Log-Harnack Inequality and Bismut Formula for Singular McKean-Vlasov SDEs

Xing Huang\textsuperscript{a),}\ Feng-Yu Wang\textsuperscript{a),b)}

\textsuperscript{a)}Center for Applied Mathematics, Tianjin University, Tianjin 300072, China
xinghuang@tju.edu.cn

\textsuperscript{b)}Department of Mathematics, Swansea University, Bay Campus, SA1 8EN, United Kingdom
wangfy@tju.edu.cn

July 26, 2022

Abstract

The log-Harnack inequality and Bismut formula are established for McKean-Vlasov SDEs with singularities in all (time, space, distribution) variables, where the drift satisfies a standard integrability condition in time-space, and may be discontinuous in the distance induced by any Dini function. The main results extend existing ones derived for the case where the drift is $L$-differentiable and Lipchitz continuous in distribution with respect to the quadratic Wasserstein distance.

AMS subject Classification: 60H10, 60B05.

Keywords: McKean-Vlasov SDEs, Log-Harnack inequality, Bismut formula, Dini function, Wasserstein distance.

1 Introduction

Let $\mathcal{P}$ be the set of all probability measures on $\mathbb{R}^d$ equipped with the weak topology, and let $W_t$ be an $m$-dimensional Brownian motion on a complete filtration probability space $(\Omega, \{\mathcal{F}_t\}_{t \geq 0}, \mathcal{F}, \mathbb{P})$. Consider the following McKean-Vlasov SDE on $\mathbb{R}^d$:

\begin{equation}
(1.1) \quad dX_t = b_t(X_t, \mathcal{L}X_t)dt + \sigma_t(X_t)dW_t, \quad t \in [0, T],
\end{equation}

*Supported in part by NNSFC (11801406, 11831014, 11921001).
where \( T > 0 \) is a fixed time, \( \mathcal{L}_X \) is the distribution of \( X_t \), and
\[
b : [0, T] \times \mathbb{R}^d \times \mathcal{P} \to \mathbb{R}^d, \quad \sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^m
\]
are measurable. Because of its wide applications, this type SDE has been intensively investigated, see for instance [3, 4, 7, 8, 14, 15, 18] and the survey [9].

In this paper, we study the regularity of (1.1) for distributions in
\[
\mathcal{P}_k := \{ \mu \in \mathcal{P} : \| \mu \|_k := \mu(| \cdot |^k)^\frac{1}{k} < \infty \}, \quad k \in (1, \infty).
\]
Note that \( \mathcal{P}_k \) is a Polish space under the Wasserstein distance
\[
\mathbb{W}_k(\mu, \nu) = \inf_{\pi \in \mathcal{C}(\mu, \nu)} \left( \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^k \pi(dx, dy) \right)^\frac{1}{k},
\]
where \( \mathcal{C}(\mu, \nu) \) is the set of all couplings of \( \mu \) and \( \nu \). The SDE (1.1) is called well-posed for distributions in \( \mathcal{P}_k \), if for any initial value \( X_0 \) with \( \mathcal{L}_{X_0} \in \mathcal{P}_k \) (respectively, any initial distribution \( \gamma \in \mathcal{P}_k \)), it has a unique solution (respectively, a unique weak solution) \( X_t = (X_t)_{t \in [0, T]} \) such that \( \mathcal{L}_X := (\mathcal{L}_X)_t \in C([0, T]; \mathcal{P}_k) \). In this case, for any \( \gamma \in \mathcal{P}_k \), let \( P_t^* \gamma = \mathcal{L}_{X_t} \) for the solution \( X_t \) with \( \mathcal{L}_{X_0} = \gamma \). We study the regularity of the map
\[
\mathcal{P}_k \ni \gamma \mapsto P_t f(\gamma) := \mathbb{E}[f(X_t^\gamma)] = \int_{\mathbb{R}^d} f(d\{P_t^* \gamma\})
\]
for \( t \in (0, T] \) and \( f \in \mathcal{B}^+_b(\mathbb{R}^d) \), where \( \mathcal{B}_b(\mathbb{R}^d) \) is the space of bounded measurable functions on \( \mathbb{R}^d \).

As powerful tools charactering the regularity in distribution for stochastic systems, the dimension-free Harnack due to [19], the log-Harnack inequality introduced in [20], and the Bismut formula developed from [5] have been intensively investigated, see the monograph [21] for the study on SPDEs and applications. In recent years, the log-Harnack inequality and Bismut formula have also been established for McKean-Vlasov SDEs with drifts \( L \)-differentiable and \( W_2 \)-Lipschitz continuous in distribution. Below we present a brief summary.

When \( b_t(x, \mu) = b_t^{(0)}(x) + b_t^{(1)}(x, \mu) \), where \( b_t^{(0)} \) satisfies some integrability condition on \( (t, x) \) and \( b_t^{(1)} \) is Lipschitz continuous in \( (x, \mu) \in \mathbb{R}^d \times \mathcal{P}_2 \), the log-Harnack inequality
\[
P_t \log f(\tilde{\gamma}) \leq \log P_t f(\gamma) + \frac{c}{t} \mathbb{W}_2(\gamma, \tilde{\gamma})^2, \quad t \in (0, T], \ f \in \mathcal{B}_b^+(\mathbb{R}^d), \ \gamma, \tilde{\gamma} \in \mathcal{P}_2
\]
has been established in [24], where \( c > 0 \) is a constant and \( \mathcal{B}_b^+(\mathbb{R}^d) \) is the space of positive elements in \( \mathcal{B}_b(\mathbb{R}^d) \). This is equivalent to the entropy-cost inequality
\[
\text{Ent}(P_t^* \gamma | P_t^* \tilde{\gamma}) \leq \frac{c}{t} \mathbb{W}_2(\gamma, \tilde{\gamma})^2, \quad t \in (0, T], \gamma, \tilde{\gamma} \in \mathcal{P}_2,
\]
where \( \text{Ent} \) is the relative entropy, i.e. for any \( \mu, \nu \in \mathcal{P} \), \( \text{Ent}(\nu | \mu) := \infty \) if \( \nu \) is not absolutely continuous with respect to \( \mu \), while

\[
\text{Ent}(\nu | \mu) := \mu(\rho \log \rho) = \int_{\mathbb{R}^d} (\rho \log \rho) \, d\mu, \quad \text{if } \rho := \frac{d\nu}{d\mu} \text{ exists.}
\]

See also [12, 22, 23] for log-Harnack inequalities with more regular \( b^{(0)} \), and see [16] for the dimension-free Harnack inequality with power.

When \( b^{(1)}_t(x, \mu) \) is \( L \)-differentiable in \( \mu \in \mathcal{P}_k \) (see Definition 3.1), the following type Bismut formula is established in [24]:

\[
D^t P_t f(\mu) = \mathbb{E} \left[ f(X^\mu_t) M^{\mu, \phi}_t \right], \quad t \in (0, T], f \in \mathcal{B}_b(\mathbb{R}^d), \mu \in \mathcal{P}_k, \phi \in L^k(\mathbb{R}^d \to \mathbb{R}^d; \mu),
\]

where \( M^{\mu, \phi}_t \) is an explicit martingale. See [22, 17, 2, 10] for earlier results with more regular \( b^{(0)} \). We note that a derivative estimate is presented in [6] for the heat kernel when the drift is of type \( b_t(x, \mu(V)) \) for some Hölder continuous function \( \phi \), where \( \mu(V) := \int_{\mathbb{R}^d} V \, d\mu \).

In this case the drift is Lipschitz continuous in distribution with respect to the distance induced by \( \varepsilon \)-Hölder continuous functions for some \( \varepsilon \in (0, 1) \):

\[
\mathbb{W}_\varepsilon(\mu, \nu) := \sup \left\{ |\mu(f) - \nu(f)| : |f(x) - f(y)| \leq |x - y|^\varepsilon \right\}.
\]

In this paper, we establish the log-Harnack inequality and Bismut formula for the drift being only Lipschitz continuous in distribution with respect to

\[
\mathbb{W}_\alpha(\mu, \nu) := \sup \left\{ |\mu(f) - \nu(f)| : |f(x) - f(y)| \leq \alpha(|x - y|) \right\},
\]

where \( \alpha \) is the square root of a Dini function, i.e. \( \alpha \) is concave with \( \alpha(0) = 0, \alpha'(s) > 0 \) for \( s > 0 \), and \( \int_0^1 \frac{\alpha(s)^2}{s} \, ds < \infty \) (Dini condition for \( \alpha^2 \)). Thus, the continuity of drift in distribution is even weaker than the Dini continuity, i.e. it may be discontinuous in the distance induced by Dini functions.

In the existing study of dimension-free/log Harnack inequalities and Bismut formulas for singular SDEs, a key technique is to regularize the drift in the spatial variable \( x \) by using Zvonkin’s transform [28]. The idea of the present paper is to realize this type regularization also in the distribution variable, by using a-priori derivative estimates and Bismut formula for (distribution independent) singular SDEs.

## 2 Log-Harnack Inequality

Since \( \mathbb{W}_2 \) is involved in the log-Harnack inequality, in this section we mainly consider (1.1) for distributions in \( \mathcal{P}_2 \), but the drift may be discontinuous in \( \mathbb{W}_k \) for any \( k > 0 \). We first state the concrete assumption and the main result on the log-Harnack inequality, then present a complete proof in a separate subsection.
2.1 Assumption and main result

We will allow \( b_t(x, \cdot) \) to be merely Lipschitz continuous in the sum of \( W_2 \) and the Wasserstein distance induced by the square root of a Dini function.

Let \( \alpha \) be in the following class

\[
\mathcal{A} := \left\{ \alpha : [0, \infty) \to [0, \infty) \text{ is increasing and concave}, \alpha(0) = 0, \int_0^1 \frac{\alpha(r)^2}{r} dr \in (0, \infty) \right\},
\]

where \( \int_0^1 \frac{\alpha(r)^2}{r} dr < \infty \) is the Dini condition for \( \alpha^2 \). For a (real or Banach valued) function \( f \), let

\[
[f]_\alpha := \sup_{x \neq y} \frac{|f(x) - f(y)|}{\alpha(|x - y|)}
\]

be its continuity modulus in \( \alpha \). Define the Wasserstein distance induced by \( \alpha \):

\[
W_\alpha(\mu, \nu) := \sup_{[f]_\alpha \leq 1} |\mu(f) - \nu(f)|, \quad \mu, \nu \in \mathcal{P}_\alpha := \left\{ \mu \in \mathcal{P} : \mu(\alpha(|\cdot|)) < \infty \right\},
\]

where \( f \) are real functions and \( \mu(f) = \int_{\mathbb{R}^d} f \, d\mu \).

By the concavity of \( \alpha \), \( W_\alpha \) is a complete distance on \( \mathcal{P}_\alpha \), \( \mathcal{P}_k \subset \mathcal{P}_\alpha \) for \( k \geq 1 \), and

\[
(2.1) \quad \alpha(s + t) \leq \alpha(s) + \alpha(t), \quad \alpha(rt) \leq r \alpha(t), \quad s, t > 0, r \geq 1.
\]

These inequalities follow from \( \alpha(0) = 0 \) and the decreasing monotonicity of \( \alpha' \) such that

\[
\alpha'(s + t) \leq \alpha'(s), \quad \frac{d}{dt} \alpha(rt) = r \alpha'(rt) \leq r \alpha'(t), \quad s, t \geq 0, r \geq 1.
\]

Since \( \alpha \) is increasing with \( \int_0^1 \frac{\alpha(s)^2}{s} ds > 0 \), we have \( \alpha(1) > 0 \) so that the second estimate in (2.1) with \( r = t^{-1} \) yields

\[
(2.2) \quad \alpha(t) \geq \alpha(1)t > 0, \quad t \in (0, 1).
\]

To measure the singularity in \( (t, x) \in [0, T] \times \mathbb{R}^d \), we recall locally integrable functional spaces presented in [25]. 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

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

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

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

We take \((p, q)\) from the space

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

and make the following assumption where \( \nabla \) is the gradient in \( x \in \mathbb{R}^d \).
(A) There exist \( \alpha \in \mathcal{A} \), \( k \in (1, \infty) \), \( \kappa \in [0, \infty) \), \( K \in (0, \infty) \), \( l \in \mathbb{N} \), and 
\[
1 \leq f_i \in \tilde{L}^n_{p_0}(T), \quad (p_i, q_i) \in \mathcal{K}, \quad 0 \leq i \leq l
\]
such that the following conditions hold.

\((A_1)\) \((\sigma_1 \sigma_i^*)(x)\) is invertible and \( \sigma_i(x) \) is weakly differentiable in \( x \) such that
\[
\|\sigma \sigma^*\| \infty + \|(\sigma \sigma^*)^{-1}\| \infty < \infty, \quad |\nabla \sigma| \leq \sum_{i=1}^l f_i,
\]

\[
\lim_{\varepsilon \downarrow 0} \sup_{t \in [0, T], |x-x'| \leq \varepsilon} \|((\sigma_i \sigma_i^*)(x) - (\sigma_i \sigma_i^*)(x'))\| = 0.
\]

\((A_2)\) \( b_t(x, \mu) = b_t^{(0)}(x) + b_t^{(1)}(x, \mu) \), where for any \( t \in [0, T] \), \( x, y \in \mathbb{R}^d \), \( \mu, \nu \in \mathcal{P}_k \),
\[
|b_t^{(0)}(x)| \leq f_0(t, x), \quad |b_t^{(1)}(x, \mu)| \leq K + \kappa|x| + \kappa||\mu||_k,
\]
\[
|b_t^{(1)}(x, \mu) - b_t^{(1)}(y, \nu)| \leq K \{ |x-y| + \mathbb{W}_\alpha(\mu, \nu) + \mathbb{W}_k(\mu, \nu) \}.
\]

We first observe that (A) implies the well-posedness of (1.1) for distributions in \( \mathcal{P}_k \).
Since \( \alpha(0) = 0 \) and \( \alpha \) is concave, there exists a constant \( c > 0 \) such that
\[
\sup_{|f|_{\alpha} \leq 1} |f(x) - f(0)| \leq \alpha(|x|) \leq c + c|x|^k, \quad x \in \mathbb{R}^d.
\]
Thus,
\[
\frac{1}{c} \mathbb{W}_\alpha(\mu, \nu) \leq \mathbb{W}_{k, var}(\mu, \nu) := \sup_{|f| \leq 1 + |x|^k} |\mu(f) - \nu(f)|.
\]
So, by [23, Theorem 3.1(1)] for \( D = \mathbb{R}^d \), under assumption (A), (1.1) is well-posed for distributions in \( \mathcal{P}_k \), and for any \( n \geq 1 \) there exists a constant \( c_0 > 0 \) such that
\[
\mathbb{E} \left[ \sup_{t \in [0, T]} |X_t|^n \right| \mathcal{F}_0 \right] \leq c_0(1 + |X_0|^n).
\]
Consequently,
\[
\sup_{t \in [0, T]} \|P_\gamma^* l\|^n \leq \mathbb{E} \left[ \sup_{t \in [0, T]} |X_t|^n \right] \leq c_n(1 + \|\gamma\|^n), \quad n \geq 1.
\]

**Theorem 2.1.** Assume (A) with \( k = 2 \), let
\[
\tilde{\alpha}(r) := \left( \int_0^r \frac{\alpha(t)^2}{t} dt \right)^{\frac{1}{2}}, \quad r \geq 0.
\]
Then there exists a constant \( c > 0 \) such that for any \( t \in (0, T] \) and \( \gamma, \tilde{\gamma} \in \mathcal{P}_2 \),
\[
\text{Ent}(P_\gamma^* \gamma | P_t^* \tilde{\gamma}) \leq \mathbb{W}_2(\gamma, \tilde{\gamma})^2 \left\{ \frac{c}{t} + \tilde{\alpha}(1 + \kappa\|\gamma\|_2 + \kappa\|\tilde{\gamma}\|_2) \right\}.
\]
If in particular \( \kappa = 0 \) (i.e. \( b^{(1)} \) is bounded), there exists a constant \( c > 0 \) such that
\[
\text{Ent}(P_\gamma^* \gamma | P_t^* \tilde{\gamma}) \leq \frac{c}{t} \mathbb{W}_2(\gamma, \tilde{\gamma})^2, \quad t \in (0, T], \gamma, \tilde{\gamma} \in \mathcal{P}_2.
\]
2.2 Proof of Theorem 2.1

