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
\[
\mathcal{P}_O := \{ \mu \text{ is a measure on } O, \mu(O) \leq 1 \}
\]
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 \( O \):
\[
L_{t,\mu} := \text{tr}\{(\sigma_t \sigma_t^*)(\cdot,\mu)\nabla^2\} + \nabla b_t(\cdot,\mu), \quad t \in [0,T], \mu \in \mathcal{P}_O,
\]
where \( T > 0 \) is a fixed constant, \( \sigma^* \) is the transposition of \( \sigma \), \( \nabla^2 \) is the Hessian operator, \( \nabla b := b \cdot \nabla \) is the derivative along \( b \), and for some \( m \in \mathbb{N} \),
\[
b : [0,T] \times O \times \mathcal{P}_O \to \mathbb{R}^d, \quad \sigma : [0,T] \times O \times \mathcal{P}_O \to \mathbb{R}^d \otimes \mathbb{R}^m
\]
are measurable such that
\[
(1.1) \quad \int_0^T dt \int_O \{ |b_t(x,\mu_t)| + \|\sigma_t(x,\mu_t)\|^2 \} \mu_t(dx) < \infty, \quad \mu = (\mu_t)_{t \in [0,T]} \in C([0,T];\mathcal{P}_O).
\]

To introduce the nonlinear Dirichlet problem for \( L_{t,\mu} \) on \( \mathcal{P}_O \), let \( C^2_D(O) \) be the class of \( f \in C^2_b(O) \) with Dirichlet condition \( f|_{\partial O} = 0 \), where \( f \in C^2_b(O) \) means that \( f \) is a bounded \( C^2 \) function on \( O \) with bounded first and second order derivatives. For any \( t \in [0,T] \) and \( \mu,\nu \in \mathcal{P}_O \) such that
\[
\int_O \{ |b_t(x,\mu)| + \|\sigma_t(x,\mu)\|^2 \} \nu(dx) < \infty,
\]
define the linear functional on \( C^2_D(O) \):
\[
L^D_{t,\mu}^* \nu : C^2_D(O) \ni f \mapsto (L^D_{t,\mu}^* \nu)(f) := \int_O L_{t,\mu} f d\nu \in \mathbb{R}.
\]
The corresponding nonlinear Dirichlet problem for \( L_{t,\mu} \) is the equation
\[
(1.2) \quad \partial_t \mu_t = L^D_{t,\mu} \mu_t, \quad t \in [0,T]
\]
for \( \mu : [0,T] \to \mathcal{P}_O \). We call \( \mu \in C([0,T];\mathcal{P}_O) \) a solution to (1.2), if
\[
\mu_t(f) = \mu_0(f) + \int_0^t \mu_s(L_{s,\mu_s} f) ds, \quad t \in [0,T], f \in C^2_D(O),
\]
where \( \mu(f) := \int f d\mu \) for a measure \( \mu \) and \( f \in L^1(\mu) \).

When \( \mu_t(dx) = \rho_t(x) dx \), (1.2) reduces to the nonlinear Dirichlet problem
\[
\partial_t \rho_t = L^D_{t,\rho_t} \rho_t, \quad t \in [0,T],
\]
where \( L_{t,\rho_t} := L_{t,\rho_t(x)}dx \), in the sense that

\[
\int_{O} (f \rho_t)(x)dx = \int_{O} (f \rho_0)(x)dx + \int_{0}^{t} ds \int_{O} (\rho_s L_s \rho_s f)(x)dx, \quad t \in [0, T], f \in C^2(D)(O).
\]

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

\[
(1.3) \quad dX_t = 1_{\{t < \tau(X)\}} \{ b_t(X_t, \mathcal{L}^O_{X_t})dt + \sigma_t(X_t, \mathcal{L}^O_{X_t})dW_t \}, \quad t \in [0, T],
\]

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

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

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

\[
\mathcal{L}^O_{\xi} := \mathbb{P}(\xi \in O \cap \cdot)
\]

is the distribution of \( \xi \) restricted to \( O \), which we call the \( O \)-distribution of \( \xi \). When different probability spaces are concerned, we denote \( \mathcal{L}^O_{\xi} \) by \( \mathcal{L}^O_{\xi|\mathcal{P}} \) to emphasize the reference probability measure.

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

\[
\int_{0}^{T \wedge \tau(X)} \left\{ \|b_t(X_t, \mathcal{L}^O_{X_t})\| + \|\sigma_t(X_t, \mathcal{L}^O_{X_t})\|^2 \right\} dt < \infty
\]

and

\[
X_t = X_0 + \int_{0}^{t \wedge \tau(X)} \{ b_s(X_s, \mathcal{L}^O_{X_s})ds + \sigma_s(X_s, \mathcal{L}^O_{X_s})dW_s \}, \quad t \in [0, T].
\]

We call \((\tilde{X}_t, \tilde{W}_t)\) a weak solution to (1.3), if there exists a complete filtration probability space \((\tilde{\Omega}, \{\tilde{\mathcal{F}}_t\}_{t \in [0,T]}, \tilde{\mathbb{P}})\) 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), \( \mu_t := \mathcal{L}^O_{X_t} \) solves the nonlinear Dirichlet problem (1.2). Indeed, since \( dX_t = 0 \) for \( t \geq \tau(X) \), we have

\[
X_t = X_{t \wedge \tau(X)}, \quad \mathcal{L}^O_{X_t}(dx) = \mathbb{P}(t < \tau(X), X_t \in dx), \quad t \in [0, T].
\]

By this and Itô’s formula, for any \( f \in C^2_D(O) \) we have

\[
\mu_t(f) = \mathbb{E}[1_{O}f(X_t)] = \mathbb{E}[f(X_t)]
\]

\[
= \mathbb{E}[f(X_0)] + \mathbb{E} \int_{0}^{t} 1_{\{s < \tau(X)\}} L_s \mu_s f(X_s)ds
\]
\[ = \mu_0(f) + \int_0^t \mu_s(L_{s,\mu_0} f) \, ds, \quad t \in [0, T]. \]

(2) An alternative model to (1.3) is

\[ dX_t = 1_O(X_t) \left\{ b_t(X_t, \mathcal{L}_{X_t}^O) \, dt + \sigma_t(X_t, \mathcal{L}_{X_t}^O) \, dW_t \right\}, \quad t \in [0, T]. \]

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 \( \sigma_t(x, \mu) = 2x, b_t(x, \mu) = 2\sqrt{\mu} \). Let \( Y_t \) solve the SDE

\[ dY_t = Y_t \, dW_t + \left( 1 - \frac{1}{2} Y_t \right) \, dt, \quad Y_0 = 0. \]

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 \( \mu \).

(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 \( \mathcal{P} \) be the space of probability measures on \( \mathbb{R}^d \), and define

\[ \tilde{b}_t(x, \mu) := 1_O(x) b_t(x, \mu(O \cap \cdot)), \quad \tilde{\sigma}_t(x, \mu) := 1_O(x) \sigma_t(x, \mu(O \cap \cdot)) \]

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

\[ dX_t = \tilde{b}_t(X_t, \mathcal{L}_{X_t}) \, dt + \tilde{\sigma}_t(X_t, \mathcal{L}_{X_t}) \, dW_t, \quad t \in [0, T]. \]

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

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

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

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

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

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

We will study the well-posedness under the following assumption.
(H) For any $\mu \in C([0, T]; \hat{P}_O)$, the killed SDE
\[
(1.5) \quad dX^\mu_t = 1_{\{t < \tau(X^\mu)\}} \left\{ b_t(X^\mu_t, \mu_t)dt + \sigma_t(X^\mu_t, \mu_t)dW_t \right\}, \quad t \in [0, T]
\]
is well-posed for initial value $X^\mu_0$ with $L^O_{X^\mu_0} = \mu_0$, and $L^O_{X^\mu} \in C([0, T]; \hat{P}_O)$.

Under this assumption, we define a map
\[
(1.6) \quad C([0, T]; \hat{P}_O) \ni \mu \mapsto \Phi \mu := L^O_{X^\mu} := (L^O_{X^\mu_t})_{t \in [0, T]} \in C([0, T]; \hat{P}_O).
\]

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

**Theorem 1.1.** Assume (H). If for any $\gamma \in \hat{P}_O$, $\Phi$ has a unique fixed point in
\[
\mathcal{C}^\gamma := \{\mu \in C([0, T]; \hat{P}_O), \mu_0 = \gamma\},
\]
then (1.3) is well-posed for $O$-distributions in $\hat{P}_O$.

