Space-Distribution PDEs for Path Independent Additive Functionals of McKean-Vlasov SDEs*

Panpan Ren[b)], **Feng-Yu Wang**[a,b)]

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

[b)] Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, United Kingdom

673788@swansea.ac.uk, wangfy@tju.edu.cn, F.-Y.Wang@swansea.ac.uk

July 17, 2018

Abstract

Let $\mathscr{P}_2(\mathbb{R}^d)$ be the space of probability measures on $\mathbb{R}^d$ with finite second moment. The path independence of additive functionals of McKean-Vlasov SDEs is characterized by PDEs on the product space $\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$ equipped with the usual derivative in space variable and Lions' derivative in distribution. These PDEs are solved by using probabilistic arguments developed from [2]. As consequence, the path independence of Girsanov transformations are identified with nonlinear PDEs on $\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$ whose solutions are given by probabilistic arguments as well. In particular, the corresponding results on the Girsanov transformation killing the drift term derived earlier for the classical SDEs are recovered as special situations.

AMS subject Classification: 60J60, 58J65.
Keywords: McKean-Vlasov SDEs, additive functional, Girsanov transformation, L-derivative.

1 Introduction

In recent years, McKean-Vlasov stochastic differential equations (SDEs), also called distribution dependent or mean field SDEs, have received increasing attentions for their theoretically

*Supported in part by NNSFC (11771326, 11431014). Financial support by the DFG through the CRC 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications" is acknowledged.

importance in characterizing non-linear Fokker-Planck equations from physics. On the other hand, SDEs have been developed as crucial mathematical tools modelling economic and finance systems. In the real world, the evolution of these systems is not only driven by micro actions (drift and noise), but also relies on the macro environment (in mathematics, distribution of the systems). So, it is reasonable to characterize economic and finance systems by using distribution dependent SDEs.

Let $\mathscr{P}(\mathbb{R}^d)$ be the space of all probability measures on $\mathbb{R}^d$, and let

$$\mathscr{P}_2(\mathbb{R}^d) = \Big\{\mu \in \mathscr{P}(\mathbb{R}^d):\ \mu(|\cdot|^2) := \int_{\mathbb{R}^d} |x|^2\mu(\mathrm{d}x) < \infty\Big\}.$$

Then $\mathscr{P}_2(\mathbb{R}^d)$ is a Polish space under the Wasserstein distance

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

where $\mathscr{C}(\mu,\nu)$ is the set of couplings for μ and ν; that is, $\pi\in\mathscr{C}(\mu,\nu)$ is a probability measure on $\mathbb{R}^d\times\mathbb{R}^d$ such that $\pi(\cdot\times\mathbb{R}^d)=\mu$ and $\pi(\mathbb{R}^d\times\cdot)=\nu$.

Let W_t be an m-dimensional Brownian motion on a standard filtered probability space $(\Omega,\mathscr{F},\{\mathscr{F}_t\}_{t\geq 0},\mathbb{P})$, and denote by $\mathscr{L}_\xi$ the distribution of a random variable ξ on $\mathbb{R}^d$. Consider the following McKean-Vlasov SDE on $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(t,X_t,\mathscr{L}_{X_t})\mathrm{d}t + \sigma(t,X_t,\mathscr{L}_{X_t})\mathrm{d}W_t, \tag{1.1}$$

where

$$\sigma:[0,\infty)\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d)\to\mathbb{R}^{d\otimes m},\quad b:[0,\infty)\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d)\to\mathbb{R}^d$$

are continuous such that for some increasing function $K:[0,\infty)\to[0,\infty)$ there holds

$$\begin{aligned}&|b(t,x,\mu)-b(t,y,\nu)| + \|\sigma(t,x,\mu)-\sigma(t,y,\nu)\|_{HS}\\ &\leq K(t)\big(|x-y|+\mathbb{W}_2(\mu,\nu)\big),\quad t\geq 0, x,y\in\mathbb{R}^d,\mu,\nu\in\mathscr{P}_2(\mathbb{R}^d)\end{aligned} \tag{1.2}$$

and

$$\|\sigma(t,\mathbf{0},\delta_{\mathbf{0}})\|_{HS} + |b(t,\mathbf{0},\delta_{\mathbf{0}})| \leq K(t),\quad t\geq 0, \tag{1.3}$$

where $\delta_{\mathbf{0}}$ is the Dirac measure at $\mathbf{0}\in\mathbb{R}^d$. For any $t\geq 0$, let $L^2(\Omega\to\mathbb{R}^d,\mathscr{F}_t,\mathbb{P})$ be the class of $\mathscr{F}_t$-measurable square integrable random variables on $\mathbb{R}^d$. By (1.2) and (1.3), for any $s\geq 0$ and $X_s\in L^2(\Omega\to\mathbb{R}^d,\mathscr{F}_s,\mathbb{P})$, (1.1) has a unique solution $(X_t)_{t\geq s}$ with

$$\sup_{t\in[s,T]}\mathbb{E}|X_t|^2<\infty,\quad T\geq s. \tag{1.4}$$

See [11] for more results on gradient estimates and Harnack inequalities of the associated nonlinear semigroup, and [7, 8] and references within for the existence and uniqueness under weaker conditions.

In this paper, we aim to characterize the path independence of the additive functional

$$A^{\mathbf{f},\mathbf{g}}_{s,t} := \int_s^t \mathbf{f}(r, X_r, \mathscr{L}_{X_r})\mathrm{d}r + \int_s^t \langle \mathbf{g}(r, X_r, \mathscr{L}_{X_r}), \mathrm{d}W_r\rangle, \quad 0 \le s \le t, \tag{1.5}$$

where

$$\mathbf{f} : [0,\infty) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}, \quad \mathbf{g} : [0,\infty) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}^m$$

are continuous, so that $A^{\mathbf{f},\mathbf{g}}_{s,t}$ for $t \ge s$ is a well-defined local semi-martingale. Throughout the paper we consider the time interval $[0, T]$ for a fixed constant $T \in [0,\infty]$. When $T = \infty$ we regard $[0, T]$ as $[0,\infty)$.

Definition 1.1. The additive functional $(A^{\mathbf{f},\mathbf{g}}_{s,t})_{T\ge t\ge s\ge 0}$ is called path independent, if there exists a measurable function

$$V : [0, T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$$

such that for any $s \in [0, T)$ and $X_s \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_s, \mathbb{P})$, the solution $(X_t)_{t\in[s,T]}$ to the SDE (1.1) from time s satisfies

$$A^{\mathbf{f},\mathbf{g}}_{s,t} = V(t, X_t, \mathscr{L}_{X_t}) - V(s, X_s, \mathscr{L}_{X_s}), \quad 0 \le s \le t < T. \tag{1.6}$$

The motivation of the study comes from mathematical statement of equilibrium financial market. In their seminal paper [1] Black and Scholes described the price dynamics (or the wealth growth) by using SDEs under a so-called real world probability measure. But for an equilibrium financial market there exists a so-called risk neutral measure having a path independent density with respect to the real world probability, see [6]. That is, under the risk neutral measure the solution of (1.1) becomes a martingale, and the density of the neutral probability with respect to the real world one depends only on the initial and current states but not those in between.

