X_0^2|] = \mathbb{W}_1(\mu_0^1,\mu_0^2). \tag{2.32}$$

Letting X_t^i solve (1.3) with initial value X_0^i, then $\mu^i := (P_t^{D*}\mu_0^i)_{t\in[0,T]}$ is the unique fixed point of Φ in $\mathscr{C}^{\mu_0^i}$, so that

$$\mu_t^i = \mathscr{L}^O_{X_t^i} = \Phi\mu_t^i = P_t^{D*}\mu_0^i, \quad i = 1,2, t \in [0,T]. \tag{2.33}$$

When α is bounded, (2.31) holds for some constant $c_3 > 0$ independent of N, which together with (2.32) yields

$$\begin{aligned}&\mathbb{W}_1(\mu_t^1,\mu_t^2) = \mathbb{W}_1((\Phi\mu^1)_t, (\Phi\mu^2)_t) \le c_3\mathbb{E}[|X_0^1 - X_0^2|] + c_3\bigg(\int_0^t K(s)\mathbb{W}_1(\mu_s^1,\mu_s^2)^2\mathrm{d}s\bigg)^{\frac12}\\ &= c_3\mathbb{W}_1(\mu_0^1,\mu_0^2) + c_3\bigg(\int_0^t K(s)\mathbb{W}_1(\mu_s^1,\mu_s^2)^2\mathrm{d}s\bigg)^{\frac12}, \quad t\in[0,T].\end{aligned}$$

By Gronwall's inequality and (2.33), we obtain

$$\mathbb{W}_1(P_t^{D*}\mu_0^1, P_t^{D*}\mu_0^2)^2 = \mathbb{W}_1(\mu_t^1,\mu_t^2)^2 \le 2c_3^2\mathbb{W}_1(\mu_0^1,\mu_0^2)^2\mathrm{e}^{2c_3^2\int_0^t K(s)\mathrm{d}s}, \quad t \in [0,T].$$

Then the proof is finished. □

3 Singular case with distribution dependent noise

In this part, we assume that σ and b are extended to $[0,T]\times\mathbb{R}^d\times\mathscr{P}_O$ but may be singular in the space variable. To measure the singularity, we recall locally integrable functional spaces introduced in [12]. For any $t > s \geq 0$ and $p, q \in (1, \infty)$, we write $f \in \tilde{L}_p^q([s,t])$ if $f : [s,t] \times \mathbb{R}^d \to \mathbb{R}$ is measurable with

$$\|f\|_{\tilde{L}_p^q([s,t])} := \sup_{z \in \mathbb{R}^d} \left\{ \int_s^t \left(\int_{B(z,1)} |f(u,x)|^p \mathrm{d}x \right)^{\frac{q}{p}} \mathrm{d}u \right\}^{\frac{1}{q}} < \infty,$$

where $B(z,1) := \{x \in \mathbb{R}^d : |x - z| \leq 1\}$ is the unit ball centered at point z. When $s = 0$, we simply denote

$$\tilde{L}_p^q(t) = \tilde{L}_p^q([0,t]), \quad \|f\|_{\tilde{L}_p^q(t)} = \|f\|_{\tilde{L}_p^q([0,t])}. \tag{3.1}$$

We will take (p,q) from the space

$$\mathscr{K} := \left\{ (p,q) : p, q > 2, \frac{d}{p} + \frac{2}{q} < 1 \right\}. \tag{3.2}$$

For any $\mu \in C([0,T]; \mathscr{P}_O)$, let

$$\sigma_t^\mu(x) := \sigma_t(x, \mu_t), \quad b_t^\mu(x) := b_t(x, \mu_t) = b_t^{\mu,0}(x) + b_t^{(1)}(x), \quad (t,x) \in [0,T] \times \mathbb{R}^d, \tag{3.3}$$

where $b_t^{\mu,0}(\cdot)$ is singular and $b_t^{(1)}(\cdot)$ is Lipschitz continuous.

As in the last section, we consider (1.3) for O-distributions in $\mathscr{P}_O$ and $\mathscr{P}_O^1$ respectively.

3.1 For O-distributions in $\mathscr{P}_O$

(C) There exist $K \in (0, \infty)$, $l \in \mathbb{N}$, $\{(p_i, q_i) : 0 \leq i \leq l\} \subset \mathscr{K}$ and $1 \leq f_i \in \tilde{L}_{p_i}^{q_i}(T)$ for $0 \leq i \leq l$, such that σ^μ and b^μ in (3.3) satisfy the following conditions.

(C_1) For any $\mu \in C([0,T]; \mathscr{P}_O)$, $a^\mu := \sigma^\mu (\sigma^\mu)^*$ is invertible with $\|a^\mu\|_\infty + \|(a^\mu)^{-1}\|_\infty \leq K$ and

$$\lim_{\varepsilon \downarrow 0} \sup_{\mu \in C([0,T]; \mathscr{P}_O)} \sup_{t \in [0,T], |x-y| \leq \varepsilon} \|a_t^\mu(x) - a_t^\mu(y)\| = 0.$$

(C_2) $b^{(1)}(0)$ is bounded on $[0,T]$, σ_t^μ is weakly differentiable for $\mu \in C([0,T]; \mathscr{P}_O)$, and

$$|b_t^{\mu,0}(x)| \leq f_0(t,x), \quad \|\nabla \sigma_t^\mu(x)\| \leq \sum_{i=1}^l f_i(t,x),$$
$$|b_t^{(1)}(x) - b_t^{(1)}(y)| \leq K|x-y|, \quad t \in [0,T], x, y \in \mathbb{R}^d.$$

(C_3) For any $t \in [0,T], x \in \mathbb{R}^d$ and $\mu, \nu \in \mathscr{P}_O$,

$$\|\sigma_t(x,\mu) - \sigma_t(x,\nu)\| + |b_t(x,\mu) - b_t(x,\nu)| \leq \hat{\mathbb{W}}_1(\mu,\nu) \sum_{i=0}^l f_i(t,x).$$

Theorem 3.1. *Assume* **(C)** *and* (A_2). *Then the following assertions hold.*

(1) (1.3) *is well-posed for O-distributions in* $\mathscr{P}_O$.