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:

\[
\hat{W}_1(\mu, \nu) := \inf_{\pi \in \mathcal{C}_O(\mu, \nu)} \int_{\hat{O} \times \hat{O}} |x - y| \pi(dx, dy),
\]
\[
\hat{W}_1(\mu, \nu) := \inf_{\pi \in \mathcal{C}_O(\mu, \nu)} \int_{\hat{O} \times \hat{O}} (1 \wedge |x - y|) \pi(dx, dy), \quad \mu, \nu \in \mathcal{P}_O,
\]

where $\pi \in \mathcal{C}_O(\mu, \nu)$ means that $\pi$ is a probability measure on $\hat{O} \times \hat{O}$ such that
\[
\pi(\{\cdot \cap O\} \times \hat{O}) = \mu, \quad \pi(\hat{O} \times \{\cdot \cap O\}) = \nu.
\]
2.1 Monotonicity in $\hat{W}_1$

(A1) For any $\mu \in C([0, T]; \mathcal{P}_O)$, $b_t(x, \mu_t)$ and $\sigma_t(x, \mu_t)$ are continuous in $x \in O$ such that for any $N \geq 1$ and $O_N := \{x \in O : |x| \leq N\}$,

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

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

$$2\langle b_t(x, \mu) - b_t(y, \nu), x - y \rangle + \|\sigma_t(x, \mu) - \sigma_t(y, \nu)\|^2_{HS} \leq K(t)\{|x - y|^2 + \hat{W}_1(\mu, \nu)^2\},$$

$$2\langle b_t(x, \mu), x \rangle + \|\sigma_t(x, \mu)\|^2_{HS} \leq K(t)(1 + |x|^2), \quad t \in [0, T].$$

(A2) There exists $r_0 \in (0, 1]$ such that the distance function $\rho_0$ to $\partial O$ is $C^2$-smooth in

$$\partial_\rho O := \{x \in \bar{O} : \rho_0(x) \leq r_0\},$$

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

$$|\sigma_t(x, \mu)^* \nabla \rho_0(x)|^{-2} \leq \alpha, \quad L_{i, \mu} \rho_0(x) \leq \alpha, \quad x \in \partial_\rho O, t \in [0, T].$$

**Theorem 2.1.** Assume (A1) and (A2). Then the following assertions hold.

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

$$\mathbb{E}\left[\sup_{t \in [0, T]} |X_t|^p \mid \mathcal{F}_0\right] \leq c(1 + |X_0|^p).$$

2. There exists a constant $c > 0$ such that

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

Under assumption (A1), for any $\mu \in C([0, T]; \mathcal{P}_O)$ 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 $\mu$ such that

$$\mathbb{E}\left[\sup_{t \in [0, T]} |X^\mu_{t \wedge \bar{\tau}}|^p \mid \mathcal{F}_0\right] = \mathbb{E}\left[\sup_{t \in [0, T]} |X^\mu_{t \wedge (\tau(X^\mu) \wedge \bar{\tau})}|^p \mid \mathcal{F}_0\right] \leq c(1 + |X_0^\mu|^p)$$

holds for any solution $X^\mu_t$ of (1.5) and any stopping time $\bar{\tau}$.

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

$$\Phi \mu := L^{O}_X = (L^{O}_X)^{t \in [0, T]}, \quad \mu \in C([0, T]; \mathcal{P}_O)$$
has a unique fixed point in
\[ C^\gamma := \{ \mu \in C([0, T]; \mathcal{P}_O) : \mu_0 = \gamma \}. \]

To this end, for \( i = 1, 2 \), let \( \mu^i \in C([0, T]; \mathcal{P}_O) \), and let \( X^i_t \) solve (2.1) for \( \mu^i \) replacing \( \mu \) with \( \mathcal{L}^O_{X^i_0} = \mu^i_0 \), i.e.
\[(2.5) \quad dX^i_t = 1_{\{t < \tau(X^i)\}} \{ b_t(X^i_t, \mu^i_t)dt + \sigma_t(X^i_t, \mu^i_t)dW^i_t \}, \quad t \in [0, T], \mathcal{L}^O_{X^i_0} = \mu^i_0. \]

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

Since
\[ \Gamma := \{(x, y) : x \in O, y \in \partial O, |x - y| = \rho_0(x)\} \]
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 \( P_\theta : O \rightarrow \partial O \) such that
\[(2.6) \quad |P_\theta x - x| = \rho_\theta(x), \quad x \in O. \]

We will use the following coupling by projection.

**Definition 2.1.** The coupling by projection \((\bar{X}^1_t, \bar{X}^2_t)\) for \((X^1_t, X^2_t) = (X^1_{t \wedge \tau_1}, X^2_{t \wedge \tau_2})\) is defined as
\[(2.7) \quad (X^1_t, X^2_t) := \begin{cases} (X^1_t, X^2_t), & \text{if } t \leq \tau_{1,2}, \\ (X^1_t, P_\theta X^1_t), & \text{if } \tau_2 < t \wedge \tau_1, \\ (P_\theta X^2_t, X^2_t), & \text{otherwise.} \end{cases} \]

It is easy to see that \( \mathcal{L}^O_{X^i_1} = \mathcal{L}^O_{X^i_2} = (\Phi \mu^i)_t \) for \( i = 1, 2 \); i.e. the distribution \( \mathcal{L}((\Phi \mu^1)_t, (\Phi \mu^2)_t) \) of the coupling by projection \((\bar{X}^1_t, \bar{X}^2_t)\) satisfies
\[ \mathcal{L}((\Phi \mu^1)_t, (\Phi \mu^2)_t) \in \mathcal{C}_O((\Phi \mu^1)_t, (\Phi \mu^2)_t). \]

Thus, by (2.1) and Definition 2.1
\[(2.8) \quad \hat{W}_1((\Phi \mu^1)_t, (\Phi \mu^2)_t) \leq E[1 \wedge |\bar{X}^1_t - \bar{X}^2_t|] \leq E[1 \wedge |X^1_{t \wedge \tau_1, 2} - X^2_{t \wedge \tau_1, 2}|] + r_0^{-1}E[\{r_0 \wedge \rho_0(X^1_t)\}1_{(t \wedge \tau_1 \geq \tau_2)}] + r_0^{-1}E[\{r_0 \wedge \rho_0(X^2_t)\}1_{(t \wedge \tau_2 \geq \tau_1)}]. \]

**Lemma 2.2.** Assume \((A_1)\). Then there exists a constant \( c > 1 \) such that for any \( t \in [0, T] \) and \( \mu^1, \mu^2 \in C([0, T]; \mathcal{P}_O), \)
\[(2.9) \quad E[|X^1_{t \wedge \tau_{1,2}} - X^2_{t \wedge \tau_{1,2}}|^2 | \mathcal{F}_0] \leq c |X^1_0 - X^2_0|^2 + c \int_0^t K(s) \hat{W}_1((\mu^1_s, \mu^2_s)^2) ds. \]

Consequently, for any \( t \in [0, T], \)
\[(2.10) \quad E[1 \wedge |X^1_{t \wedge \tau_{1,2}} - X^2_{t \wedge \tau_{1,2}}|] \leq \sqrt{c}E[1 \wedge |X^1_0 - X^2_0|] + \left( c \int_0^t K(s) \hat{W}_1((\mu^1_s, \mu^2_s)^2) ds \right)^{\frac{1}{2}}. \]
Proof. It suffices to prove (2.9), which implies (2.10) due to Jensen’s inequality.

By (A1) and Itô’s formula, we obtain

\[
d|X_t^1 - X_t^2|^2 \leq K(t) \{ |X_t^1 - X_t^2|^2 + \tilde{W}_t(\mu_1^2, \mu_2^2) \} dt + dM_t, \quad t \in [0, T \land \tau_{1,2}]
\]

for some local martingale $M_t$. This and (2.3) imply that

\[
\beta_t := \mathbb{E}[|X_{t \land \tau_{1,2}}^1 - X_{t \land \tau_{1,2}}^2|^2 | \mathcal{F}_0]
\]

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

\[
\beta_t \leq \beta_0 + \int_0^t K(s) \{ \beta_s + \tilde{W}_s(\mu_1^2, \mu_2^2) \} ds, \quad t \in [0, T].
\]

By Gronwall’s inequality, we prove (2.9). \qed

Lemma 2.3. Assume (A2). Then there exists a constant $c > 1$ independent of $\mu$ such that for any solution $X_t^{\mu}$ to (1.5) and any stopping time $\tau$,

\[
1_{\{t \land \tau(X^{\mu}) \geq \hat{\tau}\}} \mathbb{E}[r_0 \land \rho_\theta(X_t^{\mu}) | \mathcal{F}_{\hat{\tau}}] \leq c 1_{\{t \land \tau(X^{\mu}) \geq \hat{\tau}\}} \rho_\theta(X_t^{\mu}), \quad t \in [0, T].
\]

Proof. By the strong Markov property of $X_t^{\mu}$ which is implied by the well-posedness of (1.5), we may and do assume that $\hat{\tau} = 0$ and $x = X^\mu_0 \in O$, such that the desired estimate becomes

\[
\mathbb{E}^x[r_0 \land \rho_\theta(X_t^{\mu})] \leq c \rho_\theta(x), \quad t \in [0, T],
\]

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

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

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

By (A2),

\[
dh(\rho_\theta(X_t^{\mu})) \leq \alpha dt + dM_t, \quad t \in [0, T \land \tau(X^{\mu})],
\]

where $M_t$ is a martingale with

\[
d\langle M \rangle_t \geq \alpha^{-1} dt, \quad t \leq \hat{\tau} := \inf\{ t \geq 0 : \rho_\theta(X_t^{\mu}) \geq r_0/2 \}.
\]

