Bismut Formula for Intrinsic/Lions Derivatives of Distribution Dependent SDEs with Singular Coefficients*

Xing Huang [a)], **Yulin Song** [b)], **Feng-Yu Wang** [a),c)]

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

b)Department of Mathematics, Nanjing University, Nanjing, 210093, China

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

xinghuang@tju.edu.cn, ylsong@nju.edu.cn, wangfy@tju.edu.cn

April 19, 2022

Abstract

By using distribution dependent Zvonkin's transforms and Malliavin calculus, the Bismut type formula is derived for the intrinisc/Lions derivatives of distribution dependent SDEs with singular drifts, which generalizes the corresponding results derived for classical SDEs and regular distribution dependent SDEs.

AMS subject Classification: 60H1075, 60G44.
Keywords: Distribution dependent SDEs, intrinsic/Lions derivative, Zvonkin's transform, Bismut formula.

1 Introduction

Due to wide applications in the study of nonlinear PDEs and particle systems, distribution dependent stochastic differential equations (DDSDEs for short), also called McKean-Vlasov or mean-field SDEs, have been intensively investigated, see for instance [7, 10, 11, 4, 5, 16, 6, 12, 14, 17, 18, 19, 20, 21, 24] among many other references.

*X. Huang is supported by NSFC (No.11801406), Y. Song is supported by NSFC (No.11971227, 11790272) and F.-Y. Wang is supported by NSFC (No.11771326, 11831014, 11921001)

To characterize the regularity of DDSDEs, Bismut formula and derivative estimates have been presented for the distribution of solutions with respect to initial data, see for instance [27, 2, 13, 23, 3, 25]. See also [12] for the study of decoupled SDEs where the distribution parameter is fixed as the law of the associated DDSDE, and the resulting regularity estimates apply to the DDSDEs as well (see Remark 2.2 below).

In this paper, we aim to establish Bismut formula for the Lions derivative of singular DDSDEs, such that existing results derived in more regular situations are extended. This type formula was first found by Bismut [8] in 1984 using Malliavin calculus for diffusion semigroups on manifolds, then reproved by Elworthy-Li [15] in 1974 using martingale arguments. Since then the formula has been widely developed and applied for SDEs/SPDEs driven by Gaussian or Lévy noises. Recently, Bismut formula was established in [28] for SDEs with singular drifts by using Zvonkin's transform [29], which is a powerful tool in regularizing singular SDEs. In this paper, we aim to extend this result for singular DDSDEs.

Let $\mathscr{P}$ be the set of all probability measures on $\mathbb{R}^d$. Consider the following distribution-dependent SDE on $\mathbb{R}^d$:

$$\mathrm{d}X_t = (B_t + b_t)(X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t, \tag{1.1}$$

where W_t is the d-dimensional Brownian motion on a complete filtration probability space $(\Omega, \mathscr{F}, \mathbb{P}; \{\mathscr{F}_t\}_{t\geq 0})$, $\mathscr{L}_{X_t}$ is the law of X_t under $\mathbb{P}$, and

$$B, b : \mathbb{R}_+ \times \mathbb{R}^d \times \mathscr{P} \to \mathbb{R}^d, \quad \sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$$

are measurable. We will consider the SDE (1.1) with initial distributions in the class

$$\mathscr{P}_2 := \big\{\mu \in \mathscr{P} : \mu(|\cdot|^2) < \infty\big\}.$$

It is well known that $\mathscr{P}_2$ is a Polish space under the Wasserstein distance

$$\mathbb{W}_2(\mu,\nu) := \inf_{\pi\in\mathscr{C}(\mu,\nu)} \bigg(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \pi(\mathrm{d}x,\mathrm{d}y)\bigg)^{\frac{1}{2}}, \quad \mu,\nu \in \mathscr{P}_2,$$

where $\mathscr{C}(\mu,\nu)$ is the set of all couplings of μ and ν. In the following we will assume that B is regular and b is singular in the space variable.

We call (1.1) strong (resp. weak) well-posed for distributions in $\mathscr{P}_2$, i.e. for any initial value $X_0 \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_0; \mathbb{P})$ (resp. initial distribution $\mu \in \mathscr{P}_2$), if (1.1) has a unique strong (resp. weak) solution with $X_\cdot \in C([0,\infty); \mathscr{P}_2)$. When (1.1) is both strong and weak well-posed (note that unlike in the classical setting, the strong well-posedness does not imply the weak one), we call it well-posed. In this case, for any $\mu \in \mathscr{P}_2$, denote $(P_t^*\mu = \mathscr{L}_{X_t})_{t\geq 0}$ for the solution $(X_t)_{t\geq 0}$ with initial distribution $\mathscr{L}_{X_0} = \mu \in \mathscr{P}_2$. For any $f \in \mathscr{B}_b(\mathbb{R}^d)$, the class of bounded measurable functions on $\mathbb{R}^d$, we aim to establish Bismut formulas for $P_t f(\mu)$ in $\mu \in \mathscr{P}_2$, where

$$P_t f(\mu) := (P_t^*\mu)(f) := \int_{\mathbb{R}^d} f(y)(P_t^*\mu)(\mathrm{d}y), \quad t > 0.$$

To this end, we first recall the intrinsic/Lions derivatives for real functions on $\mathscr{P}_2$.

Definition 1.1. Let $f : \mathscr{P}_2 \to \mathbb{R}$.

(1) If for any $\phi \in L^2(\mathbb{R}^d \to \mathbb{R}^d; \mu)$,

$$D^I_\phi f(\mu) := \lim_{\varepsilon \downarrow 0} \frac{f(\mu \circ (\mathrm{Id} + \varepsilon\phi)^{-1}) - f(\mu)}{\varepsilon} \in \mathbb{R}$$

exists, and is a bounded linear functional in ϕ, we call f intrinsic differentiable at μ. In this case, there exists a unique $D^I f(\mu) \in L^2(\mathbb{R}^d \to \mathbb{R}^d; \mu)$ such that

$$\langle D^I f(\mu), \phi \rangle_{L^2(\mu)} = D^I_\phi f(\mu), \quad \phi \in L^2(\mathbb{R}^d \to \mathbb{R}^d; \mu).$$

We call $D^I f(\mu)$ the intrinsic derivative of f at μ. If f is intrinsic differentiable at all $\mu \in \mathscr{P}_2$, we call it intrinsic differentiable on $\mathscr{P}_2$ and denote

$$\|D^I f(\mu)\| := \|D^I f(\mu)\|_{L^2(\mu)} = \bigg(\int_{\mathbb{R}^d} |D^I f(\mu)|^2 \mathrm{d}\mu \bigg)^{\frac{1}{2}}.$$

(2) If f is intrinsic differentiable and for any $\mu \in \mathscr{P}_2$,

$$\lim_{\|\phi\|_{L^2(\mu)} \to 0} \frac{f(\mu \circ (\mathrm{Id} + \phi)^{-1}) - f(\mu) - D^I_\phi f(\mu)}{\|\phi\|_{L^2(\mu)}} = 0,$$

we call f L-differentiable on $\mathscr{P}_2$. In this case, $D^I f(\mu)$ is also denoted by $D^L f(\mu)$, and is called the L-derivative of f at μ.

Then intrinsic derivative was first introduced in [1] on the configuration space over a Riemannian manifold, while the L-derivative appeared in the Lecture notes [9] for the study of mean field games and is also called Lions derivative in references.

Note that the derivative $D^I f(\mu) \in L^2(\mathbb{R}^d \to \mathbb{R}^d; \mu)$ is μ-a.e. defined. In applications, we take its continuous version if exists. The following classes of L-differentiable functions are often used in analysis:

(a) $f \in C^1(\mathscr{P}_2)$: if f is L-differentiable such that for every $\mu \in \mathscr{P}_2$, there exists a μ-version $D^L f(\mu)(\cdot)$ such that $D^L f(\mu)(x)$ is jointly continuous in $(x, \mu) \in \mathbb{R}^d \times \mathscr{P}_2$.

(b) $f \in C^1_b(\mathscr{P}_2)$: if $f \in C^1(\mathscr{P}_2)$ and $D^L f(\mu)(x)$ is bounded.

(c) $f \in C^{1,1}(\mathbb{R}^d \times \mathscr{P}_2)$: if f is a continuous function on $\mathbb{R}^d \times \mathscr{P}_2$ such that $f(\cdot, \mu) \in C^1(\mathbb{R}^d)$, $f(x, \cdot) \in C^1(\mathscr{P}_2)$ with $\nabla f(\cdot, \mu)(x)$ and $D^L f(x, \cdot)(\mu)(y)$ jointly continuous in $(x, y, \mu) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathscr{P}_2$. If moreover these derivatives are bounded, we denote $f \in C^{1,1}_b(\mathbb{R}^d \times \mathscr{P}_2)$.

We will state the main result in Section 2 and prove it in Section 3.

2 The main result

We will assume that $b_t(\cdot,\mu)$ is Dini continuous for which we introduce the following class as in [26]:

$$\mathscr{D}=\Big\{\varphi:[0,+\infty)\to[0,+\infty)|\varphi^2 \text{ is concave and } \varphi \text{ is increasing with } \int_0^1\frac{\varphi(s)}{s}\mathrm{d}s<\infty\Big\}.$$

The condition $\int_0^1\frac{\varphi(s)}{s}\mathrm{d}s<\infty$ is known as the Dini condition. Clearly, for any $\alpha\in(0,\frac{1}{2})$ the function $\varphi_1(s)=s^\alpha$ is in $\mathscr{D}$. Let $\varphi_2(s):=\frac{1}{\log^{1+\delta}(c+s^{-1})}$ for constants $\delta>0$ and $c>0$ large enough such that φ_2^2 is concave, then φ_2 is also in $\mathscr{D}$. For any (real or $\mathbb{R}^d$-valued) function f on $\mathbb{R}^d$, let

$$[f]_\varphi:=\sup_{x\neq y\in\mathbb{R}^d}\Big\{|f(x)|+\frac{|f(x)-f(y)|}{\varphi(|x-y|)}\Big\},\quad \varphi\in\mathscr{D}.$$

For a function $f:[0,\infty)\times E\to\mathbb{R}$, where E is an abstract space, we denote

$$\|f\|_{T,\infty}:=\sup_{t\in[0,T]\times E}|f_t(x)|,\quad T>0.$$

Throughout this paper, we make the following assumption.