(2) *For any* $p \geq 1$, *there exists a constant* $c_p > 0$ *such that for any solution* X_t *to* (1.3) *for O-distributions in* $\mathscr{P}_O$,

$$\mathbb{E}\bigg[\sup_{t\in[0,T]} |X_t|^p \bigg| \mathscr{F}_0\bigg] = \mathbb{E}\bigg[\sup_{t\in[0,T]} |X_{t\wedge\tau(X)}|^p \bigg| \mathscr{F}_0\bigg] \leq c_p\big(1+|X_0|^p\big). \tag{3.4}$$

(3) *There exists a constant* $c > 0$ *such that* (2.3) *holds.*

For any $\mu \in C([0,T];\mathscr{P}_O)$, instead of (1.5) we consider the following SDE on $\mathbb{R}^d$:

$$\mathrm{d}X_t^\mu = b_t^\mu(X_t^\mu)\mathrm{d}t + \sigma_t^\mu(X_t^\mu)\mathrm{d}W_t, \quad t \in [0,T]. \tag{3.5}$$

Noting that $\tilde{X}_t^\mu := X^\mu_{t\wedge\tau(X^\mu)}$ solves (1.5), the map Φ in (1.6) is given by

$$(\Phi\mu)_t := \mathscr{L}^O_{X^\mu_{t\wedge\tau(X^\mu)}}, \quad t \in [0,T].$$

So, (2.8) and (2.27) remain true for X_t^i solving (3.5) with $\mu = \mu^i \in C([0,T];\mathscr{P}_O), i = 1,2$.

By [7, Theorem 2.1], see also [11, Theorem 1.1] for the distribution dependent setting, (C_1) and (C_2) imply that this SDE is well-posed, and for any $p \geq 1$ there exists a constant $c_p > 0$ such that

$$\mathbb{E}\bigg[\sup_{t\in[0,T]} |X_t^\mu|^p \bigg| \mathscr{F}_0\bigg] \leq c_p\big(1+|X_0^\mu|^p\big), \quad \mu \in C([0,T];\mathscr{P}_O). \tag{3.6}$$

We have the following lemma.

Lemma 3.2. *Assume* **(C)**. *Then for any* $j \geq 1$ *there exists a constant* $c > 0$ *and a function* $\varepsilon : [1,\infty) \to (0,\infty)$ *with* $\varepsilon(\theta) \downarrow 0$ *as* $\theta \uparrow \infty$, *such that for any* $\mu^1, \mu^2 \in C([0,T];\mathscr{P}_O)$ *and any* X_t^i *solving* (3.5) *with* $\mu = \mu^i, i = 1,2$,

$$\mathbb{E}\Big[\sup_{s\in[0,t]} |X_s^1 - X_s^2|^j \big| \mathscr{F}_0\Big] \leq c|X_0^1 - X_0^2|^j + \varepsilon(\theta)\mathrm{e}^{j\theta t}\hat{\mathbb{W}}_{1,\theta}(\mu^1,\mu^2)^j, \quad \theta \geq 1.$$

Proof. The assertions follows from the proof of [5, Lemma 2.1] for $\mu^i = \nu^i$ and for $\hat{\mathbb{W}}_1$ replacing $\mathbb{W}_k$ and $\mathbb{W}_{k,var}$. We figure it out for completeness.

By [13, Theorem 2.1], (C_1) and (C_2) imply that for large enough $\lambda \geq 1$, the PDE

$$\Big(\partial_t + \frac{1}{2}\mathrm{tr}\{a_t^{\nu^1}\nabla^2\} + b_t^{\mu^1}\cdot\nabla\Big)u_t = \lambda u_t - b_t^{\mu^1,0}, \quad t \in [0,T], u_T = 0 \tag{3.7}$$

for $u : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ has a unique solution such that

$$\|\nabla^2 u\|_{\tilde{L}^{q_0}_{p_0}(T)} \leq c_0, \quad \|u\|_\infty + \|\nabla u\|_\infty \leq \frac{1}{2}. \tag{3.8}$$

Let $Y_t^i := \Theta_t(X_t^i), i = 1,2, \Theta_t := id + u_t$. By Itô's formula we obtain

$$\mathrm{d}Y_t^1 = \big\{b_t^{(1)} + \lambda u_t\big\}(X_t^1)\mathrm{d}t + (\{\nabla\Theta_t\}\sigma_t^{\nu^1})(X_t^1)\,\mathrm{d}W_t,$$

$$\begin{aligned}
\mathrm{d}Y_t^2 &= \Big\{\big\{b_t^{(1)} + \lambda u_t + (\nabla\Theta_t)(b_t^{\mu^2} - b_t^{\mu^1})\big\}(X_t^2) \\
&+ \frac{1}{2}\big[\mathrm{tr}\{(a_t^{\nu^2} - a_t^{\nu^1})\nabla^2 u_t\}\big](X_t^2)\Big\}\mathrm{d}t + (\{\nabla\Theta_t\}\sigma_t^{\nu^2})(X_t^2)\,\mathrm{d}W_t.
\end{aligned}$$

Let $\eta_t := |X_t^1 - X_t^2|$ and

$$\begin{aligned}
g_r &:= \sum_{i=0}^{l} f_i(r, X_r^2), \quad \tilde{g}_r := g_r \|\nabla^2 u_r(X_r^2)\|, \\
\bar{g}_r &:= \sum_{i=1}^{2} \|\nabla^2 u_r\|(X_r^i) + \sum_{j=1}^{2}\sum_{i=0}^{l} f_i(r, X_r^j), \quad r \in [0,T].
\end{aligned}$$