By (2.12) we obtain

\[
\mathbb{E}^x[r_0 \land \rho_\theta(X_t^{\mu})] \leq 2\mathbb{E}^x[h(\rho_\theta(X_{t \land \tau(X^{\mu})}^{\mu}))] \leq 2\rho_\theta(x) + 2\alpha \mathbb{E}^x[t \land \tau(X^{\mu})].
\]

On the other hand, let

\[
\eta_t := \int_0^{\rho_\theta(X_t^{\mu})} e^{-2\alpha s} ds \int_s^t e^{2\alpha \theta} d\theta, \quad t \in [0, T \land \tau(X^{\mu}) \land \hat{\tau}].
\]
Since \( h(r) = r \) for \( r \leq \frac{m}{2} \), by (2.12), (2.13) and Itô’s formula, we find a martingale \( \hat{M}_t \) such that
\[
d\eta_t \leq -dt + d\hat{M}_t, \quad t \in [0, T \wedge \tau(X^\mu) \wedge \hat{\tau}].
\]
Consequently,
\[
\E^x[t \wedge \tau(X^\mu) \wedge \hat{\tau}] \leq \eta_0 \leq c_1 \rho_\theta(x)
\]
holds for some constant \( c_1 > 0 \). Therefore,
\[
\E^x[t \wedge \tau(X^\mu)] \leq \E^x[t \wedge \tau(X^\mu) \wedge \hat{\tau}] + T \E^x[1_{\{t \wedge \tau(X^\mu) \wedge \hat{\tau} \}}]
\]
\[
\leq c_1 \rho_\theta(x) + T \E^x(t \wedge \tau(X^\mu) > \hat{\tau}), \quad t \in [0, T].
\]
To estimate the second term, let
\[
\xi_t := \int_0^{\rho_\theta(X^\mu)} e^{-2\alpha^2 s} ds, \quad t \in [0, T \wedge \tau(X^\mu) \wedge \hat{\tau}].
\]
By \( h(r) = r \) for \( r \leq \frac{m}{2} \), (2.12), (2.13) and Itô’s formula, we see that \( \xi_t \) is a sup-martingale, so that
\[
\rho_\theta(x) \geq \xi_0 \geq \E^x[\xi_{t \wedge \tau(X^\mu) \wedge \hat{\tau}}] \geq \P^x(t \wedge \tau(X^\mu) \geq \hat{\tau}) \int_0^{r_0/2} e^{-2\alpha^2 s} ds.
\]
Combining this with (2.14) and (2.16), we prove (2.11) for some constant \( c > 0 \).

**Proof of Theorem 2.7.** (a) Well-posedness. Let \( \gamma := \mathcal{L}_X \), and consider
\[
\mathcal{C}_\gamma := \{ \mu \in C([0, T]; \mathcal{P}_0) : \mu_0 = \gamma \}.
\]
We intend to prove that \( \Phi \) is contractive in \( \mathcal{C}_\gamma \) under the complete metric
\[
\mathcal{W}_1,\theta(a, b) := \sup_{t \in [0, T]} e^{-\theta t} W_1(a, b)
\]
for large enough \( \theta > 0 \). Then \( \Phi \) has a unique fixed point in \( \mathcal{C}_\gamma \), so that the well-posedness follows from Theorem 1.1.

To this end, let \( \mu^i \in \mathcal{C}_\gamma \) and let \( X^i_t \) solve (1.5) with \( \mu = \mu^i \) and \( X^i_0 = X_0, i = 1, 2 \). By \( r_0 \leq 1 \), Lemma 2.3 and noting that
\[
1_{\{t \wedge \tau \geq \tau_1\}} \rho_\theta(X^2_{t \wedge \tau_1, 2}) \leq 1_{\{t \wedge \tau \geq \tau_1\}} |X^2_{t \wedge \tau_1, 2} - X^1_{t \wedge \tau_1, 2}|,
\]
we obtain
\[
\E \left[ 1_{\{t \wedge \tau \geq \tau_1\}} \{r_0 \wedge \rho_\theta(X^2_{t \wedge \tau_1, 2})\} \right] = \E \left( 1_{\{t \wedge \tau \geq \tau_1\}} \E \left[ \{r_0 \wedge \rho_\theta(X^2_{t \wedge \tau, 2})\} | \mathcal{F}_{\tau_1} \right] \right)
\]
\[
\leq c \E \left[ 1_{\{t \wedge \tau \geq \tau_1\}} \{r_0 \wedge \rho_\theta(X^2_{t \wedge \tau_1, 2})\} \right] \leq c \E \left[ 1 \wedge |X^1_{t \wedge \tau_1, 2} - X^2_{t \wedge \tau_1, 2}| \right].
\]
By symmetry, the same estimate holds for $E \left[ 1_{\{t \wedge \tau_2 \geq \tau_2\}} \{r_0 \wedge \rho_0 (X_{t \wedge \tau_2}) \} \right]$. 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

$$\tilde{W}_1((\Phi \mu^1)_t, (\Phi \mu^2)_t) \leq c_1 \left( \int_0^t K(s) \tilde{W}_1(\mu^1_s, \mu^2_s)^2 \, ds \right)^{\frac{1}{2}}, \quad t \in [0, T].$$

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

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

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

$$\mathcal{L}_{X_0}^O = \mu, \quad \mathcal{L}_{X_0}^O = \nu, \quad E[1 \wedge |X_0^1 - X_0^2|] = \tilde{W}_1(\mu, \nu).$$

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

$$\mu_t^1 := \mathcal{L}_{X_t^1}^O = P_{t}^{D_s} \mu, \quad \mu_t^2 := \mathcal{L}_{X_t^2}^O = P_{t}^{D_s} \nu,$$

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

$$\tilde{W}_1(P_{t}^{D_s} \mu, P_{t}^{D_s} \nu) = \tilde{W}_1((\Phi \mu^1)_t, (\Phi \mu^2)_t) \leq c_2 \tilde{W}_1(\mu, \nu) + \left( c_2 \int_0^t K(s) \tilde{W}_1(\mu_s^{D_s}, \nu_s^{D_s})^2 \, ds \right)^{\frac{1}{2}}, \quad t \in [0, T].$$

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

### 2.2 Monotonicity in $W_1$

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

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

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

$$\int_0^T \sup_{O_N} \{ |b_i(\cdot, \mu_t)| + \| \sigma_i(\cdot, \mu_t) \|_2 \} \, dt < \infty.$$

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

$$2\langle b_i(x, \mu) - b_i(y, \nu), x - y \rangle + \| \sigma_i(x, \mu) - \sigma_i(y, \nu) \|_H^2 \leq K(t) \{ |x - y|^2 + \tilde{W}_1(\mu, \nu)^2 \},$$

$$2\langle b_i(x, \mu), x \rangle + \| \sigma_i(x, \mu) \|_H^2 \leq K(t) \{ 1 + |x|^2 + \| \mu \|_1^2 \}, \quad t \in [0, T].$$
(B2) There exists \( r_0 > 0 \) such that \( \rho_0 \in C^2(\partial r_0 O) \), and there exists an increasing function \( \alpha : [0, \infty) \to [1, \infty) \) such that

\[
(2.21) \quad |\sigma_t(x, \mu)^* \nabla \rho_0|^2 \leq \alpha(\|\mu\|_1), \quad L_{t, \mu} \rho_0(x) \leq \alpha(\|\mu\|_1), \quad x \in \partial r_0 O,
\]

\[
(2.22) \quad 2\langle b_t(x, \mu), x - y \rangle + \|\sigma_t(x, \mu)\|^2_{HS} \leq K(t)\alpha(\|\mu\|_1)(1 + |x - y|^2), \quad t \in [0, T], y \in \partial O, x \in O.
\]

**Theorem 2.4.** Assume (B1) and (B2). Then the following assertions hold.

1. (1.3) is well-posed for \( O \)-distributions in \( \mathcal{P}_O^1 \). 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 \( \mathcal{P}_O^1 \),

\[
(2.23) \quad \mathbb{E}\left[ \sup_{t \in [0, T]} |X_t|^p \big| \mathcal{F}_0 \right] \leq c\left( 1 + |X_0| + \mathbb{E}[1_O(X_0)|X_0]| \right)^p.
\]

2. If \( \alpha \) is bounded, then there exists a constant \( c > 0 \) such that

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

It is standard that (B1) and (B2) imply the well-posedness of (1.5) for \( \mu \in C([0, T]; \mathcal{P}_O^1) \), and instead of (2.4), for any \( p \geq 1 \) there exists a constant \( c > 0 \) such that

\[
(2.25) \quad \mathbb{E}\left[ \sup_{t \in [0, T]} |X_t|^p \big| \mathcal{F}_0 \right] \leq c\left( 1 + |X_0|^p \right) + c \int_0^t K(s)\|\mu_s\|^p ds, \quad t \in [0, T], \mu \in C([0, T]; \mathcal{P}_O^1).
\]

Let \( i \in \mathbb{C}(0, T]; \mathcal{P}_O^1) \), \( i = 1, 2 \), let \( X_t^i \) solve (1.5) for \( \mu^i \) replacing \( \mu \) with \( \mathcal{L}O^1 \) for \( \mu^i \), and denote as before

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

Using (B1) replacing (A1), the proof of (2.9) leads to

\[
(2.26) \quad \mathbb{E} \left[ |X_{t \land \tau_1,2}^1 - X_{t \land \tau_1,2}^2|^2 \big| \mathcal{F}_0 \right] \leq c|X_0^1 - X_0^2|^2 + c \int_0^t K(s)\mathcal{W}_1(\mu_s^1, \mu_s^2)^2 ds, \quad t \in [0, T],
\]

and instead of (2.8), we have

\[
(2.27) \quad \mathcal{W}_1((\Phi \mu_1)_t, (\Phi \mu_2)_t) \leq \mathbb{E} \left[ |X_{t}^1 - X_{t}^2|^2 \right] \leq \mathbb{E} \left[ |X_{t \land \tau_1,2}^1 - X_{t \land \tau_1,2}^2|^2 \right] + \mathbb{E} \left[ \rho_0(X_t^1)1_{\{t \land \tau_1 \geq \tau_2\}} + \rho_0(X_t^2)1_{\{t \land \tau_2 \geq \tau_1\}} \right].
\]

The following lemma is analogous to Lemma 2.3.