(H) For each $t\geq 0$ and $x\in\mathbb{R}^d$, $b_t(x,\cdot)\in C^1(\mathscr{P}_2)$ with $D^Lb_t(x,\mu)(y)$ continuous in $(x,y,\mu)\in\mathbb{R}^d\times\mathbb{R}^d\times\mathscr{P}_2$, $B_t\in C^{1,1}(\mathbb{R}^d\times\mathscr{P}_2)$, $\sigma_t\in C^1(\mathbb{R}^d)$ is invertible, such that for any $T>0$,

$$\big\|(\|\sigma\|+\|\sigma^{-1}\|+|B(0,\delta_0)|+[b]_\varphi+\|\nabla\sigma\|+\|D^Lb\|+\|\nabla B\|+\|D^LB\|)\big\|_{T,\infty}<\infty$$

holds for some $\varphi\in\mathscr{D}$, where δ_0 is the Dirac measure at $0\in\mathbb{R}^d$, $[\cdot]_\varphi$ is the modulus of continuity in $x\in\mathbb{R}^d$, and D^L is Lion's derivative in $\mu\in\mathscr{P}_2$.

Remark 2.1. *With the second inequality in* (3.6) *replacing* [19, (27)], *and with t replacing A_t in* [19, (35)], *the proof of* [19, Theorem 1.1(2)] *yields that* **(H)** *implies the well-posedness of* (1.1) *for distributions in $\mathscr{P}_2$. We will show that this assumption also ensures the intrinsic differentiability of P_Tf for $T>0$ and $f\in\mathscr{B}_b(\mathbb{R}^d)$. To prove the L-differentiability of P_Tf, we make the following additional assumption.*

(C) For any $T>0$, there exists $p>2$ such that

$$\sup_{(t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathscr{P}_2}\int_{\mathbb{R}^d}|D^L(b+B)_t(x,\mu)(y)|^p\mu(\mathrm{d}y)<\infty.$$

Obviously, **(C)** holds if $D^L(b_t+B_t)(x,\mu)(y)$ is bounded in $(t,x,y,\mu)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathscr{P}_2$.

For any $T>0$, let $\mathscr{C}_T:=C([0,T];\mathbb{R}^d)$ be equipped with the uniform norm. For $\varepsilon\in(0,1)$ and $\eta,X_0\in L^2(\Omega\to\mathbb{R}^d,\mathscr{F}_0;\mathbb{P})$, let $\{X_t^{\eta,\varepsilon}\}_{t\geq 0}$ solve (1.1) with initial value $X_0+\varepsilon\eta$. The main result of the paper is the following.

Theorem 2.2. *Assume* **(H)**. *Then the following statements hold.*

(1) *For any $T>0$, the limit*

$$\nabla_\eta X_t:=\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}(X_t^{\eta,\varepsilon}-X_t) \tag{2.1}$$

exists in $L^2(\Omega\to\mathscr{C}_T;\mathbb{P})$, and there exists a constant $C_T>0$ such that

$$\mathbb{E}\left(\sup_{t\in[0,T]}|\nabla_\eta X_t|^2\right)\leq C_T\mathbb{E}|\eta|^2,\quad \eta,X_0\in L^2(\Omega\to\mathbb{R}^d,\mathscr{F}_0;\mathbb{P}). \tag{2.2}$$

(2) *P_Tf is intrinsically differentiable for any $T>0$ and $f\in\mathscr{B}_b(\mathbb{R}^d)$, and*

$$D^I_\phi(P_Tf)(\mu)=\mathbb{E}\left(f(X_T)\int_0^T\langle\zeta_t^\phi,\mathrm{d}W_t\rangle\right),\quad \mu\in\mathscr{P}_2,\phi\in L^2(\mathbb{R}^d\to\mathbb{R}^d;\mu) \tag{2.3}$$

holds for X_t solving (1.1) with $\mathscr{L}_{X_0}=\mu$, and

$$\zeta_t^\phi:=\sigma_t(X_t)^{-1}\Big\{g_t'\nabla_{\phi(X_0)}X_t+g_t\mathbb{E}\big[D^L(B_t+b_t)(y,\mathscr{L}_{X_t})(X_t)\nabla_{\phi(X_0)}X_t\big]\big|_{y=X_t}\Big\}$$

for $t\in[0,T]$ and $g\in C_b^1([0,T])$ with $g_0=0,g_T=1$. Consequently, there exists an increasing function $C:[0,\infty)\to(0,\infty)$ such that

$$\|(D^I_\phi P_tf)(\mu)\|\leq\frac{C_t}{\sqrt{t}}\big\{P_tf^2(\mu)-(P_tf(\mu))^2\big\}^{\frac{1}{2}},\quad t>0. \tag{2.4}$$

(3) *Assume further* **(C)**. *Then P_Tf is L-differentiable for any $T>0$ and $f\in\mathscr{B}_b(\mathbb{R}^d)$. As a result, for any $t>0$ and $\mu,\nu\in\mathscr{P}_2$,*

$$\|P_t^*\mu-P_t^*\nu\|_{var}:=\sup_{\|f\|_\infty\leq 1}|(P_t^*\mu)(f)-(P_t^*\nu)(f)|\leq\frac{C_t}{\sqrt{t}}\mathbb{W}_2(\mu,\nu). \tag{2.5}$$

Remark 2.1. When $b=0$, the Bismut formula and the L-differentiability of P_Tf for any $f\in\mathscr{B}_b(\mathbb{R}^d)$ have been proved in [23]. This is now included in Theorem 2.2 as a special case. When $b\neq 0$, to manage this singular term we have to use Zvonkin's transforms depending on the parameter in initial distributions.

Remark 2.2. For fixed $\mu \in \mathscr{P}_2$, consider the following decoupled SDE:

$$\mathrm{d}X_t^{x,\mu} = b_t(X_t^{x,\mu}, P_t^*\mu)\mathrm{d}t + \sigma_t(X_t^{x,\mu}, P_t^*\mu)\mathrm{d}W_t, \quad X_0^{x,\mu} = x \in \mathbb{R}^n.$$

Let $p_t^\mu(x,y)$ be the distribution density function of $X_t^{x,\mu}$. Derivatives of $p_t^\mu(x,y)$ in both x and μ have been presented in [12], where b and σ are assumed to be η-Hölder continuous for some $\eta \in (0,1]$ with respect to the spatial variable. In particular, these estimates imply estimates on $D^L P_t f(\mu)$ for $f \in \mathscr{B}_b(\mathbb{R}^d)$. In fact, let P_t^μ be the transition semigroup of $X_t^{x,\mu}$. Then we have

$$P_t f(\mu) = \int_{\mathbb{R}^n} P_t^\mu f(x)\mu(\mathrm{d}x) = \int_{\mathbb{R}^n\times\mathbb{R}^n} f(y)p_t^\mu(x,y)\mathrm{d}y\mu(\mathrm{d}x).$$

Consequently,

$$D^L P_t f(\mu)(z) = \int_{\mathbb{R}^n\times\mathbb{R}^n} f(y)\left\{D^L p_t^\mu(x,y)(z)\right\}\mathrm{d}y\mu(\mathrm{d}x) + \int_{\mathbb{R}^n} f(y)\nabla_z p_t^\mu(z,y)\mathrm{d}y.$$

This combined with [12, (3.9),Theorem 3.6] yields

$$\|D^L P_t f(\mu)(\cdot)\|_\infty \le C\|f\|_\infty \left(t^{-\frac{1}{2}} \vee t^{-\frac{1-\eta}{2}}\right).$$

3 Proof of Theorem 2.2

By **(H)**, (2.3) implies that $P_T f$ is intrinsically differentiable for $T>0$ and $f \in \mathscr{B}_b(\mathbb{R}^d)$, and (2.4) holds for some increasing function C. Then (2.5) follows from the estimate (see [9]):

$$|f(\mu) - f(\nu)| \le \sup_{\gamma\in\mathscr{P}_2} \|D^L f(\gamma)\|\mathbb{W}_2(\mu,\nu), \quad \mu,\nu \in \mathscr{P}_2. \tag{3.1}$$

To prove (3.1), let X, Y be two random variables such that

$$\mathscr{L}_X = \mu, \quad \mathscr{L}_Y = \nu, \quad \mathbb{E}|X-Y|^2 = \mathbb{W}_2(\mu,\nu)^2.$$

Then by taking $X_s = (1-s)X + sY$ for $s \in [0,1]$, (3.1) follows from the following chain rule for distributions of random variables, which is taken from [3, Theorem 2.1], see also [9, Theorem 6.5] and [23, Proposition 3.1] for earlier results under stronger conditions.

Lemma 3.1. *Let $\{X_\varepsilon\}_{\varepsilon\in[0,1]}$ be a family of random variables on $\mathbb{R}^d$ such that $\dot{X}_0 := \lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}(X_\varepsilon - X_0)$ exists in $L^2(\Omega\to\mathbb{R}^d;\mathbb{P})$. Let f be a real function on $\mathscr{P}_2$. If either $\mu := \mathscr{L}_{X_0}$ is atomless and f is L-differentiable at μ, or $f \in C^1$ in a neighborhood U of μ such that*

$$|D^L f(\mu)(x)| \le c(1+|x|), \quad x \in \mathbb{R}^d, \mu \in U$$

holds for some constant $c>0$, then

$$\lim_{\varepsilon\downarrow 0}\frac{f(\mathscr{L}_{X_\varepsilon}) - f(\mu)}{\varepsilon} = \mathbb{E}\langle D^L f(\mu)(X_0), \dot{X}_0\rangle.$$

Therefore, it suffices to prove Theorem 2.2(1), the formula (2.3), and Theorem 2.2(3).

3.1 Proof of Theorem 2.2(1)

We first explain that we may assume

$$B_t(\cdot,\mu)\in C^2(\mathbb{R}^d),\ \ \|\nabla B\|_{T,\infty}+\|\nabla^2 B\|_{T,\infty}<\infty,\ \ T>0. \tag{3.2}$$

Indeed, for $0\le \rho\in C_0^\infty(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d}\rho(x)\mathrm{d}x=1$, let

$$\tilde B_t(x,\mu):=\int_{\mathbb{R}^d} B_t(y,\mu)\rho(x-y)\mathrm{d}y,\ \ t\ge 0, x\in\mathbb{R}^d,\mu\in\mathscr{P}_2.$$

By **(H)**, this implies (3.2) for $\tilde B$ replacing B, and that $B_t-\tilde B_t$ is bounded with $\|\nabla(B-\tilde B)\|_{T,\infty}<\infty$. By combining $B_t-\tilde B_t$ with b, we may and do assume that B satisfies (3.2).