Although in Theorem 2.1 we assume \( A \) for \( k = 2 \), for later use we will also consider general \( k \in [1, \infty) \). For any \( \gamma \in \mathcal{P}_k \), consider the decoupled SDE of (1.1):

\[
dX_t^{x,\gamma} = b_t(X_t^{x,\gamma}, P_t^\gamma)dt + \sigma_t(X_t^{x,\gamma})dW_t, \quad X_0^{x,\gamma} = x.
\]

By \( A \), this SDE is well-posed, so that the solution is a standard Markov process, see [24, Theorem 2.1]. Let \( P_t^\gamma \) be the associated Markov semigroup, i.e.

\[
P_t^\gamma f(x) := \mathbb{E}[f(X_t^{x,\gamma})], \quad t \in [0, T], x \in \mathbb{R}^d, f \in \mathcal{B}_b(\mathbb{R}^d).
\]

Noting that \( X_t^{x,\gamma} \) solves (1.1) if the initial value \( x \) is random with distribution \( \gamma \), by the standard Markov property of \( X_t^{x,\gamma} \), we have

\[
P_t f(\gamma) := \int_{\mathbb{R}^d} f(x) P_t^\gamma(dx) = \int_{\mathbb{R}^d} P_t^\gamma f(x)(dx), \quad f \in \mathcal{B}_b(\mathbb{R}^d).
\]

Lemma 2.2. Let \( \sigma \) and \( b \) satisfy \( A \). For any \( p \geq 1 \), there exists a constant \( c_p > 0 \) such that

\[
\mathbb{E}[|X_t^{x,\gamma} - x|^p] \leq c_p (1 + \kappa |x|^p + \kappa \|\gamma\|_k^p) t^\frac{p}{2}, \quad t \in [0, T], x \in \mathbb{R}^d, \gamma \in \mathcal{P}_k,
\]

(2.5)

\[
\mathbb{E} \left[ \sup_{t \in [0, T]} |X_t^{x,\gamma} - X_t^{y,\gamma}|^p \right] \leq c_p |x - y|^p, \quad x, y \in \mathbb{R}^d, \gamma \in \mathcal{P}_k.
\]

(2.6)

Proof. By [23, Theorem 2.1(1)], \( A \) implies (2.6). It remains to prove (2.5). To this end, we use Zvonkin’s transform.

For any \( \lambda > 0 \), consider the following PDE for \( u : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \):

\[
\frac{\partial}{\partial t} u_t(x) + (\mathcal{L}_t^\gamma u_t)(x) + b_t^{(0)}(x) = \lambda u_t(x), \quad u_T = 0,
\]

(2.7)

where

\[
\mathcal{L}_t^\gamma = \frac{1}{2} \text{tr} \{ (\sigma_t \sigma_t^*) \nabla^2 \} + \nabla b_t(\cdot, P_t^\gamma).
\]

(2.8)

By [24] and \( A \), for large enough constants \( \lambda, c > 0 \) independent of \( \gamma \), (2.7) has a unique solution \( u^{\lambda,\gamma} \) satisfying

\[
\|u^{\lambda,\gamma}\|_\infty + \|\nabla u^{\lambda,\gamma}\|_\infty \leq \frac{1}{5}, \quad \|\nabla^2 u^{\lambda,\gamma}\|_{L^5(T)} \leq c.
\]

(2.9)

So, for any \( t \in [0, T] \),

\[
x \mapsto \Theta_t^{\lambda,\gamma}(x) := x + u_t^{\lambda,\gamma}(x), \quad x \in \mathbb{R}^d
\]

(2.10)
is a homeomorphism on \( \mathbb{R}^d \). By Itô’s formula (see [24]), we derive

\[
(2.11) \quad d\Theta^\lambda_\gamma(X^\gamma_t) = \{ \lambda u^\lambda_t(X^\gamma_t) + b^\lambda_t(X^\gamma_t, P_t^\gamma) \} dt + \{(\nabla \Theta^\lambda_t)^{\gamma}\} (X^\gamma_t) dW_t.
\]

By (A), (2.4) and (2.9), there exists a constant \( C > 1 \) such that

\[
C^{-1}|X^\gamma_t| - C \leq |\Theta^\lambda_\gamma(X^\gamma_t)| \leq C|X^\gamma_t|,
\]

\[
|\lambda u^\lambda_t(X^\gamma_t) + b^\lambda_t(X^\gamma_t, P_t^\gamma)| \leq C(1 + \kappa|X^\gamma_t| + \kappa\|\gamma\|),
\]

\[
\|\{(\nabla \Theta^\lambda_t)^{\gamma}\} (X^\gamma_t)\| \leq C, \quad (t, x, \gamma) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_k.
\]

This together with (2.11) implies (2.5).

\[\square\]

**Lemma 2.3.** Assume (A). Then there exists a constant \( c > 0 \) such that

\[
(2.12) \quad \mathbb{W}_k(P^*_\gamma, P^*_\tilde{\gamma}) \leq c\mathbb{W}(\gamma, \tilde{\gamma}) + c \int_0^t \mathbb{W}_\alpha(P^*_{s\gamma}, P^*_{s\tilde{\gamma}}) ds, \quad t \in [0, T], \gamma, \tilde{\gamma} \in \mathcal{P}_k.
\]

**Proof.** We take \( \mathcal{F}_0 \)-measurable random variables \( X^\gamma_0, X^\tilde{\gamma}_0 \) such that

\[
(2.13) \quad \mathcal{L}_{X^\gamma_0} = \gamma, \quad \mathcal{L}_{X^\tilde{\gamma}_0} = \tilde{\gamma}, \quad \mathbb{W}_k(\gamma, \tilde{\gamma})^k = \mathbb{E}[|X^\gamma_0 - X^\tilde{\gamma}_0|^k].
\]

By (2.7), (2.8) and Itô’s formula, we derive

\[
(2.14) \quad d\Theta^\lambda_\gamma(X^\gamma_t) = \{ \lambda u^\lambda_t(X^\gamma_t) + b^\lambda_t(X^\gamma_t, P_t^\gamma) \} dt
\]

\[
+ \nabla_{b_t(X^\gamma_t, P^*\gamma) - b_t(X^\gamma_t, P_t^\gamma)} \Theta^\lambda_\gamma(X^\gamma_t) dt + \{(\nabla \Theta^\lambda_t)^{\gamma}\} (X^\gamma_t) dW_t.
\]

Combining this with (2.11) and (A), we prove the desired estimate by using the maximal functional inequality, Khasminskii’s estimate and stochastic Gronwall’s inequality, see for instance the proof of [13, Lemma 2.1] for details. Below we simply outline the procedure.

By (A2) we have

\[
|b_t(X^\gamma_t, P^*_\gamma) - b_t(X^\gamma_t, P_t^\gamma)| + |b^\lambda_t(X^\gamma_t, P_t^\gamma) - b^\lambda_t(X^\gamma_t, P^*_\gamma)|
\]

\[
\leq K \{ |X^\gamma_t - X^\tilde{\gamma}_t| + \mathbb{W}_\alpha(P^*_\gamma, P_t^\gamma) + \mathbb{W}_k(P_t^\gamma, P^*_\gamma) \}.
\]

Combining this with (2.11), (2.14), (A1), the maximal functional inequality and Khasminskii’s estimate (see [25, Lemma 2.1 and Lemma 4.1]), we derive

\[
d|\Theta^\lambda_\gamma(X^\gamma_t) - \Theta^\lambda_\gamma(X^\tilde{\gamma}_t)|^{k+1} \leq dM_t + |X^\gamma_t - X^\tilde{\gamma}_t|^{k+1} d\mathcal{L}_t
\]

\[
+ c_1 \{ \mathbb{W}_\alpha(P^*_\gamma, P_t^\gamma) + \mathbb{W}_k(P_t^\gamma, P^*_\gamma) \} |\Theta^\lambda_\gamma(X^\gamma_t) - \Theta^\lambda_\gamma(X^\tilde{\gamma}_t)|^k dt,
\]

where \( c_1 > 0 \) is a constant, \( \mathcal{L}_t \) is an adapted increasing process with \( \mathbb{E}[^{\lambda \mathcal{L}_\tau}] < \infty \) for any \( \lambda > 0 \), and \( M_t \) is a local martingale. Since (2.9) implies

\[
\frac{1}{2} |X^\gamma_t - X^\tilde{\gamma}_t| \leq |\Theta^\lambda_\gamma(X^\gamma_t) - \Theta^\lambda_\gamma(X^\tilde{\gamma}_t)| \leq 2|X^\gamma_t - X^\tilde{\gamma}_t|,
\]

\[
\frac{1}{2} |X^\gamma_t - X^\tilde{\gamma}_t| \leq |\Theta^\lambda_\gamma(X^\gamma_t) - \Theta^\lambda_\gamma(X^\tilde{\gamma}_t)| \leq 2|X^\gamma_t - X^\tilde{\gamma}_t|,
\]

\[
\frac{1}{2} |X^\gamma_t - X^\tilde{\gamma}_t| \leq |\Theta^\lambda_\gamma(X^\gamma_t) - \Theta^\lambda_\gamma(X^\tilde{\gamma}_t)| \leq 2|X^\gamma_t - X^\tilde{\gamma}_t|,
\]
by the stochastic Gronwall inequality (see [26, Lemma 3.7]), we find a constant $c_2 > 1$ such that
\[
\left\{ \mathbb{E} \left[ \sup_{s \in [0,t]} |X_s^{\tilde{\gamma}} - X_s^{\gamma}|^k \big| \mathcal{F}_0 \right] \right\}^{1+k^{-1}} - c_2 |X_0^{\tilde{\gamma}} - X_0^{\gamma}|^{k+1} 
\leq c_2 \int_0^t \{ \mathbb{W}_k(P_s^{\gamma}, P_s^{\tilde{\gamma}}) + \mathbb{W}_k(P_s^{\gamma}, P_s^{\gamma}) \} \mathbb{E} \left[ |X_s^{\tilde{\gamma}}|^k \big| \mathcal{F}_0 \right] ds, \ t \in [0,T].
\]

So, there exists a constant $c_3 > 0$ such that for any $t \in [0,T],$
\[
\mathbb{E} \left[ \sup_{s \in [0,t]} |X_s^{\tilde{\gamma}} - X_s^{\gamma}|^k \big| \mathcal{F}_0 \right] - c_2 |X_0^{\tilde{\gamma}} - X_0^{\gamma}|^k 
\leq c_2 \left( \int_0^t \{ \mathbb{W}_k(P_s^{\gamma}, P_s^{\tilde{\gamma}}) + \mathbb{W}_k(P_s^{\gamma}, P_s^{\gamma}) \} \mathbb{E} \left[ |X_s^{\tilde{\gamma}}|^k \big| \mathcal{F}_0 \right] ds \right)^{k+1} 
\leq \frac{1}{2} \mathbb{E} \left[ \sup_{s \in [0,t]} |X_s^{\tilde{\gamma}} - X_s^{\gamma}|^k \big| \mathcal{F}_0 \right] + c_3 \left( \int_0^t \{ \mathbb{W}_k(P_s^{\gamma}, P_s^{\tilde{\gamma}}) + \mathbb{W}_k(P_s^{\gamma}, P_s^{\gamma}) \} ds \right)^k.
\]

This together with (2.13) yields
\[
\mathbb{W}_k(P_s^{\gamma}, P_s^{\tilde{\gamma}}) \leq \sup_{s \in [0,t]} \left( \mathbb{E} \left[ |X_s^{\tilde{\gamma}} - X_s^{\gamma}|^k \right] \right)^{\frac{1}{k}} 
\leq (2c_2)^k \mathbb{W}_k(\gamma, \tilde{\gamma}) + (2c_3)^k \int_0^t \{ \mathbb{W}_k(P_s^{\gamma}, P_s^{\tilde{\gamma}}) + \mathbb{W}_k(P_s^{\gamma}, P_s^{\gamma}) \} ds, \ t \in [0,T].
\]

By Gronwall’s inequality, this implies the desired estimate for some constant $c > 0.$

We will need the following Hölder inequality for concave functions.

**Lemma 2.4.** Let $\alpha : [0, \infty) \to [0, \infty)$ be concave. For any non-negative random variables $\xi$ and $\eta,$
\[
(2.15) \quad \mathbb{E}[\alpha(\xi)\eta] \leq \|\eta\|_{L^p(\mathbb{P})} \alpha \left( \|\xi\|_{L^\frac{p}{p-1}(\mathbb{P})} \right), \ p \geq 1.
\]

Consequently, for any random variable $\xi$ on $\mathbb{R}^d,$ $f \in C(\mathbb{R}^d; \mathbb{B})$ for a Banach space $(\mathbb{B}, \| \cdot \|_\mathbb{B})$ with $[f]_\alpha < \infty,$ and real random variable $\eta$ with $\mathbb{E}[\eta] = 0,$
\[
(2.16) \quad \|\mathbb{E}[f(\xi)\eta]\|_\mathbb{B} \leq [f]_\alpha \|\eta\|_{L^p(\mathbb{P})} \alpha \left( \|\xi - x\|_{L^\frac{p}{p-1}(\mathbb{P})} \right), \ p \geq 1, x \in \mathbb{R}^d.
\]

**Proof.** It suffices to prove for $\mathbb{E}[|\eta|^p] \in (0, \infty).$ Let $\mathbb{Q} := \frac{\mathbb{P}[\eta]}{\mathbb{E}[\eta]} \mathbb{P}.$ By Jensen’s and Hölder’s inequalities, and using (2.1), we obtain
\[
\mathbb{E}[\alpha(\xi)\eta] = \mathbb{E}[\mathbb{E}[\alpha(\xi)\eta] \big| \mathbb{Q}] \leq \mathbb{E}[|\eta|] \alpha(\mathbb{E}[\xi] \big| \mathbb{Q}) \leq \mathbb{E}[|\eta|] \alpha \left( \left( \frac{\mathbb{E}[|\eta|^p]}{\mathbb{E}[|\eta|]} \right)^{\frac{1}{p}} \right)^{\frac{p}{p-1}} = \mathbb{E}[|\eta|^p] \alpha \left( \left( \frac{\mathbb{E}[|\eta|^p]}{\mathbb{E}[|\eta|]} \right)^{\frac{1}{p}} \right)^{\frac{p}{p-1}}.
\]
Then the second inequality follows by noting that $E[\eta] = 0$ implies
\[ \|E[f(\xi)\eta]\|_B = \|E[\{f(\xi) - f(x)\}\eta]\|_B \leq [f]_aE[|\xi - x|][\eta]. \]

\[ \Box \]

**Lemma 2.5.** Assume (A). If $\alpha_k(s) := \alpha(s^{1/(k-1)})$ is concave in $s \geq 0$, then there exists a constant $c > 0$ such that
\[
\mathcal{W}_a(P_t^\gamma, P_t^{\tilde{\gamma}}) \leq c \mathcal{W}_k(\tilde{\gamma}, \gamma) \left\{ \frac{\alpha((1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)\sqrt{t})}{\sqrt{t}} + \tilde{\alpha}(1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)e^{\alpha(1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)^2}} \right\}, \quad t \in (0, T], \gamma, \tilde{\gamma} \in \mathcal{P}_k.
\]

Consequently, there exists a constant $c > 0$ such that for any $\gamma, \tilde{\gamma} \in \mathcal{P}_k$,
\[
(2.17) \quad \sup_{t \in [0, T]} \mathcal{W}_k(P_t^\gamma, P_t^{\tilde{\gamma}}) \leq \tilde{\alpha}(1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)e^{\alpha(1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)^2}\mathcal{W}_k(\gamma, \tilde{\gamma}).
\]