**Lemma 2.5.** Assume (B2). Then there exists an increasing function \( \psi : [0, \infty) \to (0, \infty) \) which is bounded if so is \( \alpha \), such that for any \( \mu \in C([0, T]; \mathcal{P}_O^1) \) and any solution \( X_t^\mu \) to (1.5) and any stopping time \( \bar{\tau} \),

\[
1_{\{t \land \tau(X^\mu) \geq \bar{\tau}\}} \mathbb{E} \left[ \rho_0(X_t^\mu) \big| \mathcal{F}_{\bar{\tau}} \right] \leq 1_{\{t \land \tau(X^\mu) \geq \bar{\tau}\}} \psi(\|\mu\|_{1, T}) \rho_0(X_{\bar{\tau}}^\mu).
\]

11
Proof. By the strong Markov property, we may assume that $\tilde{\tau} = 0$ and $x = X^\mu_0 \in O$, so that it suffices to prove

$$\Gamma_t(x) := \mathbb{E}^x[\rho_\theta(X^\mu_t)] \leq \psi(\|\mu\|_{1,T})\rho_\theta(x), \quad x \in O, t \in [0,T].$$

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

$$d|X^\mu_t - y|^2 \leq K(t)\alpha(\|\mu\|_{1,T})(1 + |X^\mu_t - y|^2)dt + dM_t, \quad t \in [0,T \land \tau(X^\mu)]$$

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

$$\mathbb{E}^x[|X^\mu_t - y|^2] \leq \rho_\theta(x)^2 + \alpha(\|\mu\|_{1,T}) \int_0^t K(s)ds$$

$$+ \int_0^t K(s)\alpha(\|\mu\|_{1,T})\mathbb{E}^x[|X^\mu_t - y|^2]ds, \quad t \in [0,T].$$

By Gronwall’s inequality and $\rho_\theta(x) \geq \frac{\alpha}{2}$, we find an increasing function $\psi_1 : [0, \infty) \to (0, \infty)$ which is bounded if so is $\alpha$, such that

$$\mathbb{E}^x[|X^\mu_t - y|^2] \leq \left\{ \rho_\theta(x)^2 + \alpha(\|\mu\|_{1,T}) \int_0^T K(s)ds \right\} e^{\alpha(\|\mu\|_{1,T}) \int_0^T K(s)ds}$$

$$\leq \left\{ \psi_1(\|\mu\|_{1,T})\rho_\theta(x) \right\}^2, \quad t \in [0,T].$$

Combining this with Jensen’s inequality, we prove (2.28) with $\psi = \psi_1$ holds for $\rho_\theta(x) \geq \frac{\alpha}{2}$.

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

$$\tilde{\tau} := \inf\{t \geq 0 : \rho_\theta(X^\mu_t) \geq r_0\}. $$

By (B2) and Itô’s formula, we obtain

$$d\rho_\theta(X^\mu_t) \leq \alpha dt + dM_t, \quad t \in [0,T \land \tau(X^\mu) \land \tilde{\tau}]$$

for some martingale satisfying (2.13). So,

$$\mathbb{E}^x[\rho_\theta(X^\mu_{t \land \tau(X^\mu) \land \tilde{\tau}})] \leq \alpha \mathbb{E}^x[t \land \tau(X^\mu) \land \tilde{\tau}].$$

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

$$\mathbb{E}^x[\rho_\theta(X^\mu_t)] = \mathbb{E}^x[\rho_\theta(X^\mu_{t \land \tau(X^\mu)})] \leq \mathbb{E}^x[\rho_\theta(X^\mu_{t \land \tau(X^\mu) \land \tilde{\tau}})] + \mathbb{E}^x\left[ \mathbf{1}_{\{t \land \tau(X^\mu) \geq \tilde{\tau}\}} \Gamma_{t \land \tilde{\tau}}(X^\mu_{\tilde{\tau}}) \right]$$

$$\leq \alpha \mathbb{E}^x[t \land \tau(X^\mu) \land \tilde{\tau}] + \mathbb{E}^x(\tilde{\tau} \land \tau(X^\mu) \geq \tilde{\tau})\psi_1(\|\mu\|_{1,T})r_0.$$

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 $\alpha$.  

\[ \square \]
Proof of Theorem 2.4. Let $X_t$ solve (1.3) for $O$-distributions in $\mathcal{P}_O^1$. Then $X_t = X_t^\mu$ for $\mu := \mathcal{L}_X^O$, so that
$$\|\mu_s\|_1 = \mathbb{E}[1_O(X_s)|X_s] = \mathbb{E}[1_{(t<\tau(X))}|X_s], \quad s \in [0, T].$$
Combining this with (2.25), we obtain
$$\|\mu_t\|^2 \leq \left(\mathbb{E}\sqrt{\mathbb{E}[1_{(t<\tau(X))}|X_t]^2|\mathcal{F}_0}\right)^2 \leq 2\left(\mathbb{E}\sqrt{c(1+1_O(X)|X_0|^2)}\right)^2 + 2c \int_0^t K(s)\|\mu_s\|^2 ds$$
$$\leq 2c(1+\mathbb{E}[1_O(X)|X_0|])^2 + 2c \int_0^t K(s)\|\mu_s\|^2 ds, \quad t \in [0, T].$$
By Gronwall’s inequality, we find a constant $c_1 > 0$ such that
$$\sup_{t \in [0,T]} \|\mu_t\|^2 \leq c_1(1 + \mathbb{E}[1_O(X)|X_0|]).$$
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 $\gamma := \mathcal{L}_X^O \in \mathcal{P}_O^1$. For any $N > 0$, let
(2.29) $\mathcal{C}_N^\gamma := \left\{ \mu \in C([0, T]; \mathcal{P}_O^1) : \mu_0 = \gamma, \sup_{t \in [0, T]} e^{-Nt}\|\mu_t\|_1 \leq N \right\}.$

We first observe that for some constant $N_0 > 0$,
(2.30) $\Phi\mathcal{C}_N^\gamma \subset \mathcal{C}_N^\gamma, \quad N \geq N_0.$

Let $\mu \in \mathcal{C}_N^\gamma$ and let $X_t^\mu$ solve (1.5) for $X_0^\mu = X_0$. Then $(\Phi\mu)_t = \mathcal{L}_X^O$. By (2.25) and
$$\|\Phi\mu_t\|_1 \leq \mathbb{E}\sqrt{\mathbb{E}[1_O(X)|X_{t\tau(X)}|^2|\mathcal{F}_0]}$$
we find a constant $c_1 > 0$ such that
$$\|\Phi\mu_t\|_1 \leq c_1(1 + \|\gamma\|_1) + c_1 \left( \int_0^t \|\mu_s\|^2 ds \right)^{\frac{1}{2}}, \quad t \in [0, T].$$
Then for any $N \geq N_0 := c_1 + 2c_1(1 + \|\gamma\|_1)$, we have
$$\sup_{t \in [0, T]} e^{-Nt}\|\Phi\mu_t\|_1 \leq c_1(1 + \|\gamma\|_1) + c_1 \sup_{t \in [0, T]} \left( \int_0^t e^{-2Ns}\|\mu_s\|^2 e^{-2N(t-s)} ds \right)^{\frac{1}{2}}$$
$$\leq c_1(1 + \|\gamma\|_1) + c_1 N \sup_{t \in [0, T]} \left( \int_0^t e^{-2N(t-s)} ds \right)^{\frac{1}{2}}$$
$$\leq c_1(1 + \|\gamma\|_1) + c_1 N \leq N.$$
Next, for any $N \geq N_0$, we intend to prove that $\Phi$ is contractive in $\mathcal{C}_N^\gamma$ under the complete metric
$$\bar{W}_{1,\theta}(\mu^1, \mu^2) := \sup_{t \in [0, T]} e^{-\theta t}\bar{W}_1(\mu^1_t, \mu^2_t)$$
for large enough $\theta > 0$, so that $\Phi$ has a unique fixed point in $C^\gamma = \bigcup_{N \geq N_0} C^\gamma_N$, hence the well-posedness follows from Theorem 1.1.