Since $b_t^{(1)} + \lambda u_t$ is Lipschitz continuous uniformly in $t \in [0,T]$, by **(C)** and the maximal functional inequality in [12, Lemma 2.1], there exists a constant $c_1 > 0$ such that

$$\begin{aligned}
&\big|\big\{b_r^{(1)} + \lambda u_r\big\}(X_r^1) - \big\{b_r^{(1)} + \lambda u_r\big\}(X_r^2)\big| \le c_1 \eta_r, \\
&\big|\big\{(\nabla\Theta_r)(b_r^{\mu^2} - b_r^{\mu^1})\big\}(X_r^2)\big| \le c_1 g_r \hat{\mathbb{W}}_1(\mu_r^1, \mu_r^2), \\
&\big|\big[\mathrm{tr}\{(a_r^{\nu^2} - a_r^{\nu^1})\nabla^2 u_r\}\big](X_r^2)\big| \le c_1 \tilde{g}_r \hat{\mathbb{W}}_1(\mu_r^1, \mu_r^2), \\
&\big\|\big\{(\nabla\Theta_r)\sigma_r^{\nu^1}\big\}(X_r^1) - \big\{(\nabla\Theta_r)\sigma_r^{\mu^2}\big\}(X_r^2)\big\| \\
&\le c_1 \bar{g}_r \eta_r + c_1 g_r \hat{\mathbb{W}}_1(\mu_r^1, \mu_r^2), \quad r \in [0,T].
\end{aligned}$$

So, by Itô's formula, for any $j \ge k$ we find a constant $c_2 > 1$ such that

$$\mathrm{d}|Y_t^1 - Y_t^2|^{2j} \le c_2 \eta_t^{2j}\mathrm{d}A_t + c_2(g_t^2 + \tilde{g}_t)\hat{\mathbb{W}}_1(\mu_t^1, \mu_t^2)^{2j}\mathrm{d}t + \mathrm{d}M_t \tag{3.9}$$

holds for some martingale M_t with $M_0 = 0$ and

$$A_t := \int_0^t \big\{1 + g_s^2 + \tilde{g}_s + \bar{g}_s^2\big\}\mathrm{d}s, \quad t \in [0,T].$$

Since $\|\nabla u\|_\infty \le \frac{1}{2}$ implies $|Y_t^1 - Y_t^2| \ge \frac{1}{2}\eta_t$, this implies

$$\begin{aligned}
\eta_t^{2j} &\le 2^{2j} M_t + 2^{2j}\eta_0^{2j} + 2^{2j} c_2 \int_0^t \eta_r^{2j}\mathrm{d}A_r \\
&+ 2^{2j} c_2 \int_0^t (g_s^2 + \tilde{g}_s)\hat{\mathbb{W}}_1(\mu_s^1, \mu_s^2)^{2j}\mathrm{d}s, \quad t \in [0,T]
\end{aligned} \tag{3.10}$$

for some constant $c_2 > 0$. By (3.8), $f_i \in \tilde{L}_{p_i}^{q_i}(T)$ for $(p_i, q_i) \in \mathscr{K}$, Krylov's estimate (see [13, Theorem 3.1]) which implies Khasminskii's estimate (see [12, Lemma 4.1(ii)]), we find an increasing function $\psi : (0,\infty) \to (0,\infty)$ and a decreasing function $\varepsilon : (0,\infty) \to (0,\infty)$ with $\varepsilon(\theta) \downarrow 0$ as $\theta \uparrow \infty$, such that

$$\begin{aligned}
&\mathbb{E}[\mathrm{e}^{rA_T}|\mathscr{F}_0] \le \psi(r), \quad r > 0, \\
&\sup_{t\in[0,T]} \mathbb{E}\bigg(\int_0^t \mathrm{e}^{-2k\theta(t-r)}(g_r^2 + \tilde{g}_r)\mathrm{d}r\bigg|\mathscr{F}_0\bigg) \le \varepsilon(\theta), \quad \theta > 0.
\end{aligned}$$

By the stochastic Gronwall inequality and the maximal inequality (see [12]), we find a constant $c_3 > 0$ depending on N such that (3.10) yields

$$\begin{aligned}&\Big\{\mathbb{E}\Big(\sup_{s\in[0,t]}\eta_s^j\Big|\mathscr{F}_0\Big)\Big\}^2\\&\le c_3\mathbb{E}\bigg(\eta_0^{2j}+\int_0^t(g_s^2+\tilde{g}_s)\hat{\mathbb{W}}_1(\mu_s^1,\mu_s^2)^{2j}\mathrm{d}s\bigg|\mathscr{F}_0\bigg)\\&\le c_3\eta_0^{2j}+c_3\mathrm{e}^{2j\theta t}\varepsilon(\theta)\hat{\mathbb{W}}_1(\mu^1,\mu^2)^{2j},\ \ t\in[0,T],\theta>0.\end{aligned}$$

This finishes the proof. □

Proof of Theorem 3.1. Let X_t solve (1.3). We have $X_t = X^\mu_{t\wedge\tau(X^\mu)}$ for X^μ_t solving (3.5) with

$$X_0^\mu = X_0,\ \ \mu_t := \mathscr{L}^O_{X_t},\ \ t\in[0,T].$$

So, (3.4) follows from (3.6). It remains to prove the well-posedness and estimate (2.3).

(a) Well-posedness. Let X_0 be an $\mathscr{F}_0$-measurable random variable on $\bar{O}$, and let $\mathscr{C}^\gamma$ be in (2.18) for $\gamma = \mathscr{L}^O_{X_0}$. By Theorem 1.1, it suffices to prove that Φ is contractive in $\mathscr{C}^\gamma$ under $\hat{\mathbb{W}}_{1,\theta}$ for large enough $\theta > 0$.

By (2.8), (2.19) and Lemma 3.2 for $X_0^1 = X_0^2 = X_0$, we find a constant $c_1 > 0$ such that

$$\hat{W}_1((\Phi\mu^1)_t,(\Phi\mu^2)_t)\le c_1\varepsilon(\theta)\hat{\mathbb{W}}_1(\mu^1,\mu^2),\quad \mu^1,\mu^2\in\mathscr{C}^\gamma.$$

Since $\varepsilon(\theta)\to 0$ as $\theta\to\infty$, Φ is $\hat{\mathbb{W}}_{1,\theta}$-contractive for large enough $\theta > 0$.

(b) Estimate (2.3). Let X_t^1, X_t^2 solve (1.3) with X_0^1, X_0^2 satisfying (2.20). Then

$$(\Phi\mu^i)_t = \mu_t^i := \mathscr{L}^O_{X^i} = P_t^{D*}\mu^i,\quad i=1,2,$$

so that (2.8), (2.19) and Lemma 3.2 imply

$$\hat{W}_1(\mu^1,\mu^2)=\hat{W}_1((\Phi\mu^1)_t,(\Phi\mu^2)_t)\le c_1\hat{\mathbb{W}}_1(\mu_0^1,\mu_0^2)+c_1\varepsilon(\theta)\hat{\mathbb{W}}_1(\mu^1,\mu^2),\ \ t\in[0,T]$$

for some constant $c_1 > 0$. Taking $\theta > 0$ large enough such that $\varepsilon(\theta)\le\frac{1}{2c_1}$, we derive (2.3) for some constant $c > 0$.