**Proof.** It suffices to prove the first estimate, since it implies (2.17) according to Lemma 2.3.

Let $X_0^\gamma$ and $X_0^{\tilde{\gamma}}$ be in (2.13). For any $\varepsilon \in [0, 2]$, let
\[ X_0^{\varepsilon} = X_0^\gamma + \varepsilon(X_0^{\tilde{\gamma}} - X_0^\gamma), \quad \gamma^\varepsilon := \mathcal{L}X_0^{\varepsilon}, \]
and let $X_t^{\gamma^\varepsilon}$ solve (1.1) with initial value $X_0^{\gamma^\varepsilon}$. Then
\[
(2.18) \quad \gamma^\varepsilon(\|\cdot\|) \leq 2\|\gamma\|_k + 2\|\tilde{\gamma}\|_k, \quad \varepsilon \in [0, 2],
\]
\[
(2.19) \quad \mathcal{W}_k(\gamma^\varepsilon, \gamma^{\varepsilon+r})^k \leq \mathbb{E}[\|X_0^{\gamma^\varepsilon} - X_0^{\gamma^{\varepsilon+r}}\|^k] = r^k\mathcal{W}_k(\gamma, \tilde{\gamma})^k, \quad \varepsilon, r \in [0, 1].
\]

For any $\varepsilon \geq 0$, consider the SDE
\[
(2.20) \quad dX_t^{x,\gamma^\varepsilon} = b_t(X_t^{x,\gamma^\varepsilon}, P_t^{x,\gamma^\varepsilon})dt + \sigma_t(X_t^{x,\gamma^\varepsilon})dW_t, \quad X_0^{x,\gamma^\varepsilon} = x, t \in [0, T].
\]

For any $r \in (0, 1)$, let
\[ \eta^{\varepsilon,r}_t = [\sigma_t(\sigma_t^{-1})(X_t^{x,\gamma^\varepsilon}[b_t(X_t^{x,\gamma^\varepsilon}, P_t^{x,\gamma^\varepsilon+r}) - b_t(X_t^{x,\gamma^\varepsilon}, P_t^{x,\gamma^\varepsilon})], \quad t \in [0, T].
\]

By (A), there exists a constant $c_1 > 0$ such that
\[
(2.21) \quad \sup_{t \in [0, T]} |\eta^{\varepsilon,r}_t| \leq c_1\left\{ \mathcal{W}_a(P_t^{x,\gamma^\varepsilon}, P_t^{x,\gamma^{\varepsilon+r}}) + \mathcal{W}_k(P_t^{x,\gamma^\varepsilon}, P_t^{x,\gamma^{\varepsilon+r}}) \right\}, \quad r, \varepsilon \in [0, 1].
\]
By Girsanov’s theorem,

\[ R^\varepsilon_t := \exp \left\{ \int_0^t \langle \eta^\varepsilon_s, dW_s \rangle - \frac{1}{2} \int_0^t |\eta^\varepsilon_s|^2 ds \right\}, \quad t \in [0, T] \]

is a martingale, and

\[ W^\varepsilon_t = W_t - \int_0^t \eta^\varepsilon_s ds, \quad t \in [0, T] \]

is a Brownian motion under the probability \( Q^\varepsilon := R^\varepsilon_T \mathbb{P} \). Rewrite (2.20) as

\[ dX^{x,\varepsilon}_t = b_t(X^{x,\varepsilon}_t, P^*_s \varepsilon^{\varepsilon+r})dt + \sigma_t(X^{x,\varepsilon}_t)dW^{\varepsilon}_t, \quad X^{x,\varepsilon}_0 = x, \quad t \in [0, T]. \]

By the weak uniqueness we obtain

\[ \mathcal{L}_{\{X^{x,\varepsilon}_t\}_{t \in [0,T]|Q^\varepsilon_r}} = \mathcal{L}_{\{X^{x,\varepsilon+r}_t\}_{t \in [0,T]}}, \]

where \( \mathcal{L}_{Q^\varepsilon} \) is the law under \( Q^\varepsilon \), so that

\[ P^{\varepsilon+r}_t f(x) - P^\varepsilon_t f(x) = \mathbb{E} \left[ f(X^{x,\varepsilon}_T)(R^\varepsilon_T - 1) \right], \quad f \in \mathcal{B}_b(\mathbb{R}^d), \varepsilon, r \in (0, 1]. \]

Hence,

\[
\begin{align*}
P^\varepsilon_t (\varepsilon^{\varepsilon+r}) - P^\varepsilon_t (\varepsilon) &= \varepsilon^{\varepsilon+r} (P^{\varepsilon+r}_t f) - \varepsilon (P^\varepsilon_t f) \\
&= \varepsilon^{\varepsilon+r} (P^{\varepsilon+r}_t f - P^\varepsilon_t f) + \varepsilon^{\varepsilon+r} (P^\varepsilon_t f) - \varepsilon (P^\varepsilon_t f) \\
&= \int_{\mathbb{R}^d} \mathbb{E} \left[ f(X^{x,\varepsilon}_T)(R^\varepsilon_T - 1) \right] \varepsilon^{\varepsilon+r} (dx) + \mathbb{E} \left[ P^{\varepsilon+r}_t f(X^{\varepsilon+r}_0) - P^\varepsilon_t f(X^\varepsilon_0) \right],
\end{align*}
\]

so that

\[ \mathbb{W}_\alpha (P^*_t \varepsilon^{\varepsilon+r}, P^*_t \varepsilon^\varepsilon)^2 = \sup_{\|f\|_1 \leq 1} \left| P^\varepsilon_t f(\varepsilon^{\varepsilon+r}) - P^\varepsilon_t f(\varepsilon) \right|^2 \leq I_1 + I_2, \]

\[
\begin{align*}
I_1 &:= 2 \sup_{\|f\|_1 \leq 1} \int_{\mathbb{R}^d} \mathbb{E} \left[ f(X^{x,\varepsilon}_T)(R^\varepsilon_T - 1) \right] \varepsilon^{\varepsilon+r} (dx) \\
I_2 &:= 2 \left| \mathbb{E} \left[ \int_0^r \frac{d}{d\theta} P^\varepsilon_t f(X^{\varepsilon+r}_0) \right] \right|^2 = 2 \left| \mathbb{E} \left[ \int_0^r \left\{ \nabla X^{\varepsilon+r}_0 \cdot P^\varepsilon_t f(X^{\varepsilon+r}_0) \right\} \right] \right|^2.
\end{align*}
\]

Below we estimate \( I_1 \) and \( I_2 \) respectively.

By (2.21), we obtain

\[
\begin{align*}
\mathbb{E} |R^\varepsilon_T - 1|^2 &= \mathbb{E} \left[ (R^\varepsilon_T - 1)^2 \right] \leq \text{esssup}_\Omega (e^{\int_0^T |\eta^\varepsilon_s|^2 ds} - 1) \\
&\leq \text{esssup}_\Omega \left( e^{\int_0^T |\eta^\varepsilon_s|^2 ds} \int_0^T |\eta^\varepsilon_s|^2 ds \right) \\
&\leq \psi(\varepsilon, r) \int_0^T \left\{ \mathbb{W}_\alpha (P^*_s \varepsilon^\varepsilon, P^*_s \varepsilon^{\varepsilon+r})^2 + \mathbb{W}_k (P^*_s \gamma^\varepsilon, P^*_s \gamma^{\varepsilon+r})^2 \right\} ds,
\end{align*}
\]
where for $c_2 := 2c_1^2$,

\begin{equation}
(2.24) \\
\psi(\varepsilon, r) := c_2 e^{c_2 \int_0^T (W_\alpha(P_2 \gamma, P_2 \gamma^+)^2 + W_k(P_2 \gamma, P_2 \gamma^+)^2) ds}.
\end{equation}

By (2.3) and (2.4) for $n = k$, we have

\begin{equation}
(2.25) \\
\bar{\psi}(\varepsilon, \gamma) := \sup_{\varepsilon, \gamma \in [0, 1]} \psi(\varepsilon, \gamma) < \infty.
\end{equation}

Combining this with (2.1), (2.5), (2.23), (2.18) and Lemma 2.4, we can find constants $k_1, k_2 > 1$ such that

\[
\left( \int_{\mathbb{R}^d} \sup_{[f]_{\alpha} \leq 1} \left| \mathbb{E} \left[ f(X_t^{x, \gamma})(R_t^{\gamma, r} - 1) \right] \right| \gamma^{\varepsilon + r}(dx) \right)^2 \\
\leq \left( \int_{\mathbb{R}^d} \alpha \left( k_1(1 + |x|) + \kappa \| \gamma \|_k + \kappa \| \tilde{\gamma} \|_k t^{\frac{1}{2}} \right) \sup_x \left( \mathbb{E}[|R_t^{\gamma, r} - 1|^2] \right)^{\frac{1}{2}} \gamma^{\varepsilon + r}(dx) \right)^2 \\
\leq \alpha \left( k_1(1 + \kappa \gamma^{\varepsilon + r}(| \cdot |) + \kappa \| \gamma \|_k + \kappa \| \tilde{\gamma} \|_k t^{\frac{1}{2}} \right)^2 \sup_x \mathbb{E}[|R_t^{\gamma, r} - 1|^2] \\
\leq k_2 \alpha \left( \left( 1 + \kappa \| \gamma \|_k + \kappa \| \tilde{\gamma} \|_k t^{\frac{1}{2}} \right) \psi(\varepsilon, r) \right)^2 \\
\times \int_0^t \left\{ W_\alpha(P_2 \gamma, P_2 \gamma^+)^2 + W_k(P_2 \gamma, P_2 \gamma^+)^2 \right\} ds, \quad t \in [0, T].
\]

Combining this with (2.1), (2.12), (2.19), (2.18), and letting

$$
\Gamma_t(\varepsilon, r) := W_\alpha(P_2 \gamma, P_2 \gamma^+)^2 + \int_0^t W_\alpha(P_2 \gamma, P_2 \gamma^+)^2 ds,
$$

we find a constant $c_4 > 0$ such that

\begin{equation}
(2.26) \\
I_1 \leq c_4 \alpha \left( 1 + \kappa \| \gamma \|_k + \kappa \| \tilde{\gamma} \|_k t^{\frac{1}{2}} \right)^2 \psi(\varepsilon, r) \left( t^2 \mathbb{E}(\gamma, \tilde{\gamma})^2 + \int_0^t \Gamma_s(\varepsilon, r) ds \right), \quad t \in [0, T].
\end{equation}

Next, by [24, Theorem 2.1 (2)], we have

$$
\nabla_v P_t^{x, \gamma} f(x) = \mathbb{E} \left[ f(X_t^{x, \gamma}) \int_0^t \frac{1}{t} \left( \sigma_s^* (\sigma_s^*)^{-1} \right) (X_s^{x, \gamma}) \nabla_v X_s^{x, \gamma}, dW_s \right], \quad v \in \mathbb{R}^d,
$$

and for any $j \geq 1$ there exists a constant $C_j > 0$ such that

$$
\mathbb{E} \left[ \sup_{s \in [0, T]} |\nabla_v X_s^{x, \gamma}|^j \right] \leq C_j |v|^j, \quad v \in \mathbb{R}^d, \quad \varepsilon \in [0, 1].
$$

Combining this with (2.5), (A) and Lemma 2.4, we find constants $c_5 > 0$ such that

$$
\sup_{[f]_{\alpha} \leq 1} |\nabla P_t^{x, \gamma} f| (x) \leq \frac{c_5}{\sqrt{t}} \alpha \left( (1 + \kappa |x|) + \kappa \| \gamma \|_k + \kappa \| \tilde{\gamma} \|_k t^{\frac{1}{2}} \right).
$$
Combining this with Lemma 2.4 for \( \alpha_k(s) := \alpha(s/(k-1)) \) replacing \( \alpha \), and using (2.1), we find a constant \( c_0 > 0 \) such that

\[
I_2 \leq 2 \left( \mathbb{E} \left[ |X_0^\gamma - X_0^\tilde{\gamma}| \right] \right)^2 \leq \frac{c_0}{t} \left( \int_0^r \|X_0^\gamma - X_0^\tilde{\gamma}\|_{L^k(\mathbb{P})} \right)^2 \times \alpha_k \left( \frac{t^{1/k}}{t} \right) (1 + \kappa \|X_0^{\gamma_k+\tilde{\gamma}_k}\|_{L^k(\mathbb{P})} + \kappa \|\gamma\|_k + \kappa \|\tilde{\gamma}\|_k)^{-1} \right)^2 \leq \frac{c_0 r^2}{t} \mathbb{E}[|X_0^\gamma - X_0^\tilde{\gamma}|^2] \alpha \left(1 + 2\kappa \|\gamma\|_k + 2\kappa \|\tilde{\gamma}\|_k \right)^2.
\]

By (2.1), we find some constant \( c' > 0 \) such that

\[
\int_0^T \frac{\alpha(r^{1/t})^2}{t} \, dt = 2 \int_0^{T^{1/2}} \frac{\alpha(s)^2}{s} \, ds \leq c' \tilde{\alpha}(r)^2 < \infty, \quad r \geq 1.
\]

So, (2.27) together with (2.22) and (2.26) yields that for some constant \( c_7 > 0 \),

\[
\Gamma_t(\varepsilon, r) \leq c_7 r^2 \mathbb{W}_k(\gamma, \tilde{\gamma})^2 H_t(\varepsilon, r) + c_7 F(\varepsilon, r) \int_0^t \Gamma_s(\varepsilon, r) \, ds,
\]

\[
(2.29)
\]

\[
H_t(\varepsilon, r) := \alpha (1 + \kappa \|\gamma\|_k + \kappa \|\tilde{\gamma}\|_k)^2 \psi(\varepsilon, r) + \tilde{\alpha} (1 + \kappa \|\gamma\|_k + \kappa \|\tilde{\gamma}\|_k)^2
\]

\[
+ \alpha ((1 + \kappa \|\gamma\|_k + \kappa \|\tilde{\gamma}\|_k) t^{1/2})^2 t^{-1},
\]

\[
F(\varepsilon, r) := \alpha (1 + \kappa \|\gamma\|_k + \kappa \|\tilde{\gamma}\|_k)^2 \psi(\varepsilon, r), \quad \varepsilon, r \in [0, 1], t \in [0, T].
\]

By Gronwall’s inequality and (2.29), for any \( \varepsilon, r \in [0, 1] \) we have

\[
\mathbb{W}_k(P_t^s \gamma^\varepsilon, P_t^s \gamma^\varepsilon + r) \leq \Gamma_t(\varepsilon, r)
\]

\[
\leq c_7 r^2 \mathbb{W}_k(\gamma, \tilde{\gamma})^2 \left\{ H_t(\varepsilon, r) + c_7 F(\varepsilon, r) e^{c_7 F(\varepsilon, r) T} \int_0^t H_s(\varepsilon, r) \, ds \right\}, \quad t \in [0, T].
\]

This together with (2.12) and (2.24)-(2.25) implies that \( \psi(\varepsilon, r) \) is bounded in \( [\varepsilon, r] \in [0, 1]^2 \) with \( \psi(\varepsilon, r) \to c_2 \) as \( r \to 0 \), so that by the dominated convergence theorem we find a constant \( c > 0 \) such that

\[
\limsup_{r \downarrow 0} \frac{\mathbb{W}_k(\gamma, \tilde{\gamma})}{r} \leq c \mathbb{W}_k(\tilde{\gamma}, \gamma) \left\{ \alpha ((1 + \kappa \|\gamma\|_k + \kappa \|\tilde{\gamma}\|_k) t^{1/2}) \right. \]

\[
+ \tilde{\alpha} (1 + \kappa \|\gamma\|_k + \kappa \|\tilde{\gamma}\|_k) e^{c_\alpha (1 + \kappa \|\gamma\|_k + \kappa \|\tilde{\gamma}\|_k)^2} \right\},
\]

\[
(2.30)
\]
where we have used the fact that for some constant $C > 1$,
\[
\alpha(1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)e^{\alpha(1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)^2} \leq e^{C\alpha(1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)^2}.
\]