To this end, let $\mu^i \in C^\gamma_N$ and $X^i_t$ solve (1.3) for $\mu = \mu^i$ and $X^i_t = X_0, i = 1, 2$. By Lemma 2.5 and noting that $\rho_\theta(x) \leq |x - y|$ for $x \in O$ and $y \in \partial O$, we find a constant $c_2 > 0$ depending on $N$ such that for any $\mu^1, \mu^2 \in C^\gamma_N$,

\[
\mathbb{E}[\rho_\theta(X^1_t)1_{t \wedge \tau_2 \geq t} + \rho_\theta(X^2_t)1_{t \wedge \tau_2 \geq t}] \\
\leq c_2\mathbb{E}[\rho_\theta(X^1_{t \wedge \tau_2}, \mu^1)1_{t \wedge \tau_2 \geq t} + \rho_\theta(X^2_{t \wedge \tau_2}, \mu^2)1_{t \wedge \tau_2 \geq t}] \leq 2c_2\mathbb{E}[|X^1_{t \wedge \tau_2} - X^2_{t \wedge \tau_2}|].
\]

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

\[
\mathbb{W}_1(((\Phi\mu^1)_t, (\Phi\mu^2)_t) \leq c_3\mathbb{E}[|X^1_0 - X^2_0|] + c_3\left(\int_0^t K(s)\mathbb{W}_1(\mu^1_s, \mu^2_s)^2ds\right)^{\frac{1}{2}}, \ \mu^1, \mu^2 \in C^\gamma_N.
\]

Since $X^1_0 = X^2_0 = X_0$, this implies the contraction of $\Phi$ in $\mathbb{W}_{1, \theta}$ for large enough $\theta > 0$.

(b) Estimate (2.24). Now, for $\mu^1_0, \mu^2_0 \in \mathcal{B}_O$, let $X^1_0, X^2_0$ be $\mathcal{F}_0$-measurable random variables on $\bar{O}$ such that

\[
\mathcal{L}^0_{X^1_0} = \mu^1_0, \ \mathcal{L}^0_{X^2_0} = \mu^2_0, \ \mathbb{E}[|X^1_0 - X^2_0|] = \mathbb{W}_1(\mu^1_0, \mu^2_0).
\]

Letting $X^i_t$ solve (1.3) with initial value $X^i_0$, then $\mu^i := (P^D_x\mu^i_0)_{t \in [0, T]}$ is the unique fixed point of $\Phi$ in $C^\mu$, so that

\[
\mu^i = \mathcal{L}^0_{X^i_t} = \Phi\mu^i_t = P^D_x\mu^i_0, \ \ i = 1, 2, t \in [0, T].
\]

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

\[
\mathbb{W}_1(\mu^1_t, \mu^2_t) = \mathbb{W}_1(((\Phi\mu^1)_t, (\Phi\mu^2)_t) \leq c_3\mathbb{E}[|X^1_0 - X^2_0|] + c_3\left(\int_0^t K(s)\mathbb{W}_1(\mu^1_s, \mu^2_s)^2ds\right)^{\frac{1}{2}}
\]

\[
= c_3\mathbb{W}_1(\mu^1_0, \mu^2_0) + c_3\left(\int_0^t K(s)\mathbb{W}_1(\mu^1_s, \mu^2_s)^2ds\right)^{\frac{1}{2}}, \ t \in [0, T].
\]

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

\[
\mathbb{W}_1(P^D_x\mu^1_0, P^D_x\mu^2_0)^2 = \mathbb{W}_1(\mu^1_t, \mu^2_t)^2 \leq 2c_3^2\mathbb{W}_1(\mu^1_0, \mu^2_0)^2e^{2c_3^2\int_0^T K(s)ds}, \ t \in [0, T].
\]

Then the proof is finished. 

\[\square\]

3 Singular case with distribution dependent noise

In this part, we assume that $\sigma$ and $b$ are extended to $[0, T] \times \mathbb{R}^d \times \mathcal{P}_O$ 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 $f \in \tilde{L}_p^\#([s, t])$ if $f : [s, t] \times \mathbb{R}^d \to \mathbb{R}$ is measurable with

$$\|f\|_{L_p^\#([s, t])} := \sup_{z \in \mathbb{R}^d} \left\{ \int_s^t \left( \int_{B(z, 1)} |f(u, x)|^p dx \right)^{\frac{2}{p}} du \right\}^{\frac{1}{q}} < \infty,$$

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

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

We will take $(p, q)$ from the space

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

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

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

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

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

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

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

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

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

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

$$|b_t^{\mu_0}(x)| \leq f_0(t, x), \quad |\nabla \sigma^\mu_t(x)| \leq \sum_{i=1}^l f_i(t, x),$$

$$|b_t^{(1)}(x) - b_t^{(1)}(y)| \leq K|x - y|, \quad t \in [0, T], x, y \in \mathbb{R}^d.$$

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

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

**Theorem 3.1.** Assume (C) and $(A_2)$. Then the following assertions hold.
(1) \((1.3)\) is well-posed for \(O\)-distributions in \(\mathcal{P}_O\).

(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 \(\mathcal{P}_O\),
\[
\mathbb{E}
\left[
\sup_{t \in [0,T]} |X_t|^p \right] \leq \mathbb{E}
\left[
\sup_{t \in [0,T]} |X_{t\cap \tau}|^p \right] \leq c_p (1 + |X_0|^p).
\]

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

For any \(\mu \in C([0,T]; \mathcal{P}_O)\), instead of \((1.5)\) we consider the following SDE on \(\mathbb{R}^d\):
\[
dX_t^\mu = b_t^\mu(X_t^\mu)dt + \sigma_t^\mu(X_t^\mu)dW_t, \quad t \in [0,T].
\]

Noting that \(\bar{X}_t^\mu := X_{t\cap \tau}^\mu\) solves \((1.5)\), the map \(\Phi\) in \((1.6)\) is given by
\[
(\Phi\mu)_t := \mathcal{L}_{\bar{X}_t^\mu}^\mu, \quad t \in [0,T].
\]

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

By \([7, \text{Theorem 2.1}]\), see also \([11, \text{Theorem 1.1}]\) 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
\[
\mathbb{E}
\left[
\sup_{t \in [0,T]} |X_t|^p \right] \leq c_p (1 + |X_0|^p), \quad \mu \in C([0,T]; \mathcal{P}_O).
\]

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 \(\mu^1, \mu^2 \in C([0,T]; \mathcal{P}_O)\) and any \(X_t^i\) solving \((3.5)\) with \(\mu = \mu^i, i = 1, 2\),
\[
\mathbb{E}
\left[
\sup_{s \in [0,t]}
\left|
X_s^1 - X_s^2\right|^j \right] \leq c |X_0^1 - X_0^2|^j + \varepsilon(\theta)e^{\theta t} \mathbb{V}_{1,0}(\mu^1, \mu^2)^j, \quad \theta \geq 1.
\]

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

By \([13, \text{Theorem 2.1}]\), \((C_1)\) and \((C_2)\) imply that for large enough \(\lambda \geq 1\), the PDE
\[
\left(\partial_t + \frac{1}{2} \text{tr} \left\{ a_t^{\nu^i} \nabla^2 \right\} + b_t^{\nu^i} \cdot \nabla \right) u_t = \lambda u_t - b_t^{\nu^i,0}, \quad t \in [0,T], u_T = 0
\]
for \(u : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d\) has a unique solution such that
\[
\|\nabla^2 u\|_{L^\infty_{\text{loc}}(\mathbb{R}^d)} \leq c_0, \quad \|u\|_{\infty} + \|\nabla u\|_{\infty} \leq \frac{1}{2}.
\]
Let \(Y_t^i := \Theta_t(X_t^i), i = 1, 2, \Theta_t := id + u_t\). By Itô’s formula we obtain
\[
dY_t^1 = \left\{ b_t^{(1)} + \lambda u_t \right\} (X_t^1) dt + \left\{ (\nabla \Theta_t) \sigma_t^{\nu^1} \right\} (X_t^1) dW_t,
\]
\[
\frac{dY_t^2}{dt} = \left\{ \left\{ b_t^{(1)} + \lambda u_t + (\nabla \Theta_t)(b_t^{u^2} - b_t^{u^1}) \right\}(X_t^2) \right. \\
\left. + \frac{1}{2} \left\{ \text{tr}\left\{ \left( a_t^{u^2} - a_t^{u^1} \right) \nabla^2 u_t \right\} \right\}(X_t^2) \right\} dt + \left\{ \nabla \Theta_t \sigma_t^{u^2} \right\}(X_t^2) dW_t.
\]

Let \( \eta_t := |X_t^1 - X_t^2| \) and
\[
g_r := \sum_{i=0}^{l} f_i(r, X_r^2), \quad \tilde{g}_r := g_r \| \nabla^2 u_r(X_r^2) \|, \\
\tilde{g}_r := \sum_{i=1}^{2} \| \nabla^2 u_r(X_r^1) \| + \sum_{j=1}^{2} \sum_{i=0}^{l} f_i(r, X_r^j), \quad r \in [0, T].
\]

