PATH INDEPENDENCE OF THE ADDITIVE FUNCTIONALS FOR MCKEAN-VLASOV STOCHASTIC DIFFERENTIAL EQUATIONS WITH JUMPS*

HUIJIE QIAO[1,2] AND JIANG-LUN WU[3]

1. School of Mathematics, Southeast University
Nanjing, Jiangsu 211189, China
2. Department of Mathematics, University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
hjqiaogean@seu.edu.cn
3. Department of Mathematics, Computational Foundry, Swansea University
Bay Campus, Swansea SA1 8EN, UK
j.l.wu@swansea.ac.uk

Abstract. In this article, the path independent property of additive functionals of McKean-Vlasov stochastic differential equations with jumps is characterised by nonlinear partial integro-differential equations involving L-derivatives with respect to probability measures introduced by P.-L. Lions. Our result extends the recent work [16] by Ren and Wang where their concerned McKean-Vlasov stochastic differential equations are driven by Brownian motion.

1. Introduction

Since the seminal work [12, 11], there have been substantial interests to study McKean-Vlasov stochastic differential equations, which are stochastic differential equations whose coefficients depend on the law of the solution, which are also referred as mean-field stochastic differential equations, see, e.g., [2] and most recently [1, 9] (and references therein). Very recently, Ren and Wang [16] explored an interesting result characterising the path independent additive functionals of McKean-Vlasov stochastic differential equations driven by Brownian motion by space-distribution partial differential equations, which extends the earlier work [18, 19] on this direction.

The object of this paper is to extend [16] to the same type of equations driven by compensated Poisson martingale measures (and Brownian motion). We aim to characterise the path-independence of additive functionals of McKean-Vlasov stochastic differential equations with jumps by certain partial integro-differential equations involving L-derivatives with respect to probability measures, following our previous work [14, 15] where therein stochastic differential equations with jumps in finite and infinite dimensions were studied, respectively. Let us also mention further interesting work [17, 10], where characterisation theorems for the path independence of additive functionals of stochastic differential equations driven by G-Brownian motion as well as for stochastic differential equations driven

AMS Subject Classification(2010): 60H30; 60J75; 35R04.

Keywords: McKean-Vlasov stochastic differential equations with jumps; Itô formula; additive functionals, partial integro-differential equations.

*This work was partly supported by NSF of China (No. 11001051, 11371352).

by Brownian motion with non-Markovian coefficients (i.e. random coefficients) are established, respectively. It is of course very interesting to extend the two cases to the situation of the concerned equations with jumps, which we will consider in our forthcoming work.

It is worthwhile to mention our results. We prove the Itô formula for McKean-Vlasov stochastic differential equations. And its proof is more simple than that in [5] and [9]. Moreover, it does not contain any abstract probability space. Therefore, it is more applicable. Besides, we compare our main result with that in [16] and [14]. And we find that it is indeed more general.

The rest of the paper is organized as follows. In the next section, we will set up the framework for introducing the McKean-Vlasov stochastic differential equations. In Section 3, we will first derive Itô formula for the solutions of our concerned McKean-Vlasov stochastic differential equations (and the proof is given in the Appendix at the end of our paper), and then we prove our characterisation theorem.

2. Preliminary

2.1. **Notations and notions.** In the subsection, we introduce notations and notions used in the sequel.

For convenience, we shall use $|\cdot|$ and $\|\cdot\|$ for norms of vectors and matrices, respectively. Furthermore, let $\langle\cdot\ ,\ \cdot\rangle$ denote the scalar product in $\mathbb{R}^d$. Let A^* denote the transpose of the matrix A.

Let $\mathscr{B}(\mathbb{R}^d)$ be the Borel σ-algebra on $\mathbb{R}^d$ and $\mathcal{M}(\mathbb{R}^d)$ be the space of all probability measures defined on $\mathscr{B}(\mathbb{R}^d)$ carrying the usual topology of weak convergence. Let $\mathcal{M}_2(\mathbb{R}^d)$ be the collection of all the probability measures μ on $\mathscr{B}(\mathbb{R}^d)$ satisfying

$$\int_{\mathbb{R}^d} |\ x\ |^2\ \mu(\mathrm{d}x) < \infty.$$

We put on $\mathcal{M}_2(\mathbb{R}^d)$ a topology induced by the following metric:

$$\rho^2(\mu_1,\mu_2) := \inf_{\pi\in\mathscr{C}(\mu_1,\mu_2)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\pi(\mathrm{d}x,\mathrm{d}y), \quad \mu_1,\mu_2\in\mathcal{M}_2(\mathbb{R}^d),$$

where $\mathscr{C}(\mu_1,\mu_2)$ denotes the set of all the probability measures whose marginal distributions are μ_1,μ_2, respectively. Thus, $(\mathcal{M}_2(\mathbb{R}^d),\rho)$ is a Polish space.

2.2. **McKean-Vlasov stochastic differential equations with jumps.** In the subsection, we introduce McKean-Vlasov stochastic differential equations with jumps and path-independence for a type of additive functionals.