Next, we make the following distribution dependent Zvonkin's transform to regularize the SDE (1.1). For any $\lambda\ge 0$, $T>0$ and $\hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}:=C([0,T];\mathscr{P}_2)$, let

$$b_t^{\hat\mu}(x):=b_t(x,\hat\mu_t),\ \ B_t^{\hat\mu}(x):=B_t(x,\hat\mu_t)\ \ t\in[0,T],x\in\mathbb{R}^d,$$

and consider the following PDE for $u^{\lambda,\hat\mu}:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$:

$$\partial_t u_t^{\lambda,\hat\mu}+\frac{1}{2}\mathrm{Tr}(\sigma_t\sigma_t^*\nabla^2 u_t^{\lambda,\hat\mu})+\nabla_{b_t^{\hat\mu}+B_t^{\hat\mu}}u_t^{\lambda,\hat\mu}+b_t^{\hat\mu}=\lambda u_t^{\lambda,\hat\mu},\ \ u_T^{\lambda,\hat\mu}=0. \tag{3.3}$$

To solve this equation, we consider the flow induced by $B_t^{\hat\mu}$:

$$\partial_t\psi_t=-B_t^{\hat\mu}\circ\psi_t,\ \ \psi_T(x)=x,\ \ t\in[0,T],x\in\mathbb{R}^d.$$

By (3.2), ψ_t is a diffeomorphism on $\mathbb{R}^d$ and

$$\sup_{t\in[0,T],\hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}}\left\{\|\nabla\psi_t\|_\infty+\|\nabla^2\psi_t\|_\infty+\|\nabla\psi_t^{-1}\|_\infty+\|\nabla^2\psi_t^{-1}\|_\infty\right\}<\infty. \tag{3.4}$$

By [26], **(H)** implies that the PDE

$$\partial_t\bar u_t^{\lambda,\hat\mu}+\frac{1}{2}\mathrm{Tr}(\sigma_t\sigma_t^*\nabla^2\bar u_t^{\lambda,\hat\mu})+\nabla_{b_t^{\hat\mu}}\bar u_t^{\lambda,\hat\mu}+b_t^{\hat\mu}\circ\psi_t^{-1}=\lambda\bar u_t^{\lambda,\hat\mu},\ \ \bar u_T^{\lambda,\hat\mu}=0 \tag{3.5}$$

has a unique solution with

$$\lim_{\lambda\to\infty}\sup_{\hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}}\|\nabla\bar u^{\lambda,\hat\mu}\|_{T,\infty}=0,\ \ \ \sup_{\lambda\ge0,\hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}}\|\nabla^2\bar u^{\lambda,\hat\mu}\|_{T,\infty}<\infty.$$

So, $u_t^{\lambda,\hat\mu}:=\bar u_t^{\lambda,\hat\mu}\circ\psi_t$ solves (3.3) with

$$\lim_{\lambda\to\infty}\sup_{\hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}}\|\nabla u^{\lambda,\hat\mu}\|_{T,\infty}=0,\ \ \ \sup_{\lambda\ge0,\hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}}\|\nabla^2 u^{\lambda,\hat\mu}\|_{T,\infty}<\infty. \tag{3.6}$$

By the uniqueness of (3.5) and that a solution $u_t^{\lambda,\hat\mu}$ to (3.3) also gives a solution $\bar u_t^{\lambda,\hat\mu} := u_t^{\lambda,\hat\mu}\circ\psi_t^{-1}$ to (3.5), (3.3) has a unique solution. By (3.4) and (3.6), there exists a universal constant $\lambda_0>0$ such that

$$\sup_{\hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}} \|\nabla u^{\lambda,\hat\mu}\|_{T,\infty}\le \frac{1}{5},\quad \sup_{\hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}} \|\nabla^2 u^{\lambda,\hat\mu}\|_{T,\infty}\le \lambda_0,\quad \lambda\ge\lambda_0. \tag{3.7}$$

For any $\mu\in\mathscr{P}_2$, let X_0 be $\mathscr{F}_0$-measurable with $\mathscr{L}_{X_0}=\mu$, and simply denote

$$u_t^{\lambda,\mu}:=u_t^{\lambda,\hat\mu} \text{ for } \hat\mu_t=P_t^*\mu,\ \ b_t^\mu(x):=b_t(x,P_t^*\mu),\ \ B_t^\mu(x):=B_t(x,P_t^*\mu),\ \ (t,x)\in[0,T]\times\mathbb{R}^d.$$

Let $\theta_t^{\lambda,\mu}(x)=x+u_t^{\lambda,\mu}(x)$. By (3.3) and Itô's formula, we derive

$$\mathrm{d}\theta_t^{\lambda,\mu}(X_t)=\big\{B_t^\mu(X_t)+\lambda u_t^{\lambda,\mu}(X_t)\big\}\mathrm{d}t+\big\{(\nabla\theta_t^{\lambda,\mu})\sigma_t\big\}(X_t)\,\mathrm{d}W_t. \tag{3.8}$$

Then

$$Y_t:=\theta_t^{\lambda,\mu}(X_t),\quad t\in[0,T] \tag{3.9}$$

solves the SDE

$$\mathrm{d}Y_t=\tilde b_t^\mu(Y_t)\mathrm{d}t+\tilde\sigma_t^\mu(Y_t)\,\mathrm{d}W_t,\quad Y_0=\theta_0^{\lambda,\mu}(X_0) \tag{3.10}$$

where

$$\tilde b_t^\mu:=(B_t^\mu+\lambda u_t^{\lambda,\mu})\circ(\theta_t^{\lambda,\mu})^{-1},\quad \tilde\sigma_t^\mu:=\big\{(\nabla\theta_t^{\lambda,\mu})\sigma_t\big\}\circ(\theta_t^{\lambda,\mu})^{-1},\quad t\in[0,T]. \tag{3.11}$$

By **(H)** and (3.7), we have

$$\sup_{\mu\in\mathscr{P}_2}\big\{\|\nabla\tilde b^\mu\|_{T,\infty}+\|\nabla\tilde\sigma^\mu\|_{T,\infty}\big\}<\infty. \tag{3.12}$$

Let $\eta, X_0\in L^2(\Omega\to\mathbb{R}^d,\mathscr{F}_0;\mathbb{P})$. For any $\varepsilon\ge0$, let X_t^ε solve the SDE

$$\mathrm{d}X_t^\varepsilon=(B_t+b_t)(X_t^\varepsilon,\mathscr{L}_{X_t^\varepsilon})\mathrm{d}t+\sigma_t(X_t^\varepsilon)\mathrm{d}W_t,\quad X_0^\varepsilon=X_0+\varepsilon\eta. \tag{3.13}$$

By Itô's formula and (3.11), $Y_t^\varepsilon:=\theta_t^{\lambda,\mu}(X_t^\varepsilon)$ solves the SDE

$$\begin{aligned}\mathrm{d}Y_t^\varepsilon=&\Big\{\tilde b_t^\mu(Y_t^\varepsilon)+\big[\nabla\theta_t^{\lambda,\mu}[(B_t+b_t)(\cdot,\mathscr{L}_\cdot)-(B_t^\mu+b_t^\mu)]\big]\big((\theta_t^{\lambda,\mu})^{-1}(Y_t^\varepsilon)\big)\Big\}\mathrm{d}t\\&+\tilde\sigma_t^\mu(Y_t^\varepsilon)\mathrm{d}W_t,\quad t\in[0,T], Y_0^\varepsilon=\theta_0^{\lambda,\mu}(X_0+\varepsilon\eta).\end{aligned} \tag{3.14}$$

Lemma 3.2. *Under* (**H**), *the family* $\{\xi_t^\varepsilon:=\varepsilon^{-1}(Y_t^\varepsilon-Y_t)\}_{t\in[0,T],\varepsilon\in(0,1]}$ *is L^2-uniformly integrable, i.e.*

$$\lim_{n\to\infty}\sup_{\varepsilon\in(0,1]}\mathbb{E}\left(\sup_{t\in[0,T]}|\xi_t^\varepsilon|^2 1_{\{\sup_{t\in[0,T]}|\xi_t^\varepsilon|^2\ge n\}}\right)=0. \tag{3.15}$$

Proof. By (3.1), (3.11), (3.12) and (3.14), we find a constant $c > 0$ such that for any $\varepsilon \in (0, 1]$, the Itô's formula gives

$$\text{(3.16)} \qquad \mathrm{d}|\xi^\varepsilon_t|^2 \le c\big(|\xi^\varepsilon_t|^2 + \mathbb{E}|\xi^\varepsilon_t|^2\big)\mathrm{d}t + c|\xi^\varepsilon_t|^2 \mathrm{d}M^\varepsilon_t, \quad t \in [0, T], |\xi^\varepsilon_0|^2 \le c|\eta|^2,$$

for some martingale M^ε_t with $\mathrm{d}\langle M^\varepsilon\rangle_t \le \mathrm{d}t$. By BDG's inequality, this implies

$$\text{(3.17)} \qquad \mathbb{E}\Big[\sup_{t\in[0,T]} |\xi^\varepsilon_t|^2\Big] \le c_1 \mathbb{E}|\eta|^2, \quad \varepsilon \in [0, 1]$$

for some constant $c_1 > 0$. Combining this with (3.16) we obtain

$$\mathrm{d}\big\{|\xi^\varepsilon_t|^2 \mathrm{e}^{-c(t+M^\varepsilon_t)}\big\} \le cc_1(\mathbb{E}|\eta|^2)\mathrm{e}^{-c(t+M^\varepsilon_t)}\mathrm{d}t,$$

so that for

$$N_\varepsilon := \frac{1}{2}\sup_{s\in[0,T]} \mathrm{e}^{2cM^\varepsilon_s} + \frac{1}{2}\sup_{s\in[0,T]} \mathrm{e}^{-2cM^\varepsilon_s},$$

we find a constant $c_2 > 0$ such that

$$\text{(3.18)} \qquad \begin{aligned} D_\varepsilon : &= \sup_{t\in[0,T]} |\xi^\varepsilon_t|^2 \le \mathrm{e}^{cT}|\eta|^2 N_\varepsilon + cc_1(\mathbb{E}|\eta|^2)T\mathrm{e}^{cT}N_\varepsilon \\ &\le c_2(|\eta|^2 + \mathbb{E}|\eta|^2)N_\varepsilon, \ \varepsilon \in (0, 1]. \end{aligned}$$