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
\[
\left| \left\{ b_t^{(1)} + \lambda u_t \right\}(X_t^1) - \left\{ b_t^{(1)} + \lambda u_t \right\}(X_t^2) \right| \leq c_1 \eta_t, \\
\left| \left\{ \left( \nabla \Theta_t \right)(b_t^{u^2} - b_t^{u^1}) \right\}(X_t^2) \right| \leq c_1 g_r \tilde{W}_1(\mu_r^1, \mu_r^2), \\
\left| \text{tr}\left\{ \left( a_t^{u^2} - a_t^{u^1} \right) \nabla^2 u_r \right\} \right|(X_t^2) \leq c_1 \tilde{g}_r \tilde{W}_1(\mu_r^1, \mu_r^2), \\
\left| \left\{ \left( \nabla \Theta_t \right)\sigma_t^{u^2} \right\}(X_t^1) - \left\{ \left( \nabla \Theta_t \right)\sigma_t^{u^2} \right\}(X_t^2) \right| \leq c_1 \tilde{g}_r \eta_t + c_1 g_r \tilde{W}_1(\mu_r^1, \mu_r^2), \quad r \in [0, T].
\]

So, by Itô’s formula, for any \( j \geq k \) we find a constant \( c_2 > 1 \) such that
\[
d|Y_t^j - Y_t^k|^2 \leq c_2 \eta_t^{2j} dA_t + c_2 (g_t^2 + \tilde{g}_t) \tilde{W}_1(\mu_r^1, \mu_r^2)^2 dt + dM_t
\]
holds for some martingale \( M_t \) with \( M_0 = 0 \) and
\[
A_t := \int_0^t \left\{ 1 + g_s^2 + \tilde{g}_s + \tilde{g}_s^2 \right\} ds, \quad t \in [0, T].
\]

Since \( \| \nabla u \|_\infty \leq \frac{1}{2} \) implies \( |Y_t^1 - Y_t^2| \geq \frac{1}{2} \eta_t \), this implies
\[
\eta_t^{2j} \leq 2^{2j} M_t + 2^{2j} \eta_0^{2j} + 2^{2j} c_2 \int_0^t \eta_r^{2j} dA_r \\
+ 2^{2j} c_2 \int_0^t (g_s^2 + \tilde{g}_s) \tilde{W}_1(\mu_s^1, \mu_s^2)^2 ds, \quad t \in [0, T]
\]
for some constant \( c_2 > 0 \). By \( 3.8 \), \( f_i \in \tilde{L}_{p_i}^q(T) \) for \( (p_i, q_i) \in K \), Krylov’s estimate (see [13, Theorem 3.1]) which implies Khasminskii’s estimate (see [12, Lemma 4.1(ii)]), we find an increasing function \( \psi : (0, \infty) \rightarrow (0, \infty) \) and a decreasing function \( \varepsilon : (0, \infty) \rightarrow (0, \infty) \) with \( \varepsilon(\theta) \downarrow 0 \) as \( \theta \uparrow \infty \), such that
\[
\mathbb{E}[e^{rA_T}|\mathcal{F}_0] \leq \psi(r), \quad r > 0, \\
\sup_{t \in [0,T]} \mathbb{E}\left( \int_0^t e^{-2k\theta(t-r)}(g_r^2 + \tilde{g}_r) dr \left| \mathcal{F}_0 \right\} \right) \leq \varepsilon(\theta), \quad \theta > 0.
\]
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
\[
\left\{ \mathbb{E} \left( \sup_{s \in [0,t]} \eta_s^2 \bigg| \mathcal{F}_0 \right) \right\}^2 
\leq c_3 \mathbb{E} \left( \eta_0^{2j} + \int_0^t (g_s^2 + \bar{g}_s) \tilde{W}_1(\mu_s^1, \mu_s^2)^{2j} ds \bigg| \mathcal{F}_0 \right) 
\leq c_3 \eta_0^{2j} + c_3 e^{2j\alpha \varepsilon(\theta) \tilde{W}_1(\mu^1, \mu^2)^{2j}}, \quad t \in [0, T], \theta > 0.
\]
This finishes the proof. \( \square \)

**Proof of Theorem 3.7** Let \(X_t\) solve (1.1). We have \(X_t = X_{t \wedge \tau(\gamma)}^\mu\) for \(X_t^\mu\) solving (3.5) with
\[
X_0^\mu = X_0, \quad \mu_t := \mathcal{L}_{X_t}^0, \quad t \in [0, T].
\]
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 \(\mathcal{F}_0\)-measurable random variable on \(\bar{O}\), and let \(\mathcal{G} = \mathcal{L}_X^0\). By Theorem 1.1, it suffices to prove that \(\Phi\) is contractive in \(\mathcal{G}\) under \(\tilde{W}_1(\cdot, \cdot)\) 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
\[
\tilde{W}_1((\Phi \mu^1)_t, (\Phi \mu^2)_t) \leq c_1 \varepsilon(\theta) \tilde{W}_1(\mu^1, \mu^2), \quad \mu^1, \mu^2 \in \mathcal{G}.
\]
Since \(\varepsilon(\theta) \to 0\) as \(\theta \to \infty\), \(\Phi\) is \(\tilde{W}_1, \theta\)-contractive for large enough \(\theta > 0\).

(b) Estimate (2.3). Let \(X^1_t, X^2_t\) solve (1.3) with \(X_0^1, X_0^2\) satisfying (2.20). Then
\[
(\Phi \mu^i)_t = \mu^i_t := \mathcal{L}_{X_t}^0 = P^\mu_t \mu^i, \quad i = 1, 2,
\]
so that (2.8), (2.19) and Lemma 3.2 imply
\[
\tilde{W}_1(\mu^1, \mu^2) = \tilde{W}_1((\Phi \mu^1)_t, (\Phi \mu^2)_t) \leq c_1 \tilde{W}_1(\mu^1_0, \mu^2_0) + c_1 \varepsilon(\theta) \tilde{W}_1(\mu^1, \mu^2), \quad t \in [0, T]
\]
for some constant \(c_1 > 0\). Taking \(\theta > 0\) large enough such that \(\varepsilon(\theta) \leq \frac{1}{2c_1}\), we derive (2.3) for some constant \(c > 0\). \( \square \)

### 3.2 For \(\mathcal{O}\)-distributions in \(\mathcal{P}^1_{\mathcal{O}}\)

**(D)** There exist an increasing function \(\alpha : [0, \infty) \to (0, \infty)\), constants \(K > 0, l \in \mathbb{N}\), \{\((p_i, q_i) : 0 \leq i \leq l\} \subset \mathcal{H}\) and functions \(1 \leq f_i \in L^q_{p_i}(T)\) for \(0 \leq i \leq l\) such that \(\sigma^\mu\) and \(b^\mu\) in (3.3) satisfy the following conditions.

**(D)\(_1\)** For any \(\mu \in C([0, T]; \mathcal{P}^1_{\mathcal{O}}), a^\mu := \sigma^\mu(\sigma^\mu)^*\) is invertible with
\[
\|a^\mu\|_\infty + \|(a^\mu)^{-1}\|_\infty \leq \alpha(\|\mu\|_{1,T}), \quad \lim_{\varepsilon \downarrow 0} \sup_{\mu \in C([0, T]; \mathcal{P}^1_{\mathcal{O}})} \sup_{t \in [0, T]} \|a^\mu_t(x) - a^\mu_t(y)\| = 0.
\]
\( b^{(1)}(0) \) is bounded on \([0, T]\), \( \sigma_i^\mu \) is weakly differentiable for \( \mu \in C([0, T]; \mathcal{P}_O^1) \), and
\[
|b_i^{\mu,0}(x)| \leq f_0(t, x) + \alpha(|\mu|_1), \quad \|\nabla \sigma_i^\mu(x)\| \leq \sum_{i=1}^l f_i(t, x) + \alpha(|\mu|_1),
\]
\[
|b_i^{(1)}(x) - b_i^{(1)}(y)| \leq K|x - y|, \quad t \in [0, T], x, y \in \mathbb{R}^d.
\]

(D3) For any \( t \in [0, T], x \in \mathbb{R}^d \) and \( \mu, \nu \in \mathcal{P}_O \),
\[
\|\sigma_i(x, \mu) - \sigma_i(x, \nu)\| + |b_i(x, \mu) - b_i(x, \nu)| \leq \mathbb{W}_1(\mu, \nu) \sum_{i=0}^l f_i(t, x).
\]

(D4) There exists \( r_0 \in (0, 1] \) such that \( \rho_0 \in C_0^2(\partial_r O) \), and for any \( \mu \in C([0, T]; \mathcal{P}_O^1) \),
\[
\langle b_i^\mu(x), \nabla \rho_0(x) \rangle \leq \alpha(|\mu|_1), \quad x \in \partial_r O,
\]
\[
\langle b_i^\mu(x, x - y) \rangle \leq \alpha(|\mu|_1,T)(f_0(t, x, x^2) + |x - y|^2), \quad x \in O, y \in \partial O, t \in [0, T].
\]

Note that when \( b^{(1)} = 0 \), (3.11) is implied by the first condition in (D2).

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

1. (1.3) is well-posed for \( O \)-distributions in \( \mathcal{P}_O^1 \).