Let $(\Omega,\mathscr{F},\mathbb{P};(\mathscr{F}_t)_{t\geqslant0})$ be a complete, filtered probability space. Let (B_t) be a m-dimensional $(\mathscr{F}_t)_{t\geqslant0}$-Brownian motion. Let $(\mathbb{U},\|\cdot\|_{\mathbb{U}})$ be a finite dimensional normed space with its Borel σ-algebra $\mathscr{U}$. Let ν be a σ-finite measure defined on $(\mathbb{U},\mathscr{U})$. We fix $\mathbb{U}_0=\{\|u\|_{\mathbb{U}}\leqslant\alpha\}$, where $\alpha>0$ is a constant, with $\nu(\mathbb{U}\backslash\mathbb{U}_0)<\infty$ and $\int_{\mathbb{U}_0}\|u\|_{\mathbb{U}}^2\,\nu(\mathrm{d}u)<\infty$. Following e.g. [7, 8], there exists an integer-valued $(\mathscr{F}_t)_{t\geqslant0}$-Poisson random measure $N(\mathrm{d}t,\mathrm{d}u)$ on $(\Omega,\mathcal{F},\mathbb{P};(\mathscr{F}_t)_{t\geqslant0})$ with intensity $\mathbb{E}N(\mathrm{d}t,\mathrm{d}u)=\mathrm{d}t\nu(\mathrm{d}u)$. Denote

$$\tilde{N}(\mathrm{d}t,\mathrm{d}u) := N(\mathrm{d}t,\mathrm{d}u) - \mathrm{d}t\nu(\mathrm{d}u),$$

that is, $\tilde{N}(\mathrm{d}t,\mathrm{d}u)$ stands for the compensated $(\mathscr{F}_t)_{t\geqslant0}$-predictable martingale measure of $N(\mathrm{d}t,\mathrm{d}u)$. Moreover, B_t and $N(\mathrm{d}t,\mathrm{d}u)$ are mutually independent.

Now, fix $T > 0$ and consider the following McKean-Vlaso stochastic differential equation with jumps on $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(t, X_t, \mathcal{L}_{X_t})\mathrm{d}t + \sigma(t, X_t, \mathcal{L}_{X_t})\mathrm{d}B_t + \int_{\mathbb{U}_0} f(t, X_{t-}, \mathcal{L}_{X_t}, u)\tilde{N}(\mathrm{d}t, \mathrm{d}u), \qquad t \in [0, T], \tag{1}$$

where $\mathcal{L}_{X_t}$ denotes the distribution of X_t under $\mathbb{P}$. Here the coefficients $b : [0, T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R}^{d\times m}$ and $f : [0, T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \times \mathbb{U}_0 \mapsto \mathbb{R}^d$ are all Borel measurable. We assume:

($\mathbf{H}_1$) (i) b and σ are bounded,
(ii) There exists a constant $L_1 \geqslant 0$ such that for $t_1, t_2 \in [0, T]$, $x_1, x_2 \in \mathbb{R}^d$, $\mu_1, \mu_2 \in \mathcal{M}_2(\mathbb{R}^d)$,

$$|b(t_1, x_1, \mu_1) - b(t_2, x_2, \mu_2)| + \|\sigma(t_1, x_1, \mu_1) - \sigma(t_2, x_2, \mu_2)\| \leqslant L_1(|x_1 - x_2| + \rho(\mu_1, \mu_2)).$$

($\mathbf{H}_2$) There exists a constant $L_2 \geqslant 0$ such that for $t_1, t_2 \in [0, T]$, $x_1, x_2 \in \mathbb{R}^d$, $\mu_1, \mu_2 \in \mathcal{M}_2(\mathbb{R}^d)$, $u \in \mathbb{U}_0$,

$$|f(t_1, x_1, \mu_1, u) - f(t_2, x_2, \mu_2, u)| \leqslant L_2\|u\|_{\mathbb{U}}(|x_1 - x_2| + \rho(\mu_1, \mu_2)),$$

and for $t \in [0, T], x \in \mathbb{R}^d, \mu \in \mathcal{M}_2(\mathbb{R}^d)$, $u \in \mathbb{U}_0$,

$$|f(t, x, \mu, u)| \leqslant L_2\|u\|_{\mathbb{U}}.$$

Under ($\mathbf{H}_1$) ($\mathbf{H}_2$), based on [5, Theorem 3.1, Page 7](Although the theorem is proved in [5] for $b(x, \mu), \sigma(x, \mu), f(x, \mu, u)$, the proof is right for the time inhomogeneous case), it holds that for any $s \in [0, T]$ and $X_s \in L^2(\Omega, \mathscr{F}_s, \mathbb{P}; \mathbb{R}^d)$, Eq.(1) has a unique solution $(X_t)_{t\geqslant s}$ with

$$\mathbb{E}\Big(\sup_{t\in[s,T]} |X_t|^2\Big) < \infty. \tag{2}$$

And then we introduce the following additive functional

$$\begin{aligned} F_{s,t} \ :=\ & \int_s^t g_1(r, X_r, \mathcal{L}_{X_r})\mathrm{d}r + \int_s^t \langle g_2(r, X_r, \mathcal{L}_{X_r}), \mathrm{d}B_r\rangle + \int_s^t \int_{\mathbb{U}_0} g_3(r, X_r, \mathcal{L}_{X_r}, u)\tilde{N}(\mathrm{d}r, \mathrm{d}u) \\ & + \int_s^t \int_{\mathbb{U}_0} g_4(r, X_r, \mathcal{L}_{X_r}, u)\nu(\mathrm{d}u)\mathrm{d}r, \quad 0 \leqslant s \leqslant t \leqslant T, \end{aligned} \tag{3}$$

where

$$g_1 : [0, T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R}, \qquad g_2 : [0, T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R}^m,$$
$$g_3 : [0, T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \times \mathbb{U}_0 \mapsto \mathbb{R}, g_4 : [0, T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \times \mathbb{U}_0 \mapsto \mathbb{R},$$

are Borel measurable, $g_1(t, x, \mu), g_2(t, x, \mu), g_3(t, x, \mu, u)$ are continuous in (t, x, μ) and $\int_{\mathbb{U}_0} g_4(t, x, \mu, u)\nu(\mathrm{d}u)$ is continuous in (t, x, μ), so that $F_{s,t}$ is a well-defined local semi-martingale.