For instance, let $\mathbf{f} = \frac{1}{2}|\mathbf{g}|^2$. Then $A^{\mathbf{f},\mathbf{g}}_{s,t}$ becomes

$$A^{\mathbf{g}}_{s,t} := \frac{1}{2}\int_s^t |\mathbf{g}(r, X_r, \mathscr{L}_{X_r})|^2\mathrm{d}r + \int_s^t \langle \mathbf{g}(r, X_r, \mathscr{L}_{X_r}), \mathrm{d}W_r\rangle, \quad 0 \le s \le t. \tag{1.7}$$

By the Girsanov theorem, when

$$\mathbb{E}\mathrm{e}^{\frac{1}{2}\int_s^t |\mathbf{g}(r,X_r,\mathscr{L}_{X_r})|^2\mathrm{d}r} < \infty, \tag{1.8}$$

$\mathrm{d}\mathbb{Q}^{\mathbf{g}}_{s,t} := \mathrm{e}^{-A^{\mathbf{g}}_{s,t}}\mathrm{d}\mathbb{P}$ is a probability measure. So, to adopt $\mathbb{Q}^{\mathbf{g}}_{s,t}$ as a risk neutral measure, we need to verify the path independence of the additive functional $A^{\mathbf{g}}_{s,t}$ in the sense of (1.6). In particular, when

$$b = \sigma\tilde{b} \text{ for some measurable } \tilde{b} : [0,\infty) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}^m, \tag{1.9}$$

and (1.8) holds for $\mathbf{g} := \tilde{b}$, let

$$A_{s,t} := \frac{1}{2}\int_s^t |\tilde{b}(r, X_r, \mathscr{L}_{X_r})|^2\mathrm{d}r + \int_s^t \langle \tilde{b}(r, X_r, \mathscr{L}_{X_r}), \mathrm{d}W_r\rangle, \quad 0 \le s \le t. \tag{1.10}$$

Then $\mathrm{d}\mathbb{Q}_{s,t} := \mathrm{e}^{-A_{s,t}}\mathrm{d}\mathbb{P}$ is a probability measure such that

$$\tilde{W}_r := W_r + \int_s^r \tilde{b}(u, X_u, \mathscr{L}_{X_u})\mathrm{d}u, \quad r \in [s,t]$$

is an m-dimensional Brownian motion, and hence

$$X_r = X_s + \int_s^r \sigma(u, X_u, \mathscr{L}_{X_u})\mathrm{d}\tilde{W}_u, \quad r \in [s,t]$$

is a $\mathbb{Q}_{s,t}$-martingale as required for an equilibrium financial market. We would like to investigate the path independence of the additive functional $A_{s,t}$ such that $\mathbb{Q}_{s,t}$ is a risk neutral measure.

In Section 2, we will characterize the path independence of $A_{s,t}^{\mathbf{f},\mathbf{g}}$ using PDEs on $\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$, see Theorem 2.2 below for details. Following the idea of [2], such type PDEs are solved using solutions of an associated SDE, see Theorem 2.3 for details. As a consequence, the path independence of $A_{s.t}^{\mathbf{g}}$ in (1.7) and $A_{s,t}$ in (1.10) is identified with nonlinear PDEs on $\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$, see Corollaries 2.4 and 2.5 below. When the SDE is distribution independent, i.e. $b(t,x,\mu)$ and $\sigma(t,x,\mu)$ do not depend on $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, Corollary 2.5 recovers the corresponding existing results derived in [10, 13, 14], see also [9, 12] for extensions to SDEs with jumps and semi-linear SPDEs. Finally, complete proofs of these results are presented in Section 3.

2 Main results

To state our results, we first recall the definition of L-derivative for functions on $\mathscr{P}_2(\mathbb{R}^d)$, which was introduced by P.-L. Lions in his lectures [3] at College de France, see also [2, 5]. In the following we introduce a straightforward definition without using abstract probability spaces as in previous references. Let ∂_t denote the partial differential in time parameter $t \geq 0$, ∂_x or ∂_y the gradient operator in variables x or $y \in \mathbb{R}^d$, and ∂_x^2 the Hessian operator in $x \in \mathbb{R}^d$. Let $\mathrm{Id}: \mathbb{R}^d \to \mathbb{R}^d$ be the identity map, i.e. $\mathrm{Id}(x) = x$ for $x \in \mathbb{R}^d$. It is easy to see that for any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\phi \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$, we have $\mu \circ (\mathrm{Id} + \phi)^{-1} \in \mathscr{P}_2(\mathbb{R}^d)$.

Definition 2.1. Let $T \in (0, \infty]$, and set $[0,T] = [0,\infty)$ when $T = \infty$.

(1) A function $f : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is called L-differentiable at $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, if the functional

$$L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu) \ni \phi \mapsto f(\mu \circ (\mathrm{Id} + \phi)^{-1})$$

is Fréchet differentiable at $\mathbf{0} \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$; that is, there exists (hence, unique) $\xi \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$ such that

$$\lim_{\mu(|\phi|^2)\to 0} \frac{f(\mu \circ (\mathrm{Id} + \phi)^{-1}) - f(\mu) - \mu(\langle \xi, \phi \rangle)}{\sqrt{\mu(|\phi|^2)}} = 0. \tag{2.1}$$

In this case, we denote $\partial_\mu f(\mu) = \xi$ and call it the L-derivative of f at μ.

(2) A function $f : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is called L-differentiable on $\mathscr{P}_2(\mathbb{R}^d)$ if the L-derivative $\partial_\mu f(\mu)$ exists for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. If moreover $(\partial_\mu f(\mu))(y)$ has a version differentiable in $y \in \mathbb{R}^d$ such that $(\partial_\mu f(\mu))(y)$ and $\partial_y(\partial_\mu f(\mu))(y)$ are jointly continuous in $(\mu, y) \in \mathscr{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, we denote $f \in C^{(1,1)}(\mathscr{P}_2(\mathbb{R}^d))$.

(3) A function $f : [0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is said to be in the class $C^{1,2,(1,1)}([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$, if the derivatives

$$\partial_t f(t,x,\mu),\ \partial_x f(t,x,\mu),\ \partial_x^2 f(t,x,\mu),\ \partial_\mu f(t,x,\mu)(y),\ \partial_y \partial_\mu f(t,x,\mu)(y)$$

exist and are jointly continuous in the corresponding arguments (t,x,μ) or (t,x,μ,y). If $f \in C^{1,2,(1,1)}([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ with all these derivatives bounded on $[0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$, we denote $f \in C_b^{1,2,(1,1)}([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$.

(4) Finally, we write $f \in \mathscr{C}([0,\infty) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$, if $f \in C^{1,2,(1,1)}([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ and the function

$$(t,x,\mu) \mapsto \int_{\mathbb{R}^d} \big\{\|\partial_y \partial_\mu f\| + \|\partial_\mu f\|^2\big\}(t,x,\mu)(y)\mu(\mathrm{d}y)$$

is locally bounded, i.e. it is bounded on compact subsets of $[0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$.

For readers' understanding of the L-derivative, we present below an example for a class of functions inducing the Borel σ-algebra on $\mathscr{P}_2(\mathbb{R}^d)$. See [2, Example 2.2] for concrete choices of F and h_i.