For any $n > m > 1$, by BDG's inequality and $\mathrm{d}\langle M^\varepsilon\rangle_t \le \mathrm{d}t$, we find a constant $K > 0$ such that this implies

$$\begin{aligned} &\mathbb{E}(D_\varepsilon - c_2 n)^+ \le c_2\mathbb{E}\big[((|\eta|^2 + \mathbb{E}|\eta|^2)N_\varepsilon - n)^+\big] \\ &\le c_2 m\mathbb{E}\big[1_{\{|\eta|^2+\mathbb{E}|\eta|^2\le m\}}(N_\varepsilon - n/m)^+\big] + c_2\mathbb{E}\big[(|\eta|^2 + \mathbb{E}|\eta|^2)1_{\{|\eta|^2+\mathbb{E}|\eta|^2>m\}}\mathbb{E}(N_\varepsilon|\mathscr{F}_0)\big] \\ &\le \frac{c_2 m}{n/m}\mathbb{E}(N^2_\varepsilon) + K\mathbb{E}\big[(|\eta|^2 + \mathbb{E}|\eta|^2)1_{\{|\eta|^2+\mathbb{E}|\eta|^2>m\}}\big], \quad \varepsilon \in (0, 1]. \end{aligned}$$

Taking supremum with respect to $\varepsilon \in (0, 1]$ on both sides, and letting first $n \to \infty$ then $m \to \infty$ we finish the proof. □

For (L-)differentiable (real, vector, or matrix valued) functions f on $\mathbb{R}^d$ and g on $\mathscr{P}_2$, let

$$\text{(3.19)} \qquad \tilde{\Xi}^\varepsilon_g(t) := \frac{1}{\varepsilon}(g(\mathscr{L}_{X^\varepsilon_t}) - g(\mathscr{L}_{X_t})) - \mathbb{E}\langle D^L g(\mathscr{L}_{X_t})(X_t), \nabla(\theta^{\lambda,\mu}_t)^{-1}(Y_t)\xi^\varepsilon_t\rangle,$$

$$\text{(3.20)} \qquad \Xi^\varepsilon_f(t) := \frac{1}{\varepsilon}(f(Y^\varepsilon_t) - f(Y_t)) - \nabla_{\xi^\varepsilon_t} f(Y_t), \quad \varepsilon > 0, t \in [0, T].$$

The following lemma can be proved by using Lemma 3.1, (3.2) and the argument in the proof of [23, Lemma 3.4], we omit the details to save space.

Lemma 3.3. *Assume* (**H**)*. For any function* $f \in C^1(\mathbb{R}^d)$ *and* $g \in C_b^1(\mathscr{P}_2)$ *with* $\|\nabla f\|_\infty + \sup_{\mu\in\mathscr{P}_2} \|D^L g(\mu)\| < \infty$*, there exists a constant* $C > 0$ *such that*

$$|\Xi_f^\varepsilon(t)|^2 \le C\|\nabla f\|_\infty^2 |\xi_t^\varepsilon|^2, \quad |\tilde{\Xi}_g^\varepsilon(t)|^2 \le C\|D^L g\|_\infty^2 \mathbb{E}|\xi_t^\varepsilon|^2, \quad t \in [0,T],$$

and

$$\lim_{\varepsilon\to 0}(\mathbb{E}|\Xi_f^\varepsilon(t)|^2 + |\tilde{\Xi}_g^\varepsilon(t)|^2) = 0, \quad t \in [0,T].$$

Lemma 3.4. *Assume* (**H**)*. Then the limit*

$$\nabla_\eta Y_t := \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(Y_t^\varepsilon - Y_t) \tag{3.21}$$

exists in $L^2(\Omega \to \mathscr{C}_T; \mathbb{P})$*, and is the unique solution to the linear equation*

$$\begin{aligned} v_t^\eta = &[\nabla\theta_0^{\lambda,\mu}](X_0)\eta + \int_0^t \Big\{\nabla_{v_s^\eta}\tilde{b}_s^\mu(Y_s) + F_s^\mu\big(X_s\big)v_s^\eta\Big\}\mathrm{d}s \\ &+ \int_0^t \nabla_{v_s^\eta}\tilde{\sigma}_s^\mu(Y_s)\,\mathrm{d}W_s, \quad t \in [0,T], \end{aligned} \tag{3.22}$$

where for any random variables X, v *on* $\mathbb{R}^d$ *and* $s \in [0,T]$*,*

$$F_s^\mu(X)v := \{\nabla\theta_s^{\lambda,\mu}\}(X)\mathbb{E}\big[D^L(B_s + b_s)(y, \mathscr{L}_X)(X)(\nabla\theta_s^{\lambda,\mu})^{-1}(X)v\big]\Big|_{y=X}. \tag{3.23}$$

Consequently, for any $p \ge 1$ *there exists a constant* $c > 0$ *such that*

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |\nabla_\eta Y_t|^p \Big| \mathscr{F}_0\Big) \le c\big(|\eta|^p + (\mathbb{E}|\eta|^2)^{\frac{p}{2}}\big), \quad \eta, X_0 \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_0; \mathbb{P}). \tag{3.24}$$

Proof. For the existence of $\nabla_\eta Y_t$ in $L^2(\Omega \to \mathscr{C}_T; \mathbb{P})$, it suffices to verify

$$\lim_{\varepsilon,\delta\to 0} \mathbb{E}\Big(\sup_{t\in[0,T]} |\xi_t^\varepsilon - \xi_t^\delta|^2\Big) = 0. \tag{3.25}$$

By (3.23), (3.14) and (3.20), we obtain

$$\xi_t^\varepsilon = \frac{1}{\varepsilon}(\theta_0^{\lambda,\mu}(X_0 + \varepsilon\eta) - \theta_0^{\lambda,\mu}(X_0)) + \int_0^t A_s^\varepsilon \mathrm{d}s + \int_0^t B_s^\varepsilon \mathrm{d}W_s, \tag{3.26}$$

where for $(\Xi, \tilde{\Xi})$ in (3.19) and (3.20),

$$\begin{aligned} A_s^\varepsilon &:= \frac{1}{\varepsilon}\Big\{(\tilde{b}_s^\mu(Y_s^\varepsilon) - \tilde{b}_s^\mu(Y_s)) + [\nabla\theta_s^{\lambda,\mu}[(B_s + b_s)(\cdot, \mathscr{L}_\cdot) - B_s^\mu - b_s^\mu)]]((\theta_s^{\lambda,\mu})^{-1}(Y_s^\varepsilon))\Big\} \\ &= \Xi_{\tilde{b}_s^\mu}^\varepsilon(s) + \nabla_{\xi_s^\varepsilon}\tilde{b}_s^\mu(Y_s) + F_s^\mu(X_s^\varepsilon)\xi_s^\varepsilon + \{\nabla\theta_s^{\lambda,\mu}(X_s^\varepsilon)\}\tilde{\Xi}_{(B_s+b_s)(X_s^\varepsilon,\cdot)}^\varepsilon(s), \end{aligned}$$

$$B^\varepsilon_s := \frac{1}{\varepsilon}(\tilde\sigma^\mu_s(Y^\varepsilon_s) - \tilde\sigma^\mu_s(Y_s)) = \Xi^\varepsilon_{\tilde\sigma^\mu_s}(s) + \nabla_{\xi^\varepsilon_s}\tilde\sigma^\mu_s(Y_s).$$

Since $\|\nabla^2\theta^{\lambda,\mu}_0\|_\infty < \infty$ due to (3.7), we have

$$\lim_{\varepsilon,\delta\downarrow 0}\mathbb{E}\bigg|\frac{\theta^{\lambda,\mu}_0(X_0+\varepsilon\eta)-\theta^{\lambda,\mu}_0(X_0)}{\varepsilon} - \frac{\theta^{\lambda,\mu}_0(X_0+\delta\eta)-\theta^{\lambda,\mu}_0(X_0)}{\delta}\bigg|^2 = 0.$$

Moreover, by Lemma 3.3, **(H)** and (3.7) imply

$$\begin{aligned}\lim_{\varepsilon,\delta\to 0}\mathbb{E}\big\{&|\tilde\Xi^\varepsilon_{(B_s+b_s)(X^\varepsilon_s,\cdot)}(s)|^2 + |\Xi^\varepsilon_{\tilde\sigma^\mu_s}(s)|^2 + |\Xi^\varepsilon_{\tilde b^\mu_s}(s)|^2\\ &+ \tilde\Xi^\delta_{(B_s+b_s)(X^\delta_s,\cdot)}(s)|^2 + |\Xi^\delta_{\tilde\sigma^\mu_s}(s)|^2 + |\Xi^\delta_{\tilde b^\mu_s}(s)|^2 + |F^\mu_s(X^\varepsilon_s) - F^\mu_s(X^\delta_s)|^2\big\} = 0.\end{aligned}$$

Combining these with (3.26) and applying BDG's inequality, we find a constant $c>0$ such that

$$\limsup_{\varepsilon,\delta\downarrow 0}\mathbb{E}\Big[\sup_{s\in[0,t]}|\xi^\varepsilon_s - \xi^\delta_s|^2\Big] \le c\int_0^t \limsup_{\varepsilon,\delta\downarrow 0}\mathbb{E}\Big[\sup_{s\in[0,r]}|\xi^\varepsilon_s - \xi^\delta_s|^2\Big]\mathrm{d}r,\quad t\in[0,T].$$

Since Lemma 3.2 ensures

$$\limsup_{\varepsilon,\delta\downarrow 0}\mathbb{E}\sup_{s\in[0,T]}\Big[|\xi^\varepsilon_s - \xi^\delta_s|^2\Big] < \infty,$$

by Gronwall's lemma we prove (3.25).