Definition 2.1. *The additive functional $F_{s,t}$ is called path independent, if there exists a function*

$$V : [0, T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R},$$

such that for any $s \in [0, T]$ and $X_s \in L^2(\Omega, \mathscr{F}_s, \mathbb{P}; \mathbb{R}^d)$, the solution $(X_t)_{t\in[s,T]}$ of Eq.(1) satisfies

$$F_{s,t} = V(t, X_t, \mathcal{L}_{X_t}) - V(s, X_s, \mathcal{L}_{X_s}). \tag{4}$$

2.3. **L-derivative for functions on $\mathcal{M}_2(\mathbb{R}^d)$.** In the subsection we recall the definition of L-derivative for functions on $\mathcal{M}_2(\mathbb{R}^d)$. And the definition was first introduced by Lions [2]. Moreover, he used some abstract probability spaces to describe the L-derivatives. Here, for the convenience to understand the definition, we apply a straight way to state it ([16]). Let I be the identity map on $\mathbb{R}^d$. For $\mu \in \mathcal{M}_2(\mathbb{R}^d)$ and $\phi \in L^2(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d), \mu; \mathbb{R}^d)$, $\mu(\phi) := \int_{\mathbb{R}^d} \phi(x)\mu(\mathrm{d}x)$. Moreover, by simple calculation, it holds that $\mu \circ (I+\phi)^{-1} \in \mathcal{M}_2(\mathbb{R}^d)$.

Definition 2.2. *(i) A function $h : \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R}$ is called L-differentiable at $\mu \in \mathcal{M}_2(\mathbb{R}^d)$, if the functional*

$$L^2(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d), \mu; \mathbb{R}^d) \ni \phi \mapsto h(\mu \circ (I+\phi)^{-1})$$

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

$$\lim_{\mu(|\phi|^2) \to 0} \frac{h(\mu \circ (I+\phi)^{-1}) - h(\mu) - \mu(\langle \xi, \phi \rangle)}{\sqrt{\mu(|\phi|^2)}} = 0.$$

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

(ii) A function $h : \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R}$ is called L-differentiable on $\mathcal{M}_2(\mathbb{R}^d)$ if L-derivative $\partial_\mu h(\mu)$ exists for all $\mu \in \mathcal{M}_2(\mathbb{R}^d)$.

(iii) By the same way, $\partial^2_\mu h(\mu)(y)$ for $y \in \mathbb{R}^d$ can be defined.

Next, we introduce some related spaces.

Definition 2.3. *The function h is said to be in $C^2(\mathcal{M}_2(\mathbb{R}^d))$, if $\partial_\mu h$ is continuous, for any $\mu \in \mathcal{M}_2(\mathbb{R}^d)$, $\partial_\mu h(\mu)(\cdot)$ is differentiable, and its derivative $\partial_y \partial_\mu h : \mathcal{M}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is continuous, and for any $y \in \mathbb{R}^d$, $\partial_\mu h(\cdot)(y)$ is differentiable, and its derivative $\partial^2_\mu h : \mathcal{M}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is continuous.*

Definition 2.4. *(i) The function $h : [0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R}$ is said to be in $C^{1,2,2}([0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$, if $h(t,x,\mu)$ is C^1 in $t \in [0,T]$, C^2 in $x \in \mathbb{R}^d$ and $\mu \in \mathcal{M}_2(\mathbb{R}^d)$ respectively, and its derivatives*

$$\partial_t h(t,x,\mu), \partial_x h(t,x,\mu), \partial^2_x h(t,x,\mu), \partial_\mu h(t,x,\mu)(y), \partial_y \partial_\mu h(t,x,\mu)(y), \partial^2_\mu h(t,x,\mu)(y,y')$$

are jointly continuous in the corresponding variable family (t,x,μ), (t,x,μ,y) or (t,x,μ,y,y').

(ii) The function $h : [0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R}$ is said to be in $C^{1,2,2}_b([0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$, if $h \in C^{1,2,2}([0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$ and all its derivatives are uniformly bounded on $[0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d)$. If $h \in C^{1,2,2}([0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$ or $h \in C^{1,2,2}_b([0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$ and h is independent of t, we write $h \in C^{2,2}(\mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$ or $h \in C^{2,2}_b(\mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$.

(iii) The function $h : [0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d) \mapsto \mathbb{R}$ is said to be in $C^{1,2,2,1}_b([0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$, if $h \in C^{1,2,2}_b([0,T] \times \mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$ and all its derivatives are Lipschitz continuous. In addition, if h is independent of t, we write $h \in C^{2,2,1}_b(\mathbb{R}^d \times \mathcal{M}_2(\mathbb{R}^d))$.

3. Main results and their proofs

In the section, we state and prove the main results. First of all, we prove the Itô formula which is an important tool in our following proofs.

Proposition 3.1. (The Itô formula) *Suppose that* ($\mathbf{H}_1$) ($\mathbf{H}_2$) *hold. Then for any* $h \in C_b^{1,2,2}([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d))$, *it holds that for* $t \geqslant 0$,