□

3.2 For O-distributions in $\mathscr{P}_O^1$

(D) There exist an increasing function $\alpha : [0,\infty)\to(0,\infty)$, constants $K > 0, l\in\mathbb{N}$, $\{(p_i,q_i): 0\le i\le l\}\subset\mathscr{K}$ and functions $1\le f_i\in\tilde{L}^{q_i}_{p_i}(T)$ for $0\le i\le l$ such that σ^μ and b^μ in (3.3) satisfy the following conditions.

(D_1) For any $\mu\in C([0,T];\mathscr{P}^1_O)$, $a^\mu := \sigma^\mu(\sigma^\mu)^*$ is invertible with

$$\begin{aligned}&\|a^\mu\|_\infty+\|(a^\mu)^{-1}\|_\infty\le\alpha(\|\mu\|_{1,T}),\\&\lim_{\varepsilon\downarrow 0}\sup_{\mu\in C([0,T];\mathscr{P}^1_O)}\sup_{t\in[0,T],|x-y|\le\varepsilon}\|a_t^\mu(x)-a_t^\mu(y)\|=0.\end{aligned}$$

(D_2) $b^{(1)}(0)$ is bounded on $[0,T]$, σ_t^μ is weakly differentiable for $\mu \in C([0,T];\mathscr{P}_O^1)$, and

$$|b_t^{\mu,0}(x)| \le f_0(t,x) + \alpha(\|\mu\|_{1,T}), \quad \|\nabla \sigma_t^\mu(x)\| \le \sum_{i=1}^l f_i(t,x) + \alpha(\|\mu\|_{1,T}),$$
$$|b_t^{(1)}(x) - b_t^{(1)}(y)| \le K|x-y|, \quad t\in[0,T], x,y\in\mathbb{R}^d.$$

(D_3) For any $t\in[0,T], x\in\mathbb{R}^d$ and $\mu,\nu\in\mathscr{P}_O$,

$$\|\sigma_t(x,\mu)-\sigma_t(x,\nu)\| + |b_t(x,\mu)-b_t(x,\nu)| \le \mathbb{W}_1(\mu,\nu)\sum_{i=0}^l f_i(t,x).$$

(D_4) There exists $r_0\in(0,1]$ such that $\rho_\partial \in C_b^2(\partial_{r_0}O)$, and for any $\mu\in C([0,T];\mathscr{P}_O^1)$,

$$\langle b_t^\mu(x), \nabla\rho_\partial(x)\rangle \le \alpha(\|\mu\|_1), \quad x\in\partial_{r_0}O, \tag{3.11}$$

$$\langle b_t^\mu(x), x-y\rangle \le \alpha(\|\mu\|_{1,T})(f_0(t,x)^2+|x-y|^2), \quad x\in O, y\in\partial O, t\in[0,T]. \tag{3.12}$$

Note that when $b^{(1)}=0$, (3.11) is implied by the first condition in (D_2).

Theorem 3.3. *Assume* **(D)**. *Then the following assertions hold.*

1. (1.3) *is well-posed for O-distributions in $\mathscr{P}_O^1$.*
2. *For any $p\ge1$, there exists a constant $c_p>0$ such that for any solution X_t to* (1.3) *for O-distributions in $\mathscr{P}_O$,*

$$\mathbb{E}\bigg[\sup_{t\in[0,T]}|X_t|^p\bigg|\mathscr{F}_0\bigg] \le c_p\big\{1+|X_0|^p+\big(\mathbb{E}[\mathbf{1}_O(X_0)|X_0|]\big)^p\big\}. \tag{3.13}$$

3. *If α is bounded, then there exists a constant $c>0$ such that* (2.24) *holds.*

By the proof of [5, (2.17)], **(D)** implies that for any $\mu\in C([0,T];\mathscr{P}_O^1)$, the SDE (3.5) is well-posed, and for any $p\ge1$ there exists a constant $c_p>0$ such that

$$\mathbb{E}\bigg[\sup_{t\in[0,T]}|X_t^\mu|^{2p}\bigg|\mathscr{F}_0\bigg] \le c_p\bigg\{1+|X_0^\mu|^{2p}+\int_0^T\|\mu_s\|_1^{2p}\mathrm{d}s\bigg\}, \quad \mu\in C([0,T];\mathscr{P}_O). \tag{3.14}$$

For any $\mu^1,\mu^2\in\mathscr{P}_O^1$, let X_t^i solve (3.5) for $\mu=\mu^i, i=1,2$.

For any $N>0$ and $\gamma\in\mathscr{P}_O^1$, let $\mathscr{C}_N^\gamma$ be in (2.29). Since restricting to $\mu,\nu\in\mathscr{C}_N^\gamma$ the conditions in **(D)** hold for a constant α_N replacing the function α, by repeating the proof of Lemma 3.2 with $\mathbb{W}$ replacing $\hat{\mathbb{W}}$, we prove that the following result.

Lemma 3.4. *Assume* **(D)**. *For any $N > 0$ and $j \geq 1$, there exists a constant $c > 0$ and a function $\varepsilon : [1, \infty) \to (0, \infty)$ with $\varepsilon(\theta) \downarrow 0$ as $\theta \uparrow \infty$, such that for any $\mu^1, \mu^2 \in \mathscr{C}_N^\gamma$ and any X_t^i solving* (3.5) *with $\mu = \mu^i, i = 1, 2$,*

$$\mathbb{E}\Big[\sup_{s\in[0,t]} |X_s^1 - X_s^2|^j \Big| \mathscr{F}_0\Big] \leq c|X_0^1 - X_0^2|^j + \varepsilon(\theta) \mathrm{e}^{j\theta t} \hat{\mathbb{W}}_{1,\theta}(\mu^1, \mu^2)^j, \quad \theta \geq 1.$$

When α is bounded, the constant c does not depend on N.