Example 2.1. *Let $n \in \mathbb{N}, \{h_i\}_{1 \le i \le n} \subset C^2(\mathbb{R}^d)$ with $\|\partial_x^2 h_i\|_\infty < \infty$ and let $F \in C^1(\mathbb{R}^n)$. Then the function*

$$\mathscr{P}_2(\mathbb{R}^d) \ni \mu \mapsto f(\mu) := F(\mu(h_1), \cdots, \mu(h_n))$$

is in $C^{(1,1)}(\mathscr{P}_2(\mathbb{R}^d))$ with

$$\partial_\mu f(\mu)(y) = \sum_{i=1}^n (\partial_i F)(\mu(h_1), \cdots, \mu(h_n))\partial_y h_i(y), \quad \mu \in \mathscr{P}_2(\mathbb{R}^d), y \in \mathbb{R}^d.$$

Proof. By the chain rule it suffices to prove for $f(\mu) := \mu(h_1)$, i.e. $n = 1$ and $F(r) = r$. Since $\|\partial_x^2 h_i\|_\infty < \infty$, there exists a constant $C > 0$ such that

$$|h_i(x)| + |\partial_x h_i(x)|^2 \le C(1 + |x|^2), \quad x \in \mathbb{R}^d,$$

so that $h_1 \in L^1(\mu)$ and $\partial_x h_1 \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$ for $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. Then, for any $\phi \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$, by Taylor's expansion we have

$$\begin{aligned}
&\lim_{\|\phi\|_{L^2(\mu)} \to 0} \frac{|f(\mu \circ (\mathrm{Id} + \phi)^{-1}) - f(\mu) - \mu(\langle \partial h_1, \phi \rangle)|}{\|\phi\|_{L^2(\mu)}} \\
&= \lim_{\|\phi\|_{L^2(\mu)} \to 0} \frac{1}{\|\phi\|_{L^2(\mu)}} \bigg| \int_{\mathbb{R}^d} \big\{h_1(x + \phi(x)) - h_1(x) - \langle \partial_x h_1(x), \phi(x) \rangle\big\}\mu(\mathrm{d}x)\bigg|
\end{aligned}$$

$$\leq \lim_{\|\phi\|_{L^2(\mu)}\to 0} \frac{\|\partial_x^2 h_1\|_\infty}{2\|\phi\|_{L^2(\mu)}} \int_{\mathbb{R}^d} |\phi(x)|^2 \mu(\mathrm{d}x) \leq \lim_{\|\phi\|_{L^2(\mu)}\to 0} \|\partial_x^2 h_1\|_\infty \|\phi\|_{L^2(\mu)} = 0.$$

So, by definition, $\partial_\mu f(\mu)(y) = \partial_y h_1(y)$. □

Let us explain that the above definition of L-derivative coincides with the Wasserstein derivative introduced by P.-L. Lions using probability spaces. Given $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, let $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}}) = (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d), \mu)$ and $X = \mathrm{Id}$. Then X is a random variable with $\mathscr{L}_X|_{\tilde{\mathbb{P}}} = \mu$. For any square integrable random variable Y, we have $\phi := Y \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$. Moreover, since $X = \mathrm{Id}$, for any $A \in \mathscr{B}(\mathbb{R}^d)$,

$$\begin{aligned}
(\mu \circ (\mathrm{Id} + \phi)^{-1})(A) &= \mu(\{x : (\mathrm{Id} + \phi)(x) \in A\}) = \mu(\{x : x + \phi(x) \in A\}) \\
&= \tilde{\mathbb{P}}(\{x : X(x) + Y(x) \in A\}) = \tilde{\mathbb{P}}(X + Y \in A) \\
&= \mathscr{L}_{X+Y}|_{\tilde{\mathbb{P}}}(A).
\end{aligned}$$

So, $(\mu \circ (\mathrm{Id} + \phi)^{-1}) = (\mathscr{L}_{X+Y}|_{\tilde{\mathbb{P}}})$, and (2.1) means that

$$L^2(\tilde{\Omega} \to \mathbb{R}^d, \tilde{\mathbb{P}}) \ni Y \mapsto f(\mathscr{L}_{X+Y}|_{\tilde{\mathbb{P}}})$$

is Fréchet differentiable with derivative $\partial_\mu f(\mu) := \xi$, which coincides with [3, Definition 6.1] given by P.-L. Lions. Note that the atomless restriction on the probability space therein is to ensure the existence of a random variable with distribution μ. It is crucial that (see [5, Proposition A.2]) the definition of $\partial_\mu f(\mu) \in L^2(\mathbb{R}^d \to \mathbb{R}^d, \mu)$ dose not depend on the choice of probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ and random variable X with $\mathscr{L}_X|_{\tilde{\mathbb{P}}} = \mu$. So, in particular, we may take the above specific choice $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}}) = (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d), \mu)$ and $X = \mathrm{Id}$.

The following differential operator on $[0, \infty) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$ associated with the SDE (1.1) has been introduced in [2]: for any $V \in C^{1,2,(1,1)}([0,\infty) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ and $(t, x, \mu) \in [0, \infty) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$, let

$$\begin{aligned}
\mathbf{L}_{\sigma,b} V(t,x,\mu) &= \frac{1}{2}\mathrm{tr}\big(\sigma\sigma^* \partial_x^2 V\big)(t,x,\mu) + \langle b, \partial_x V\rangle(t,x,\mu) \\
&\quad + \int_{\mathbb{R}^d} \Big[\frac{1}{2}\mathrm{tr}\big\{(\sigma\sigma^*)(t,y,\mu)\partial_y \partial_\mu V(t,x,\mu)(y)\big\} \\
&\quad + \big\langle b(t,y,\mu), \partial_\mu V(t,x,\mu)(y)\big\rangle\Big]\mu(\mathrm{d}y).
\end{aligned} \tag{2.2}$$

Our first result is the following characterization on the path independence of the functional $A_{s,t}^{\mathbf{f},\mathbf{g}}$ in (1.5).

Theorem 2.2. *Assume that σ and b satisfy* (1.2) *and* (1.3) *for some locally bounded function K. Let $T \in [0, \infty]$, $\mathbf{f} \in C([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ and $\mathbf{g} \in C([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}^m)$. For any $V \in \mathscr{C}([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$, $A_{s,t}^{\mathbf{f},\mathbf{g}}$ is path independent in the sense of* (1.6) *if and only if*

$$\begin{cases} (\partial_t + \mathbf{L}_{\sigma,b})V(t,x,\mu) = \mathbf{f}(t,x,\mu), \\ (\sigma^* \partial_x V)(t,x,\mu) = \mathbf{g}(t,x,\mu), \end{cases} \qquad t \in [0,T], x \in \mathbb{R}^d, \mu \in \mathscr{P}_2(\mathbb{R}^d). \tag{2.3}$$

To provide a class of $(\mathbf{f},\mathbf{g})$ such that the additive functional $A^{\mathbf{f},\mathbf{g}}_{s,t}$ is path independent in the sense of (1.6), we adopt the idea of [2] to solve the PDE (2.3) using an SDE accompying with (1.1). To state this accompying SDE, for any $\mu\in\mathscr{P}_2(\mathbb{R}^d)$ and $s\ge 0$, let $(X^\mu_{s,t})_{t\ge s}$ solve (1.1) from time s with $\mathscr{L}_{X^\mu_{s,s}}=\mu$. Let

$$P^*_{s,t}\mu=\mathscr{L}_{X^\mu_{s,t}},\quad t\ge s,\quad \mu\in\mathscr{P}_2(\mathbb{R}^d). \tag{2.4}$$

As shown in [11] that $P^*_{s,t}$ is a nonlinear semigroup satisfying

$$P^*_{t,r}P^*_{s,t}=P^*_{s,r},\quad 0\le s\le t\le r. \tag{2.5}$$

Now, for any $x\in\mathbb{R}^d,\mu\in\mathscr{P}_2(\mathbb{R}^d)$ and $s\ge0$, let $(X^{x,\mu}_{s,t})_{t\ge s}$ solve the SDE

$$\mathrm{d}X^{x,\mu}_{s,t}=b(t,X^{x,\mu}_{s,t},P^*_{s,t}\mu)\mathrm{d}t+\sigma(t,X^{x,\mu}_{s,t},P^*_{s,t}\mu)\mathrm{d}W_t,\quad X^{x,\mu}_{s,s}=x. \tag{2.6}$$

We have the following result.