$$
\begin{aligned}
\mathrm{d}h(t,X_t,\mathcal{L}_{X_t}) &= (\partial_t+\mathbf{L}_{b,\sigma,f})h(t,X_t,\mathcal{L}_{X_t})\mathrm{d}t+\langle(\sigma^*\partial_x h)(t,X_t,\mathcal{L}_{X_t}),\mathrm{d}B_t\rangle\\
&\quad+\int_{\mathbb{U}_0}\left[h(t,X_t+f(t,X_t,\mathcal{L}_{X_t},u),\mathcal{L}_{X_t})-h(t,X_t,\mathcal{L}_{X_t})\right]\tilde{N}(\mathrm{d}t,\mathrm{d}u), \quad (5)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{L}_{b,\sigma,f}h(t,x,\mu) &:= \langle b,\partial_x h\rangle(t,x,\mu)+\frac{1}{2}\mathrm{tr}\Big((\sigma^*\sigma)\partial_x^2 h\Big)(t,x,\mu)\\
&\quad+\int_{\mathbb{R}^d}\langle b(t,y,\mu),(\partial_\mu h)(t,x,\mu)(y)\rangle\mu(\mathrm{d}y)\\
&\quad+\frac{1}{2}\int_{\mathbb{R}^d}\mathrm{tr}\Big((\sigma^*\sigma)(t,y,\mu)\partial_y\partial_\mu h(t,x,\mu)(y)\Big)\mu(\mathrm{d}y)\\
&\quad+\int_{\mathbb{U}_0}\Bigg[h\Big(t,x+f(t,x,\mu,u),\mu\Big)-h(t,x,\mu)\\
&\qquad\qquad-\langle f(t,x,\mu,u),\partial_x h(t,x,\mu)\rangle\Bigg]\nu(\mathrm{d}u)\\
&\quad+\int_{\mathbb{U}_0}\int_0^1\int_{\mathbb{R}^d}\Big\langle\partial_\mu h(t,x,\mu)\Big(y+\eta f(t,y,\mu,u)\Big)\\
&\qquad\qquad-\partial_\mu h(t,x,\mu)(y),f(t,y,\mu,u)\Big\rangle\mu(\mathrm{d}y)\mathrm{d}\eta\nu(\mathrm{d}u). \quad (6)
\end{aligned}
$$

Because its proof is too long, we place it to the Appendix so as to make the context more compact. Now, it is the position to state and prove our main result.

Theorem 3.2. *Assume that* b,σ,f *satisfy* ($\mathbf{H}_1$) ($\mathbf{H}_2$). *Then for* $V\in C_b^{1,2,2}([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d))$, $g_1\in C([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\mapsto\mathbb{R})$, $g_2\in C([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\mapsto\mathbb{R}^m)$, $g_3(\cdot,\cdot,\cdot,u)\in C([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\mapsto\mathbb{R})$ *and* $\int_{\mathbb{U}_0}g_4(\cdot,\cdot,\cdot,u)\nu(\mathrm{d}u)\in C([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\mapsto\mathbb{R})$, $F_{s,t}$ *is path independent in the sense of (4) if and only if* (V,g_1,g_2,g_3,g_4) *satisfies the integral-partial differential equation*

$$
\left\{
\begin{array}{l}
(\partial_t+\mathbf{L}_{b,\sigma,f})V(t,x,\mu)=g_1(t,x,\mu)+\int_{\mathbb{U}_0}g_4(t,x,\mu,u)\nu(\mathrm{d}u),\\
(\sigma^*\partial_x V)(t,x,\mu)=g_2(t,x,\mu),\\
V\Big(t,x+f(t,x,\mu,u),\mu\Big)-V(t,x,\mu)=g_3(t,x,\mu,u),\\
t\in[0,T],x\in\mathbb{R}^d,\mu\in\mathcal{M}_2(\mathbb{R}^d),u\in\mathbb{U}_0.
\end{array}
\right. \quad (7)
$$

Proof. First, we prove sufficiency. For $V\in C_b^{1,2,2}([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d))$, based on Proposition 3.1, it holds that

$$
\begin{aligned}
\mathrm{d}V(t,X_t,\mathcal{L}_{X_t}) &= (\partial_t+\mathbf{L}_{b,\sigma,f})V(t,X_t,\mathcal{L}_{X_t})\mathrm{d}t+\langle(\sigma^*\partial_x V)(t,X_t,\mathcal{L}_{X_t}),\mathrm{d}B_t\rangle\\
&\quad+\int_{\mathbb{U}_0}\left[V(t,X_t+f(t,X_t,\mathcal{L}_{X_t},u),\mathcal{L}_{X_t})-V(t,X_t,\mathcal{L}_{X_t})\right]\tilde{N}(\mathrm{d}t,\mathrm{d}u). \quad (8)
\end{aligned}
$$

Inserting (7) in (8), we have

$$
\mathrm{d}V(t,X_t,\mathcal{L}_{X_t}) = g_1(t,X_t,\mathcal{L}_{X_t})\mathrm{d}t+\int_{\mathbb{U}_0}g_4(t,X_t,\mathcal{L}_{X_t},u)\nu(\mathrm{d}u)\mathrm{d}t+\langle g_2(t,X_t,\mathcal{L}_{X_t}),\mathrm{d}B_t\rangle
$$

$$+\int_{\mathbb{U}_0} g_3(t,X_t,\mathcal{L}_{X_t},u)\tilde{N}(\mathrm{d}t,\mathrm{d}u).$$

By integrating the above equality from s to t, one can obtain (4). That is, $F_{s,t}$ is path independent.

Next, let us show necessity. On one hand, since $F_{s,t}$ is path independent, it follows from Definition 2.1 that

$$\begin{aligned}
V(t,X_t,\mathcal{L}_{X_t})-V(0,X_0,\mathcal{L}_{X_0}) &= \int_0^t g_1(r,X_r,\mathcal{L}_{X_r})\mathrm{d}r+\int_0^t\langle g_2(r,X_r,\mathcal{L}_{X_r}),\mathrm{d}B_r\rangle\\
&\quad+\int_0^t\int_{\mathbb{U}_0} g_3(r,X_r,\mathcal{L}_{X_r},u)\tilde{N}(\mathrm{d}r,\mathrm{d}u)\\
&\quad+\int_0^t\int_{\mathbb{U}_0} g_4(r,X_r,\mathcal{L}_{X_r},u)\nu(\mathrm{d}u)\mathrm{d}r,\quad t\geqslant 0. \qquad (9)
\end{aligned}$$