Moreover, we need the following result analogous to Lemma 2.5.

Lemma 3.5. *Assume* **(D)**. *Then the assertion in Lemma 2.5 holds.*

Proof. It suffices to prove (2.28) for some increasing function ψ which is bounded if so is α.

(a) Let $\rho_\partial(x) \geq \frac{r_0}{2}$ and $y \in \partial O$ such that $\rho_\partial(x) = |y - x|$. By (3.12) and (D_2), we find an increasing function $\psi_1 : [0, \infty) \to (0, \infty)$ which is bounded if so is α, such that

$$\mathrm{d}|X_t^\mu - y|^2 \leq \psi_1(\|\mu\|_{1,T}) \Big(\sum_{i=0}^l f_i(t, X_t^\mu)^2 + |X_t^\mu - y|^2 \Big) \mathrm{d}t + \mathrm{d}M_t, \quad t \in [0, T \wedge \tau(X^\mu)]$$

for some martingale M_t. Next, by [13, Theorem 3.1], **(D)** implies that for some increasing function $\psi_2 : [0, \infty) \to (0, \infty)$ which is bounded if so is α, the following Krylov's estimate holds:

$$\mathbb{E}\bigg(\int_0^T f_i(t, X_t^\mu)^2 \mathrm{d}t \bigg| \mathscr{F}_0 \bigg) \leq \psi_2(\|\mu\|_{1,T}) \|f_i\|_{\tilde{L}_{q_i}^{p_i}(T)}^2, \quad 0 \leq i \leq l.$$

Combining these with $|x - y| = \rho_\partial(x)$, we derive

$$\begin{aligned} \mathbb{E}[|X_t^\mu - y|^2 | \mathscr{F}_0] \leq \rho_\partial(x)^2 &+ \psi_1(\|\mu\|_{1,T}) \psi_2(\|\mu\|_{1,T}) \sum_{i=0}^l \|f_i\|_{\tilde{L}_{q_i}^{p_i}(T)}^2 \\ &+ \psi_1(\|\mu\|_{1,T}) \int_0^t \mathbb{E}[|X_s^\mu - y|^2 | \mathscr{F}_0] \mathrm{d}s, \quad t \in [0, T]. \end{aligned}$$

By Gronwall's inequality and $\rho_\partial(x) \geq \frac{r_0}{2}$, we find an increasing function $\psi : [0, \infty) \to (0, \infty)$ which is bounded if so is α, such that

$$\mathbb{E}[|X_t^\mu - y|^2 | \mathscr{F}_0] \leq \psi(\|\mu\|_{1,T}) \rho_\partial(x).$$

Since $\rho_\partial(X_t^\mu) \leq |X_t^\mu - y|$, we prove (2.28) for $\rho_\partial(x) \geq \frac{r_0}{2}$.

(b) Let $\rho_\partial(x) < \frac{r_0}{2}$. By (D_1), (3.11) and $\rho_\partial \in C_b^2(\partial_{r_0} O)$, (2.21) holds for some different increasing function α which is bounded if so is the original one. Then step (b) in proof of Lemma 2.5 implies the desired estimate. □

Proof of Theorem 3.3. Let X_t solve (1.3) for O-distributions in $\mathscr{P}_O^1$. We have $X_t = X_{t\wedge\tau(X^\mu)}^\mu$ for X_t^μ solving (3.5) with

$$X_0^\mu = X_0, \quad \mu_t = \mathscr{L}_{X_t}^O, \quad t \in [0, T].$$

So, as explained in the beginning of proof of Theorem 2.4 that (3.13) follows from (3.14). It suffices to prove the well-posedness and estimate (2.24).

Let X_0 be an $\mathscr{F}_0$-measurable random variable with $\mathscr{L}_{X_0}^O \in \mathscr{P}_O^1$, and let $\mathscr{C}_N^\gamma$ be in (2.29) for $N > 0$. By the proof of [5, Lemma 2.2(1)], there exists $N_0 > 0$ such that $\Phi\mathscr{C}_N^\gamma \subset \mathscr{C}_N^\gamma$ for any $N \geq N_0$. For the well-posedness, it suffices to prove that for any $N \geq N_0$, Φ is contractive in $\mathscr{C}_N^\gamma$ under the metric $\mathbb{W}_{1,\theta}$ for large enough $\theta > 0$. This follows from (2.27), Lemma 3.4 and Lemma 3.5.

Finally, by using $\mathbb{W}_1$ replacing $\hat{\mathbb{W}}_1$ in step (b) in the proof of Theorem 3.1, (2.24) follows from Lemma 3.4 with c independent of N.

□

4 Singular case with distribution independent noise

In this part, we let $\sigma_t(x,\mu) = \sigma_t(x)$ do not depend on μ, so that (1.3) becomes

$$\mathrm{d}X_t = \mathbf{1}_{\{t<\tau(X)\}}\big\{b_t(X_t, \mathscr{L}_{X_t}^O)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t\big\}, \quad t \in [0,T]. \tag{4.1}$$

In this case, we are able to study the well-posedness of the equation on an arbitrary connected open domain O, for which we only need $b_t(x,\cdot)$ to be Lipschitz continuous with respect to a weighted variation distance.

For a measurable function $V : O \to [1,\infty)$, let

$$\mathscr{P}_O^V := \Big\{\mu \in \mathscr{P}_O : \ \mu(V) := \int_O V \mathrm{d}\mu < \infty\Big\}.$$

This is a Polish space under the weighted variation distance

$$\|\mu - \nu\|_V := \sup_{|f|\leq V} |\mu(f) - \nu(f)|, \quad \mu,\nu \in \mathscr{P}_O^V. \tag{4.2}$$

When $V \equiv 1$, $\|\cdot\|_V$ reduces to the total variation norm. We will take V from the class $\mathscr{V}$ defined as follows.

Definition 4.1. We denote $V \in \mathscr{V}$, if $1 \leq V \in C^2(\mathbb{R}^d)$ such that the level set $\{V \leq r\}$ for $r > 0$ is compact, and there exist constants $K, \varepsilon > 0$ such that for any $x \in O$,

$$\sup_{y\in B(x,\varepsilon)} \big\{|\nabla V(y)| + \|\nabla^2 V(y)\|\big\} \leq KV(x),$$

where $B(x,\varepsilon) := \{y \in \mathbb{R}^d : |y - x| < \varepsilon\}$.