Theorem 2.3. *Assume that* $b_i,\sigma_{ij},\mathbf{f}\in C_b^{1,2,(1,1)}([0,T]\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d)),\ 1\le i\le d,1\le j\le m.$ *Then for any* $\Phi\in C_b^{2,(1,1)}(\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d))$,

$$V(t,x,\mu):=\mathbb{E}\bigg(\Phi(X^{x,\mu}_{t,T},P^*_{t,T}\mu)-\int_t^T\mathbf{f}(r,X^{x,\mu}_{t,r},P^*_{t,r}\mu)\mathrm{d}r\bigg),\quad t\in[0,T],x\in\mathbb{R}^d,\mu\in\mathscr{P}_2(\mathbb{R}^d)$$

is the unique solution to the first PDE in (2.3) *with* $V(T,\cdot,\cdot)=\Phi(\cdot,\cdot)$.

Consequently, for such a function V, $A^{\mathbf{f},\mathbf{g}}_{s,t}$ *is path independent in the sense of* (1.6) *if and only if*

$$\mathbf{g}(t,x,\mu)=\sigma^*\partial_x\mathbb{E}\bigg(V(T,X^{x,\mu}_{t,T},P^*_{t,T}\mu)-\int_t^T\mathbf{f}(r,X^{x,\mu}_{t,r},P^*_{t,r}\mu)\mathrm{d}r\bigg). \tag{2.7}$$

Next, we consider $\mathbf{f}:=\frac{1}{2\beta}|\mathbf{g}|^2$ for a constant $\beta\ne0$. Then the additive functional $A^{\mathbf{f},\mathbf{g}}_{s,t}$ reduces to

$$A^{\mathbf{g};\beta}_{s,t}:=\frac{1}{2\beta}\int_s^t|\mathbf{g}(r,X_r,\mathscr{L}_{X_r})|^2\mathrm{d}r+\int_s^t\langle\mathbf{g}(r,X_r,\mathscr{L}_{X_r}),\mathrm{d}W_r\rangle,\quad 0\le s\le t. \tag{2.8}$$

This covers $A^{\mathbf{g}}_{s,t}$ in (1.7) for $\beta=1$. As a consequence of Theorems 2.2 and 2.3, we have the following result on the path independence of $A^{\mathbf{g};\beta}_{s,t}$ and the corresponding nonlinear PDE:

$$(\partial_t+\mathbf{L}_{\sigma,b})V(t,x,\mu)=\frac{1}{2\beta}|\sigma^*\partial_xV|^2(t,x,\mu),\quad (t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d). \tag{2.9}$$

Corollary 2.4. *Assume that* σ *and* b *satisfy* (1.2) *and* (1.3) *for some locally bounded function* K. *Let* $T\in(0,\infty]$ *and* $0\ne\beta\in\mathbb{R}$.

(1) *Let* $V\in\mathscr{C}([0,T]\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d))$. *Then* $A^{\mathbf{g};\beta}_{s,t}$ *is path independent in the sense of* (1.6) *if and only if* V *solves the nonlinear PDE* (2.9) *and* $\mathbf{g}=\sigma^*\partial_xV$.

(2) *Let* $b_i, \sigma_{ij} \in C_b^{1,2,(1,1)}([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ *for* $1 \le i \le d, 1 \le j \le m$. *For any* $\Phi \in C_b^{2,(1,1)}(\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ *with* $\inf \Phi > 0$,

$$V(t,x,\mu) := -\beta \log \big\{ \mathbb{E}\Phi(X_{t,T}^{x,\mu}, P_{t,T}^*\mu) \big\}, \quad (t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d) \tag{2.10}$$

is the unique solution to the nonlinear PDE (2.9) *with*

$$V(T,x,\mu) = -\beta \log \Phi(x,\mu), \quad (x,\mu) \in \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d).$$

Finally, we consider the path independence of the functional $A_{s,t}$ in (1.10). Let

$$\begin{aligned} \mathbf{L}_\sigma V(t,x,\mu) =& \frac{1}{2}\mathrm{tr}\big(\sigma\sigma^* \partial_x^2 V\big)(t,x,\mu) + \frac{1}{2}|\sigma^*\partial_x V|^2(t,x,\mu) \\ &+ \int_{\mathbb{R}^d} \Big[\frac{1}{2}\mathrm{tr}\big\{(\sigma\sigma^*)(t,y,\mu)\partial_y \partial_\mu V(t,x,\mu)(y)\big\} \\ &\quad + \big\langle (\sigma\sigma^*\partial_y V)(t,y,\mu), \partial_\mu V(t,x,\mu)(y) \big\rangle \Big] \mu(\mathrm{d}y). \end{aligned} \tag{2.11}$$

Corollary 2.5. *Assume that* σ *and* b *satisfy* (1.2) *and* (1.3) *for some locally bounded function* K. *Let* $T \in (0,\infty]$.

(1) *Let* $V \in \mathscr{C}([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$. *Then* (1.9) *holds for* $\tilde{b} := \sigma^*\partial_x V$ *and* $A_{s,t}$ *in* (1.10) *is path independent in the sense of* (1.6) *if and only if*

$$\begin{cases} (\partial_t + \mathbf{L}_\sigma)V(t,x,\mu) = 0, \\ b(t,x,\mu) = (\sigma\sigma^*\partial_x V)(t,x,\mu), \end{cases} \quad t \in [0,T], x \in \mathbb{R}^d, \mu \in \mathscr{P}_2(\mathbb{R}^d). \tag{2.12}$$

(2) *A function* $V \in C_b^{1,2,(1,1)}([0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ *solves* (2.12) *if and only if there exists* $\Phi \in C_b^{2,(1,1)}(\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ *with* $\inf \Phi > 0$ *such that*

$$\begin{cases} V(t,x,\mu) = -\frac{1}{2}\mathbb{E}\big\{ \log \Phi(X_{t,T}^{x,\mu}, P_{t,T}^*\mu)\big\}, \\ b(t,x,\mu) = (\sigma\sigma^*\partial_x V)(t,x,\mu), \quad t \in [0,T], x \in \mathbb{R}^d, \mu \in \mathscr{P}_2(\mathbb{R}^d). \end{cases}$$

Since $b(t,x,\mu) = (\sigma\sigma^*\partial_x V)(t,x,\mu)$ implies that both $X_{t,T}^{x,\mu}$ and $P_{t,T}^*\mu$ may depend on V, unlike Theorem 2.3 and Corollary 2.4(2) providing solutions of (2.3) and (2.9) respectively, Corollary 2.5(2) only gives an alternative version of (2.12) but not solutions. To construct a nontrivial solution of (2.12), the nonlinear term

$$\int_{\mathbb{R}^d} \big\langle (\sigma\sigma^*\partial_y V)(t,y,\mu), \partial_\mu V(t,x,\mu)(y) \big\rangle \mu(\mathrm{d}y)$$

in $\mathbf{L}_\sigma$ causes an essential difficulty. To overcome this difficulty, many other things have to be treated. So, we would like to leave this problem to a forthcoming paper.