On the other hand, by integrating (8) from 0 to t, we get that

$$\begin{aligned}
V(t,X_t,\mathcal{L}_{X_t})-V(0,X_0,\mathcal{L}_{X_0}) &= \int_0^t(\partial_r+\mathbf{L}_{b,\sigma,f})V(r,X_r,\mathcal{L}_{X_r})\mathrm{d}r\\
&\quad+\int_0^t\langle(\sigma^*\partial_x V)(r,X_r,\mathcal{L}_{X_r}),\mathrm{d}B_r\rangle\\
&\quad+\int_0^t\int_{\mathbb{U}_0}\Big[V(r,X_r+f(r,X_r,\mathcal{L}_{X_r},u),\mathcal{L}_{X_r})\\
&\qquad\qquad -V(r,X_r,\mathcal{L}_{X_r})\Big]\tilde{N}(\mathrm{d}r,\mathrm{d}u). \qquad (10)
\end{aligned}$$

Thus, $V(t,X_t,\mathcal{L}_{X_t})-V(0,X_0,\mathcal{L}_{X_0})$ has two expressions. Since $V(t,X_t,\mathcal{L}_{X_t})-V(0,X_0,\mathcal{L}_{X_0})$ is a semimartingale, by uniqueness for decomposition of the semimartingale it holds that

$$\begin{aligned}
&g_1(r,X_r,\mathcal{L}_{X_r})+\int_{\mathbb{U}_0} g_4(r,X_r,\mathcal{L}_{X_r},u)\nu(\mathrm{d}u)=(\partial_r+\mathbf{L}_{b,\sigma,f})V(r,X_r,\mathcal{L}_{X_r}),\\
&g_2(r,X_r,\mathcal{L}_{X_r})=(\sigma^*\partial_x V)(r,X_r,\mathcal{L}_{X_r}),\\
&g_3(r,X_r,\mathcal{L}_{X_r},u)=V(r,X_r+f(r,X_r,\mathcal{L}_{X_r},u),\mathcal{L}_{X_r})-V(r,X_r,\mathcal{L}_{X_r}), r\in[0,T].
\end{aligned}$$

And then for any $s\in[0,T]$ and $\mu=\mathcal{L}_{X_s}\in\mathcal{M}_2(\mathbb{R}^d)$, we know that

$$\begin{aligned}
&g_1(s,X_s,\mu)+\int_{\mathbb{U}_0} g_4(s,X_s,\mu,u)\nu(\mathrm{d}u)=(\partial_r+\mathbf{L}_{b,\sigma,f})V(s,X_s,\mu),\\
&g_2(s,X_s,\mu)=(\sigma^*\partial_x V)(s,X_s,\mu),\\
&g_3(s,X_s,\mu,u)=V(s,X_s+f(s,X_s,\mu,u),\mu)-V(s,X_s,\mu),
\end{aligned}$$

and then

$$\begin{aligned}
&g_1(s,x,\mu)+\int_{\mathbb{U}_0} g_4(s,x,\mu,u)\nu(\mathrm{d}u)=(\partial_r+\mathbf{L}_{b,\sigma,f})V(s,x,\mu),\\
&g_2(s,x,\mu)=(\sigma^*\partial_x V)(s,x,\mu),\\
&g_3(s,x,\mu,u)=V(s,x+f(s,x,\mu,u),\mu)-V(s,x,\mu), x\in supp(\mu).
\end{aligned}$$

To show (7), we replace μ by $\mu^n=\mu*\mathcal{N}_d(0,\frac{1}{n}I_d)$ in the above equality, where $\mathcal{N}_d(0,\frac{1}{n}I_d)$ denotes the d-dimensional Gaussian distribution with mean 0 and covariance matrix $\frac{1}{n}I_d$.

Note that $supp(\mathcal{N}_d(0,\frac{1}{n}I_d))=\mathbb{R}^d$, $supp(\mu^n)=\mathbb{R}^d$ and $\mu^n\to\mu$ as $n\to\infty$, which together with continuity of all the related functions in μ yields (7). The proof is completed. $\square$

Next, we give out a solution of the partial integro-differential equation in (7). To do this, we introduce two McKean-Vlaso stochastic differential equations with jumps on $\mathbb{R}^d$: for any $\xi\in L^2(\Omega,\mathscr{F},\mathbb{P};\mathbb{R}^d)$ and $x\in\mathbb{R}^d$,

$$
\begin{aligned}
X_t^{s,\xi} &= \xi+\int_s^t b(r,X_r^{s,\xi},\mathcal{L}_{X_r^{s,\xi}})\mathrm{d}r+\int_s^t \sigma(r,X_r^{s,\xi},\mathcal{L}_{X_r^{s,\xi}})\mathrm{d}B_r \\
&\quad +\int_s^t\int_{\mathbb{U}_0} f(r,X_{r-}^{s,\xi},\mathcal{L}_{X_r^{s,\xi}},u)\tilde{N}(\mathrm{d}r,\mathrm{d}u), \quad 0\leqslant s<t\leqslant T, \qquad (11)\\
X_t^{s,x,\xi} &= x+\int_s^t b(r,X_r^{s,x,\xi},\mathcal{L}_{X_r^{s,\xi}})\mathrm{d}r+\int_s^t \sigma(r,X_r^{s,x,\xi},\mathcal{L}_{X_r^{s,\xi}})\mathrm{d}B_r \\
&\quad +\int_s^t\int_{\mathbb{U}_0} f(r,X_{r-}^{s,x,\xi},\mathcal{L}_{X_r^{s,\xi}},u)\tilde{N}(\mathrm{d}r,\mathrm{d}u), \quad 0\leqslant s<t\leqslant T, \qquad (12)
\end{aligned}
$$

and a backward McKean-Vlasov stochastic differential equation

$$
\left\{\begin{array}{l}
\mathrm{d}Y_t^{s,x,\xi}=g_1(t,X_t^{s,x,\xi},\mathcal{L}_{X_t^{s,\xi}})\mathrm{d}t+\int_{\mathbb{U}_0} g_4(t,X_t^{s,x,\xi},\mathcal{L}_{X_t^{s,\xi}},u)\nu(\mathrm{d}u)\mathrm{d}t,\\
Y_T^{s,x,\xi}=\Phi(X_T^{s,x,\xi},\mathcal{L}_{X_T^{s,\xi}}).
\end{array}\right. \qquad (13)
$$