By the triangle inequality,
\[
\left|\mathbb{W}_\alpha(P_t^*\gamma, P_t^*\tilde{\gamma}) - \mathbb{W}_\alpha(P_t^*\gamma, P_t^*\gamma^\varepsilon)\right| \leq \mathbb{W}_\alpha(P_t^*\gamma^\varepsilon, P_t^*\tilde{\gamma}^\varepsilon), \quad \varepsilon, r \in [0, 1],
\]
so that the (2.30) implies that $\mathbb{W}_\alpha(P_t^*\gamma, P_t^*\tilde{\gamma})$ is Lipschitz continuous (hence a.e. differentiable) in $\varepsilon \in [0, 1]$ for any $t \in (0, T]$, and
\[
\left|\frac{d}{d\varepsilon}\mathbb{W}_\alpha(P_t^*\gamma, P_t^*\tilde{\gamma})\right| \leq \limsup_{\varepsilon \to 0} \frac{\mathbb{W}_\alpha(P_t^*\gamma^\varepsilon, P_t^*\tilde{\gamma}^\varepsilon)}{\varepsilon}
\leq c\mathbb{W}_k(\tilde{\gamma}, \gamma)\left\{\frac{\alpha((1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)\gamma^2)}{\sqrt{t}} \right.
\left. + \tilde{\alpha}(1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)e^{\alpha(1 + \kappa\|\gamma\|_k + \kappa\|\tilde{\gamma}\|_k)^2}\right\}, \quad \varepsilon \in [0, 1].
\]

Noting that $\gamma^1 = \tilde{\gamma}$, this implies the desired estimate. \hfill \Box

**Proof of Theorem 2.1.** Let $k = 2$ so that $\alpha_k = \alpha$ is concave as needed by Lemma 2.5. According to [23, Theorem 2.3] for $D = \mathbb{R}^d$, see also [27, Theorem 4.1], (A) implies the following log-Harnack inequality for some constant $c_0 > 0$ and any $\gamma \in \mathcal{P}_2$:
\[
P_t^\gamma \log f(x) \leq \log P_t^\gamma f(y) + \frac{c_0}{t}|x - y|^2, \quad x, y \in \mathbb{R}^d, t \in (0, T], f \in \mathcal{B}_b^+(\mathbb{R}^d).
\]

Then by [23, (4.13)], see also [11, Theorem 2.1], it suffices to find a constant $c > 0$ such that
\[
\sup_{t \in (0, T]} \log \mathbb{E}[|R_t^{\gamma, \tilde{\gamma}}|^2] \leq c\tilde{\alpha}(1 + \kappa\|\gamma\|_2 + \kappa\|\tilde{\gamma}\|_2)^2e^{\alpha(1 + \kappa\|\gamma\|_2 + \kappa\|\tilde{\gamma}\|_2)^2}\mathbb{W}_2(\gamma, \tilde{\gamma}), \quad \gamma, \tilde{\gamma} \in \mathcal{P}_2,
\]
where
\[
R_t^{\gamma, \tilde{\gamma}} := e^{\int_0^t (\eta_s^{\gamma, \tilde{\gamma}}, dW_s) - \frac{1}{2} \int_0^t |\eta_s^{\gamma, \tilde{\gamma}}|^2 ds},
\eta_s^{\gamma, \tilde{\gamma}} := \left\{\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\right\}(X^\gamma)\left\{b_s(X_s^\gamma, P_s^\gamma) - b_s(X_s^\gamma, P_s^\gamma)\right\}, \quad s \leq t \leq T.
\]
Noting that (A) implies
\[
|\eta_s^{\gamma, \tilde{\gamma}}|^2 \leq c_1\left\{\mathbb{W}_\alpha(P_s^\gamma, \tilde{\gamma})^2 + \mathbb{W}_2(P_s^\gamma, \tilde{\gamma})^2\right\}, \quad s \in [0, T]
\]
for some constant $c_1 > 0$, we have
\[
\mathbb{E}[|R_t^{\gamma, \tilde{\gamma}}|^2] \leq e^{c_1 \int_0^t [\mathbb{W}_\alpha(P_s^\gamma, \tilde{\gamma})^2 + \mathbb{W}_2(P_s^\gamma, \tilde{\gamma})^2] ds}.
\]
Moreover, by (2.28) and Lemma 2.5, there exists a constant $c > 0$ such that

$$
\sup_{t \in [0,T]} \int_0^t \left\{ \mathbb{W}_\alpha(P^*_s \gamma, P^*_s \tilde{\gamma})^2 + \mathbb{W}_2(P^*_s \gamma, P^*_s \tilde{\gamma})^2 \right\} ds \\
\leq c\tilde{\alpha}(1 + \kappa \|\gamma\|_2 + \kappa \|\tilde{\gamma}\|_2)^2 e^{c\alpha(1 + \kappa \|\gamma\|_2 + \kappa \|\tilde{\gamma}\|_2)^2} \mathbb{W}_2(\gamma, \tilde{\gamma})^2.
$$

Therefore, (2.31) holds for some constant $c > 0$. \qed

### 3 Bismut Formula

Let $k \in (1, \infty)$ and denote $k^* := \frac{k}{k-1}$. In this part, we establish the Bismut formula for the intrinsic derivative of $P_t f(\gamma)$ for $\gamma \in \mathcal{P}_k$. To this end, we assume

(3.1) \quad b_t(x, \mu) = b_t^0(x) + B_t(x, \mu, \mu(V)), \quad t \in [0,T], x \in \mathbb{R}^d, \mu \in \mathcal{P},

where for a Banach space $(\mathbb{B}, \| \cdot \|_\mathbb{B})$,

$$
V : \mathbb{R}^d \to \mathbb{B}, \quad B : [0,T] \times \mathbb{R}^d \times \mathcal{P} \times \mathbb{B} \to \mathbb{R}^d
$$

are measurable such that $[V]_\alpha \leq 1$ for some $\alpha \in \mathcal{A}$, i.e. $V$ is only square root Dini continuous and hence $\mu \mapsto B_t(x, \mu, \mu(V))$ may be not intrinsically differentiable. The following definition is taken from [2] using the idea of [1].

**Definition 3.1.** Let $f \in C(\mathcal{P}_k; \mathbb{B})$ for a Banach space $\mathbb{B}$. $f$ is called intrinsically differentiable at a point $\mu \in \mathcal{P}_k$, if

$$
T_{\mu,k} := L^k(\mathbb{R}^d \to \mathbb{R}^d, \mu) \ni \phi \mapsto D_{\phi}^I f(\mu) := \lim_{\varepsilon \downarrow 0} \frac{f(\mu \circ (id + \varepsilon \phi)^{-1}) - f(\mu)}{\varepsilon} \in \mathbb{B}
$$

is a well defined bounded linear operator. In this case, the norm of the intrinsic derivative $D^I f(\mu)$ is given by

$$
\|D^I f(\mu)\|_{L^k(\mu)} := \sup_{\|\phi\|_{L^k(\mu)} \leq 1} \|D^I_{\phi} f(\mu)\|_\mathbb{B}.
$$

If moreover

$$
\lim_{\|\phi\|_{\mathcal{L}^k(\mu)} \downarrow 0} \frac{\|f(\mu \circ (id + \phi)^{-1}) - f(\mu) - D^I_{\phi} f(\mu)\|_\mathbb{B}}{\|\phi\|_{T_{\mu,k}}} = 0,
$$

then $f$ is called $L$-differentiable at $\mu$. The function $f$ is called intrinsically differentiable or $L$-differentiable on $\mathcal{P}_k$, if it is so at any $\mu \in \mathcal{P}_k$. 

14
Recall that a real function $f$ on a Banach space $B$ is called Gateaux differentiable, if for any $z \in B$,

$$v \mapsto \nabla^B_v f(z) := \lim_{\varepsilon \downarrow 0} \frac{f(z + \varepsilon v) - f(z)}{\varepsilon}$$

is a well-defined bounded linear functional. In this case we denote

$$\|\nabla^B f(z)\|_{B^*} := \sup_{\|v\|_B \leq 1} |\nabla^B_v f(z)|.$$  
Moreover, $f$ is called Fréchet differentiable if it is Gateaux differentiable and

$$\lim_{\|v\|_B \downarrow 0} \frac{|f(z + v) - f(z) - \nabla^B_v f(z)|}{\|v\|_B} = 0, \quad z \in B.$$  

It is well-known that a Gateaux differentiable function $f$ is Fréchet differentiable provided $\nabla^B_v f(z)$ is continuous in $(v, z) \in B \times B$.

When $B = \mathbb{R}^l$ for some $l \geq 1$, we simply denote $\nabla^B = \nabla$.

### 3.1 Main result

We will establish Bismut formula for the intrinsic derivative of $P_t f$ under the following assumption.

- **(B)** Let $k \in (1, \infty)$ and let $b$ in (3.1).

- **(B1)** $b(0)$ and $\sigma$ satisfy the corresponding conditions in (A).

- **(B2)** There exists $\alpha \in \mathcal{A}$ such that $\alpha_k(s) := \alpha(s^{1/k})$ is concave in $s \geq 0$, and

$$[V]_\alpha := \sup_{x \neq y} \frac{\|V(x) - V(y)\|_B}{\alpha(|x - y|)} \leq 1.$$  

- **(B3)** For any $t \in [0, T]$, $B_t \in C(\mathbb{R}^d \times \mathcal{P}_k \times \mathbb{B})$, $B_t(x, \mu, z)$ is differentiable in $x$, $L$-differentiable in $\mu \in \mathcal{P}_k$, and Fréchet differentiable in $z \in \mathbb{B}$ such that $\nabla^B_v B_t(x, \mu, z)$ is continuous in $(v, z) \in \mathbb{B} \times \mathbb{B}$. Next, there exist constants $K > 0$ and $\kappa \geq 0$ such that

$\begin{align*}
|B_t(x, \mu, z)| &\leq K + \kappa(|x| + \|\mu\|_k + \|z\|_B), \\
|\nabla B_t(\cdot, \mu, z)(x)| + \|D^L B_t(x, \cdot, \mu)(\cdot)(\mu)\|_{L^k(\mu)} + \|\nabla^B B_t(x, \mu, \cdot)(z)\|_{B^*} &\leq K, \\
(t, x, \mu, z) &\in [0, T] \times \mathbb{R}^d \times \mathcal{P}_k \times \mathbb{B}.
\end{align*}$

Moreover, for any $(t, x, \mu, z) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_k \times \mathbb{B}$, there exists a constant $c(t, x, \mu, z) > 0$ such that

$$|\{D^L B_t(x, \cdot, z)(\mu)(y)| \leq c(t, x, \mu, z)(1 + |y|^{k-1}), \quad y \in \mathbb{R}^d.$$
Since $\alpha$ is concave, the concavity of $\alpha_k$ holds for $k \geq 2$. When $\alpha(s) = s^\varepsilon$ for some $\varepsilon \in (0, 1)$, $\alpha_k$ is concave for $k \geq 1 + \varepsilon$. Since (B) implies (A), as explained before that under this assumption (1.1) is well-posed for distributions in $\mathcal{P}_k$.

For $\mu \in \mathcal{P}_k$, consider the decoupled SDE

$$dX_t^{x,\mu} = \left\{ b_t^{(0)}(X_t^{x,\mu}) + B_t(X_t^{x,\mu}, P_t^*\mu, P_t V(\mu)) \right\} dt + \sigma_t(X_t^{x,\mu}) dW_t,$$

$$X_0^{x,\mu} = x, t \in [0, T].$$

(3.2)

According to [24, Theorem 1.1(1) and Theorem 1.2], (B) implies that (3.2) is well-posed,

$$\nabla_vX_t^{x,\mu} := \lim_{\varepsilon \downarrow 0} \frac{X_t^{x+\varepsilon v,\mu} - X_t^{x,\mu}}{\varepsilon}, \quad t \in [0, T]$$

exists in $L^p(\Omega \to C([0, T]; \mathbb{R}^d); \mathbb{P})$ for any $p \geq 1$, and there exists a constant $c_p > 0$ such that

(3.3)

$$\mathbb{E}\left[ \sup_{t \in [0, T]} |\nabla_vX_t^{x,\mu}|^p \right] \leq c_p |v|^p, \quad v \in \mathbb{R}^d, \mu \in \mathcal{P}_k, x \in \mathbb{R}^d.$$

To state the Bismut formula for $P_t f$, we introduce the following $I_t^f$ which comes from the Bismut formula presented in [24]: for fixed $t \in (0, T]$, let

(3.4)

$$I_t^f(\mu, \phi) := \frac{1}{t} \int_{\mathbb{R}^d} \mathbb{E}\left[ f(X_t^{x,\mu}) \int_0^t \left\langle \zeta_s(X_s^{x,\mu}) \nabla \phi(x), X_s^{x,\mu} \right\rangle dW_s \right] \mu(dx),$$

$$\zeta_s := \sigma^*_s(\sigma_s\sigma^*_s)^{-1}, \quad s \in [0, t], \mu \in \mathcal{P}_k, \phi \in T_{\mu,k}.$$

By $(B_1)$ and (3.3), we find a constant $c > 0$ such that

(3.5)

$$|I_t^f(\mu, \phi)| \leq \frac{c}{\sqrt{t}} (P_t|f|^{k^*}(\mu))^{\frac{1}{k^*}} \|\phi\|_{L_{k}(\mu)}, \quad \mu \in \mathcal{P}_k, \phi \in T_{\mu,k}.$$

Next, let $X_0^\mu$ be $\mathcal{F}_0$-measurable such that $\mathbb{L}X_0^\mu = \mu$, and let $X_t^\mu$ solve (1.1) with initial value $X_0^\mu$. For any $\varepsilon \geq 0$, denote

$$\mu_\varepsilon := \mu \circ (id + \varepsilon \phi)^{-1}, \quad X_0^{\mu_\varepsilon} := X_0^\mu + \varepsilon \phi(X_0^\mu).$$

Let $X_t^{\mu_\varepsilon}$ solve (1.1) with initial value $X_0^{\mu_\varepsilon}$. So,

$$X_t^\mu = X_t^{i_0^\mu}, \quad P_t^*\mu_\varepsilon = \mathbb{L}X_t^{\mu_\varepsilon}, \quad t \in [0, T], \varepsilon \geq 0.$$

We will study

(3.6)

$$\nabla_\phi X_t^{\mu} := \lim_{\varepsilon \downarrow 0} \frac{X_t^{\mu_\varepsilon} - X_t^{\mu}}{\varepsilon}, \quad t \in [0, T].$$
Theorem 3.1. Assume (B) and let $\zeta_s$ and $I_t^f$ be in (3.4). Then the following assertions hold.

(1) For any $t \in (0, T]$, $P_t \nu$ is intrinsically differentiable on $\mathcal{P}_k$, and there exists a constant $c > 0$ such that
\[
\|D^f P_t \nu(\mu)\|_{L^{k^*}(\mu)} \leq \frac{c \alpha((1 + \kappa \|\nu\|_k) \eta_{t^2})}{\sqrt{t}} e^{c \tilde{\alpha}(1 + \kappa \|\nu\|_k)^2}, \quad t \in (0, T], \mu \in \mathcal{P}_k.
\]

(2) The limit in (3.6) exists in $L^k(\Omega \to C([0, T], \mathbb{R}^d), \mathbb{P})$, and there exists a constant $c > 0$ such that
\[
\mathbb{E}\left[ \sup_{t \in [0, T]} |\nabla_\phi X^\mu_{t^k}| \right] \leq c \|\phi\|_{L^k(\mu)}, \quad \mu \in \mathcal{P}_k.
\]

(3) For any $t \in (0, T]$ and $f \in \mathcal{B}_b(\mathbb{R}^d)$, $P_t f$ is intrinsically differentiable on $\mathcal{P}_k$. Moreover, for any $\mu \in \mathcal{P}_k$ and $\phi \in T_{\mu,k}$,
\[
D^f P_t f(\mu) = f_t^f(\mu, \phi) + \mathbb{E}\left[ f(X^\mu_s) \int_t^t \left\{ \zeta_s(X^\mu_s) \{ N_s + \tilde{N}_s \} \right\} \mathrm{d}W_s \right],
\]
\[
N_s := \left\{ \nabla_{D^f P_t \nu(\mu)} B_s(X^\mu_s, P_s^\mu, \cdot) \right\}(P_s \nu(\mu)),
\]
\[
\tilde{N}_s := \mathbb{E}\left[ \left\{ D^f B_s(y, \cdot, P_s \nu(\mu)) \right\}(P_s^\mu)(X^\mu_s, \nabla_\phi X^\mu_s) \right|_{y = X^\mu_t},
\]
where $X^\mu_t$ solves (1.1) with initial distribution $\mathcal{L}_{X_0} = \mu$.