3 Proofs

We need the following Itô's formula for distribution dependent functionals, see [2, Proposition 6.1] or [5, Proposition A.8] under stronger conditions on σ and f.

Lemma 3.1 (Itô's formula for distribution dependent functional)**.** *For any $f \in \mathscr{C}([0,\infty)\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d))$, $f(t,X_t,\mathscr{L}_{X_t})$ is a semi-martingale with*

$$\mathrm{d}f(t,X_t,\mathscr{L}_{X_t}) = (\partial_t+\mathbf{L}_{\sigma,b})f(t,X_t,\mathscr{L}_{X_t})\mathrm{d}t + \langle(\sigma^*\partial_x f)(t,X_t,\mathscr{L}_{X_t}),\mathrm{d}W_t\rangle, \tag{3.1}$$

where $\mathbf{L}_{\sigma,b}$ is in (2.2).

Proof. Let $\mu_t=\mathscr{L}_{X_t}$ and

$$\bar b(t,x)=b(t,x,\mu_t),\ \ \bar\sigma(t,x)=\sigma(t,x,\mu_t),\ \ \ \bar f(t,x)=f(t,x,\mu_t),\ \ \ t\ge 0, x\in\mathbb{R}^d.$$

Then $(X_t)_{t\ge0}$ solves the classical SDE

$$\mathrm{d}X_t=\bar b(t,X_t)\mathrm{d}t+\bar\sigma(t,X_t)\mathrm{d}W_t.$$

By the definition 2.1 (4), $f\in\mathscr{C}([0,\infty)\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d))$ implies that $\bar f(t,x)$ is C^2-smooth in $x\in\mathbb{R}^d$. So, if $\bar f(t,x)$ is C^1 in $t\ge0$, we will be able to apply the classical Itô's formula to derive

$$\begin{aligned}\mathrm{d}f(t,X_t,\mathscr{L}_{X_t}) &= \mathrm{d}\bar f(t,X_t)=\langle(\bar\sigma^*\partial_x\bar f)(t,X_t),\mathrm{d}W_t\rangle\\ &\quad+\Big\{\partial_t\bar f+\frac12\sum_{i,j=1}^d(\bar\sigma\bar\sigma^*)_{ij}\partial_{x_i}\partial_{x_j}\bar f+\sum_{i=1}^d\bar b_i\partial_{x_i}\bar f\Big\}(t,X_t)\mathrm{d}t\\ &=\langle(\sigma^*\partial_x f)(t,X_t,\mathscr{L}_{X_t}),\mathrm{d}W_t\rangle+(\partial_t\bar f)(t,X_t)\mathrm{d}t\\ &\quad+\Big\{\frac12\sum_{i,j=1}^d(\sigma\sigma^*)_{ij}\partial_{x_i}\partial_{x_j}f+\sum_{i=1}^d b_i\partial_{x_i}f\Big\}(t,X_t,\mathscr{L}_{X_t})\mathrm{d}t.\end{aligned}$$

Therefore, to finish the proof, it suffices to show that $\bar f(t,x)$ is differentiable in $t\ge0$ and

$$\begin{aligned}(\partial_t\bar f)(t,x)&=\partial_t f(t,x,\nu)|_{\nu=\mu_t}+\partial_t f(s,x,\mu_t)|_{s=t}\\ &=:(\partial_t f)(t,x,\mu_t)+(\partial_t^\mu f)(t,x,\mu_t),\end{aligned}\tag{3.2}$$

where

$$\begin{aligned}(\partial_t^\mu f)(t,x,\mu_t):=&\frac12\sum_{i,j=1}^d\int_{\mathbb{R}^d}\Big[(\sigma\sigma^*)_{ij}(t,y,\mu_t)\partial_{y_j}\{(\partial_\mu f)_i(t,x,\mu_t)\}(y)\Big]\mu_t(\mathrm{d}y)\\ &+\sum_{i=1}^d\int_{\mathbb{R}^d}\Big[b_i(t,y,\mu_t)(\partial_\mu f)_i(t,x,\mu_t)\Big]\mu_t(\mathrm{d}y),\end{aligned}\tag{3.3}$$

is continuous in $(t,x) \in [0,\infty) \times \mathbb{R}^d$, since $f \in \mathscr{C}([0,\infty) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ and for any $T \in (0,\infty)$, $\{\mu_t : t \in [0,T]\}$ is a compact set in $\mathscr{P}_2(\mathbb{R}^d)$. Below we prove (3.2) by two steps.

(a) According to [5, Proposition A.6], if

$$\mathbb{E}\int_0^T \big\{|b(t,X_t,\mu_t)|^2 + \|\sigma(t,X_t,\mu_t)\|_{HS}^4\big\}\mathrm{d}t < \infty, \tag{3.4}$$

then for any $f \in \mathscr{C}([0,\infty) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$,

$$\begin{aligned} \bar{f}(t,x,\mu_{t+s}) - f(t,x,\mu_t) = \int_t^{t+s} \mathrm{d}r \int_{\mathbb{R}^d} \Big[\frac{1}{2}\sum_{i,j=1}^d (\sigma\sigma^*)_{ij}(r,x,\mu_r)\partial_{y_j}\{(\partial_\mu f)_i(r,y,\mu_r)\}(y) \\ + \sum_{i=1}^d b_i(r,y,\mu_r)(\partial_\mu f)_i(r,x,\mu_r)(y)\Big]\mu_r(\mathrm{d}y), \quad s>0. \end{aligned} \tag{3.5}$$

By conditions on b,σ and f, this implies (3.2).

(b) In general, let $T>0$ be fixed. By (1.2) and (1.3) we have

$$|b(t,x,\mu_t)|^2 + \|\sigma(t,x,\mu_t)\|_{HS}^2 \le C\Big(1+|x|^2 + \mathbb{W}_2^2(\mu_t,\delta_{\mathbf{0}})\Big), \quad x \in \mathbb{R}^d, t \in [0,T] \tag{3.6}$$

for some constant $C>0$. This, together with (1.4), implies

$$\mathbb{E}\int_0^T \big\{|b(t,X_t,\mu_t)|^2 + \|\sigma(t,X_t,\mu_t)\|_{HS}^2\big\}\mathrm{d}t < \infty.$$

So, to verify (3.4), we need to make approximations on σ. For any $k \in \mathbb{N}$, let

$$\phi_k(x) = \big(\{x_i \wedge k\} \vee \{-k\}\big)_{1\le i \le d}, \quad x \in \mathbb{R}^d.$$

Let $\sigma^{(k)}(t,x,\mu) = \sigma(t,\phi_k(x),\mu)$, and let $X_t^{(k)}$ solve the SDE

$$\mathrm{d}X_t^{(k)} = b(t,X_t^{(k)},\mathscr{L}_{X_t^{(k)}})\mathrm{d}t + \sigma^{(k)}(t,X_t^{(k)},\mathscr{L}_{X_t^{(k)}})\mathrm{d}W_t, \quad X_0^{(k)} = X_0. \tag{3.7}$$

Then as explained in (a), $\mu_t^{(k)} := \mathscr{L}_{X_t^{(k)}}$ satisfies

$$\begin{aligned} & f(t,x,\mu_{t+s}^{(k)}) - f(t,x,\mu_t^{(k)}) \\ & = \int_t^{t+s} \mathrm{d}r \int_{\mathbb{R}^d} \Big[\frac{1}{2}\sum_{i,j=1}^d (\sigma^{(k)}(\sigma^{(k)})^*)_{ij}(r,x,\mu_r^{(k)})\partial_{y_j}\{(\partial_\mu f)_i(r,y,\mu_r^{(k)})\}(y) \\ & \quad + \sum_{i=1}^d b_i(r,y,\mu_r^{(k)})(\partial_\mu f)_i(r,x,\mu_r^{(k)})(y)\Big]\mu_r^{(k)}(\mathrm{d}y), \quad s>0. \end{aligned} \tag{3.8}$$

We intend to show that with $k \to \infty$ this implies (3.5) and hence, completes the proof.