Finally, by **(H)**, Lemma 3.3 and the definitions of $F^\mu_s, \nabla_\eta Y_s, A^\varepsilon_s, B^\varepsilon_s$ and the fact $X_s = (\theta^{\lambda,\mu}_s)^{-1}(Y_s)$, we may let $\varepsilon\downarrow 0$ in (3.26) to derive (3.22), which together with **(H)** implies

$$\mathbb{E}|v^\eta_t|^2 \le c_1\mathbb{E}|\eta|^2,\quad t\in[0,T].$$

Combining this with (3.22), **(H)** and BDG's inequality, we prove (3.24) for some constant $c>0$. □

Proof of Theorem 2.2(1). By (3.21), Lemma 3.3 and (3.7), we can find a constant $c>0$ such that

$$\begin{aligned}&\lim_{\varepsilon\to 0}\mathbb{E}\left(\sup_{t\in[0,T]}\left|\frac{1}{\varepsilon}(X^\varepsilon_t - X_t) - \{\nabla(\theta^{\lambda,\mu}_t)^{-1}\}(Y_t)\nabla_\eta Y_t\right|^2\right)\\ =&\lim_{\varepsilon\to 0}\mathbb{E}\left(\sup_{t\in[0,T]}\left|\frac{1}{\varepsilon}((\theta^{\lambda,\mu}_t)^{-1}(Y^\varepsilon_t) - (\theta^{\lambda,\mu}_t)^{-1}(Y_t)) - \{\nabla(\theta^{\lambda,\mu}_t)^{-1}\}(Y_t)\nabla_\eta Y_t\right|^2\right)\\ \le& c\lim_{\varepsilon\to 0}\mathbb{E}\sup_{t\in[0,T]}|\xi^\varepsilon_t - \nabla_\eta Y_t|^2 = 0.\end{aligned}$$

This implies

$$\nabla_\eta X_t := \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}(X^\varepsilon_t - X_t) = \big\{\nabla(\theta^{\lambda,\mu}_t)^{-1}\big\}(Y_t)\nabla_\eta Y_t \tag{3.27}$$

exists in $L^2(\Omega\to\mathscr{C}_T;\mathbb{P})$, together with (3.7) and (3.24), yields (2.2). □

3.2 Proof of (2.3)

Lemma 3.5. *Assume* **(H)**. *Let* $k \in [0,T)$ *and* $g \in C_b^1([0,T])$ *with* $g_k = 0$ *and* $g_T = 1$. *Then*

$$\begin{aligned}\nabla_\eta \mathbb{E}(f(Y_T)|\mathscr{F}_k) &:= \lim_{\varepsilon\to 0} \frac{\mathbb{E}(f(Y_T^\varepsilon) - f(Y_T)|\mathscr{F}_k)}{\varepsilon} \\ &= \mathbb{E}\bigg(f(Y_T) \int_k^T \langle \zeta_t^\eta, \mathrm{d}W_t \rangle \bigg| \mathscr{F}_k \bigg), \quad f \in \mathscr{B}_b(\mathbb{R}^d), \eta \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_0; \mathbb{P}),\end{aligned} \tag{3.28}$$

where

$$\zeta_t^\eta := \sigma_t(X_t)^{-1} \Big\{ g_t' \nabla_\eta X_t + g_t \mathbb{E}[D^L(B_t + b_t)(y, \mathscr{L}_{X_t})(X_t) \nabla_\eta X_t]|_{y=X_t} \Big\}. \tag{3.29}$$

Consequently, it holds

$$\begin{aligned}\nabla_\eta \mathbb{E}(f(X_T)|\mathscr{F}_k) &:= \lim_{\varepsilon\to 0} \frac{\mathbb{E}(f(X_T^\varepsilon) - f(X_T)|\mathscr{F}_k)}{\varepsilon} \\ &= \mathbb{E}\bigg(f(X_T) \int_k^T \langle \zeta_t^\eta, \mathrm{d}W_t \rangle \bigg| \mathscr{F}_k \bigg), \qquad f \in \mathscr{B}_b(\mathbb{R}^d), \eta \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_0; \mathbb{P}).\end{aligned} \tag{3.30}$$

Proof. Simply denote $v_t = g_t \nabla_\eta Y_t$ for $t \in [0,T]$. By Itô's formula, (3.22) and (3.27), we obtain

$$\begin{aligned}\mathrm{d}v_t =& \nabla_{v_t} \tilde{b}_t^\mu(Y_t)\mathrm{d}t + \nabla_{v_t} \tilde{\sigma}_t^\mu(Y_t)\, \mathrm{d}W_t + g_t' v_t \mathrm{d}t \\ &+ g_t \nabla \theta_t^{\lambda,\mu}(X_t) \mathbb{E}[D^L(B_t + b_t)(y, \mathscr{L}_{X_t})(X_t) \nabla_\eta X_t]|_{y=X_t} \mathrm{d}t, \quad v_k = 0, t \in [k,T].\end{aligned} \tag{3.31}$$

On the other hand, let $h_t = \int_k^t \zeta_s^\eta \mathrm{d}s$ for $t \in [k,T]$. By (3.12) and [22, Theorem 2.2.1], the Malliavin derivative $D_h Y_t$ of Y_t along h satisfies

$$\mathrm{d}D_h Y_t = \nabla_{D_h Y_t} \tilde{b}_t^\mu(Y_t)\mathrm{d}t + \nabla_{D_h Y_t} \tilde{\sigma}_t^\mu(Y_t)\, \mathrm{d}W_t + \tilde{\sigma}_t^\mu(Y_t)\, \mathrm{d}h_t, \quad D_h Y_k = 0, t \in [k,T].$$

By the definition of h we see that $D_h Y_t$ solves (3.31), so that the uniqueness implies $v_t = D_h Y_t$. In particular, $\nabla_\eta Y_T = v_T^\eta = D_h Y_T$. Thus, for any $f \in C_b^1(\mathbb{R}^d)$, by the dominated convergence theorem due to (3.24), and the integration by parts formula for Malliavin derivative ([22, Lemma 1.2.1]), we obtain

$$\begin{aligned}\nabla_\eta \mathbb{E}(f(Y_T)|\mathscr{F}_k) &= \mathbb{E}(\nabla_{v_T^\eta} f(Y_T)|\mathscr{F}_k) \\ &= \mathbb{E}(\nabla_{D_h Y_T} f(Y_T)|\mathscr{F}_k) \\ &= \mathbb{E}(D_h f(Y_T)|\mathscr{F}_k) = \mathbb{E}\Big(f(Y_T) \int_k^T \langle \zeta_t^\eta, \mathrm{d}W_t \rangle |\mathscr{F}_k \Big).\end{aligned}$$

So, (3.28) holds for $f \in C_b^1(\mathbb{R}^d)$. By an approximation argument (see [23, Page 4764]), the formula also holds for $f \in \mathscr{B}_b(\mathbb{R}^d)$. Since $\theta_T^{\lambda,\mu} = \mathrm{Id}$, we have $(X_T, X_T^\varepsilon) = (Y_T, Y_T^\varepsilon)$ so that this implies (3.30).

□

Proof of (2.3). Let $\eta=\phi(X_0)$. We have

$$\mathscr{L}_{X_0^\varepsilon}=\mathscr{L}_{X_0+\varepsilon\phi(X_0)}=\mu\circ(\mathrm{Id}+\varepsilon\phi)^{-1},\ \ \varepsilon\in[0,1].$$

Moreover, (3.29) with $\eta=\phi(X_0)$ implies $\zeta_t^\eta=\zeta_t^\phi$ for ζ_t^ϕ in (2.3). So, letting $k=0$ in (3.30) and taking expectation on both sides, we prove (2.3). □

3.3 Proof of Theorem 2.2(3)

Let $u^{\lambda,\hat\mu}$ solve (3.3) for $\hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}$. We first characterize the Lipschitz continuity of $u^{\lambda,\mu}$ in μ.

Lemma 3.6. *Assume* **(H)** *and let $T>0$. There exists a constant $c>0$ such that for any $\hat\mu,\hat\nu\in C([0,T];\mathscr{P}_2)$,*

$$\|u_s^{\lambda,\hat\mu}-u_s^{\lambda,\hat\nu}\|_\infty+\|\nabla u_s^{\lambda,\hat\mu}-\nabla u_s^{\lambda,\hat\nu}\|_\infty\le c\int_s^T\frac{\mathrm{e}^{-\lambda(t-s)}}{\sqrt{t-s}}\mathbb{W}_2(\hat\mu_t,\hat\nu_t)\mathrm{d}t,\ \ s\in[0,T].\tag{3.32}$$

Proof. Let $P_{s,t}^{\hat\nu}$ be the Markov semigroup associated with the SDE:

$$\mathrm{d}X_{s,t}=\{B_t^{\hat\nu}+b_t^{\hat\nu}\}(X_{s,t})\mathrm{d}t+\sigma_t(X_{s,t})\mathrm{d}W_t,\ \ t\ge s\ge0.$$

By (2.4) with $B+b\equiv B^{\hat\nu}+b^{\hat\nu}$ and $\mu=\delta_x$ for $x\in\mathbb{R}^d$, we find a constant $c_1>0$ such that

$$\|\nabla P_{s,t}^{\hat\nu}f\|_\infty\le\frac{c_1}{\sqrt{t-s}}\|f\|_\infty,\ \ f\in\mathscr{B}_b(\mathbb{R}^d),0\le s<t\le T.\tag{3.33}$$

Next, by Duhamel's formula, the unique solution to (3.3) satisfies

$$u_s^{\lambda,\hat\mu}=\int_s^T\mathrm{e}^{-\lambda(t-s)}P_{s,t}^{\hat\nu}(\nabla_{B_t^{\hat\mu}+b_t^{\hat\mu}-B_t^{\hat\nu}-b_t^{\hat\nu}}u_t^{\lambda,\hat\mu}+b_t^{\hat\mu})\mathrm{d}t,\ \ \hat\mu\in\mathscr{C}_{T,\mathscr{P}_2}.$$

Moreover, by (3.1) and **(H)**, we find a constant $c>0$ such that

$$\|B_t^{\hat\mu}+b_t^{\hat\mu}-B_t^{\hat\nu}-b_t^{\hat\nu}\|_\infty\le c\mathbb{W}_2(\hat\mu_t,\hat\nu_t),\ \ t\in[0,T].$$