Under $(\mathbf{H}_1)$ $(\mathbf{H}_2)$, based on [5, Theorem 3.1, Page 7], it holds that the above equations (11) (12) have unique solutions $X_t^{s,\xi}$, $X_t^{s,x,\xi}$, respectively. If we more assume g_1, $\int_{\mathbb{U}_0} g_4(\cdot,\cdot,\cdot,u)\nu(\mathrm{d}u)$, Φ are bounded, the above equation (13) also has a unique solution. For $\Phi\in C_b^{2,2,1}(\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d))$, $g_1(t,\cdot,\cdot)\in C_b^{2,2,1}(\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d))$, $\int_{\mathbb{U}_0} g_4(t,\cdot,\cdot,u)\nu(\mathrm{d}u)\in C_b^{2,2,1}(\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d))$ and $g_1(\cdot,x,\mu)\in C([0,T])$, $\int_{\mathbb{U}_0} g_4(\cdot,x,\mu,u)\nu(\mathrm{d}u)\in C([0,T])$, set

$$
\begin{aligned}
V(t,x,\mu) &:= \mathbb{E}\Big[\Phi(X_T^{t,x,\xi},\mathcal{L}_{X_T^{t,\xi}})-\int_t^T g_1(r,X_r^{t,x,\xi},\mathcal{L}_{X_r^{t,\xi}})\mathrm{d}r\\
&\qquad -\int_t^T\int_{\mathbb{U}_0} g_4(r,X_r^{t,x,\xi},\mathcal{L}_{X_r^{t,\xi}},u)\nu(\mathrm{d}u)\mathrm{d}r\Big], \quad \mu=\mathcal{L}_\xi, \qquad (14)
\end{aligned}
$$

and then by [9, Theorem 9.2, Page 3159], it holds that $V(t,x,\mu)\in C_b^{1,2,2}([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d))$ is the unique solution of the following nonlocal integral-partial differential equation

$$
\left\{\begin{array}{l}
(\partial_t+\mathbf{L}_{b,\sigma,f})V(t,x,\mu)=g_1(t,x,\mu)+\int_{\mathbb{U}_0} g_4(t,x,\mu,u)\nu(\mathrm{d}u), \quad t\in[0,T], x\in\mathbb{R}^d, \mu\in\mathcal{M}_2(\mathbb{R}^d),\\
V(T,x,\mu)=\Phi(x,\mu).
\end{array}\right.
$$

Thus, by combining Theorem 3.2 with [9, Theorem 9.2, Page 3159], one can have the following result.

Corollary 3.3. *Assume that $(\mathbf{H}_1)$ $(\mathbf{H}_2)$ hold, $(b(t,x,\mu),\sigma(t,x,\mu))\in C_b^{1,2,2,1}([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\mapsto\mathbb{R}^d\times\mathbb{R}^{d\times m})$, and all the derivatives of $f(t,x,\mu,u)$ in t order 1 and in x,μ up to order 2 are bounded by $L\|u\|_{\mathbb{U}}$ and Lipschitz continuous with a Lipschitz factor $L\|u\|_{\mathbb{U}}$. Then for $V(t,x,\mu)$ defined in (14), $g_2\in C([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\mapsto\mathbb{R}^m)$ and $g_3(\cdot,\cdot,\cdot,u)\in C([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\mapsto\mathbb{R})$, $F_{s,t}$ is path independent in the sense of (4)* *if and only if* V, g_2, g_3 *satisfy*

$$\left\{\begin{array}{l}(\sigma^*\partial_x V)(t,x,\mu)=g_2(t,x,\mu),\\ V\Big(t,x+f(t,x,\mu,u),\mu\Big)-V(t,x,\mu)=g_3(t,x,\mu,u),\\ t\in[0,T], x\in\mathbb{R}^d, \mu\in\mathcal{M}_2(\mathbb{R}^d), u\in\mathbb{U}_0.\end{array}\right.$$

In the following we analysis some special cases of $F_{s,t}$. If $g_1=0, g_4=0$, $F_{s,t}$ reduces to

$$F^{g_2,g_3}_{s,t}:=\int_s^t\langle g_2(r,X_r,\mathcal{L}_{X_r}),\mathrm{d}B_r\rangle+\int_s^t\int_{\mathbb{U}_0}g_3(r,X_r,\mathcal{L}_{X_r},u)\tilde{N}(\mathrm{d}r,\mathrm{d}u).$$

We follow up the above deduction to get that for $V(t,x,\mu):=\mathbb{E}[\Phi(X^{t,x,\xi}_T,\mathcal{L}_{X^{t,\xi}_T})]$, $g_2\in C([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\mapsto\mathbb{R}^m)$ and $g_3(\cdot,\cdot,\cdot,u)\in C([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\mapsto\mathbb{R})$, $F^{g_2,g_3}_{s,t}$ is path independent in the sense of (4) if and only if V, g_2, g_3 satisfy

$$\left\{\begin{array}{l}(\sigma^*\partial_x V)(t,x,\mu)=g_2(t,x,\mu),\\ V\Big(t,x+f(t,x,\mu,u),\mu\Big)-V(t,x,\mu)=g_3(t,x,\mu,u),\\ t\in[0,T], x\in\mathbb{R}^d, \mu\in\mathcal{M}_2(\mathbb{R}^d), u\in\mathbb{U}_0.\end{array}\right.$$