4.1 Main result

(E) σ has an extension to $[0,T]\times\mathbb{R}^d$ which is weakly differentiable in $x \in \mathbb{R}^d$, and b has a decomposition $b_t(x,\mu) = b_t^{(0)}(x) + b_t^{(1)}(x,\mu)$, such that the following conditions hold.

(E_1) $a := \sigma\sigma^*$ is invertible with $\|a\|_\infty + \|a^{-1}\|_\infty < \infty$ and

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

(E_2) there exist $l \in \mathbb{N}$ and $1 \le f_i \in \tilde{L}^{p_i}_{q_i}(T)$ with $(p_i, q_i) \in \mathscr{K}, 0 \le i \le l$, such that

$$|\mathbf{1}_O b^{(0)}| \le f_0, \quad \|\nabla\sigma\| \le \sum_{i=1}^{l} f_i.$$

(E_3) there exists $V \in \mathscr{V}$ such that for any $\mu \in \mathscr{C}_V := C([0,T]; \mathscr{P}_O^V)$, $\mathbf{1}_O(x) b_t^{(1)}(x, \mu_t)$ is locally bounded in $(t,x) \in [0,T] \times \mathbb{R}^d$. Moreover, there exist constants $K, \varepsilon > 0$ such that

$$\begin{aligned}&\langle b^{(1)}(x,\mu), \nabla V(x)\rangle + \varepsilon |b^{(1)}(x,\mu)| \sup_{B(x,\varepsilon)} \big\{|\nabla V| + |\nabla^2 V|\big\} \\ &\le K\big\{V(x) + \mu(V)\big\} \quad x \in O, \mu \in \mathscr{P}_O^V.\end{aligned}$$

(E_4) there exists a constant $\kappa > 0$ such that

$$\sup_{x\in O} |b_t(x,\mu) - b_t(x,\nu)| \le \kappa \|\mu - \nu\|_V, \quad \mu, \nu \in \mathscr{P}_O^V. \tag{4.3}$$

Theorem 4.1. *Assume* **(E)**. *Then* (4.1) *is well-posed for O-distributions in* $\mathscr{P}_O^V$, *and for any* $p \ge 1$, *there exists a constant* $c_p > 0$ *such that any solution* X_t *of* (4.1) *for O-distributions in* $\mathscr{P}_O^V$ *satisfies*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} V(X_t)^p \Big| \mathscr{F}_0\Big] \le c_p V(X_0)^p. \tag{4.4}$$

Proof. Let $\mathscr{P}^V(\bar{O})$ be the space of all probability measures μ on $\bar{O}$ with $\mu(V) < \infty$, which is a Polish space under the weighted variation distance defined in (4.2) for $\mu, \nu \in \mathscr{P}^V(\bar{O})$. We extend $b_t(x, \cdot)$ from $\mathscr{P}_O^V$ to $\mathscr{P}^V(\bar{O})$ by setting

$$b_t(x,\mu) := b_t(x, \mu(O \cap \cdot)), \quad \mu \in \mathscr{P}^V(\bar{O}).$$

Then **(E)** implies the same assumption for $\mathscr{P}^V(\bar{O})$ replacing $\mathscr{P}_O^V$. So, the desired assertions follow from Theorem 4.2 presented in the next subsection. □

4.2 An extension of Theorem 4.1

Consider the following SDE on $\bar{O}$:

$$\mathrm{d}X_t = \mathbf{1}_{\{t<\tau(X)\}}\big\{b_t(X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t\big\}, \quad t \in [0,T], \tag{4.5}$$

where $\tau(X) := \inf\{t \ge 0 : X_t \in \partial O\}$ as before, and $\mathscr{L}_{X_t}$ is the distribution of X_t.

The strong/weak solution of (4.5) is defined as in Definition 1.1 with $\mathscr{L}$ replacing $\mathscr{L}^O$. We call this equation well-posed for distributions in $\mathscr{P}^V(\bar{O})$, if for any $\mathscr{F}_0$-measurable random variable X_0 on $\bar{O}$ with $\mathscr{L}_{X_0} \in \mathscr{P}^V(\bar{O})$ (respectively, any $\mu_0 \in \mathscr{P}^V(\bar{O})$), (4.5) has a unique solution starting at X_0 (respectively, a unique weak solution with initial distribution μ_0) such that $\mathscr{L}_X = (\mathscr{L}_{X_t})_{t\in[0,T]} \in C([0,T]; \mathscr{P}^V(\bar{O}))$.

Theorem 4.2. *Assume that* **(E)** *holds for* $\mathscr{P}^V(\bar O)$ *replacing* $\mathscr{P}^V_O$. *Then* (4.5) *is well-posed for distributions in* $\mathscr{P}^V(\bar O)$ *and* (4.4) *holds.*

Proof. (1) Let X_0 be an $\mathscr{F}_0$-measurable random variable on $\bar O$ with

$$\gamma := \mathscr{L}_{X_0} \in \mathscr{P}_V(\bar O).$$

Let

$$\mathscr{C}_V^\gamma(\bar O) := \big\{\mu \in C([0,T];\mathscr{P}_V(\bar O)):\ \mu_0 = \gamma\big\}.$$

For any $\mu \in \mathscr{C}_V^\gamma(\bar O)$, let X_t^μ solve (4.1) with $X_0^\mu = X_0$, i.e.

$$\mathrm{d}X_t^\mu = \mathbf{1}_{\{t<\tau(X^\mu)\}}\big\{b_t(X_t^\mu,\mu_t)\mathrm{d}t + \sigma_t(X_t^\mu)\mathrm{d}W_t\big\},\quad X_0^\mu = X_0, t\in[0,T]. \tag{4.6}$$

Let $(\Phi\mu)_t := \mathscr{L}_{X_t^\mu}, t\in[0,T]$. Then it suffices to prove that Φ has a unique fixed point in $\mathscr{C}_V^\gamma(\bar O)$. To this end, for any $N\ge 1$, let

$$\mathscr{C}_{V,N}^\gamma(\bar O) := \Big\{\mu\in\mathscr{C}_V^\gamma(\bar O):\ \sup_{t\in[0,T]} \mathrm{e}^{-Nt}\mu_t(V) \le N\gamma(V)\Big\}.$$

It suffices to find a constant $N_0>0$ such that for any $N\ge N_0$, Φ has a unique fixed point in $\mathscr{C}_{V,N}^\gamma(\bar O)$. We finish the proof by two steps.