Combining these with (3.7) and (3.33), we find a constant $c_2>0$ such that

$$\begin{aligned}|u_s^{\lambda,\hat\mu}-u_s^{\lambda,\hat\nu}|&\le\int_s^T\big|\mathrm{e}^{-\lambda(t-s)}P_{s,t}^{\hat\nu}(\nabla_{B_t^{\hat\mu}+b_t^{\hat\mu}-B_t^{\hat\nu}-b_t^{\hat\nu}}u_t^{\lambda,\hat\nu}+b_t^{\hat\mu}-b_t^{\hat\nu})\big|\mathrm{d}t\\&\le c_2\int_s^T\mathrm{e}^{-\lambda(t-s)}\mathbb{W}_2(\hat\mu_t,\hat\nu_t)\mathrm{d}t,\\|\nabla u_s^{\lambda,\hat\mu}-\nabla u_s^{\lambda,\hat\nu}|&\le\int_s^T\Big|\mathrm{e}^{-\lambda(t-s)}\nabla P_{s,t}^{\hat\nu}\big(\nabla_{B_t^{\hat\mu}+b_t^{\hat\mu}-B_t^{\hat\nu}-b_t^{\hat\nu}}u_t^{\lambda,\hat\nu}+b_t^{\hat\mu}-b_t^{\hat\nu}\big)\Big|\mathrm{d}t\\&\le c_2\int_s^T\frac{\mathrm{e}^{-\lambda(t-s)}}{\sqrt{t-s}}\mathbb{W}_2(\hat\mu_t,\hat\nu_t)\mathrm{d}t,\ \ t\in[0,T].\end{aligned}$$

Therefore, (3.32) holds for some constant $c>0$. □

For $r\in[0,1],\mu\in\mathscr{P}_2, X_0\in L^2(\Omega\to\mathbb{R}^d,\mathscr{F}_0;\mathbb{P})$ and $\phi\in L^2(\mathbb{R}^d\to\mathbb{R}^d;\mu)$, let X_t^r be the solution of (1.1) with initial value $X_0^r=X_0+r\phi(X_0)$, and denote

$$\mu_t^r:=\mathscr{L}_{X_t^r},\quad t\in[0,T]. \tag{3.34}$$

We have

$$\mu_t^0=\mathscr{L}_{X_t},\quad \mu_0^r=\mathscr{L}_{X_0+r\phi(X_0)}=\mu\circ(\mathrm{Id}+r\phi)^{-1},\quad r\in[0,1],t\in[0,T]. \tag{3.35}$$

Let $\theta_t^{\lambda,\mu^r}:=\mathrm{Id}+u_t^{\lambda,\mu^r}$ for u_t^{λ,μ^r} solving (3.3) with μ^r replacing $\hat\mu$, and let

$$Y_t^{r,\varepsilon}:=\theta_t^{\lambda,\mu^r}(X_t^{r+\varepsilon}),\quad r,\varepsilon\in[0,1],t\in[0,T]. \tag{3.36}$$

Then

$$Y_t:=Y_t^{0,0}=\theta_t^{\lambda,\mu}(X_t),\quad \tilde Y_t^r:=Y_t^{r,0}=\theta_t^{\lambda,\mu^r}(X_t^r),\quad r\in[0,1],t\in[0,T]. \tag{3.37}$$

By (3.22) and (3.27) for the solution to (1.1) with initial value $X_0+r\phi(X_0)$ and $\eta=\phi(X_0)$, we obtain

$$\nabla_{\phi(X_0)}X_t^r=(\nabla\theta_t^{\lambda,\mu^r})^{-1}(X_t^r)v_t^{\phi,r},\quad v_t^{\phi,r}:=\nabla_{\phi(X_0)}\tilde Y_t^r, \tag{3.38}$$

$$\begin{aligned} v_t^{\phi,r}=&[\nabla\theta_0^{\lambda,\mu^r}](X_0+r\phi(X_0))\phi(X_0)+\int_0^t\nabla_{v_s^{\phi,r}}\tilde\sigma_s^{\mu^r}(\tilde Y_s^r)\,\mathrm{d}W_s\\ &+\int_0^t\Big\{\nabla_{v_s^{\phi,r}}\tilde b_s^{\mu^r}(\tilde Y_s^r)+F_s^{\mu^r}\big(X_s^r\big)v_s^{\phi,r}\Big\}\mathrm{d}s,\quad t\in[0,T],\ r\in[0,1],\end{aligned} \tag{3.39}$$

where F^{μ^r} is defined in (3.23) with μ^r replacing μ.

Lemma 3.7. *Assume* **(H)**.

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

$$\sup_{r\in[0,1],t\in[0,T]}\mathbb{W}_2(\mu_t^r,\mu_t)\le c\|\phi\|_{L^2(\mu)},\quad \phi\in L^2(\mathbb{R}^d\to\mathbb{R}^d;\mu). \tag{3.40}$$

Moreover, for any $p\ge 2$ there exists a constant $c(p)>0$ such that

$$\begin{aligned}&\sup_{r\in[0,1]}\mathbb{E}\Big(\sup_{t\in[0,T]}\big(|\tilde Y_t^r-Y_t|^p+|v_t^{\phi,r}|^p+|X_t^r-X_t|^p+|\nabla_{\phi(X_0)}X_t^r|^p\big)\Big|\mathscr{F}_0\Big)\\ &\le c(p)\big(|\phi(X_0)|^p+\|\phi\|_{L^2(\mu)}^p\big),\quad \phi\in L^2(\mathbb{R}^d\to\mathbb{R}^d;\mu).\end{aligned} \tag{3.41}$$

(2) *Let $k\in[0,T)$ and $g\in C_b^1([0,T])$ with $g_k=0$ and $g_T=1$. For any $f\in\mathscr{B}_b(\mathbb{R}^d)$ and $\phi\in L^2(\mathbb{R}^d\to\mathbb{R}^d;\mu)$,*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}r}\mathbb{E}(f(X_T^r)|\mathscr{F}_k)&=\lim_{\varepsilon\to0}\frac{\mathbb{E}(f(X_T^{r+\varepsilon})-f(X_T^r)|\mathscr{F}_k)}{\varepsilon}\\ &=\mathbb{E}\bigg(f(X_T^r)\int_k^T\langle\zeta_t^{\phi,r},\mathrm{d}W_t\rangle|\mathscr{F}_k\bigg),\quad r\in[0,1]\end{aligned} \tag{3.42}$$

holds for

$$\zeta_t^{\phi,r} = N_t(X_t^r)\nabla_{\phi(X_0)}X_t^r, \quad r\in[0,1], t\in[0,T], \tag{3.43}$$

where for $X, v\in L^2(\Omega\to\mathbb{R}^d;\mathbb{P})$,

$$N_t(X)v := \sigma_t(X)^{-1}\Big\{g_t'v + g_t\mathbb{E}D^L(B_t+b_t)(y,\mathscr{L}_X)(X)v\big|_{y=X}\Big\}, \quad t\in[0,T]. \tag{3.44}$$

Consequently, it holds

$$\frac{\mathrm{d}}{\mathrm{d}r}P_Tf(\mu\circ(\mathrm{Id}+r\phi)^{-1}) = \mathbb{E}\bigg(f(X_T^r)\int_0^T\langle\zeta_t^{\phi,r},\mathrm{d}W_t\rangle\bigg), \quad r\in[0,1]. \tag{3.45}$$

Proof. (1) Recall that $Y_t^r = \theta_t^{\lambda,\mu}(X_t^r)$. Since $\mathbb{W}_2(\mu_t^r,\mu_t)^2 \le \mathbb{E}|X_t^r - X_t|^2$, (3.40) follows from (3.17) and (3.7). To prove (3.41), let $\eta = \phi(X_0)$ and denote $\mu = \mathscr{L}_{X_0}$. By (3.36) for $\varepsilon = 0$, we have

$$X_t^r = (\theta_t^{\lambda,\mu^r})^{-1}(\tilde{Y}_t^r), \quad r\in[0,1], t\in[0,T], \tag{3.46}$$

and by (3.18),

$$\varepsilon\xi_t^\varepsilon = Y_t^\varepsilon - Y_t, \quad t\in[0,T], \varepsilon\in[0,1].$$

Then (3.18), (3.32), (3.40), (3.39) and Lemma 3.4 with $\tilde{Y}_t^r$, μ^r replacing Y_t, μ respectively imply

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{t\in[0,T]}\big\{|\tilde{Y}_t^r - Y_t|^p + |v_t^{\phi,r}|^p\big\}\Big|\mathscr{F}_0\Big)\\
&\le c(p)\mathbb{E}\Big(\sup_{t\in[0,T]}\big\{|\tilde{Y}_t^r - Y_t^r|^p + |Y_t^r - Y_t|^p + |v_t^{\phi,r}|^p\big\}\Big|\mathscr{F}_0\Big)\\
&\le c(p)\big(|\phi(X_0)|^p + \|\phi\|_{L^2(\mu)}^p\big)
\end{aligned}$$

for some constant $c(p)>0$. By (3.7) and (3.46), this implies (3.41).

(2) By (3.36) and Lemma 3.5 for $\tilde{Y}_t^r$ replacing Y_t and μ^r replacing μ, we obtain

$$\begin{aligned}
&\nabla_{\phi(X_0)}\mathbb{E}(f(\tilde{Y}_T^r)|\mathscr{F}_k) := \lim_{\varepsilon\downarrow 0}\frac{\mathbb{E}(f(Y_T^{r,\varepsilon}) - \mathbb{E}f(\tilde{Y}_T^r)|\mathscr{F}_k)}{\varepsilon}\\
&= \mathbb{E}\left(f(\tilde{Y}_T^r)\int_k^T\langle\zeta_t^{\phi,r},\mathrm{d}W_t\rangle|\mathscr{F}_k\right), \quad f\in\mathscr{B}_b(\mathbb{R}^d), r\in[0,1], k\in[0,T).
\end{aligned} \tag{3.47}$$

Combining this with (3.34), (3.36) and (3.37), we derive

$$\frac{\mathrm{d}}{\mathrm{d}r}\mathbb{E}(f(X_T^r)|\mathscr{F}_k) = \lim_{\varepsilon\downarrow 0}\frac{\mathbb{E}(f(X_T^{r+\varepsilon}) - \mathbb{E}f(X_T^r)|\mathscr{F}_k)}{\varepsilon}$$

$$= \lim_{\varepsilon\downarrow 0} \frac{\mathbb{E}((f\circ(\theta_T^{\lambda,\mu^r})^{-1})(Y_T^{r,\varepsilon}) - \mathbb{E}(f\circ(\theta_T^{\lambda,\mu^r})^{-1})(\tilde{Y}_T^r)|\mathscr{F}_k)}{\varepsilon}$$
$$= \mathbb{E}\left(f(X_T^r)\int_k^T \langle \zeta_t^{\phi,r}, \mathrm{d}W_t\rangle|\mathscr{F}_k\right).$$

In particular, for $k=0$, taking expectation on both sides of (3.42), we get (3.45). Then the proof is finished. $\square$

Having Lemmas 3.6 and 3.7 in hands, we prove the L-differentiability of P_Tf as follows by modifying step (c) in the proof of [23, Theorem 2.1].