By Itô's formula and using (1.2) and (1.3), we may find out a constant $C > 0$ such that

$$\mathrm{d}|X_t - X_t^{(k)}|^2 \le \mathrm{d}M_t + C\big\{|X_t - X_t^{(k)}|^2 + \mathbb{E}|X_t - X_t^{(k)}|^2 + \|\sigma(t, X_t, \mu_t) - \sigma^{(k)}(t, X_t, \mu_t)\|_{HS}^2\big\}\mathrm{d}t, \quad t \in [0,T] \tag{3.9}$$

holds for some martingale M_t. By (1.2) and the definition of $\sigma^{(k)}$, for some constant $C' > 0$ we have

$$\|\sigma(t, X_t, \mu_t) - \sigma^{(k)}(t, X_t, \mu_t)\|_{HS}^2 \le C'|X_t - \phi_k(X_t)|^2 \le C'|X_t|^2 1_{\{|X_t| \ge k\}}.$$

Combining this with (3.9) and using Gronwall's lemma, we arrive at

$$\mathbb{E}|X_t - X_t^{(k)}|^2 \le C'\mathrm{e}^{2Ct} \int_0^t \mathbb{E}\big[|X_s|^2 1_{|X_s| \ge k}\big] ds, \quad t \in [0,T].$$

This, together with (1.4), implies

$$\lim_{k\to\infty} \sup_{t\in[0,T]} W_2(\mu_t, \mu_t^{(k)})^2 \le \lim_{k\to\infty} \sup_{t\in[0,T]} \mathbb{E}|X_t - X_t^{(k)}|^2 = 0.$$

In particular, $\{\mu_t^{(k)} : \ t \in [0,T], k \ge 1\}$ is compact in $\mathscr{P}_2(\mathbb{R}^d)$. So, from the continuity of $\sigma b, \partial_\mu f$, and $\partial_y \partial_\mu f$, the linear growth of $|\sigma b|$, and the condition $f \in \mathscr{C}([0,\infty) \times \mathbb{R}^d, \mathscr{P}_2(\mathbb{R}^d))$, it is easy to see that with $k \to \infty$ (3.8) implies (3.5). $\square$

Proof of Theorem 2.2. (1) Let $\mu_t = \mathscr{L}_{X_t}$. Applying the Itô formula (3.1) yields

$$\mathrm{d}V(t, X_t, \mu_t) = (\partial_t V + \mathbf{L}_{\sigma,b} V)(t, X_t, \mu_t)\mathrm{d}t + \langle (\sigma^* \partial_x V)(t, X_t, \mu_t), \mathrm{d}W_t \rangle. \tag{3.10}$$

This, together with (2.3), gives

$$\mathrm{d}V(t, X_t, \mu_t) = \mathbf{f}(t, X_t, \mu_t)\mathrm{d}t + \langle \mathbf{g}(t, X_t, \mu_t), \mathrm{d}W_t \rangle, \qquad t \geqslant 0. \tag{3.11}$$

Whence, (1.6) follows by integrating (3.11) from s to t.

(2) On the other hand, for any $s \in [0,T)$ and $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, let $X_s \in L^2(\Omega \to \mathbb{R}^d, \mathscr{F}_t, \mathbb{P})$ with $\mathscr{L}_{X_s} = \mu$, and let $(X_t)_{t\in[s,T]}$ solve (1.1) from time s. By combining (3.10) with (3.11) and using the uniqueness of decomposition for semi-martingale, we infer that

$$\mathbf{f}(t, X_t, \mu_t) = (\partial_t V + \mathbf{L}_{\sigma,b} V)(t, X_t, \mu_t), \qquad \mathbf{g}(t, X_t, \mu_t) = (\sigma^* \partial_x V)(t, X_t, \mu_t), \quad t \in [s,T],$$

where $\mu_t := \mathscr{L}_{X_t}$ with $\mu_s = \mu$. In particular, for any $s \in [0,T)$ and any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, we have

$$\mathbf{f}(s, X_s, \mu) = (\partial_t V + \mathbf{L}_{\sigma,b} V)(s, X_s, \mu), \qquad \mathbf{g}(s, X_s, \mu) = (\sigma^* \partial_x V)(s, X_s, \mu),$$

where $\mathscr{L}_{X_s} = \mu$. By the continuity of functions in these formulas, we obtain

$$\begin{aligned} &\mathbf{f}(s, x, \mu) = (\partial_t V + \mathbf{L}_{\sigma,b} V)(s, x, \mu), \\ &\mathbf{g}(s, x, \mu) = (\sigma^* \partial_x V)(s, x, \mu), \quad s \in [0,T], \mu \in \mathscr{P}_2(\mathbb{R}^d), x \in \mathrm{supp}\mu. \end{aligned} \tag{3.12}$$

To prove these formulas for all $x \in \mathbb{R}^d$, for any $\varepsilon > 0$ we let ν^ε be the centered Gaussian measure on $\mathbb{R}^d$ with covariance εI, where I is the identity matrix. Then

$$\mu^\varepsilon := \mu * \nu^\varepsilon \to \mu \text{ in } \mathscr{P}_2(\mathbb{R}^d) \text{ as } \varepsilon \downarrow 0. \tag{3.13}$$

Moreover, since ν^ε has full support for $\varepsilon > 0$, we have $\mathrm{supp}\mu^\varepsilon = \mathbb{R}^s$ for $\varepsilon > 0$ as well. So, by (3.12), for any $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\varepsilon > 0$, we have

$$\begin{aligned} &\mathbf{f}(s,x,\mu^\varepsilon) = (\partial_t V + \mathbf{L}_{\sigma,b} V)(s,x,\mu^\varepsilon),\\ &\mathbf{g}(s,x,\mu^\varepsilon) = (\sigma^*\partial_x V)(s,x,\mu^\varepsilon), \quad s \in [0,T], x \in \mathbb{R}^d. \end{aligned}$$

Combining this with (3.13) and noting that all functions in these formulas are continuous, by letting $\varepsilon \downarrow 0$ we prove (2.3). $\square$

To prove Theorem 2.3, we will need the following lemma, which reduces to the main result Theorem 6.2 in [2] when $b(t,x,\mu)$ and $\sigma(t,x,\mu)$ are independent of t. Since the proof of [2, Theorem 6.2] also applies to the the present time inhomogeneous situation, we skip the proof.

Lemma 3.2 ([2])**.** *In the situation of Theorem* 2.3*, let* $\Phi \in C_b^{2,(1,1)}(\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$*. Then* $V(t,x,\mu) := \mathbb{E}\Phi(X_{t,T}^{x,\mu}, P_{t,T}^*\mu)$ *is the unique solution to the PDE*

$$\begin{cases} (\partial_t + \mathbf{L}_{\sigma,b})V(t,x,\mu) = 0,\\ V(T,x,\mu) = \Phi(x,\mu), \qquad (t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d). \end{cases} \tag{3.14}$$

We will also need the following lemma for a probabilistic representation of a particular solution to the first equation in (2.3).