(a) The Φ-invariance of $\mathscr{C}_{V,N}^\gamma(\bar O)$ for large N. For any $\lambda\ge 0$and $N\ge 1$, $\mathscr{C}_{V,N}^\gamma(\bar O)$ is a complete space under the metric

$$\rho_\lambda(\mu,\nu) := \sup_{t\in[0,T]} \mathrm{e}^{-\lambda t}\|\mu_t-\nu_t\|_V,\quad \mu,\nu\in\mathscr{C}_{V,N}^\gamma(\bar O).$$

Let $\mu\in\mathscr{C}_{V,N}^\gamma(\bar O)$. By (4.6), **(E)** with $V\in\mathscr{V}$ and Itô's formula, for any $p\ge 1$ we find a constant $c_1(p)>0$ such that

$$\mathrm{d}V(X_t^\mu)^p \le \mathbf{1}_{\{t<\tau(X^\mu)\}}\big\{\mathrm{d}M_t + c_1\big\{V(X_t^\mu)^p + \mu_t(V)^p\big\}\mathrm{d}t\big\},\quad t\in[0,T],$$

where M_t is a martingale with

$$\mathrm{d}\langle M\rangle_t \le c_1 V(X_t^\mu)^p\mathrm{d}t.$$

By using BDG's and Gronwall's inequality, we find a constant $c_2(p)>0$ such that

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\in[0,t]} V(X_s^\mu)^p\Big] &= \mathbb{E}\Big[\sup_{s\in[0,t\wedge\tau(X^\mu)]} V(X_s^\mu)^p\Big]\\ &\le c_2(p)V(X_0)^p + c_2(p)\int_0^t \mu_s(V)^p\mathrm{d}s,\quad t\in[0,T].\end{aligned} \tag{4.7}$$

Consequently, for $p=1$ and $c_2=c_2(1)$ we derive

$$(\Phi\mu)_t(V) = \mathbb{E}[V(X_t^\mu)] \le c_2\gamma(V) + c_2\bigg(\int_0^t \mu_s(V)^2\mathrm{d}s\bigg)^{\frac12},$$

so that by $\mu \in \mathscr{C}^\gamma_{V,N}(\bar O)$ we obtain

$$\sup_{t\in[0,T]} \mathrm{e}^{-Nt}(\Phi\mu)_t(V) \le c_2\gamma(V) + c_2 \sup_{t\in[0,T]} \bigg(\int_0^t \mathrm{e}^{-2N(t-s)}N^2\gamma(V)^2\mathrm{d}s\bigg)^{\frac{1}{2}}$$
$$c_2\big(1+\sqrt{N}\big)\gamma(V) \le N\gamma(V)$$

provided $N \ge N_0$ for a large enough constant $N_0 \ge 1$. By the continuity of X_t^μ in t, $(\Phi\mu)_t$ is weakly continuous in t. Therefore,

$$\Phi\mathscr{C}^\gamma_{V,N}(\bar O) \subset \mathscr{C}^\gamma_{V,N}(\bar O), \quad N \ge N_0.$$

(b) Let $N \ge N_0$. It remains to show that Φ has a unique fixed point in $\mathscr{C}^\gamma_{V,N}(\bar O)$. By (4.7) with $p=2$ and $V \ge 1$, there exists a constant $c_3 > 0$ such that

$$\mathbb{E}\Big[\sup_{t\in[0,T]} V(X_t^\mu)^2\Big|\mathscr{F}_0\Big] \le c_3^2 V(X_0)^2, \quad \mu \in \mathscr{C}^\gamma_{V,N}(\bar O). \tag{4.8}$$

For any $\mu^i \in \mathscr{C}_V(\bar O), i=1,2$, we estimate $\|(\Phi\mu^1)_t - (\Phi\mu^2)_t\|_V$ by using Girsanov's theorem. Let X_t^1 be the unique solution for the SDE

$$\mathrm{d}X_t^1 = \mathbf{1}_{\{t<\tau(X^1)\}}\big\{b_t(X_t^1,\mu_t^1)\mathrm{d}t + \sigma_t(X_t^1)\mathrm{d}W_t\big\}, \quad X_0^1 = X_0. \tag{4.9}$$

By the definition of Φ, we have

$$(\Phi\mu^1)_t = \mathscr{L}_{X_t^1}, \quad t\in[0,T]. \tag{4.10}$$

To construct $(\Phi\mu^2)_t$ using Girsanov's theorem, let

$$\xi_t := \mathbf{1}_{\{t<\tau(X^1)\}}\{\sigma_t^*(\sigma_t\sigma_t^*)^{-1}\}(X_t^1)\{b_t(X_t^1,\mu_t^2) - b_t(X_t^1,\mu_t^1)\}, \quad t\in[0,T].$$

By **(E)**, there exists a constant $k>0$ such that

$$|\xi_t| \le k\|\mu_t^1 - \mu_t^2\|_V, \quad t\in[0,T]. \tag{4.11}$$

So, by Girsanov's theorem,

$$\tilde W_t := W_t - \int_0^t \xi_s\mathrm{d}s, \quad t\in[0,T]$$

is an m-dimensional Brownian motion under the probability measure $\mathbb{Q} := R_T\mathbb{P}$, where

$$R_s := \mathrm{e}^{\int_0^s\langle\xi_t,\mathrm{d}W_t\rangle - \frac{1}{2}\int_0^s|\xi_t|^2\mathrm{d}t}, \quad s\in[0,T].$$

Reformulate (4.9) as

$$\mathrm{d}X_t^1 = \mathbf{1}_{\{t<\tau(X^1)\}}\big\{b_t(X_t^1,\mu_t^2)\mathrm{d}t + \sigma_t(X_t^1)\mathrm{d}\tilde W_t\big\}, \quad X_0^1 = \tilde X_0.$$