This result also can be obtained by Theorem 3.2 and [5, Theorem 7.3, Page 47]. If $g_1=\frac{1}{2\beta}|g_2|^2, \beta\neq0$, $F_{s,t}$ reduces to

$$\begin{aligned}F^{g_2,g_3,g_4}_{s,t}\quad:=\quad&\int_s^t\frac{1}{2\beta}|g_2|^2(r,X_r,\mathcal{L}_{X_r})\mathrm{d}r+\int_s^t\langle g_2(r,X_r,\mathcal{L}_{X_r}),\mathrm{d}B_r\rangle\\ &+\int_s^t\int_{\mathbb{U}_0}g_3(r,X_r,\mathcal{L}_{X_r},u)\tilde{N}(\mathrm{d}r,\mathrm{d}u)\\ &+\int_s^t\int_{\mathbb{U}_0}g_4(r,X_r,\mathcal{L}_{X_r},u)\nu(\mathrm{d}u)\mathrm{d}r,\quad 0\leqslant s\leqslant t\leqslant T.\end{aligned}$$

Thus, by Theorem 3.2, it holds that $F^{g_2,g_3,g_4}_{s,t}$ is path independent in the sense of (4) if and only if

$$\left\{\begin{array}{l}(\partial_t+\mathbf{L}_{b,\sigma,f})V(t,x,\mu)=\frac{1}{2\beta}|\sigma^*\partial_x V|^2(t,x,\mu)+\int_{\mathbb{U}_0}g_4(t,x,\mu,u)\nu(\mathrm{d}u),\\ (\sigma^*\partial_x V)(t,x,\mu)=g_2(t,x,\mu),\\ V\Big(t,x+f(t,x,\mu,u),\mu\Big)-V(t,x,\mu)=g_3(t,x,\mu,u),\\ t\in[0,T], x\in\mathbb{R}^d, \mu\in\mathcal{M}_2(\mathbb{R}^d), u\in\mathbb{U}_0.\end{array}\right. \tag{15}$$

Finally, we discuss the relation between Theorem 3.2 and [14, Theorem 2.6] [16, Theorem 2.2]. Let $\lambda:[0,\infty)\times\mathbb{U}_0\to(0,1]$ be a measurable function. And then there exists an integer-valued $(\mathscr{F}_t)_{t\geqslant0}$-Poisson random measure $N_\lambda(\mathrm{d}t,\mathrm{d}u)$ on $(\Omega,\mathscr{F},\mathbb{P};(\mathscr{F}_t)_{t\geqslant0})$ with intensity $\mathbb{E}(N_\lambda(\mathrm{d}t,\mathrm{d}u))=\lambda(t,u)\mathrm{d}t\nu(\mathrm{d}u)$. Denote

$$\tilde{N}_\lambda(\mathrm{d}t,\mathrm{d}u):=N_\lambda(\mathrm{d}t,\mathrm{d}u)-\lambda(t,u)\mathrm{d}t\nu(\mathrm{d}u)$$

that is, $\tilde{N}_\lambda(\mathrm{d}t,\mathrm{d}u)$ stands for the compensated $(\mathscr{F}_t)_{t\geqslant0}$-predictable martingale measure of $N_\lambda(\mathrm{d}t,\mathrm{d}u)$. Moreover, $\tilde{N}_\lambda(\mathrm{d}t,\mathrm{d}u)$ is independent of B_t. We replace $\tilde{N}(\mathrm{d}t,\mathrm{d}u)$ by $\tilde{N}_\lambda(\mathrm{d}t,\mathrm{d}u)$ in Eq.(1). Thus, the solution of the new equation is denoted as X^λ_t. Besides, we assume that there exists a measurable function $\tilde{b}:[0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d)\to\mathbb{R}^m$ such that $b=\sigma\tilde{b}$. For convenience of the following deduction, we also assume:

$$
\begin{aligned}
&\mathbb{E}\Big[\exp\Big\{\frac{1}{2}\int_0^T\Big|\tilde{b}(s,X_s^\lambda,\mathcal{L}_{X_s^\lambda})\Big|^2\mathrm{d}s\Big\}\Big]<\infty,\\
&\int_0^T\int_{\mathbb{U}_0}\left(\frac{1-\lambda(s,u)}{\lambda(s,u)}\right)^2\lambda(s,u)\nu(\mathrm{d}u)\mathrm{d}s<\infty.
\end{aligned}
$$

So, set

$$
\begin{aligned}
\Gamma_t: \ = \ \exp\Big\{&-\int_0^t\langle\tilde{b}(s,X_s^\lambda,\mathcal{L}_{X_s^\lambda}),\mathrm{d}B_s\rangle-\frac{1}{2}\int_0^t\Big|\tilde{b}(s,X_s^\lambda,\mathcal{L}_{X_s^\lambda})\Big|^2\mathrm{d}s\\
&-\int_0^t\int_{\mathbb{U}_0}\log\lambda(s,u)\tilde{N}_\lambda(\mathrm{d}s,\mathrm{d}u)\\
&-\int_0^t\int_{\mathbb{U}_0}\Big(\big(\log\lambda(s,u)\big)\lambda(s,u)+\big(1-\lambda(s,u)\big)\Big)\nu(\mathrm{d}u)\mathrm{d}s\Big\},
\end{aligned}
$$

and then by the same deduction to that in [14], it holds that Γ_t is an exponential martingale. Define a new probability $\mathbb{Q}$ as