Lemma 3.3. *In the situation of Theorem* 2.3*, let*

$$V_{\mathbf{f}}(t,x,\mu) = -\mathbb{E}\int_t^T \mathbf{f}(r, X_{t,r}^{x,\mu}, P_{t,r}^*\mu)\mathrm{d}r, \quad (t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d).$$

Then $V_{\mathbf{f}}$ *is the unique solution to the PDE*

$$\begin{cases} (\partial_t + \mathbf{L}_{\sigma,b})V_{\mathbf{f}}(t,x,\mu) = \mathbf{f}(t,x,\mu),\\ V_{\mathbf{f}}(T,x,\mu) = 0, \quad (t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d). \end{cases} \tag{3.15}$$

Proof. (a) We first observe that $V_{\mathbf{f}}(t,x,\mu)$ solves (3.15). Obviously,

$$V_{\mathbf{f}}(T,x,\mu) = 0.$$

It remains to prove

$$(\partial_t + \mathbf{L}_{\sigma,b})V_{\mathbf{f}}(t,x,\mu) = \mathbf{f}(t,x,\mu). \tag{3.16}$$

By the definition of $V_{\mathbf{f}}$ and our condition on $\mathbf{f}$, we have

$$(\mathbf{L}_{\sigma,b}V_{\mathbf{f}})(t,x,\mu) = -\int_t^T \mathbf{L}_{\sigma,b}\{\mathbb{E}\mathbf{f}(r, X_{t,r}^{x,\mu}, P_{t,r}^*\mu)\}\mathrm{d}r,$$

and

$$(\partial_t V_{\mathbf{f}})(t,x,\mu) = \mathbf{f}(t,x,\mu) - \int_t^T \partial_t\{\mathbb{E}\mathbf{f}(r, X_{t,r}^{x,\mu}, P_{t,r}^*\mu)\}\mathrm{d}r.$$

So,

$$(\partial_t + \mathbf{L}_{\sigma,b})V_{\mathbf{f}}(t,x,\mu) = \mathbf{f}(t,x,\mu) - \int_t^T (\partial_t + \mathbf{L}_{\sigma,b})\{\mathbb{E}\mathbf{f}(r, X_{t,r}^{x,\mu}, P_{t,r}^*\mu)\}\mathrm{d}r.$$

On the other hand, applying Lemma 3.2 to $T = r$ and $\Phi(x,\mu) = \mathbf{f}(r,x,\mu)$, we obtain

$$(\partial_t + \mathbf{L}_{\sigma,b})\{\mathbb{E}\mathbf{f}(r, X_{t,r}^{x,\mu}, P_{t,r}^*\mu)\} = 0, \qquad r \in (t,T].$$

Therefore, (3.16) holds.

(b) We assume that $U(t,x,\mu)$ is another solution to (3.15) with $U(T,x,\mu) = 0$. By Lemma 3.1, for any $0 \leqslant t \leqslant s \leqslant T$,

$$U(s, X_{t,s}^{x,\mu}, P_{t,s}^*\mu) - \int_t^s \mathbf{f}(u, X_{t,u}^{x,\mu}, P_{t,u}^*\mu)\mathrm{d}u$$

and

$$V_{\mathbf{f}}(s, X_{t,s}^{x,\mu}, P_{t,s}^*\mu) - \int_t^s \mathbf{f}(u, X_{t,u}^{x,\mu}, P_{t,u}^*\mu)\mathrm{d}u$$

are martingales. Then

$$U(s, X_{t,s}^{x,\mu}, P_{t,s}^*\mu) - V_{\mathbf{f}}(s, X_{t,s}^{x,\mu}, P_{t,s}^*\mu)$$

is a martingale. Combining this with $U(T,x,\mu) = V_{\mathbf{f}}(T,x,\mu) = 0$, we arrive at

$$U(t,x,\mu) - V_{\mathbf{f}}(t,x,\mu) = \mathbb{E}\big(U(T, X_{t,T}^{x,\mu}, P_{t,T}^*\mu) - V_{\mathbf{f}}(T, X_{t,T}^{x,\mu}, P_{t,T}^*\mu)|\mathscr{F}_t\big) = 0.$$

Then the uniqueness is proved. □

Proof of Theorem 2.3. By Theorem 2.2, it suffices to prove the first assertion. By Lemma 3.2, we deduce that

$$V_1(t,x,\mu) := \mathbb{E}\Phi(X_{t,T}^{x,\mu}, P_{t,T}^*\mu)$$

is the unique solution to the PDE (3.14). And, according to Lemma 3.3, we know that

$$V_{\mathbf{f}}(t,x,\mu) := -\mathbb{E}\int_t^T \mathbf{f}(r, X_{t,r}^{x,\mu}, P_{t,r}^*\mu)\mathrm{d}r$$

solves (3.15). So,

$$V(t,x,\mu) := V_1(t,x,\mu) + V_{\mathbf{f}}(t,x,\mu) = \mathbb{E}\Phi(X_{t,T}^{x,\mu}, P_{t,T}^*\mu) - \mathbb{E}\int_t^T \mathbf{f}(r, X_{t,r}^{x,\mu}, P_{t,r}^*\mu)\mathrm{d}r \tag{3.17}$$

solves the first equation in (2.3).

On the other hand, let $V(t,x,\mu)$ solve the first equation in (2.3), and let

$$\Phi(x,\mu)=V(T,x,\mu),\quad (x,\mu)\in\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d).$$

It suffices to prove

$$V(t,x,\mu)=\mathbb{E}\bigg(\Phi(X^{x,\mu}_{t,T},P^*_{t,T}\mu)-\int_t^T\mathbf{f}(r,X^{x,\mu}_{t,r},P^*_{t,r}\mu)\mathrm{d}r\bigg). \tag{3.18}$$

Indeed, by (2.3) and Lemma 3.3, we have

$$(\partial_t+\mathbf{L}_{\sigma,b})(V-V_{\mathbf{f}})(t,x,\mu)=0.$$

So, Lemma 3.2 and $V_{\mathbf{f}}(T,x,\mu)=0$ imply

$$(V-V_{\mathbf{f}})(t,x,\mu)=\mathbb{E}\Phi(X^{x,\mu}_{t,T},P^*_{t,T}\mu)$$

with $(V-V_{\mathbf{f}})(T,x,\mu)=V(T,x,\mu)=\Phi(x,\mu)$. This, together with the definition of $V_{\mathbf{f}}$, implies (3.18). Then the proof is completed. □

Proof of Corollary 2.4. Assertion (1) is direct consequence of Theorem 2.2 for $\mathbf{f}=\frac{1}{2\beta}|\mathbf{g}|^2$. It remains to prove assertion (2).