By the weak uniqueness of (4.6), we obtain

$$(\Phi\mu^2)_t = \mathbb{Q}(X^1_{t\wedge\tau(X^1)} \in \mathrm{d}x) = \mathscr{L}_{X^1_t|\mathbb{Q}}.$$

Combining this with (4.8) and (4.10), we derive

$$\begin{aligned}\|(\Phi\mu^1)_t - (\Phi\mu^2)_t\|_V &\le \mathbb{E}\big[V(X^1_t)|R_t - 1|\big]\\ &\le \mathbb{E}\big[\{\mathbb{E}(V(X^1_t)^2|\mathscr{F}_0)\}^{\frac{1}{2}}\{\mathbb{E}(|R_t-1|^2|\mathscr{F}_0)\}^{\frac{1}{2}}\big]\\ &\le c_3\mathbb{E}\big[V(X_0)\{\mathbb{E}(|R_t-1|^2|\mathscr{F}_0)\}^{\frac{1}{2}}\big].\end{aligned} \tag{4.12}$$

On the other hand, by $\mu^1, \mu^2 \in \mathscr{C}^\gamma_{V,N}(\bar{O})$, (4.11), and noting that $\mathrm{e}^r - 1 \le r\mathrm{e}^r$ for $r \ge 0$, we find a constant $c > 0$ such that

$$\begin{aligned}&\mathbb{E}[|R_t - 1|^2|\mathscr{F}_0] = \mathbb{E}[\mathrm{e}^{2\int_0^t \langle \xi_s, \mathrm{d}W_s\rangle - \int_0^t |\xi_s|^2\mathrm{d}s} - 1|\mathscr{F}_0]\\ &\le \mathbb{E}\big[\mathrm{e}^{2\int_0^t \langle \xi_s, \mathrm{d}W_s\rangle - 2\int_0^t |\xi_s|^2\mathrm{d}t}\big|\mathscr{F}_0\big]\mathrm{e}^{k^2\int_0^t \|\mu^1_s - \mu^2_s\|_V^2\mathrm{d}s} - 1\\ &= \mathrm{e}^{k^2\int_0^t \|\mu^1_s - \mu^2_s\|_V^2\mathrm{d}s} - 1 \le \mathrm{e}^{k^2\int_0^t \|\mu^1_s - \mu^2_s\|_V^2\mathrm{d}s}\int_0^t k^2\|\mu^1_s - \mu^2_s\|_V^2\mathrm{d}s\\ &\le c^2\int_0^t \|\mu^1_s - \mu^2_s\|_V^2\mathrm{d}s, \quad t\in[0,T].\end{aligned}$$

Combining this with (4.12) and letting $C = cc_3\,\mathbb{E}[V(X_0)]$, we arrive at

$$\begin{aligned}\rho_\lambda(\Phi(\mu^1), \Phi(\mu^2)) &\le C\sup_{t\in[0,T]}\mathrm{e}^{-\lambda t}\bigg(\int_0^t \|\mu^1_s - \mu^2_s\|_V^2\mathrm{d}s\bigg)^{\frac{1}{2}}\\ &\le C\rho_\lambda(\mu^1,\mu^2)\bigg(\int_0^t \mathrm{e}^{-2\lambda(t-s)}\mathrm{d}s\bigg)^{\frac{1}{2}}.\end{aligned}$$

Thus, when $\lambda > 0$ is large enough, Φ is contractive in ρ_λ and hence has a unique fixed point in $C^\gamma_{V,N}(\bar{O})$.

(3) Uniqueness and (4.4). It is easy to see that for any (weak) solution X_t of (4.5) for distributions in $\mathscr{P}^V(\bar{O})$, $\mu_t := \mathscr{L}_{X_t}$ is a fixed point of Φ in $\mathscr{C}^\gamma_V(\bar{O})$. Since Φ has a unique fixed point, this implies the (weak) uniqueness of (4.1). Finally, by Gronwall's inequality, (4.4) follows from (4.8) for $X^\mu_t = X_t$ and $\mu_t := \mathscr{L}_{X_t}$, where μ is the unique fixed point of Φ. □

Acknowledgement. The author would like to thank the referees for helpful comments and corrections.

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,* Stochastic Process. Appl. 146(2022), 264–310.

[2] V. Barbu, M. Röckner, *From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE,* Ann. Probab. 48(2020),1902–1920.

[3] I. V. Evstigneev, *Measurable selection theorem and probabilistic control models in general topological spaces,* Mathematics of the USSR-Sbornik, 59(1988), 25–37.

[4] X. Huang, P. Ren, F.-Y. Wang, *Distribution dependent stochastic differential equations,* Front. Math. China 16(2021), 257–301.

[5] X. Huang, F.-Y. Wang, *Singular McKean-Vlasov (reflecting) SDEs with distribution dependent noise,* J. Math. Anal. Appl. 514(2022), 126301 21pp.

[6] J. P. McKean, *A class of Markov processes associated with nonlinear parabolic equations,* Proc. Nat. Acad. Sci. U.S.A. 56(1966), 1907–1911.

[7] P. Ren, *Singular McKean-Vlasov SDEs: Well-posedness, regularities and Wang's Hanrack inequality,* to appear in Stoch. Proc. Appl.

[8] A.-S. Sznitman, *Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated,* J. Funct. Anal. 56(1984), 311–336.

[9] A.-S. Sznitman, *Topics in propagations of chaos,* Lecture notes in Math. Vol. 1464, pp. 165–251, Springer, Berlin, 1991.

[10] F.-Y. Wang, *Distribution dependent reflecting stochastic differential equations,* to appear in Sci. China Math. arXiv:2106.12737.

[11] F.-Y. Wang, *Regularity estimates and intrinsic-Lions derivative formula for singular McKean-Vlasov SDEs,* arXiv:2109.02030.

[12] P. Xia, L. Xie, X. Zhang, G. Zhao, *$L^q(L^p)$-theory of stochastic differential equations,* Stoch. Proc. Appl. 130(2020), 5188–5211.

[13] C. Yuan, S.-Q. Zhang, *A study on Zvonkin's transformation for stochastic differential equations with singular drift and related applications,* J. Diff. Equat. 297(2021), 277–319.

[14] Y. Zhao, *Distribution dependent absorbing stochastic differential equations,* Master Thesis (2022), Tianjin University.