By (3.5) and (3.7), we find a constant $c > 0$ such that
\[
\|D^f P_t f(\mu)\|_{L^{k^*}(\mu)} \leq \frac{c \|f\|_{L^{k^*}(\mu)}}{\sqrt{t}} e^{c \tilde{\alpha}(1 + \kappa \|\nu\|_k)^2}, \quad t \in (0, T], f \in \mathcal{B}_b(\mathbb{R}^d), \mu \in \mathcal{P}_k.
\]

3.2 Proof of Theorem 3.1

By the definition of the intrinsic derivative, we intend to calculate
\[
D^f P_t f(\mu) := \lim_{\varepsilon \downarrow 0} \frac{P_t f(\mu_{\varepsilon}) - P_t f(\mu)}{\varepsilon} = \lim_{\varepsilon \downarrow 0} \frac{\mathbb{E}[f(X^\mu_t^\varepsilon) - f(X^\mu_t)]}{\varepsilon}.
\]

To this end, for any $\gamma \in \mathcal{P}_k$, consider the decoupled SDE (3.2) for $\gamma$ replacing $\mu$, i.e.
\[
dx_t^{x,\gamma} = \left\{ b_t^{(0)}(X_t^{x,\gamma}) + B_t(X_t^{x,\gamma}, P_t^\gamma, P_t V(\gamma)) \right\} \mathrm{d}t + \sigma_t(X_t^{x,\gamma}) \mathrm{d}W_t,
\]
\[
t \in [0, T], X_0^{x,\gamma} = x \in \mathbb{R}^d.
\]

Let
\[
P_t^\gamma f(x) := \mathbb{E}[f(X_t^{x,\gamma})], \quad x \in \mathbb{R}^d,
\]
\[
P_t^\gamma f(\gamma) := \int_{\mathbb{R}^d} P_t^\gamma f(\gamma) \mathrm{d}\gamma, \quad t \geq 0, f \in \mathcal{B}_b(\mathbb{R}^d), \gamma, \tilde{\gamma} \in \mathcal{P}_k.
\]
For \( \varepsilon \geq 0 \), let \( X_t^{\mu, \varepsilon} \) be the solution of (3.9) with initial value \( X_0^\varepsilon \), i.e,

\[
dX_t^{\mu, \varepsilon} = \left\{ b_t^{(0)}(X_t^{\mu, \varepsilon}) + B_t(X_t^{\mu, \varepsilon}; P_t^\gamma, P_t V(\gamma)) \right\} dt + \sigma_t(X_t^{\mu, \varepsilon}) dW_t,
\]

\( t \in [0, T], \ X_0^{\mu, \varepsilon} = X_0^\varepsilon. \)

Then \( X_t^{\mu, \varepsilon} \) solves (1.1) with initial value \( X_0^\varepsilon \), so that

\[
P_t f(\mu_\varepsilon) = P_t^{\mu_\varepsilon} f(\mu_\varepsilon) = \mathbb{E}[f(X_t^{\mu, \varepsilon})], \ \varepsilon \geq 0, t \in [0, T], f \in \mathcal{B}_b(\mathbb{R}^d).
\]

Noting that \( \mu_0 = \mu \), (3.8) reduces to

\[
D_{\phi}^t P_t f(\mu) = \lim_{\varepsilon \downarrow 0} \frac{P_t^{\mu_\varepsilon} f(\mu_\varepsilon) - P_t^\mu f(\mu)}{\varepsilon} = \lim_{\varepsilon \downarrow 0} \left\{ \frac{P_t^{\mu} f(\mu_\varepsilon)}{\varepsilon} - \frac{P_t^{\mu_\varepsilon} f(\mu)}{\varepsilon} \right\}.
\]

(3.10)

So, to calculate \( D_{\phi}^t P_t f(\mu) \), we only need to study the limits of

\[
J_1 f(t, \varepsilon) := \frac{P_t^{\mu} f(\mu_\varepsilon) - P_t^\mu f(\mu)}{\varepsilon}, \ J_2 f(t, \varepsilon) := \frac{P_t^{\mu_\varepsilon} f(\mu_\varepsilon) - P_t^{\mu} f(\mu_\varepsilon)}{\varepsilon}.
\]

By [24, Theorem 1.2(2)] for \( \beta_s = \frac{\gamma}{2} \), for any \( t \in (0, T], \varepsilon \geq 0 \) and \( f \in \mathcal{B}_b(\mathbb{R}^d) \), we have

\[
\frac{d}{d\varepsilon} P_t^{\mu} f(\mu_\varepsilon) := \lim_{r \downarrow 0} \frac{P_t^{\mu} f(\mu_{\varepsilon+r}) - P_t^{\mu} f(\mu_{\varepsilon})}{r}
\]

\[
= \int_{\mathbb{R}^d} \mathbb{E} \left[ f(X_t^{x+\varepsilon \phi(x), \mu}) \frac{1}{t} \int_0^t \langle \xi_t(X_t^{x+\varepsilon \phi(x), \mu}) \nabla \phi(x), X_t^{x+\varepsilon \phi(x), \mu} \rangle dW_s \right] \mu(dx
\]

\[
=: I_1^t(\varepsilon, \mu, \phi), \ t \in (0, T).
\]

In particular,

\[
\lim_{\varepsilon \downarrow 0} \frac{P_t^{\mu} f(\mu_\varepsilon) - P_t^{\mu} f(\mu)}{\varepsilon} = I_1^t(\mu, \phi), \ t \in (0, T).
\]

Let \( u_t = u_t^{\lambda_0, \mu}, \Theta_t = \Theta_t^{\lambda_0, \mu} \) be defined in (2.7) and (2.10) for large enough \( \lambda_0 > 0 \) such that (2.9) holds for \( \lambda = \lambda_0 \), and let

\[
C_k := \{ v \in L^k(\Omega \to C([0, T]; \mathbb{R}^d), \mathbb{P}) \text{ is } \mathcal{F}_t\text{-adapted}, v_0 = \nabla_{\phi(X_0^\mu)} \Theta_0(X_0^\mu) \}
\]

which is complete under the metric

\[
\rho_{\lambda}(v^1, v^2) := \left( \mathbb{E} \sup_{t \in [0, T]} e^{-\lambda_0 k} |v_t^1 - v_t^2|^k \right)^{\frac{1}{k}}, \ v^1, v^2 \in C_k, \lambda > 0.
\]
For any \( s \in [0, T] \), \( v \in L^k(\Omega \to \mathbb{R}^d, \mathbb{P}) \), define
\[
\psi_s(v) := \mathbb{E} \left[ \langle \{D^\mu B_s(y, \cdot, P_s V(\mu))\}(P_s^* \mu)(X_s^\mu), (\nabla \Theta_s(X_s^\mu))^{-1} v \rangle \right] |_{y = X_s^\mu}.
\]
To calculate the limit of \( J_2 f(t, \varepsilon) \) as \( \varepsilon \to 0 \), for any \((v, \bar{v}) \in C([0, T]; \mathbb{B}) \times \mathcal{C}_k \), we denote
\[
(3.14) \quad N_s(v) := \nabla^2_{\psi_s}(X_s^\mu, P_s^* \mu, \cdot)(P_s V(\mu)), \quad \tilde{N}_s(\bar{v}) := \psi_s(\bar{v}),
\]
and write
\[
\begin{align*}
\tilde{I}_t^V(\mu, v, \bar{v}) &= \tilde{I}_t^{V, 1}(\mu, v, \bar{v}) + \tilde{I}_t^{V, 2}(\mu, v, \bar{v}), \\
\tilde{I}_t^{V, 1}(\mu, v, \bar{v}) &= \mathbb{E} \left[ V(X_t^\mu) \int_0^t \frac{\alpha((1 + \kappa \|\mu\|_k) s^\frac{1}{2})}{\{\tilde{\alpha}((1 + \kappa \|\mu\|_k) t^\frac{1}{2}) t\}} \langle \zeta_s(\mu) N_s(v), dW_s \rangle \right], \\
\tilde{I}_t^{V, 2}(\mu, v, \bar{v}) &= \mathbb{E} \left[ V(X_t^\mu) \int_0^t \langle \zeta_s(\mu) \tilde{N}_s(\bar{v}), dW_s \rangle \right], \quad t \in [0, T].
\end{align*}
\]
Consider the following equation for \((v, \bar{v}) \in C([0, T]; \mathbb{B}) \times \mathcal{C}_k\):
\[
\begin{align*}
d\bar{v}_t &= \left[ \nabla(\nabla \Theta_t(X_t^\mu))^{-1} \tilde{v}_t \right] \left\{ B_t(\mu, P_t^* \mu, P_t V(\mu)) + \lambda_0 u_t \right\}(X_t^\mu) dt \\
&\quad + \nabla \psi_t(\bar{v}_t) \Theta_t(X_t^\mu) dt + \frac{\alpha((1 + \kappa \|\mu\|_k) t^\frac{1}{2})}{\{\tilde{\alpha}((1 + \kappa \|\mu\|_k) t^\frac{1}{2}) t\}} \nabla N_t(v) \Theta_t(X_t^\mu) dt \\
&\quad + \left[ \nabla(\nabla \Theta_t(X_t^\mu))^{-1} \tilde{v}_t \right] \left\{ (-\nabla \Theta_t) \sigma_t \right\}(X_t^\mu) dW_t, \quad \bar{v}_0 = \nabla \psi(X_0^\mu) \Theta_0(X_0^\mu), \\
v_t &= \frac{t\tilde{\alpha}((1 + \kappa \|\mu\|_k) t^\frac{1}{2})}{\alpha((1 + \kappa \|\mu\|_k) t^\frac{1}{2})} \left\{ tI_t^V(\mu, \phi) + \tilde{I}_t^V(\mu, v, \bar{v}) \right\}, \quad t \in [0, T].
\end{align*}
\]
We will prove that this equation has a unique solution and
\[
(3.15) \quad D_{\phi}^t P_t V(\mu) = \frac{\alpha((1 + \kappa \|\mu\|_k) t^\frac{1}{2})}{\{t\tilde{\alpha}((1 + \kappa \|\mu\|_k) t^\frac{1}{2})\}} v_t, \quad t \in (0, T],
\]
\[
\nabla \phi X_t^\mu := \lim_{\varepsilon \downarrow 0} \frac{X_t^{\mu + \varepsilon} - X_t^\mu}{\varepsilon} = (\nabla \Theta_t(X_t^\mu))^{-1} \tilde{v}_t, \quad t \in [0, T].
\]

**Lemma 3.2.** Assume (B). For any \( \mu \in \mathcal{P}_k \) and \( \phi \in T_{\mu,k} \), the equation (3.14) has a unique solution \( \{(v_t, \bar{v}_t) = (v_t(\mu, \phi), \bar{v}_t(\mu, \phi))\}_{t \in [0, T]} \), and there exists a constant \( c > 0 \) such that
\[
(3.16) \quad \sup_{\|\phi\|_{L^k(\mu)} \leq 1} \|v_t(\mu, \phi)\|_{\mathbb{B}} \leq e^{c\tilde{\alpha}(1 + \kappa \|\mu\|_k)^2} \sqrt{\tilde{\alpha}((1 + \kappa \|\mu\|_k) t^\frac{1}{2})}, \quad \mu \in \mathcal{P}_k, t \in [0, T],
\]
and for any \( j \geq 1 \),
\[
(3.17) \quad \mathbb{E} \left[ \sup_{t \in [0, T]} |\bar{v}_t(\mu, \phi)|^j |\mathcal{F}_0 \right] \leq |\phi(X_0^\mu)|^j + \|\phi\|_{L^k(\mu)}^j e^{c\tilde{\alpha}(1 + \kappa \|\mu\|_k)^2}, \quad \mu \in \mathcal{P}_k, t \in [0, T].
\]
Proof. By Jensen’s inequality, for estimate (3.17) we only need to consider $j \geq k$.

(a) As in the proof of [24, Lemma 4.1], there exist constants $\theta > 1$ and $0 < f \in \mathcal{B}(0, T \times \mathbb{R}^d)$ such that

$$\|\nabla \Theta \sigma\| \leq f, \quad \mathbb{E} \int_0^T f_s(X^\mu_s)^{2\theta} ds < \infty.$$  \hspace{1cm} (3.18)

For any $n \geq 0$, define

$$\tau_n := T \land \inf \left\{ t \geq 0 : \int_0^t f_s(X^\mu_s)^{2\theta} ds \geq n \right\},$$

$$\mathcal{C}_{k,n} := \left\{ (z \land \tau_n), \quad z \in \mathcal{C}_k \right\}.$$  \hspace{1cm} (3.19)

We have $\lim_{n \to \infty} \tau_n = T$. Let

$$\mathcal{V}_0 := \left\{ v \in C([0, T]; \mathbb{B}) : v_0 = 0 \right\}.$$  \hspace{1cm} (3.20)

By (3.18), $\|\nabla B\|_\infty < \infty$ due to (B), and

$$\int_0^t \frac{\alpha(s)\bar{t}}{s\alpha(s)} ds = 2 \int_0^{rt^\frac{1}{\bar{t}}} \frac{\alpha(s)}{s\alpha(s)} ds = 4 \int_0^{rt^\frac{1}{\bar{t}}} \bar{\alpha}(s) ds = 4 \bar{\alpha}(rt^\frac{1}{\bar{t}}), \quad r \geq 0,$$  \hspace{1cm} (3.21)

we see that for any $(v, \tilde{v}) \in \mathcal{V}_0 \times \mathcal{C}_k,$

$$H_s(v, \tilde{v}) := \tilde{v}_0 + \int_0^s \nabla \psi_t(\tilde{v}_t) \Theta_t(X^\mu_t) dt$$

$$+ \int_0^s \left\{ \nabla \Theta_t(X^\mu_t) \bar{v}_t \right\} \left( B_t(\cdot, P^*_t \mu, P_t V(\mu)) + \lambda_0 u_t \right) \Theta_t(X^\mu_t) dt$$

$$+ \int_0^s \left\{ \bar{\alpha}((1 + \kappa \| \mu \|_k) t^\frac{1}{\bar{t}}) \nabla \Theta_t(X^\mu_t) \right\} \left( B_t(\cdot, P^*_t \mu, P_t V(\mu)) \Theta_t(X^\mu_t) dt$$

$$+ \int_0^s \left\{ \nabla \Theta_t(X^\mu_t) \right\} \left( B_t(\cdot, P^*_t \mu, P_t V(\mu)) \Theta_t(X^\mu_t) dt$$

is a continuous adapted process on $\mathbb{R}^d$ such that

$$H_{\land \tau_n} : \mathcal{V}_0 \times \mathcal{C}_{k,n} \to \mathcal{C}_{k,n}, \quad n \geq 1.$$  \hspace{1cm} (3.22)

On the other hand, let

$$H_t(v, \tilde{v}) := \frac{t\bar{\alpha}((1 + \kappa \| \mu \|_k) t^\frac{1}{\bar{t}})}{\alpha((1 + \kappa \| \mu \|_k) t^\frac{1}{\bar{t}})} \left\{ I_t^\nu (\mu, \phi) + \tilde{I}_t^\nu (\mu, v, \tilde{v}) \right\}.$$  \hspace{1cm} (3.23)

We claim that it suffices to prove
(i) for each \( n \geq 1 \), the map

\[
\Phi^{(n)} := (\bar{H}, H_{\Lambda n}) : \mathcal{V}_0 \times \mathcal{C}_{k,n} \to \mathcal{V}_0 \times \mathcal{C}_{k,n}
\]

is well-defined and has a unique fixed point \((v^{(n)}, \bar{v}^{(n)})\).