Proof of Theorem 2.2(3). Let $\{\zeta_t^{\phi,r}\}_{r\in[0,1]}$ be in Lemma 3.7. By **(H)**, (3.41) and the Riesz representation theorem, there exists $\gamma\in L^2(\mathbb{R}^d\to\mathbb{R}^d;\mu)$ such that

$$\langle \gamma,\phi\rangle_{L^2(\mu)} = \mathbb{E}\bigg(f(X_T)\int_0^T \langle \zeta_t^{\phi,0}, \mathrm{d}W_t\rangle\bigg),\quad \phi\in L^2(\mathbb{R}^d\to\mathbb{R}^d;\mu).$$

By (3.45) and (3.35), we obtain

$$\begin{aligned}
&\frac{|P_Tf(\mu\circ(\mathrm{Id}+\phi)^{-1}) - P_Tf(\mu) - \langle\gamma,\phi\rangle_{L^2(\mu)}|}{\|\phi\|_{L^2(\mu)}}\\
&\le \frac{1}{\|\phi\|_{L^2(\mu)}}\int_0^1 \left|\mathbb{E}\left(f(X_T^r)\int_0^T\langle\zeta_t^{\phi,r},\mathrm{d}W_t\rangle - f(X_T)\int_0^T\langle\zeta_t^{\phi,0},\mathrm{d}W_t\rangle\right)\right|\mathrm{d}r\\
&\le \varepsilon_1(\phi)+\varepsilon_2(\phi)+\varepsilon_3(\phi),
\end{aligned}$$

where, by (3.43) and (3.38)

$$\begin{aligned}
\varepsilon_1(\phi) &:= \frac{1}{\|\phi\|_{L^2(\mu)}}\int_0^1 \bigg|\mathbb{E}\bigg(\big(f(X_T^r)-f(X_T)\big)\int_0^T\langle\zeta_t^{\phi,0},\mathrm{d}W_t\rangle\bigg)\bigg|\mathrm{d}r,\\
\varepsilon_2(\phi) &:= \frac{\|f\|_\infty}{\|\phi\|_{L^2(\mu)}}\int_0^1 \mathbb{E}\bigg|\int_0^T \big\langle\{N_t(X_t^r)-N_t(X_t)\}(\nabla\theta_t^{\lambda,\mu})^{-1}(X_t)v_t^\phi,\mathrm{d}W_t\big\rangle\bigg|\mathrm{d}r,\\
\varepsilon_3(\phi) &:= \frac{\|f\|_\infty}{\|\phi\|_{L^2(\mu)}}\int_0^1 \mathbb{E}\bigg|\int_0^T \big\langle N_t(X_t^r)\big\{(\nabla\theta_t^{\lambda,\mu^r})^{-1}(X_t^r)v_t^{\phi,r} - (\nabla\theta_t^{\lambda,\mu})^{-1}(X_t)v_t^\phi\big\},\mathrm{d}W_t\big\rangle\bigg|\mathrm{d}r.
\end{aligned}$$

So, it suffices to prove

$$\lim_{\|\phi\|_{L^2(\mu)}\to 0}\big\{\varepsilon_1(\phi)+\varepsilon_2(\phi)+\varepsilon_3(\phi)\big\}=0.$$

a) We first modify the proof of [23, (2.3)] to verify

$$\lim_{\|\phi\|_{L^2(\mu)}\to 0}\varepsilon_1(\phi)=0. \tag{3.48}$$

Denote

$$I_k := \int_0^k \big\langle \zeta_t^{\phi,0}, \mathrm{d}W_t\big\rangle, \quad I_k^{\phi,r} = \big|\mathbb{E}[I_k\{f(X_T^r) - f(X_T)\}]\big|, \quad k \in (0,T), r \in (0,1].$$

By (3.42) with $g_t := \frac{t-k}{T-k}$ for $t \in [k,T]$, **(H)**, (3.27) and (3.41), and noting that I_k is $\mathscr{F}_k$-measurable, we find constants $c_1, c_2 > 0$ such that

$$\begin{aligned}
I_k^{\phi,r} &= \bigg|\mathbb{E}\bigg[I_k \int_0^r \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathbb{E}(f(X_T^\varepsilon)|\mathscr{F}_k)\mathrm{d}\varepsilon\bigg]\bigg| \\
&\le \mathbb{E}\bigg[|I_k| \cdot \bigg|\int_0^r \mathbb{E}\bigg(f(X_T^\varepsilon)\int_k^T \langle \zeta_t^{\phi,\varepsilon}, \mathrm{d}W_t\rangle\bigg|\mathscr{F}_k\bigg)\mathrm{d}\varepsilon\bigg|\bigg] \\
&\le c_1\|f\|_\infty \int_0^r \mathbb{E}\bigg[|I_k|\bigg(\int_k^T \big\{\big|\nabla_{\phi(X_0)}X_t^\varepsilon\big|^2 + \mathbb{E}|\nabla_{\phi(X_0)}X_t^\varepsilon|^2\big\}\mathrm{d}t\bigg)^{\frac{1}{2}}\bigg]\mathrm{d}\varepsilon \\
&\le c_2\|f\|_\infty\|\phi\|_{L^2(\mu)}^2, \quad k \in [0,T),\ f \in \mathscr{B}_b(\mathbb{R}^d).
\end{aligned}$$

So,

$$\lim_{\|\phi\|_{L^2(\mu)}\to 0} \frac{1}{\|\phi\|_{L^2(\mu)}}\int_0^1 I_k^{\phi,r}\mathrm{d}r = 0, \quad k \in (0,T).$$

Combining this with (3.41) and (3.43), we obtain

$$\begin{aligned}
&\lim_{\|\phi\|_{L^2(\mu)}\to 0} \varepsilon_1(\phi) \le \lim_{\|\phi\|_{L^2(\mu)}\to 0} \frac{1}{\|\phi\|_{L^2(\mu)}}\int_0^1 \bigg|\mathbb{E}\bigg(\big(f(X_T^r) - f(X_T)\big)\int_k^T \langle \zeta_t^{\phi,0}, \mathrm{d}W_t\rangle\bigg)\bigg|\mathrm{d}r \\
&\le c\|f\|_\infty\sqrt{T-k}, \quad k \in (0,T)
\end{aligned}$$

for some constant $c > 0$. By letting $k \uparrow T$ we prove (3.48).

b) For any $\phi \in L^2(\mathbb{R}^d \to \mathbb{R}^d; \mu)$, $s \in [0,T]$ and $r \in [0,1]$, let

$$h_{s,r}(\phi) := \big(\mathbb{E}|D^L(B_s + b_s)(y, \mathscr{L}_{X_s^r})(X_s^r) - D^L(B_s + b_s)(z, \mathscr{L}_{X_s})(X_s)|^2\big)^{\frac{1}{2}}\big|_{(y,z)=(X_s^r, X_s)}.$$

Noting that $h_{s,r}(\phi)^2 < 4(\|D^L B\|^2 + \|D^L b\|^2)_{T,\infty}$ due to **(H)**, by **(C)**, (3.41) and the dominated convergence theorem, we obtain

$$\lim_{\|\phi\|_{L^2(\mu)}\to 0} h_{s,r}(\phi) = 0, \qquad \sup_{(s,r)\in[0,T]\times[0,1]}\ \sup_{\|\phi\|_{L^2(\mu)}\le 1} h_{s,r}(\phi)^2 < \infty. \tag{3.49}$$

Moreover, by **(H)** and (3.44), we find a constant $c_1 > 0$ such that

$$\begin{aligned}
&\|N_s(X_s^r)v - N_s(X_s)v\| \\
&\le c_1\big\{|X_s^r - X_s|(|v| + (\mathbb{E}|v|^2)^{\frac{1}{2}}) + h_{s,r}(\phi)(\mathbb{E}|v|^2)^{\frac{1}{2}}\big\}, \quad v \in L^2(\Omega \to \mathbb{R}^d; \mathbb{P}).
\end{aligned} \tag{3.50}$$

Combining (3.41), (3.49) and (3.50), we may find a constant $c_2 > 0$ such that the dominated convergence theorem yields

$$\begin{aligned}\limsup_{\|\phi\|_{L^2(\mu)}\to 0} \varepsilon_2(\phi) &\le \limsup_{\|\phi\|_{L^2(\mu)}\to 0} \frac{c_2}{\|\phi\|_{L^2(\mu)}}\int_0^1 \mathbb{E}\bigg(\int_0^T \big\{(|v_s^\phi|^2 + \|\phi\|^2_{L^2(\mu)})\cdot |X_s^r - X_s|^2 \\ &\qquad\qquad + h_{s,r}(\phi)^2\|\phi\|^2_{L^2(\mu)}\big\}\mathrm{d}s\bigg)^{\frac{1}{2}}\mathrm{d}r \\ &\le \limsup_{\|\phi\|_{L^2(\mu)}\to 0} \frac{c_2}{\|\phi\|_{L^2(\mu)}}\int_0^1 \bigg\{\Big(\mathbb{E}\Big[\sup_{s\in[0,T]}|v_s^\phi|^2\Big] + \|\phi\|^2_{L^2(\mu)}\Big)^{\frac{1}{2}}\Big(\mathbb{E}\int_0^T |X_s^r - X_s|^2\mathrm{d}s\Big)^{\frac{1}{2}} \\ &\qquad\qquad + \|\phi\|_{L^2(\mu)}\Big(\int_0^T \mathbb{E}h_{s,r}(\phi)^2\mathrm{d}s\Big)^{\frac{1}{2}}\bigg\}\mathrm{d}r = 0.\end{aligned}$$