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 \( \mathcal{P}_O \),
\[
\mathbb{E} \left[ \sup_{t \in [0, T]} |X_t|^p \mid \mathcal{F}_0 \right] \leq c_p \left\{ 1 + |X_0|^p + \left( \mathbb{E}[1_O(X_0)|X_0]|X_0] \right)^p \right\}.
\]

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

By the proof of [3] (2.17), (D) implies that for any \( \mu \in C([0, T]; \mathcal{P}_O^1) \), the SDE (3.5) is well-posed, and for any \( p \geq 1 \) there exists a constant \( c_p > 0 \) such that
\[
\mathbb{E} \left[ \sup_{t \in [0, T]} |X_t^\mu|^{2p} \mid \mathcal{F}_0 \right] \leq c_p \left\{ 1 + |X_0|^{2p} + \int_0^T \|\mu_s\|_{1}^{2p} \, ds \right\}, \quad \mu \in C([0, T]; \mathcal{P}_O).
\]

For any \( \mu^1, \mu^2 \in \mathcal{P}_O^1 \), let \( X^i_t \) solve (3.5) for \( \mu = \mu^i, i = 1, 2 \).

For any \( N > 0 \) and \( \gamma \in \mathcal{P}_O^1 \), let \( C_N^\gamma \) be in (2.29). Since restricting to \( \mu, \nu \in C_N^\gamma \) the conditions in (D) hold for a constant \( \alpha_N \) replacing the function \( \alpha \), by repeating the proof of Lemma 3.2 with \( \mathbb{W} \) replacing \( \tilde{\mathbb{W}} \), we prove that the following result.

19
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 $\mu^1, \mu^2 \in \mathcal{C}_N^0$ and any $X_t^\mu$ solving (3.5) with $\mu = \mu^i, i = 1, 2$,

$$
\mathbb{E} \left[ \sup_{s \in [0, t]} |X_s^1 - X_s^2|^j \bigg| \mathcal{F}_0 \right] \leq c |X_0^1 - X_0^2|^j + \varepsilon(\theta) e^{\delta t} \mathbb{E}_{1, 0} \left[ (\mu^1, \mu^2)^j \right], \ \theta \geq 1.
$$

When $\alpha$ 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 $\psi$ which is bounded if so is $\alpha$.

(a) Let $\rho_\theta(x) \geq \frac{m}{2}$ and $y \in \partial O$ such that $\rho_\theta(x) = |y - x|$. By (3.12) and (D2), we find an increasing function $\psi_1 : [0, \infty) \to (0, \infty)$ which is bounded if so is $\alpha$, such that

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

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 $\alpha$, the following Krylov’s estimate holds:

$$
\mathbb{E} \left( \int_0^T f_i(t, X_t^\mu)^2 dt \bigg| \mathcal{F}_0 \right) \leq \psi_2(\|\mu\|_{1,T}) \|f_i\|_{L_{\mu}^p(T)}^2, \quad 0 \leq i \leq l.
$$

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

$$
\mathbb{E}[|X_t^\mu - y|^2|\mathcal{F}_0] \leq \rho_\theta(x)^2 + \psi_1(\|\mu\|_{1,T}) \psi_2(\|\mu\|_{1,T}) \sum_{i=0}^l \|f_i\|_{L_{\mu}^p(T)}^2
$$

$$
+ \psi_1(\|\mu\|_{1,T}) \int_0^t \mathbb{E}[|X_s^\mu - y|^2|\mathcal{F}_0] ds, \quad t \in [0, T].
$$

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

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

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

(b) Let $\rho_\theta(x) < \frac{m}{2}$. By (D1), (3.11) and $\rho_\theta \in C^0(\partial \tilde{O})$, (2.21) holds for some different increasing function $\alpha$ 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 $\mathcal{P}_1^0$. We have $X_t = X_t^{\mu_{t \wedge \tau(X^\mu)}}$ for $X_t^\mu$ solving (3.5) with

$$
X_0^\mu = X_0, \quad \mu_t = \mathcal{L}_{X_t}^Q, \quad t \in [0, T].
$$
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 $\mathcal{F}_0$-measurable random variable with $\mathcal{L}_{X_0} \in \mathcal{P}_O^1$, and let $\mathcal{C}_N$ be in (2.29) for $N > 0$. By the proof of [5, Lemma 2.2(1)], there exists $N_0 > 0$ such that $\Phi \mathcal{C}_N \subset \mathcal{C}_N$ for any $N \geq N_0$. For the well-posedness, it suffices to prove that for any $N \geq N_0$, $\Phi$ is contractive in $\mathcal{C}_N$ under the metric $W_{1,\theta}$ for large enough $\theta > 0$. This follows from (2.27), Lemma 3.4 and Lemma 3.5.

Finally, by using $W_1$ replacing $\hat{W}_1$ 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 $\mu$, so that (1.3) becomes

\[(4.1) \quad dX_t = 1_{\{t < \tau(X)\}} \left\{ b_t(X_t, \mathcal{L}_{X_t}) dt + \sigma_t(X_t) dW_t \right\}, \quad t \in [0, T].\]

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

\[\mathcal{P}^V_O := \left\{ \mu \in \mathcal{P}_O : \mu(V) := \int_O V d\mu < \infty \right\}.\]

This is a Polish space under the weighted variation distance

\[(4.2) \quad \| \mu - \nu \|_V := \sup_{|f| \leq V} |\mu(f) - \nu(f)|, \quad \mu, \nu \in \mathcal{P}^V_O.\]

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

**Definition 4.1.** We denote $V \in \mathcal{V}$, 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 $x \in O$,

\[\sup_{y \in B(x, \varepsilon)} \left\{ |\nabla V(y)| + \| \nabla^2 V(y) \| \right\} \leq KV(x),\]

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

4.1 Main result

(E) $\sigma$ 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 \ \|\sigma\|_\infty + \|\sigma^{-1}\|_\infty < \infty \ and \\
\lim_{\varepsilon \to 0} \sup_{|x-y| \leq \varepsilon, t \in [0, T]} \|a_t(x) - a_t(y)\| = 0.

(E_2) \ there \ exist \ l \in \mathbb{N} \ and \ 1 \leq f_i \in \hat{L}^p_{\xi_i}(T) \ with \ (p_i, q_i) \in \mathcal{K}, \ 0 \leq i \leq l, \ such \ that \\
\|1_O b^{(0)}(x)\| \leq f_0, \ \|\nabla \sigma\| \leq \sum_{i=1}^l f_i.

(E_3) \ there \ exists \ V \in \mathcal{V} \ such \ that \ for \ any \ \mu \in \mathcal{C}_V := C([0, T]; \mathcal{P}^V_O), \ 1_O(x)b^{(1)}_t(x, \mu) \ is \ locally \ bounded \ in \ (t, x) \in [0, T] \times \mathbb{R}^d. \ Moreover, \ there \ exist \ constants \ K, \varepsilon > 0 \ such \ that \\
\langle b^{(1)}(x, \mu), \nabla V(x) \rangle + \varepsilon |b^{(1)}(x, \mu)| \sup_{B(x, \varepsilon)} \{ |\nabla V| + |\nabla^2 V| \} \\
\leq K \{ V(x) + \mu(V) \} \ x \in O, \mu \in \mathcal{P}^V_O.

(E_4) \ there \ exists \ a \ constant \ \kappa > 0 \ such \ that \\
\sup_{x \in O} |b_t(x, \mu) - b_t(x, \nu)| \leq \kappa \|\mu - \nu\|_V, \ \mu, \nu \in \mathcal{P}^V_O.

**Theorem 4.1.** \ Assume (E). \ Then (4.1) is well-posed for O-distributions in \mathcal{P}^V_O, \ and for any \ p \geq 1, \ there \ exists \ a \ constant \ c_p > 0 \ such \ that \ any \ solution \ X_t \ of (4.1) \ for \ O-distributions \ in \mathcal{P}^V_O \ satisfies \\
\mathbb{E} \left[ \sup_{t \in [0, T]} V(X_t)^p \right] \leq c_p V(X_0)^p.

**Proof.** \ Let \mathcal{P}^V(\bar{O}) \ be \ the \ space \ of \ all \ probability \ measures \ \mu \ 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 \mathcal{P}^V(\bar{O}). \ We \ extend \ b_t(x, \cdot) \ from \ \mathcal{P}^V_O \ to \ \mathcal{P}^V(\bar{O}) \ by \ setting \\
b_t(x, \mu) := b_t(x, \mu(O \cap \cdot)), \ \mu \in \mathcal{P}^V(\bar{O}).

Then (E) implies the same assumption for \mathcal{P}^V(\bar{O}) replacing \mathcal{P}^V_O. \ So, \ the \ desired \ assertions \ follow \ from \ Theorem 4.1 presented \ in \ the \ next \ subsection. \ \ \ \ \Box

### 4.2 \ An \ extension \ of \ Theorem 4.1

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

\begin{equation}
\frac{dX_t}{dt} = 1_{\{t < \tau(X)\}} \left\{ b_t(X_t, \mathcal{L}X_t)dt + \sigma_t(X_t)dW_t \right\}, \ \ t \in [0, T],
\end{equation}

where \ \tau(X) := \inf\{t \geq 0 : X_t \in \partial\bar{O}\} \ as \ before, \ and \ \mathcal{L}X_t \ is \ the \ distribution \ of \ X_t.