(ii) there exists a constant \( c > 0 \) such that for any \((t, \mu) \in [0, T] \times \mathcal{P}_k\) and \( n \geq 1 \),

\[
\sup_{p \in [0, T]} \|v^{(n)}_t\|_{L^p(\mu)} \leq c e^{c(1+\kappa\|\mu\|_k)^2} \sqrt{\alpha((1+\kappa\|\mu\|_k)t^2)},
\]

\[
\mathbb{E} \left[ \sup_{t \in [0, T]} |\bar{v}^{(n)}_t|^2 |\mathcal{F}_0 \right] \leq |\phi(X_0^\mu)|^2 + \|\phi\|_{L^p(\mu)}^2 e^{c(1+\kappa\|\mu\|_k)^2}.
\]

Indeed, by (i), the fixed points are consistent:

\[
(v^{(n)}, \bar{v}^{(n)})_{t \leq \tau_n} = (v^{(n+i)}, \bar{v}^{(n+i)})_{t \leq \tau_n}; \quad n, i \geq 1, t \in [0, T],
\]

so that \( \mathbb{P} \)-a.s.

\[
(v(\mu, \phi), \bar{v}(\mu, \phi)) := \lim_{n \to \infty} (v^{(n)}, \bar{v}^{(n)})_{t \leq \tau_n} = \sum_{n=1}^{\infty} 1_{[\tau_{n-1}, \tau_n)} (v^{(n)}, \bar{v}^{(n)})_{t \leq \tau_n}
\]

exists in \( C([0, T]; \mathbb{B} \times \mathbb{R}^d) \), and by (ii) the estimates (3.16) and (3.17) hold. Combining this with

\[
(3.21) \quad \Phi^{(n)}(v^{(n)}, \bar{v}^{(n)}) := (\bar{H}, H_{\Lambda n}) (v^{(n)}, \bar{v}^{(n)}) = (v^{(n)}, \bar{v}^{(n)}), \quad n \geq 1,
\]

we see that \((v(\mu, \phi), \bar{v}(\mu, \phi))\) solves (3.14), and the uniqueness follows from that of the fixed point for \( \Phi^{(n)}, n \geq 1 \).

(b) We first verify

\[
(3.22) \quad \bar{H} : \mathcal{V}_0 \times \mathcal{C}_k \to \mathcal{V}_0,
\]

so that \( \Phi^{(n)} : \mathcal{V}_0 \times \mathcal{C}_{k,n} \to \mathcal{V}_0 \times \mathcal{C}_{k,n} \) for every \( n \geq 1 \). By (3.4), we have

\[
I_t^V(\mu, \phi) = \frac{1}{t} \int_{\mathbb{R}^d} \mathbb{E} \left[ \left\{ V(X_t^\mu) \right\} \int_0^t \langle \zeta_s(x^\mu_s) \nabla \phi(x) X_s^\mu, dW_s \rangle \right] \mu(dx).
\]

By \([V]_\alpha \leq 1, \ (2.16)\) in Lemma 2.4 for \( p = 2 \), \((2.15)\) in Lemma 2.4 for \((p, \alpha, \mathbb{P}) = (k, \alpha_k, \mu)\), \((2.1)\) and \((2.5)\), we find constants \( c_1, c_2 > 0 \) such that

\[
\|I_t^V(\mu, \phi)\|_{\mathbb{B}} \leq \frac{c_1}{\sqrt{t}} \int_{\mathbb{R}^d} \alpha((1 + \kappa|x| + \kappa\|\mu\|_k) t^{\frac{d}{2}})|\phi(x)|\mu(dx)
\]

\[
= \frac{c_1}{\sqrt{t}} \int_{\mathbb{R}^d} \alpha_k((1 + \kappa |x| + \kappa\|\mu\|_k)^{k-1} l^{k-1})|\phi(x)|\mu(dx)
\]

\[
\leq \frac{c_2}{\sqrt{t}} \alpha_k((1 + \kappa\|\mu\|_k)^{k-1} l^{k-1})\|\phi\|_{L^k(\mu)}
\]

\[
= \frac{c_2}{\sqrt{t}} \alpha((1 + \kappa\|\mu\|_k)^{\frac{d}{2}})\|\phi\|_{L^k(\mu)}, \quad t \in (0, T], \mu \in \mathcal{P}_k, \phi \in T_{\mu,k}.
\]
To estimate $\tilde{H}_t^V(\mu, v, \tilde{v})$, noting that $X_t^\mu$ solves (3.9) for $\gamma = \mu$ and initial value $X_0^\mu$ replacing $x$, by the standard Markov property for solutions of (3.9), we have

$$
\tilde{I}_t^{V,1}(\mu, v, \tilde{v}) = \int_{\mathbb{R}^d} E \left[ V(X_t^x, \mu) \int_0^t \alpha((1 + \kappa \|\mu\|_k) s^\frac{1}{2}) \right. \\
\left. \times \left( \zeta_s(X_s^x, \mu) \nabla^B \bar{B}_s(X_s^x, \mu, P_s^\mu, \cdot)(P_s V(\mu), \ dW_s) \right) \right] \mu(dx),
$$

$$
\tilde{I}_t^{V,2}(\mu, v, \tilde{v}) = \int_{\mathbb{R}^d} E \left[ V(X_t^x, \mu) \int_0^t \zeta_s(X_s^x, \mu) \right. \\
\left. \times E \left[ \left\{ \{ D^2 \bar{B}_s(y, \cdot, P_s V(\mu)) \} \{ P_s^\mu(X_s^x) \} \{ \nabla \Theta_s(X_s^x)^{-1} \tilde{v}_s \} \right\} \right] \mu(dx).
$$

So, by (B2), (B3), (2.16) in Lemma 2.4 for $p = 2$, (2.5) and Jensen’s inequality, we find a constant $c_3 > 0$ such that

$$
\| \tilde{I}_t^{V,1}(\mu, v, \tilde{v}) \|_B \leq c_3 \left( \int_0^t \frac{\alpha((1 + \kappa \|\mu\|_k) s^\frac{1}{2})^2}{s \tilde{\alpha}((1 + \kappa \|\mu\|_k) s^\frac{1}{2})} \|v_s\|^2_B ds \right)^{\frac{1}{2}} \alpha((1 + \kappa \|\mu\|_k) t^\frac{1}{2}),
$$

\[ (3.24) \]

$$
\| \tilde{I}_t^{V,2}(\mu, v, \tilde{v}) \|_B \leq c_3 \alpha((1 + \kappa \|\mu\|_k) t^\frac{1}{2}) \left( \int_0^t \left( E |\tilde{v}_s|^k \right)^{\frac{1}{2}} ds \right)^{\frac{1}{2}}.
$$

Combining this with (3.19), (3.20) and (3.23), we find a constant $c_4 > 0$ such that

$$
\| \tilde{H}_t(v, \tilde{v}) \|_B \leq c_4 \|\phi\|_{L^k(\mu)} \sqrt{\tilde{\alpha}((1 + \kappa \|\mu\|_k) t^\frac{1}{2})} \\
+ c_4 \sqrt{\frac{t \tilde{\alpha}((1 + \kappa \|\mu\|_k) t^\frac{1}{2})}{\alpha((1 + \kappa \|\mu\|_k) s^\frac{1}{2})}} \left( \int_0^t \frac{\alpha((1 + \kappa \|\mu\|_k) s^\frac{1}{2})^2}{s \tilde{\alpha}((1 + \kappa \|\mu\|_k) s^\frac{1}{2})} \|v_s\|^2_B ds \right)^{\frac{1}{2}} \\
+ c_4 \sqrt{\frac{t \tilde{\alpha}((1 + \kappa \|\mu\|_k) t^\frac{1}{2})}{\alpha((1 + \kappa \|\mu\|_k) s^\frac{1}{2})}} \left( \int_0^t \left( E |\tilde{v}_s|^k \right)^{\frac{1}{2}} ds \right)^{\frac{1}{2}}.
$$

\[ (3.25) \]

Then (3.22) holds.

(c) We intend to prove that $\Phi^{(n)}$ in (i) has a unique fixed point in $\mathcal{Y}_0 \times \mathcal{C}_{k,n}$. Obviously, for any $\lambda > 0$, $\mathcal{Y}_0 \times \mathcal{C}_{k}$ is complete under the metric

$$
\rho_{\lambda, n}((v, \tilde{v}), (w, \tilde{w})) := \sup_{t \in [0,T]} e^{-\lambda t} \|v_t - w_t\|_B + \rho_{\lambda, n}^{(k)}(\tilde{v}, \tilde{w}),
$$

$$
\rho_{\lambda, n}^{(k)}(\tilde{v}, \tilde{w}) := \left( E \left[ \sup_{t \in [0,T]} e^{-\lambda k t} |\tilde{v}_t - \tilde{w}_t|^k \right] \right)^{\frac{1}{k}}, \ v, w \in \mathcal{Y}_0, \tilde{v}, \tilde{w} \in \mathcal{C}_k.
$$

So, it suffices to prove the contraction of $\Phi^{(n)}$ in $\rho_{\lambda, n}$ for large enough $\lambda > 0$. 22
By the definition of $H_t$, $\|\nabla (B + \lambda_0 u)\|_\infty < \infty$, and repeating step (2) in the proof of [24, Lemma 4.1], we find a function $c_n : (0, \infty) \rightarrow (0, \infty)$ with $c_n(\lambda) \rightarrow 0$ as $\lambda \rightarrow \infty$, such that
\[
\rho_{\lambda,n}^{(k)}(H(v, \tilde{v}), H(w, \tilde{w})) := \left( \mathbb{E} \left[ \sup_{t \in [0, \tau_n]} e^{-\lambda t} \|H_t(v, \tilde{v}) - H_t(w, \tilde{w})\|^k \right] \right)^{\frac{1}{k}}
\]
(3.26)
\[
\leq \sup_{t \in [0, T]} e^{-\lambda t} \|v_t - w_t\|_\mathcal{B} \sup_{s \in [0, T]} \int_0^s e^{-\lambda (s-t)} \frac{\alpha((1 + \kappa \|\mu\|_k)t^{\frac{\alpha}{\kappa}})}{\tilde{\alpha}((1 + \kappa \|\mu\|_k)t^{\frac{\alpha}{\kappa}})} dt + c(\lambda, n) \rho_{\lambda,n}^k(\tilde{v}, \tilde{w}), \quad n \geq 1, (v, \tilde{v}), (w, \tilde{w}) \in \mathcal{V}_0 \times \mathcal{C}_{k,n}, \; \lambda > 0.
\]
Moreover, by (3.20), (3.23) and (3.24), we find a constant $c_4 > 0$ such that
\[
\|\tilde{H}_t(v, \tilde{v}) - \tilde{H}_t(w, \tilde{w})\|_\mathcal{B} = \left\{ t \tilde{\alpha}((1 + \kappa \|\mu\|_k)t^{\frac{\alpha}{\kappa}}) \right\}^{\frac{1}{\alpha}} \|\tilde{I}_t^V(\mu, v - w, \tilde{v} - \tilde{w})\|_\mathcal{B}
\]
\[
\leq c_4 \sqrt{\frac{\alpha((1 + \kappa \|\mu\|_k)s^{\frac{\alpha}{\kappa}})}{s \tilde{\alpha}((1 + \kappa \|\mu\|_k)s^{\frac{\alpha}{\kappa}})}} \int_0^t \|v_s - w_s\|_\mathcal{B}^2 ds^{\frac{1}{2}}
\]
\[
+ c_4 \sqrt{\frac{\alpha((1 + \kappa \|\mu\|_k)s^{\frac{\alpha}{\kappa}})}{s \tilde{\alpha}((1 + \kappa \|\mu\|_k)s^{\frac{\alpha}{\kappa}})}} \left( \int_0^t (\mathbb{E}|\tilde{v}_s - \tilde{w}_s|^k)^{\frac{1}{k}} ds \right)^{\frac{1}{2}},
\]
$v, w \in \mathcal{V}_0, \tilde{v}, \tilde{w} \in \mathcal{C}_{k,n}, t \in [0, T]$. 

Combining this with (3.19) and (3.26), we conclude that $\Phi^{(n)}$ is contractive in the complete metric space $(\mathcal{V}_0 \times \mathcal{C}_{k,n}, \rho_{\lambda,n})$ for large enough $\lambda > 0$, and hence has a unique fixed point.

(d) It remains to verify (ii). Combining the second equation in (3.14), (B), Itô’s formula, the maximal inequality and Khasminskii’s estimate (see [25, Lemma 2.1 and Lemma 4.1]), for any $j \geq k$ we find a constant $c_1 > 0$ such that
\[
d|\tilde{v}^{n,j+1}_t|^{j+1}
\]
\[
\leq c_1 \left\{ |\tilde{v}^{n,j}_t|^j d\mathcal{L}_t + (\mathbb{E}|\tilde{v}^{n,j}_t|^k)^{\frac{j+1}{k}} dt
\]
\[
+ |\tilde{v}^{n,j}_t|^j \frac{\alpha((1 + \kappa \|\mu\|_k)t^{\frac{\alpha}{\kappa}})}{\{t \tilde{\alpha}((1 + \kappa \|\mu\|_k)t^{\frac{\alpha}{\kappa}})\}^{\frac{1}{2}}} \|\tilde{v}^{n,j}_t\|_\mathcal{B} dt + dM_t \right\}, \quad t \in [0, \tau_n]
\]
for some local martingale $M_t$ and an adapted increasing process $\mathcal{L}_t$ with $\mathbb{E}[e^{p \mathcal{L}_T}] < \infty$ for any $p > 0$. Then by the stochastic Gronwall inequality (see [26, Lemma 3.7]), we find
constants \( c_2, c_3 > 0 \) such that

\[
\mathbb{E}
\left[
\sup_{s \in [0,t]} |\tilde{v}_s^{(n)}|^j \right] \mathcal{F}_0 \leq c_2 |\phi(X_0^n)|^j + c_2 \left( \int_0^t \left( \mathbb{E}[|\tilde{v}_s^{(n)}|^k] \right)^{\frac{j+k}{k}} ds \right)^{\frac{k}{j+k}}
\]

\[
+ c_2 \left( \int_0^t \frac{\alpha((1 + \kappa||\mu||_k)s^{\frac{1}{2}})}{\left( 1 + \kappa||\mu||_k s^{\frac{1}{2}} \right)^{\frac{3}{2}}} |v_s^{(n)}| ds \right)^{\frac{1}{j+k}}
\]

\[
\leq c_3 |\phi(X_0^n)|^j + c_3 \left( \int_0^t \left( \mathbb{E}[|\tilde{v}_s^{(n)}|^k] \right)^{\frac{j+k}{k}} ds \right)^{\frac{k}{j+k}}
\]

\[
+ c_3 \left( \int_0^t \frac{\alpha((1 + \kappa||\mu||_k)s^{\frac{1}{2}})}{\left( s\tilde{\alpha}((1 + \kappa||\mu||_k)s^{\frac{1}{2}}) \right)^{\frac{3}{2}}} |v_s^{(n)}| ds \right)^{\frac{j}{k}} + \frac{1}{2} \mathbb{E} \left[ \sup_{s \in [0,t]} |\tilde{v}_s^{(n)}|^j \right] \mathcal{F}_0 .
\]

Noting that (3.19) and the Schwarz inequality imply

\[
\left( \int_0^t \frac{\alpha((1 + \kappa||\mu||_k)s^{\frac{1}{2}})}{\left( s\tilde{\alpha}((1 + \kappa||\mu||_k)s^{\frac{1}{2}}) \right)^{\frac{3}{2}}} |v_s^{(n)}| ds \right)^{2} \leq 4\tilde{\alpha}((1 + \kappa||\mu||_k)t^{\frac{1}{2}}) \int_0^t |v_s^{(n)}|^2 \|v_s^{(n)}\|_B ds,
\]

we find a constant \( c > 0 \) such that

\[
\mathbb{E}
\left[
\sup_{s \in [0,t]} |\tilde{v}_s^{(n)}|^j \right] \mathcal{F}_0 \leq c |\phi(X_0^n)|^j + c \left( \int_0^t \left( \mathbb{E}[|\tilde{v}_s^{(n)}|^k] \right)^{\frac{j+k}{k}} ds \right)^{\frac{k}{j+k}}
\]

\[
+ c \left( \tilde{\alpha}((1 + \kappa||\mu||_k)t^{\frac{1}{2}}) \int_0^t |v_s^{(n)}|^2 \|v_s^{(n)}\|_B ds \right)^{\frac{j}{k}}, \quad t \in [0,T].
\]