c) It remains to prove

$$\lim_{\|\phi\|_{L^2(\mu)}\to 0} \varepsilon_3(\phi) = 0. \tag{3.51}$$

Noting that

$$\begin{aligned}&|N_t(X_t^r)\big\{(\nabla\theta_t^{\lambda,\mu^r})^{-1}(X_t^r)v_t^{\phi,r} - (\nabla\theta_t^{\lambda,\mu})^{-1}(X_t)v_t^\phi\big\}| \\ &\le c[\|(\nabla\theta_t^{\lambda,\mu^r})^{-1}(X_t^r) - (\nabla\theta_t^{\lambda,\mu})^{-1}(X_t)\||v_t^\phi| + |v_t^{\phi,r} - v_t^\phi|] \\ &+ c\mathbb{E}[|D^L(B_t+b_t)(y,\mathscr{L}_{X_t^r})(X_t^r)|[\|(\nabla\theta_t^{\lambda,\mu^r})^{-1}(X_t^r) - (\nabla\theta_t^{\lambda,\mu})^{-1}(X_t)\|]|v_t^\phi|]\big|_{y=X_t^r} \\ &+ c\mathbb{E}[|D^L(B_t+b_t)(y,\mathscr{L}_{X_t^r})(X_t^r)||v_t^{\phi,r} - v_t^\phi|]\big|_{y=X_t^r},\end{aligned}$$

according to BDG's inequality, (3.41) and (3.7), it suffices to prove

$$\begin{aligned}&\lim_{\|\phi\|_{L^2(\mu)}\to 0} \frac{\mathbb{E}\sup_{t\in[0,T]}\Big\{\mathbb{E}\left[(1+|D^L(B+b)_t(y,\mathscr{L}_{X_t^r})(X_t^r)|)\sup_{t\in[0,T]}|v_t^{\phi,r} - v_t^\phi|\right]\big|_{y=X_t^r}\Big\}}{\|\phi\|_{L^2(\mu)}} \\ &= 0.\end{aligned}$$

By **(C)**, this is implied by

$$\lim_{\|\phi\|_{L^2(\mu)}\to 0} \frac{\Big(\mathbb{E}\sup_{t\in[0,T]}|v_t^{\phi,r} - v_t^\phi|^{\frac{p}{p-1}}\Big)^{\frac{p-1}{p}}}{\|\phi\|_{L^2(\mu)}} = 0. \tag{3.52}$$

To prove (3.52), we observe that

$$|F_s^{\mu^r}(X_s^r)v_s^{\phi,r} - F_s^\mu(X_s)v_s^\phi|$$

$$\leq |F_s^{\mu^r}(X_s^r)v_s^{\phi,r} - F_s^\mu(X_s^r)v_s^{\phi,r}| + |F_s^\mu(X_s^r)v_s^{\phi,r} - F_s^\mu(X_s^r)v_s^\phi| + |F_s^\mu(X_s^r)v_s^\phi - F_s^\mu(X_s)v_s^\phi|$$
$$=: J_1 + J_2 + J_3.$$

By (3.23), **(H)**, Lemma 3.6, (3.7) and (3.41), we find a constant $c_1 > 0$ such that

$$\begin{aligned}
J_1 &\leq \big|\nabla(\theta_s^{\lambda,\mu^r} - \theta_s^{\lambda,\mu})(X_s^r)\mathbb{E}[D^L(B_s+b_s)(y,\mathscr{L}_{X_s^r})(X_s^r)(\nabla\theta_s^{\lambda,\mu})^{-1}(X_s^r)v_s^{\phi,r}]|_{y=X_s^r}\big| \\
&\leq c_1\|\phi\|_{L^2(\mu)}^2, \\
J_2 &\leq c_1\mathbb{E}[|D^L(B_s+b_s)(y,\mathscr{L}_{X_s^r})(X_s^r)||v_s^{\phi,r} - v_s^\phi|]\big|_{y=X_s^r}.
\end{aligned}$$

Moreover, similarly to (3.50), by **(H)**, **(C)**, (3.41) and (3.7), we find a nonnegative random variables $\{h_{s,r}(\phi)\}_{s\in[0,T],r\in[0,1]}$ satisfying (3.49) such that

$$J_3 \leq h_{s,r}(\phi)\|\phi\|_{L^2(\mu)}.$$

So, there exists a constant $c_2 > 0$ such that

$$\begin{aligned}
&|F_s^{\mu^r}(X_s^r)v_s^{\phi,r} - F_s^\mu(X_s)v_s^\phi| \\
&\leq c_2\mathbb{E}[|D^L(B_s+b_s)(y,\mathscr{L}_{X_s^r})(X_s^r)||v_s^{\phi,r} - v_s^\phi|]\big|_{y=X_s^r} \\
&\quad + c_2(h_{s,r}(\phi) + \|\phi\|_{L^2(\mu)})\|\phi\|_{L^2(\mu)}, \quad s\in[0,T], r\in[0,1].
\end{aligned}$$

Combining this with **(H)**, (3.7), (3.32), (3.11) and (3.39), we find some constant $c_3 > 0$, a martingale M_t^r with $\mathrm{d}\langle M^r\rangle_t \leq \mathrm{d}t$ and nonnegative random variables $\{h_{s,r}(\phi)\}_{s\in[0,T],r\in[0,1]}$ satisfying (3.49) such that

$$\begin{aligned}
|v_t^{\phi,r} - v_t^\phi| &\leq c_3|\phi(X_0)|(\|\phi\|_{L^2(\mu)} + [\nabla\theta_0^{\lambda,\mu}](X_0 + r\phi(X_0)) - [\nabla\theta_0^{\lambda,\mu}](X_0)) \\
&+ c_3\int_0^t \big\{|v_s^{\phi,r} - v_s^\phi| + \mathbb{E}[|D^L(B_s+b_s)(y,\mathscr{L}_{X_s^r})(X_s^r)||v_s^{\phi,r} - v_s^\phi|]\big|_{y=X_s^r}\big\}\mathrm{d}s \\
&+ c_3\int_0^t \big\{h_{s,r}(\phi)(|v_s^\phi| + \|\phi\|_{L^2(\mu)})\big\}\mathrm{d}s \\
&+ c_3\bigg|\int_0^t \big\{|v_s^{\phi,r} - v_s^\phi| + h_{s,r}(\phi)|v_s^\phi|\big\}\mathrm{d}M_s^r\bigg|, \quad t\in[0,T], r\in[0,1].
\end{aligned} \tag{3.53}$$

By BDG's inequality, Hölder's inequality, (3.41), (3.7) and (3.49), we find a constant $c_4 > 0$ and $\varepsilon(\phi)$ with $\lim_{\|\phi\|_{L^2(\mu)}\to 0}\varepsilon(\phi) = 0$ such that

$$U_t := \sup_{s\in[0,t]}|v_s^{\phi,r} - v_s^\phi|^{\frac{p}{p-1}}, \quad t\in[0,T]$$

satisfies

$$\mathbb{E}U_t \leq \|\phi\|_{L^2(\mu)}^{\frac{p}{p-1}}\varepsilon(\phi) + c_4\int_0^t \mathbb{E}U_s\mathrm{d}s + \frac{1}{2}\mathbb{E}U_t, \quad t\in[0,T].$$

By Gronwall's lemma, we obtain (3.52) and the proof is completed. □

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] D. Baños, *The Bismut-Elworthy-Li formula for mean-field stochastic differential equations,* Ann. Inst. Henri Poincaré Probab. Stat. 54(2018), 220-233.

[3] 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.

[4] V. Barbu, M. Röckner, *Probabilistic representation for solutions to non-linear Fokker-Planck equations,* SIAM J. Math. Anal. 50(2018), 4246-4260.

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

[6] S. Benachour, B. Roynette, D. Talay, and P. Vallois, *Nonlinear selfstabilizing processes. I. Existence, invariant probability, propagation of chaos,* Stochastic Process. Appl. 75(1998), 173-201.

[7] M. Bossy and D. Talay, *A stochastic particle method for the McKean-Vlasov and the Burgers equation,* Math. Comput. 66(1997), 157-192.

[8] J. M. Bismut, *Large Deviations and the Malliavin Calculus,* Birkhäser, Boston, 1984.

[9] P. Cardaliaguet, *Notes on mean filed games* (from P.L. Lions' lectures at Collège de France), 2013.

[10] R. Carmona and F. Delarue, *Probabilistic analysis of mean-field games*, SIAM J. Control Optim. 51(2013), 2705-2734.

[11] R. Carmona and F. Delarue, *Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics,* Ann. Probab. 43(2015), 2647-2700.

[12] P.-E. Chaudry De Raynal, N. Frikha, *Well-posedness for some non-linear SDEs and related PDE on the Wasserstein space*, arXiv:1811.06904v3

[13] D. Crisan and E. McMurray, *Smoothing properties of McKean-Vlasov SDEs*, Probab. Theory Relat. Fields. 171(2018), 97-148.

[14] D. Dawson, J. Vaillancourt, *Stochastic McKean-Vlasov equations,* Nonlinear Differential Equations Appl. 2(1995), 199-229.

[15] K. D. Elworthy, X.-M. Li, *Formulae for the derivatives of the heat semigroups,* J. Funct. Anal. 125(1994), 252-286.

[16] X. Huang, M. Röckner, F.-Y. Wang, *Nonlinear Fokker–Planck equations for probability measures on path space and path-distribution dependent SDEs*, Discrete Contin. Dyn. Syst. 39(2019), 3017-3035.

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

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

[19] X. Huang, F.-Y. Wang, *McKean-Vlasov SDEs with drifts discontinuous under Wasserstein distance,* Discrete Contin. Dyn. Syst. 41(2021), 1667-1679.

[20] P. Kotelenez, T. Kurtz, *Macroscopic limits for stochastic partial differential equations of McKean-Vlasov type,* Probab. Theory Related Fields 146(2010), 189-222.

[21] Yu. S. Mishura, A. Yu. Veretennikov, *Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations,* arXiv:1603.02212.

[22] D. Nualart, *The Malliavin Calculus and Related Topics*, Second edition, Springer-Verlag Berlin Heidelberg, 2006.

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

[24] 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.

[25] Y. Song, *Gradient Estimates and Exponential Ergodicity for Mean-Field SDEs with Jumps,* J. Theoret. Probab. 33(2020), 201-238.

[26] F.-Y. Wang, *Gradient estimate and applications for SDEs in Hilbert space with multiplicative noise and Dini continuous drift,* J. Differential Equations 260(2016), 2792-2829.

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

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

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