Under the condition of assertion (2), let $\tilde V(t,x,\mu)=\mathbb{E}\Phi(X^{x,\mu}_{t,T},P^*_{t,T}\mu)$. By Lemma 3.2 we have

$$(\partial_t+\mathbf{L}_{\sigma,b})\tilde V(t,x,\mu)=0.$$

Since for V in (2.10) we have $V=-\beta\log\tilde V$, this implies

$$\begin{aligned}(\partial_t+\mathbf{L}_{\sigma,b})V(t,x,\mu)&=-\frac{\beta(\partial_t+\mathbf{L}_{\sigma,b})\tilde V}{\tilde V}(t,x,\mu)+\frac{\beta|\sigma^*\partial_x\tilde V|^2(t,x,\mu)}{2\tilde V^2(t,x,\mu)}\\&=\frac{1}{2\beta}|\sigma^*\partial_xV|^2.\end{aligned}$$

So, (2.9) holds, and the boundary condition $V(T,x,\mu)=-\beta\log\Phi(x,\mu)$ follows from (2.10) and the definition of $\tilde V$.

On the other hand, let $V\in C_b^{1,2,(1,1)}([0,T]\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d))$ solve (2.9). We take

$$\tilde V(t,x,\mu)=\exp[-\beta^{-1}V(t,x,\mu)],\quad (t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d). \tag{3.19}$$

It is easy to see that (2.9) implies

$$(\partial_t+\mathbf{L}_{\sigma,b})\tilde V(t,x,\mu)=0,\quad (t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d).$$

Therefore, by Lemma 3.2 we have

$$\tilde V(t,x,\mu)=\mathbb{E}\tilde V(T,X^{x,\mu}_{t,T},P^*_{t,T}\mu)=:\mathbb{E}\Phi(X^{x,\mu}_{t,T},P^*_{t,T}\mu)\quad (t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d).$$

Combining this with (3.19), we obtain (2.10) and hence finish the proof. □

Proof of Corollary 2.5. By (2.2) and (2.11), the definitions of $\mathbf{L}_\sigma$ and $\mathbf{L}_{\sigma,b}$, we have

$$
\begin{aligned}
&(\partial_t + \mathbf{L}_\sigma)V(t,x,\mu)\\
&= \partial_t V(t,x,\mu) + \frac{1}{2}\mathrm{tr}\big(\sigma\sigma^*\partial_x^2 V\big)(t,x,\mu) + \big\langle b, \partial_x V\big\rangle(t,x,\mu)\\
&\quad + \int_{\mathbb{R}^d}\Big[\frac{1}{2}\mathrm{tr}\big\{(\sigma\sigma^*)(t,y,\mu)\partial_y\partial_\mu V(t,x,\mu)(y)\big\} + \big\langle b(t,y,\mu), \partial_\mu V(t,x,\mu)(y)\big\rangle\Big]\mu(\mathrm{d}y)\\
&\quad + \frac{1}{2}|\sigma^*\partial_x V|^2(t,x,\mu) - \big\langle b, \partial_x V\big\rangle(t,y,\mu)\\
&\quad + \int_{\mathbb{R}^d}\big\langle(\sigma\sigma^*\partial_y V)(t,y,\mu), \partial_\mu V(t,x,\mu)(y)\big\rangle\mu(\mathrm{d}y)\\
&\quad - \int_{\mathbb{R}^d}\big\langle b(t,y,\mu), \partial_\mu V(t,x,\mu)(y)\big\rangle\mu(\mathrm{d}y)\\
&= (\partial_t + \mathbf{L}_{\sigma,b})V(t,x,\mu) + \frac{1}{2}|\sigma^*\partial_x V|^2(t,x,\mu) - \big\langle b, \partial_x V\big\rangle(t,x,\mu)\\
&\quad + \int_{\mathbb{R}^d}\big\langle(\sigma\sigma^*\partial_y V - b)(t,y,\mu), \partial_\mu V(t,x,\mu)(y)\big\rangle\mu(\mathrm{d}y).
\end{aligned}
$$

Combining this with $b(t,x,\mu) = (\sigma\sigma^*\partial_x V)(t,x,\mu)$, we obtain

$$
(\partial_t + \mathbf{L}_\sigma)V(t,x,\mu) = (\partial_t + \mathbf{L}_{\sigma,b})V(t,x,\mu) - \frac{1}{2}|\sigma^*\partial_x V|^2(t,x,\mu). \tag{3.20}
$$

We are now ready to finish the proof by using Theorem 2.2 and Corollary 2.4.

If (2.12) holds, then (1.9) holds for $\tilde b = \sigma^*\partial_x V$, and (3.20) implies (2.3) for $\mathbf{f}(t,x,\mu) = \frac{1}{2}|\tilde b|^2(t,x,\mu)$ and $\mathbf{g}(t,x,\mu) = \tilde b(t,x,\mu)$. So, by Theorem 2.2(1), $A_{s,t}$ is path independent. On the other hand, if (1.9) holds for $\tilde b = \sigma^*\partial_x V$ and $A_{s,t}$ is path independent in the sense of (1.6) for some $V \in C^{1,2,(1,1)}([0,T]\times\mathbb{R}^d\times\mathscr{P}_2(\mathbb{R}^d))$, then by Theorems 2.2(2) and (3.20), (2.12) holds. So, assertion (1) is proved.

Finally, by (3.20), the first equation in (2.12) is equivalent to (2.9) for $\beta = 1$. Then the second assertion (2) follows from Corollary 2.4(2) for $\beta = 1$. □

References

[1] F. Black, M. Scholes, *The pricing of options and corporate liabilities,* J. Political Economy 81(1973), 637-654.

[2] R. Buckdahn, J. Li, S. Peng, C. Rainer, *Mean-field stochastic differential equations and associated PDEs,* Annal. Probab. 2(2017), 824-878.

[3] P. Cardaliaguet, *Notes on mean field games,* P.-L. Lions lectures at College de France. Online at https://www.researchgate.net/publication/228702832_Notes_on_Mean_Field_Games

[4] J.-F. Chassagneux, D. Crisan, F. Delarue, *Classical solutions to the master equation for large population equilibria,* arXiv:1411.3009, 2014.

[5] W. Hammersley, D. Šiška, L. Szpruch, *McKean-Vlasov SDE under measure dependent Lyapunov conditions,* arXiv:1802.03974v1

[6] S. Hodges, A. Carverhill, *Quasi mean reversion in an efficient stock market: the characterisation of economic equilibria which support Black-Scholes option pricing,* The Economic Journal 103(1993), 395–405.

[7] X. Huang, F.-Y. Wang, *Distribution dependent SDEs with singular coefficients,* arXiv:1805.01682.

[8] Yu. S. Mishura, A. Yu. Veretennikov, *Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations,* arXiv:1603.02212v4

[9] H. Qiao, J.-L. Wu, *Characterizing the path-independence of the Girsanov transformation for non-Lipschitz SDEs with jumps,* Statist. Probab. Lett. 119(2016), 326–333.

[10] A. Truman, F.-Y. Wang, J.-L. Wu, W. Yang, *A link of stochastic differential equations to nonlinear parabolic equations,* Sci. China Math. 55(2012), 1971–1976.

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

[12] M. Wang, J.-L. Wu, *Necessary and sufficient conditions for path-independence of Girsanov transformation for infinite-dimensional stochastic evolution equations,* Front. Math. China 9(2014), 601–622.

[13] B. Wu, J.-L. Wu, *Characterizing the path-independent property of the Girsanov density for degenerated stochastic differential equations,* Stat. Probab. Letters 133(2018), 71–79.

[14] J.-L. Wu, W. Yang, *On stochastic differential equations and a generalised Burgers equation,* pp 425–435 in Stochastic Analysis and Its Applications to Finance: Essays in Honor of Prof. Jia-An Yan (eds. T S Zhang, X Y Zhou), Interdisciplinary Mathematical Sciences, Vol. 13, World Scientific, Singapore, 2012.