Choosing \( j = k \), taking expectation to both sides, and noting that

\[
c \left( \int_0^t \left( \mathbb{E}[|\tilde{v}_s^{(n)}|^k] \right)^{\frac{k+1}{k}} ds \right)^{\frac{k}{k+1}}
\]

\[
\leq c_4 \left( \int_0^t \left( \mathbb{E}[|\tilde{v}_s^{(n)}|^k] \right)^{\frac{1}{2}} ds \right)^{\frac{k}{2}} \leq \frac{1}{2} \mathbb{E} \left[ \sup_{s \in [0,t]} |\tilde{v}_s^{(n)}|^k \right] + \frac{c_4^2}{2} \int_0^t \mathbb{E}[|\tilde{v}_s^{(n)}|^k] ds
\]

holds for some constant \( c_4 > c \), we derive

\[
\mathbb{E}
\left[
\sup_{s \in [0,t]} |\tilde{v}_s^{(n)}|^k \right] \leq 2c_4 ||\phi||_{L^k}^k + c_4 \int_0^t \mathbb{E}[|\tilde{v}_s^{(n)}|^k] ds
\]

\[
+ 2c_4 \left( \tilde{\alpha}((1 + \kappa||\mu||_k)t^{\frac{1}{2}}) \int_0^t |v_s^{(n)}|^2 \|v_s^{(n)}\|_B ds \right)^{\frac{j}{k}}.
\]
Next, let
\[
\tilde{\nu}_t := \alpha \frac{((1 + \kappa \|\mu\|_k) t^{\frac{1}{2}})}{\tilde{\alpha}((1 + \kappa \|\mu\|_k) t^{\frac{1}{2}}) t} \nu_t(\mu, \phi),
\]
\[
\tilde{\nu}_t := \tilde{\nu}_t(\mu, \phi), \quad \tilde{\nu}_t^\varepsilon = \frac{\Theta_t(X_t^{\mu_\varepsilon}) - \Theta_t(X_t^{\mu})}{\varepsilon}.
\]
Then (3.15) follows from

$$\lim_{\varepsilon \to 0} \sup_{t \in (0,T]} \sqrt{t} \left\| v_t^\varepsilon - \hat{v}_t \right\|_\mathcal{B} = 0, \quad \lim_{\varepsilon \to 0} \mathbb{E} \left[ \sup_{t \in (0,T]} |\tilde{v}_t^\varepsilon - \tilde{v}_t| \right] = 0.$$ 

In the following, we first estimate $\|v_t^\varepsilon, 1\|_\mathcal{B}$, $\|v_t^\varepsilon, 2\|_\mathcal{B}$ and $|\tilde{v}_t^\varepsilon|$, then we prove these limits respectively.

(a) To calculate $v_t^\varepsilon, 1$, let

$$V_N := \frac{(N \wedge \|V\|_\mathcal{B})V}{\|V\|_\mathcal{B}}, \quad N \geq 1.$$ 

By (3.11), we have

$$\lim_{\varepsilon \downarrow 0} \frac{P_t^\mu V_N(\mu \varepsilon) - P_t^\mu V_N(\mu)}{\varepsilon} = I_{V_N}^t(\mu, \phi),$$

$$P_t^\mu V_N(\mu \varepsilon) - P_t^\mu V_N(\mu) = \int_0^\varepsilon \frac{d}{dr} P_t^\mu V_N(\mu r) dr = \int_0^\varepsilon I_{V_N}^t(r, \mu, \phi) dr, \quad N \geq 1.$$ 

Since $(B_2)$ implies

$$\|V_N\|_\mathcal{B} \leq \|V\|_\mathcal{B} \leq c(1 + |\cdot|^{k-1})$$

for some constant $c > 0$, by (3.3), (2.5), the same argument to (3.23), and dominated convergence theorem for $N \to \infty$, we find a constant $c(\mu, \phi) > 0$ depending on $\mu, \phi$ such that

$$\|v_t^\varepsilon, 1\|_\mathcal{B} = \frac{\|P_t^\mu V(\mu \varepsilon) - P_t^\mu V(\mu)\|_\mathcal{B}}{\varepsilon} \leq \frac{c(\mu, \phi) \alpha(t^{\frac{1}{2}})}{\sqrt{t}}, \quad \varepsilon \in (0, 1], t \in (0, T],$$

and

$$\lim_{\varepsilon \downarrow 0} \left\| \frac{P_t^\mu V(\mu \varepsilon) - P_t^\mu V(\mu)}{\varepsilon} - I_t^V(\mu, \phi) \right\|_\mathcal{B}$$

$$\leq \lim_{N \to \infty} \lim_{\varepsilon \downarrow 0} \left\{ \frac{|I_t^{V - V_N}(\mu, \phi)|}{\varepsilon} + \frac{\|P_t^\mu (V - V_N)(\mu \varepsilon) - P_t^\mu (V - V_N)(\mu)\|_\mathcal{B}}{\varepsilon} \right\}$$

$$\leq \lim_{N \to \infty} \frac{c(\phi) L^{k(\mu)}}{\sqrt{t}} \left( \int_{\mathbb{R}^d} \mathbb{E}[\|(V - V_N)(X_t^{x, \mu})\|_\mathcal{B}^k] \mu(dx) \right)^{\frac{1}{k}} = 0, \quad t \in (0, T].$$

Hence,

$$h_{t,1}(\varepsilon) := \left\{ v_t^\varepsilon, 1 - I_t^V(\mu, \phi) \right\} \frac{\{t \alpha(t^{\frac{1}{2}})\}^{\frac{1}{2}}}{\alpha(t^{\frac{1}{2}})}$$

26
satisfies
\[
\lim_{\varepsilon \to 0} \|h_{t,1}(\varepsilon)\|_B = 0, \quad \sup_{\varepsilon \in (0,1]} \|h_{t,1}(\varepsilon)\|_B \leq 2c(\mu, \phi)\sqrt{\alpha(t^2)}, \quad t \in [0, T].
\]

(b) We now use Girsanov's theorem to estimate \(v_t^{\varepsilon,2}\). Recall that \(X_t^{\mu,\varepsilon}\) solve (3.2) with initial value \(X_0^\varepsilon\). Let
\[
R_t^\varepsilon := e^\int_0^t \eta_s^\varepsilon dW_s - \frac{1}{2} \int_0^t |\eta_s^\varepsilon|^2 ds,
\]
\[
\eta_s^\varepsilon := \zeta_s(X_s^{\mu,\varepsilon}) 
\times \left\{ B_s(X_s^{\mu,\varepsilon}, P_s^* \mu, P_s V(\mu)) - B_s(X_s^\varepsilon \mu, P_s \mu, P_s V(\mu)) \right\}, \quad s, t \in [0, T].
\]
Since (B) implies (A) and that \(\alpha_k\) is concave, (2.17) holds so that there exists a constant \(c > 0\) such that
\[
\sup_{x \in \mathbb{R}^d, z \in \mathbb{R}, t \in [0, T]} |B_t(x, P_s^* \mu, \varepsilon) - B_t(x, P_s \mu, \varepsilon)|
\leq c(1 + \|\varepsilon\|_B) \alpha(1 + \|\varepsilon\|_B) e^{\alpha(1 + \|\varepsilon\|_B)^2} \leq c(1 + \|\varepsilon\|_B) e^{\alpha(1 + \|\varepsilon\|_B)^2}, \quad \varepsilon \in [0, 1].
\]

Hence, by (B) we find a constant \(c(\mu, \phi) > 0\) such that
\[
|\eta_s^\varepsilon| \leq K \|P_s V(\mu)\|_B - 2 \|P_s V(\mu)\|_B
+ \varepsilon \|\phi\|\int_0^t |\eta_s^\varepsilon|^2 ds, \quad s \in [0, T], \quad \varepsilon \in (0, 1].
\]

Then by Girsanov's theorem,
\[
W_t^\varepsilon := W_t - \int_0^t \eta_s^\varepsilon ds, \quad s \in [0, T]
\]
is a Brownian motion under \(Q := R_T^\varepsilon P\). Reformulate (3.9) with \(\gamma = \mu\) as
\[
dX_t^{\mu,\varepsilon} = \left\{ b_t^{(0)}(X_t^{\mu,\varepsilon}) + B_t(X_t^{\mu,\varepsilon}, P_s^* \mu, P_s V(\mu)) \right\} dt + \sigma_t(X_t^{\mu,\varepsilon}) dW_t
= \left\{ b_t^{(0)}(X_t^{\mu,\varepsilon}) + B_t(X_t^{\mu,\varepsilon}, P_s^* \mu, P_s V(\mu)) \right\} dt + \sigma_t(X_t^{\mu,\varepsilon}) dW_t^\varepsilon,
\]
\[
X_0^{\mu,\varepsilon} = X_0^{\mu,\varepsilon} = X_0^\mu = X_0^\varepsilon + \varepsilon \phi(X_0^\mu).
\]

By the weak uniqueness of (3.9) with \(\gamma = \mu\), we get
\[
v_t^{\varepsilon,2} := \frac{P_t^\mu V(\mu) - P_t^\varepsilon V(\mu)}{\varepsilon} = \frac{1}{\varepsilon} E[V(X_t^{\mu,\varepsilon})(R_t^\varepsilon - 1)], \quad t \in [0, T].
\]
Combining this with \([V]_\alpha \leq 1\), and using Lemma 2.4 for \(p = k^*\), we obtain

\[
\|v^\varepsilon_{1,2}\|_B - \frac{1}{\varepsilon} \mathbb{E}[V(X^\mu_t)(R^\varepsilon_t - 1)] \leq \frac{1}{\varepsilon} \mathbb{E}[\alpha(|X^\mu_t - X^\mu_{t+\varepsilon}|)|R^\varepsilon_t - 1|] \\
\leq \frac{1}{\varepsilon} \left(\mathbb{E}[|R^\varepsilon_t - 1|^k]\right)^{\frac{1}{k'}} \alpha\left(\mathbb{E}[|X^\mu_t - X^\mu_{t+\varepsilon}|]^k\right)^{\frac{1}{k'}}.
\]

By (2.5) and (3.33), we find a constant \(c_0(\mu, \phi) > 0\) such that

\[
\mathbb{E}[\|V(X^\mu_t)||k^*\|] \leq c_0(\mu, \phi), \quad t \in [0, T], \varepsilon \in (0, 1].
\]

On the other hand, by (3.37), for any \(p \geq 1\) there exists a constant \(c(p, \mu, \phi) > 0\) such that

\[
\mathbb{E}[|R^\varepsilon_t - 1|^p] \leq c(p, \mu, \phi)\varepsilon^p \left(\int_0^t (1 + \|v^\varepsilon_s\|_B^2)\,ds\right)^{\frac{p}{2}}, \quad t \in [0, T], \varepsilon \in (0, 1].
\]

Combining this with (3.38) and (4.10), we find a constant \(c_1(\mu, \phi) > 0\) such that

\[
\|v^\varepsilon_t\|_B \leq \left(\mathbb{E}[\|V(X^\mu_t)||k^*\|_B]\right)^{\frac{1}{k'}} \left(\mathbb{E}[|R^\varepsilon_t - 1|^k]\right)^{\frac{1}{k'}} \\
\leq c_1(\mu, \phi) \left(\int_0^t (1 + \|v^\varepsilon_s\|_B^2)\,ds\right)^{\frac{1}{2}}.
\]

This together with (3.30) and (3.34) yields that for some constant \(c_2(\mu, \phi) > 0\),

\[
\|v^\varepsilon_t\|_B^2 \leq 2\|v^\varepsilon_t\|_B^2 + 2\|v^\varepsilon_{t+\varepsilon}\|_B^2 \leq c_2(\mu, \phi) \left(\frac{\alpha(t\frac{\varepsilon^2}{t})^2}{t} + \int_0^t \|v^\varepsilon_s\|_B^2\,ds\right), \quad t \in [0, T].
\]

By Gronwall's inequality, we find a constant \(c_3(\mu, \phi) > 0\) such that

\[
\|v^\varepsilon_t\|_B^2 \leq c_3(\mu, \phi) \left(\frac{\alpha(t\frac{\varepsilon^2}{t})^2}{t} + \int_0^t \frac{\alpha(s\frac{\varepsilon^2}{s})^2}{s}e^{\alpha(s\frac{\varepsilon^2}{s})^2(s-t)}\,ds\right) \\
\leq c_3(\mu, \phi) \frac{\alpha(t\frac{\varepsilon^2}{t})^2}{t}, \quad t \in [0, T], \varepsilon \in (0, 1].
\]

(c) To estimate \(|\tilde{v}^\varepsilon_t|\), we recall that \(\Theta_t = u_t + id\) for \(u_t = u^\lambda_{0,\mu}\). Similarly to (2.11) and (2.14), by Itô's formula for \(\Theta_t(X^\mu_t) - \Theta_t(X^\mu_0)\), we obtain

\[
d\tilde{v}^\varepsilon_t = \frac{1}{\varepsilon} \left\{\left((\nabla \Theta_t)\sigma_t\right)\{X^\mu_t\} - \left((\nabla \Theta_t)\sigma_t\right)\{X^\mu_0\}\right\}\,dW_t \\
+ \frac{1}{\varepsilon} \left[\lambda_0 u_t(X^\mu_t) - \lambda_0 u_t(X^\mu_0) + \nabla B_t(X^\mu_{t+\varepsilon}, P^\mu_{t+\varepsilon}, P_V(\mu_{t+\varepsilon})) - B_t(X^\mu_t, P^\mu_t, P_V(\mu))\Theta_t(X^\mu_t)\right]dt \\
+ \frac{1}{\varepsilon} \left[B_t(X^\mu_t, P^\mu_t, P_V(\mu)) - B_t(X^\mu_t, P^\mu_t, P_V(\mu))\right]dt,
\]

\[
\tilde{v}^\varepsilon_0 = \frac{\Theta_0(X^\mu_0) - \Theta_0(X^\mu_0)}{\varepsilon}.
\]
Similarly to (3.27), by combining this with (B), (2.9), (3.42) and the stochastic Gronwall inequality, for any \( j \geq 1 \) we find constants \( c(j), c_4(\mu, \phi) > 0 \) such that

\[
\mathbb{E} \left[ \sup_{s \in [0, t]} |\tilde{v}_s^{\varepsilon}| \right] \leq c(j) |\phi(X_0^\mu)|^j + c_4(\mu, \phi)
\]

for some constant \( \tilde{c}(\mu, \phi) \) as in (3.28), by Gronwall’s inequality we find a constant \( c_5(\mu, \phi) > 0 \) such that

\[
\mathbb{E} \left[ \sup_{s \in [0, T]} |\tilde{v}_s^{\varepsilon}| \right] \leq c_5(\mu, \phi), \quad \varepsilon \in (0, 1],
\]

so that (3.44) implies

\[
\mathbb{E} \left[ \sup_{s \in [0, T]} |\tilde{v}_s^{\varepsilon}| \right] \leq c_0(\mu, \phi) + c(j) |\phi(X_0^\mu)|^j, \quad \varepsilon \in (0, 1],
\]

for some constant \( c_0(\mu, \phi) > 0 \).