The \ strong/weak \ solution \ of (4.5) \ is \ defined \ as \ in \ Definition 4.1 \ with \ \mathcal{L} \ replacing \ \mathcal{L}^O. \ We \ call \ this \ equation \ well-posed \ for \ distributions \ in \ \mathcal{P}^V(\bar{O}), \ if \ for \ any \ \mathcal{F}_0-\text{measurable} \ random \ variable \ X_0 \ on \ \bar{O} \ with \ \mathcal{L}X_0 \in \mathcal{P}^V(\bar{O}) \ (respectively, \ any \ \mu_0 \in \mathcal{P}^V(\bar{O})), \ (4.5) \ has \ a \ unique \ solution \ starting \ at \ X_0 \ (respectively, \ a \ unique \ weak \ solution \ with \ initial \ distribution \ \mu_0) \ such \ that \ \mathcal{L}X = (\mathcal{L}X_t)_{t \in [0, T]} \in C([0, T]; \mathcal{P}^V(\bar{O})).
Theorem 4.2. Assume that (E) holds for $\mathcal{P}^V(\bar{O})$ replacing $\mathcal{P}^O$. Then (4.3) is well-posed for distributions in $\mathcal{P}^V(\bar{O})$ and (4.4) holds.

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

$$\gamma := \mathcal{L}X_0 \in \mathcal{P}^V(\bar{O}).$$

Let

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

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

$$dX_t^\mu = 1_{\{t<\tau(X_\nu)\}} \left\{ b_t(X_t^\mu, \mu_t)dt + \sigma_t(X_t^\mu)dW_t \right\}, \quad X_0^\mu = X_0, t \in [0,T].$$

Let $(\Phi \mu)_t := \mathcal{L}X_t^\mu, t \in [0,T]$. Then it suffices to prove that $\Phi$ has a unique fixed point in $\mathcal{C}^V(\bar{O})$.

To this end, for any $N \geq 1$, let

$$\mathcal{C}_{V,N}(\bar{O}) := \left\{ \mu \in \mathcal{C}^V(\bar{O}) : \sup_{t \in [0,T]} e^{-Nt}\mu_t(V) \leq N\gamma(V) \right\}.$$ 

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

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

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

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

$$dV(X_t^\mu)^p \leq 1_{\{t<\tau(X_\nu)\}} \left\{ dM_t + c_1 \left\{ V(X_t^\mu)^p + \mu_t(V)^p \right\} dt \right\}, \quad t \in [0,T],$$ 

where $M_t$ is a martingale with

$$d\langle M \rangle_t \leq c_1 V(X_t^\mu)^p dt.$$ 

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

$$\mathbb{E} \left[ \sup_{s \in [0,t]} V(X_s^\mu)^p \right] = \mathbb{E} \left[ \sup_{s \in [0,\tau(X_\nu)]} V(X_s^\mu)^p \right]$$

$$\leq c_2(p) V(X_0)^p + c_2(p) \int_0^t \mu_s(V)^p ds, \quad t \in [0,T].$$

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

$$(\Phi \mu)_t(V) = \mathbb{E}[V(X_t^\mu)] \leq c_2 \gamma(V) + c_2 \left( \int_0^t \mu_s(V)^2 ds \right)^{\frac{1}{2}},$$

23
so that by $\mu \in \mathcal{C}_{V,N}^i(\bar{O})$ we obtain

\[
\sup_{t \in [0,T]} e^{-Nt}(\Phi \mu)_t(V) \leq c_2 \gamma(V) + c_2 \sup_{t \in [0,T]} \left( \int_0^t e^{-2N(t-s)} N^2 \gamma(V)^2 ds \right)^{\frac{1}{2}}
\]

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

\[
\Phi \mathcal{C}_{V,N}^i(\bar{O}) \subset \mathcal{C}_{V,N}^i(\bar{O}), \quad N \geq N_0.
\]

(b) Let $N \geq N_0$. It remains to show that $\Phi$ has a unique fixed point in $\mathcal{C}_{V,N}^i(\bar{O})$. By (4.11) with $p = 2$ and $V \geq 1$, there exists a constant $c_3 > 0$ such that

\[
\mathbb{E} \left[ \sup_{t \in [0,T]} V(X_t^\mu)^2 \right] \leq c_3^2 V(X_0)^2, \quad \mu \in \mathcal{C}_{V,N}^i(\bar{O}).
\]

For any $\mu^i \in \mathcal{C}_V(\bar{O}), i = 1, 2$, 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

\[
dX_t^1 = 1_{\{t < \tau(X_t^1)\}} \{ b_t(X_t^1, \mu_t^1) dt + \sigma_t(X_t^1) dW_t \}, \quad X_0^1 = X_0.
\]

By the definition of $\Phi$, we have

\[
(\Phi \mu^1)_t = \mathcal{L}_X^1, \quad t \in [0,T].
\]

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

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

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

\[
|\xi_t| \leq k \| \mu_t^1 - \mu_t^2 \|_V, \quad t \in [0,T].
\]

So, by Girsanov’s theorem,

\[
\tilde{W}_t := W_t - \int_0^t \xi_s ds, \quad t \in [0,T]
\]

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

\[
R_s := e^{\int_0^s \langle \xi_t, dW_t \rangle - \frac{1}{2} \int_0^s |\xi_t|^2 dt}, \quad s \in [0,T].
\]

Reformulate (4.9) as

\[
dX_t^1 = 1_{\{t < \tau(X_t^1)\}} \{ b_t(X_t^1, \mu_t^2) dt + \sigma_t(X_t^1) d\tilde{W}_t \}, \quad X_0^1 = \tilde{X}_0.
\]

24
By the weak uniqueness of (4.6), we obtain
\[(\Phi\mu^2)_t = \mathbb{Q}(X^1_{\tau_t(X^1)}) \in dX) = \mathcal{L}X^1_t|\mathbb{Q}.
\]
Combining this with (4.8) and (4.10), we derive
\[
\|(\Phi\mu^1)_t - (\Phi\mu^2)_t\|_V \leq \mathbb{E}[V(X^1_t)|R_t - 1|]
\]
\[
\leq \mathbb{E}\left[\left\{\mathbb{E}(V(X^1_t)^2|\mathcal{F}_0)\right\}^{\frac{1}{2}}\left\{\mathbb{E}(|R_t - 1|^2|\mathcal{F}_0)\right\}^{\frac{1}{2}}\right]
\]
\[
\leq c_3\mathbb{E}\left[V(X_0)\left\{\mathbb{E}(|R_t - 1|^2|\mathcal{F}_0)\right\}^{\frac{1}{2}}\right].
\]
(4.12)

On the other hand, by \(\mu^1, \mu^2 \in \mathcal{C}^\gamma_{V,N}(\tilde{O}), (4.11)\), and noting that \(e^r - 1 \leq re^r\) for \(r \geq 0\), we find a constant \(c > 0\) such that
\[
\mathbb{E}[|R_t - 1|^2|\mathcal{F}_0] = \mathbb{E}[e^2\int_0^t(\xi_s, dW_s)-\int_0^t|\xi|^2ds - 1|\mathcal{F}_0]
\]
\[
\leq \mathbb{E}\left[e^{2\int_0^t(\xi_s, dW_s)-2\int_0^t|\xi|^2ds|\mathcal{F}_0}\right]\int_0^t k^2||\mu^1_s - \mu^2_s||_V^2ds
\]
\[
\leq c^2\int_0^t ||\mu^1_s - \mu^2_s||_V^2ds, \quad t \in [0, T].
\]

Combining this with (4.12) and letting \(C = c_3\mathbb{E}[V(X_0)]\), we arrive at
\[
\rho(\Phi(\mu^1), \Phi(\mu^2)) \leq C \sup_{t \in [0, T]} e^{-\lambda t^3} \left(\int_0^t ||\mu^1_s - \mu^2_s||_V^2ds\right)^{\frac{1}{2}}
\]
\[
\leq C \rho(\mu^1, \mu^2) \left(\int_0^t e^{-2\lambda(t-s)}ds\right)^{\frac{1}{2}}.
\]

Thus, when \(\lambda > 0\) is large enough, \(\Phi\) is contractive in \(\rho\) and hence has a unique fixed point in \(\mathcal{C}^\gamma_{V,N}(\tilde{O})\).

(3) Uniqueness and (4.1). It is easy to see that for any (weak) solution \(X_t\) of (4.5) for distributions in \(\mathscr{D}V(\tilde{O})\), \(\mu_t := \mathbb{L}X_t\) is a fixed point of \(\Phi\) in \(\mathcal{C}^\gamma_{V}(\tilde{O})\). Since \(\Phi\) has a unique fixed point, this implies the (weak) uniqueness of (4.1). Finally, by Gronwall’s inequality, (4.1) follows from (4.8) for \(X^\mu_t = X_t\) and \(\mu_t := \mathbb{L}X_t\), where \(\mu\) is the unique fixed point of \(\Phi\).

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, arXiv:2005.10057

[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*, arXiv:2110.08846

[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, *$L^2(L^p)$-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.