(d) Now, let

\[
\psi_t^\varepsilon := \mathbb{E} \left[ \left\{ \{ D^L B_t(z, \cdots, P_t V(\mu)) \} (P_t^* \mu) (X_t^\mu), (\nabla \Theta_t (X_t^\mu))^{-1} \tilde{v}_t^\varepsilon \} + \{ \nabla_{\tilde{v}_t^\varepsilon} B_t(X_t^\mu, P_t^* \mu, P_t V(\mu)) \} \right\} \right]_{z = X_t^\mu}
\]

Then (3.43) can be reformulated as

\[
\tilde{v}_t^\varepsilon - \tilde{v}_0^\varepsilon = \int_0^t \left\{ \nabla \left( \nabla \Theta_s (X_s^\mu) \right)^{-1} \tilde{v}_s^\varepsilon \left[ B_s(\cdot, P_s^* \mu, P_s V(\mu)) + \lambda_0 u_s \right] (X_s^\mu) + \nabla \tilde{v}_s^\varepsilon \Theta_s (X_s^\mu) \right\} ds + \int_0^t \left\{ \nabla \left( \nabla \Theta_s (X_s^\mu) \right)^{-1} \tilde{v}_s^\varepsilon \left\{ (\nabla \Theta_s) \sigma_s \right\} (X_s^\mu) \right\} dW_s + \alpha_t^\varepsilon, \quad t \in [0, T],
\]
where for $t \in [0, T]$,

$$
\alpha_t^\varepsilon := \int_0^t \xi_s^\varepsilon \, ds + \int_0^t \eta_s^\varepsilon \, dW_s,
$$

$$
\xi_s^\varepsilon := \frac{1}{\varepsilon} \left\{ [B_s(\cdot, P^s_\mu \mu, P_s V(\mu_\varepsilon)) + \lambda_0 u_s](X_s^\mu) - [B_s(\cdot, P^s_\mu, P_s V(\mu)) + \lambda_0 u_s](X_s^\mu) \right\}
$$

$$
+ \nabla \frac{1}{\varepsilon} \left\{ [B_s(X_s^\mu, P^s_\mu \mu, P_s V(\mu_\varepsilon)) - B_s(X_s^\mu, P^s_\mu, P_s V(\mu))] u_s(X_s^\mu) \right\}
$$

$$
- \nabla \left( \nabla \Theta_s(X_s^\mu) \right)^{-1} \left. \left[ B_s(\cdot, P^s_\mu, P_s V(\mu)) + \lambda_0 u_s \right] (X_s^\mu) \right. - \nabla \psi_s \Theta_s(X_s^\mu),
$$

$$
\eta_s^\varepsilon := \left\{ (\nabla \Theta_s)^s \sigma_s \right\} (X_s^\mu) - \left\{ (\nabla \Theta_s)^s \sigma_s \right\} (X_s^\mu) - \nabla \left( \nabla \Theta_s(X_s^\mu) \right)^{-1} \left. \left( (\nabla \Theta_s)^s \sigma_s \right) (X_s^\mu) \right. \}
$$

By the same argument leading to [24, (4.23), (4.26)] and (3.27) above, (B) together with (2.9), (3.42), (3.45), (3.46) implies

$$
\lim_{\varepsilon \to 0} \mathbb{E} \left[ \sup_{s \in [0, T]} |\alpha_s(\varepsilon)|^2 \right| \mathcal{F}_0 = 0, \quad j \geq 1,
$$

and by Jensen’s inequality we find a constant $c > 0$ such that

$$
\mathbb{E} \left[ \sup_{s \in [0, t]} |\bar{v}_s^\varepsilon - \tilde{v}_s| \right| \mathcal{F}_0 \leq c \bar{v}_0^\varepsilon - \bar{v}_0 + c \left( \mathbb{E} \left[ \sup_{s \in [0, t]} |\alpha_s(\varepsilon)|^2 \right| \mathcal{F}_0 \right) \right]^{\frac{1}{2}}
$$

$$
+ c \left( \int_0^t \|v_s^\varepsilon - \tilde{v}_s\|^2 \, ds \right) + c \left( \int_0^t \left\{ \mathbb{E}[|\bar{v}_s^\varepsilon - \tilde{v}_s|^k] \right\}^2 \, ds \right)^{\frac{1}{2}}, \quad t \in [0, T].
$$

So, by (3.46) and Fatou’s lemma, we see that

$$
\theta_t := \limsup_{\varepsilon \to 0} \mathbb{E} \left[ \sup_{s \in [0, t]} |\bar{v}_s^\varepsilon - \tilde{v}_s| \right],
$$

(3.47)

$$
\tilde{\theta}_t := \limsup_{\varepsilon \to 0} \sup_{s \in [0, t]} \frac{\sqrt{s}}{s^{\frac{1}{2}}} \|v_s^\varepsilon - \tilde{v}_s\|_{\mathbb{B}}
$$

satisfy

$$
\theta_t \leq \mathbb{E} \left[ \limsup_{\varepsilon \to 0} \mathbb{E} \left[ \sup_{s \in [0, t]} |\bar{v}_s^\varepsilon - \tilde{v}_s| \right| \mathcal{F}_0 \right] \right]
$$

$$
\leq c \left( \int_0^t \frac{\alpha(s^\frac{1}{2})^2}{s} \limsup_{\varepsilon \to 0} \frac{s}{\alpha(s^\frac{1}{2})^2} \|v_s^\varepsilon - \tilde{v}_s\|^2 \, ds \right) \right]^{\frac{1}{2}}
$$

$$
+ c \left( \int_0^t \left\{ \limsup_{\varepsilon \to 0} \mathbb{E}[|\bar{v}_s^\varepsilon - \tilde{v}_s|^k] \right\}^2 \, ds \right)^{\frac{1}{2}}
$$

$$
\leq c \left( \int_0^t \frac{\alpha(s^\frac{1}{2})^2}{s} \tilde{\theta}_s^2 \, ds \right) \right]^{\frac{1}{2}} + c \left( \int_0^t \bar{\theta}_s^2 \, ds \right)^{\frac{1}{2}}, \quad t \in [0, T].
$$

(3.48)
(e) We now estimate $\|v_t^\varepsilon - \hat{v}_t\|_B$. Similarly to (d), for $N_s$ and $\bar{N}_s$ in (3.13), we have
\[
\frac{R_t^\varepsilon - 1}{\varepsilon} = \int_0^t \langle \left\langle R_s^\varepsilon \left( \varepsilon^{-1} \eta_s^\varepsilon, dW_s \right) \right\rangle \rangle,
\]
and
\[
= \int_0^t \langle \left\langle \zeta_s(X_s^\mu) \left( B_s(X_s^\mu, P_s^\mu, P_s V(\mu)) - B_s(X_s^\mu, P_s^\mu, P_s V(\mu)) \right) \right\rangle \rangle,
\]
where
\[
h_t(\varepsilon) := \int_0^t \langle \left\langle \zeta_s(X_s^\mu) \left\{ R_sB_s(X_s^\mu, P_s^\mu, P_s V(\mu)) - B_s(X_s^\mu, P_s^\mu, P_s V(\mu)) \right\} \right\rangle \rangle,
\]
satisfies
\[
\lim_{\varepsilon \to 0} \mathbb{E} \left[ \sup_{t \in [0,T]} |h_t(\varepsilon)|^p \right] = 0, \quad p \geq 1.
\]
Noting that $X_t^\mu = X_t^{\mu,\phi}$, by [24, Theorem 2.1] we find a constant $c(\mu) > 0$ such that
\[
\mathbb{E} \left[ \sup_{t \in [0,T]} |X_t^{\mu,\phi} - X_t^\mu|^k \right] \leq c(\mu)\varepsilon^k, \quad \varepsilon \in [0,1], \mu \in \mathcal{P}_k.
\]
Combining these with (3.39), (3.42) and (3.41), we find a constant $k_1(\mu, \phi) > 0$ such that
\[
\left\| v_t^{\varepsilon,2} - \frac{1}{\varepsilon} \mathbb{E} [V(X_t^\mu)(R_t^\varepsilon - 1)] \right\|_B \leq k_1(\mu, \phi)\varepsilon, \quad t \in [0,T], \varepsilon \in (0,1).
\]
By (3.14) for $v_t = v_t(\mu, \phi), \tilde{v}_t = \tilde{v}_t(\mu, \phi)$, and by the definition of $\hat{v}_t$ in (3.32), we have
\[
\hat{v}_t = I_t^V(\mu, \phi) + \mathbb{E} \left[ V(X_t^\mu) \int_0^t \left\langle \zeta_s(X_s^\mu) \left\{ N_s(\tilde{v}) + \bar{N}_s(\tilde{v}) \right\} \right\rangle \right].
\]
Combining this with (3.30), (3.35), (3.50), (3.49), (3.52) and the same argument to (3.24), we find a constant $k_2(\mu, \phi)$ and a measurable function $\bar{h} : (0,T] \times [0,1] \to (0,\infty)$ with
\[
\sup_{t \in (0,1], \varepsilon \in (0,1]} \frac{\sqrt{t}}{\alpha(t^2)} \bar{h}_t(\varepsilon) \leq k_2(\mu, \phi), \quad \lim_{\varepsilon \to 0} \sup_{t \in [0,T]} \sqrt{t} \bar{h}_t(\varepsilon) = 0,
\]
such that
\[
\|v_t^\varepsilon - \hat{v}_t\|_B \leq \bar{h}_t(\varepsilon) + \left\| \mathbb{E} \left[ V(X_t^\mu) \int_0^t \left\langle \zeta_s(X_s^\mu) \left\{ N_s(\tilde{v}) + \bar{N}_s(\tilde{v}) \right\} \right\rangle \right] \right\|_B \leq \bar{h}_t(\varepsilon) + k_2(\mu, \phi) \left( \int_0^t \left\{ \|v_s^\varepsilon - \hat{v}_s\|_B^2 + (\mathbb{E}[\tilde{v}_s^\varepsilon - \bar{v}_s^\varepsilon]^{\frac{1}{2}})^2 \right\} ds \right)^{\frac{1}{2}}, \quad t \in (0,T].
\]
Noting that (3.16), (3.31) and (3.42) imply that $\tilde{\theta}_t$ defined in (3.47) satisfies
\[
\sup_{t \in (0,T]} \tilde{\theta}_t \leq \sup_{\varepsilon \in (0,1]} \sup_{s \in [0,T]} \frac{\sqrt{s}}{\alpha(s^2)} \|v_\varepsilon^s - \hat{v}_s\|_B =: \tilde{c}(\mu, \phi) < \infty,
\]
so that by Fatou’s lemma we obtain
\[
\tilde{\theta}_t^2 \leq Ck_2(\mu, \phi)^2 \int_0^t \left\{ \frac{\alpha(s^2)}{s} \tilde{\theta}_s^2 + \hat{\theta}_s^2 \right\} ds, \quad t \in (0,T],
\]
where by (2.2),
\[
C := \sup_{t \in (0,T]} \frac{s}{t} \alpha(t^{\frac{1}{2}})^2 < \infty.
\]
Combining this with (3.48), \( \int_0^T \frac{\alpha(t^{\frac{1}{2}})^2}{t} dt < \infty \), and applying Gronwall’s inequality, we prove (3.32) and hence finish the proof.

We are now ready to prove Theorem 3.1.

**Proof of Theorem 3.1.** By (3.10) and (3.12), it suffices to prove that for any \( t \in (0,T] \) and \( f \in \mathcal{B}_b(\mathbb{R}^d) \),
\[
(3.53) \quad \lim_{\varepsilon \to 0} \frac{P_t^{\mu_{e}} f(\mu_{e})}{\varepsilon} - \frac{P_t^{\mu} f(\mu_{e})}{\varepsilon} = \mathbb{E} \left[ f(X_t^\varepsilon) \int_0^t \left\langle \zeta_s(X_s^\varepsilon) \{N_s + \tilde{N}_s\}, dW_s \right\rangle \right].
\]
Let \( R_t^\varepsilon \) be in (3.36). By (3.38) for \( f \) replacing \( V \), we obtain
\[
(3.54) \quad \frac{P_t^{\mu_{e}} f(\mu_{e})}{\varepsilon} - \frac{P_t^{\mu} f(\mu_{e})}{\varepsilon} = \frac{1}{\varepsilon} \mathbb{E} \left[ f(X_t^{\mu_{e} - \mu})(R_t^\varepsilon - 1) \right], \quad t \in (0,T].
\]
Noting that (3.51) implies
\[
\lim_{\varepsilon \to 0} \mathbb{E} \left[ \sup_{t \in [0,T]} |X_t^{\mu_{e} - \mu} - X_t^{\mu_{e} - \mu}|^k \right] = 0,
\]
while \((B_2)\), Lemma 3.3 and (3.36) lead to
\[
\lim_{\varepsilon \to 0} \frac{R_t^\varepsilon - 1}{\varepsilon} = \int_0^t \left\langle \zeta_s(X_s^\varepsilon) \{N_s + \tilde{N}_s\}, dW_s \right\rangle
\]
in \( L^2(\mathbb{P}) \), by taking \( \varepsilon \to 0 \) in (3.54) we deduce (3.53) for \( f \in C_b(\mathbb{R}^d) \). By an approximation argument as in [24, Proof of (2.3)] for \( f \in \mathcal{B}_b(\mathbb{R}^d) \), this implies (3.54) for \( f \in \mathcal{B}_b(\mathbb{R}^d) \). \( \square \)
References

[1] S. Albeverio, Y. G. Kondratiev, M. Röckner, *Differential geometry of Poisson spaces*, C R Acad Sci Paris Sér I Math. 323(1996), 1129-1134.

[2] J. Bao, P. Ren, F.-Y. Wang, *Bismut formulas for Lions derivative of McKean-Vlasov SDEs with memory*, J. Differential Equations 282(2021), 285-329.

[3] M. Bauer, T. M-Brandis, *Existence and regularity of solutions to multi-dimensional mean-field stochastic differential equations with irregular drift*, arXiv:1912.05932.

[4] M. Bauer, T. M-Brandis, F. Proske, *Strong solutions of mean-field stochastic differential equations with irregular drift*, Electron. J. Probab. 23(2018), 1-35.

[5] J. M. Bismut, *Large Deviations and the Malliavin Calculus*, Boston: Birkhäuser, MA, 1984.

[6] P.-E. Chaudru de Raynal, *Strong well-posedness of McKean-Vlasov stochastic differential equation with Hölder drift*, Stochastic Process. Appl. 130(2020), 79-107.

[7] G. Crippa, C. De Lellis, *Estimates and regularity results for the Di Perna-Lions flow*, J. Reine Angew. Math. 616(2008), 15-46.

[8] I. Csiszár, J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.

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

[10] X. Huang, Y. Song, F.-Y. Wang, *Bismut formula for intrinsic/Lions derivatives of distribution dependent SDEs with singular coefficients*, Discrete Contin. Dyn. Syst. 42(2022), 4597-4614.

[11] X. Huang, W. Lv, *Exponential Ergodicity and propagation of chaos for path-distribution dependent stochastic Hamiltonian system*, arXiv:2109.13728v3.

[12] X. Huang, F.-Y. Wang, *Distribution dependent SDEs with singular coefficients*, Stochastic Process. Appl. 129 (2019), 4747-4770.

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

[14] D. Lacker, *On a strong form of propagation of chaos for McKean-Vlasov equations*, Electron. Commun. Probab. 23(2018), 1-11.
[15] Yu. S. Mishura, A. Yu. Veretennikov, *Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations*, Theo. Probab. Math. Statist. 103(2020), 59-101.

[16] P. Ren, *Singular McKean-Vlasov SDEs: Well-posedness, regularities and Wang's Hanrack inequality*, arXiv:2110.08846.

[17] P. Ren, F.-Y. Wang, *Bismut Formula for Lions Derivative of Distribution Dependent SDEs and Applications*, J. Differential Equations 267(2019), 4745-4777.

[18] A.-S. Sznitman, *Topics in propagation of chaos*, In “École d'Été de Probabilités de Sain-Flour XIX-1989”, Lecture Notes in Mathematics 1464, p. 165-251, Springer, Berlin, 1991.

[19] F.-Y. Wang, *Logarithmic Sobolev inequalities on noncompact Riemannian manifolds*, Probab. Theory Relat. Fields 109(1997), 417-424.

[20] F.-Y. Wang, *Harnack inequalities on manifolds with boundary and applications*, J. Math. Pures Appl. 94(2010), 304-321.

[21] F.-Y. Wang, *Harnack Inequality and Applications for Stochastic Partial Differential Equations*, Springer, New York, 2013.

[22] F.-Y. Wang, *Distribution-dependent SDEs for Landau type equations*, Stochastic Process. Appl. 128(2018), 595-621.

[23] F.-Y. Wang, *Distribution dependent reflecting stochastic differential equations*, arXiv:2106.12737.

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

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

[26] L. Xie, X. Zhang, *Ergodicity of stochastic differential equations with jumps and singular coefficients*, Ann. Inst. Henri Poincaré Probab. Stat. 56(2020), 175-229.

[27] C. Yuan, S.-Q. Zhang, *A study on Zvonkin’s transformation for stochastic differential equations with singular drift and related applications*, J. Differential Equations 297(2021), 277-319.

[28] A. K. Zvonkin, *A transformation of the phase space of a diffusion process that removes the drift*, Math. Sb. 93(1974), 129-149.