$$
\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}=\Gamma_T.
$$

Thus, under $\mathbb{Q}$,

$$
\tilde{B}_t:=B_t+\int_0^t\tilde{b}(s,X_s^\lambda,\mathcal{L}_{X_s^\lambda})\mathrm{d}s
$$

is a d-dimensional Brownian motion and

$$
\bar{N}(\mathrm{d}t,\mathrm{d}u)=N_\lambda(\mathrm{d}t,\mathrm{d}u)-\mathrm{d}t\nu(\mathrm{d}u)
$$

is the compensated $(\mathscr{F}_t)_{t\geqslant0}$-predictable martingale measure of $N_\lambda(\mathrm{d}t,\mathrm{d}u)$. Moreover, Eq.(1) becomes

$$
X_t^\lambda=X_s^\lambda+\int_s^t\sigma(r,X_r^\lambda,\mathcal{L}_{X_r^\lambda})\mathrm{d}\tilde{B}_r+\int_s^t\int_{\mathbb{U}_0}f(r,X_{r-}^\lambda,\mathcal{L}_{X_r^\lambda},u)\bar{N}(\mathrm{d}r,\mathrm{d}u).
$$

That is, X_t^λ is a local martingale.

Now, take

$$
\begin{aligned}
&g_1(t,x,\mu)=\frac{1}{2}\Big|\tilde{b}(t,x,\mu)\Big|^2,\ g_2(t,x,\mu)=\tilde{b}(t,x,\mu),\\
&g_3(t,x,\mu,u)=\log\lambda(t,u),\quad g_4(t,x,\mu,u)=\log\lambda(t,u)+\big(\frac{1}{\lambda(t,u)}-1\big).
\end{aligned}
$$

And then by Theorem 3.2, we know that $F_{0,t}^\lambda=-\log\Gamma_t$ is path independent in the sense of (4) if and only if

$$
\left\{
\begin{array}{l}
(\partial_t+\mathbf{L}_{b,\sigma,f})V(t,x,\mu)=\frac{1}{2}\left|(\sigma^*\partial_xV)(t,x,\mu)\right|^2+\int_{\mathbb{U}_0}\Big(\big(\log\lambda(t,u)\big)\lambda(t,u)+\big(1-\lambda(t,u)\big)\Big)\nu(\mathrm{d}u),\\
(\sigma^*\partial_xV)(t,x,\mu)=\tilde{b}(t,x,\mu),\\
V\Big(t,x+f(t,x,\mu,u),\mu\Big)-V(t,x,\mu)=\log\lambda(t,u),\\
t\in[0,T],x\in\mathbb{R}^d,\mu\in\mathcal{M}_2(\mathbb{R}^d),u\in\mathbb{U}_0.
\end{array}
\right.
\tag{16}
$$

The equation is just right Eq.(15) with $\beta=1$. And then we rewrite Eq.(16) as

$$\left\{\begin{array}{l}\partial_t V(t,x,\mu)=\mathbf{L}_{\sigma,f}V(t,x,\mu),\\ b(t,x,\mu)=(\sigma\sigma^*\partial_x V)(t,x,\mu),\\ \lambda(t,u)=\exp\left\{V\Big(t,x+f(t,x,\mu,u),\mu\Big)-V(t,x,\mu)\right\},\\ t\in[0,T],x\in\mathbb{R}^d,\mu\in\mathcal{M}_2(\mathbb{R}^d),u\in\mathbb{U}_0,\end{array}\right.$$

where

$$\begin{aligned}\mathbf{L}_{\sigma,f}V(t,x,\mu) \;:=\; & -\frac{1}{2}\mathrm{tr}\Big((\sigma^*\sigma)\partial_x^2 V\Big)(t,x,\mu)-\frac{1}{2}\left|(\sigma^*\partial_x V)(t,x,\mu)\right|^2\\ &-\int_{\mathbb{R}^d}\langle(\sigma\sigma^*\partial_y V)(t,y,\mu),(\partial_\mu V)(t,x,\mu)(y)\rangle\mu(\mathrm{d}y)\\ &-\frac{1}{2}\int_{\mathbb{R}^d}\mathrm{tr}\Big((\sigma^*\sigma)(t,y,\mu)\partial_y\partial_\mu V(t,x,\mu)(y)\Big)\mu(\mathrm{d}y)\\ &-\int_{\mathbb{U}_0}\Big[e^{V(t,x+f(t,x,\mu,u),\mu)-V(t,x,\mu)}-1\\ &\qquad\qquad -\langle f(t,x,\mu,u),\partial_x V(t,x,\mu)\rangle e^{V(t,x+f(t,x,\mu,u),\mu)-V(t,x,\mu)}\Big]\nu(\mathrm{d}u)\\ &-\int_{\mathbb{U}_0}\int_0^1\int_{\mathbb{R}^d}\Big\langle\partial_\mu V(t,x,\mu)\Big(y+\eta f(t,y,\mu,u)\Big)\\ &\qquad\qquad -\partial_\mu V(t,x,\mu)(y),f(t,y,\mu,u)\Big\rangle\mu(\mathrm{d}y)\mathrm{d}\eta\nu(\mathrm{d}u).\end{aligned}$$

If b,σ,f,V are independent of μ, this is just right Theorem 2.6 in [14]. Moreover, if $f=0$, this is exactly Theorem 2.2 in [16]. Therefore, our result is more general.

4. Appendix

The proof of Proposition 3.1.

Set $\mu_t:=\mathcal{L}_{X_t}$, and then $h(t,X_t,\mathcal{L}_{X_t})=h(t,X_t,\mu_t)$. Define $\bar{h}(t,x):=h(t,x,\mu_t)$, and then $\bar{h}(t,X_t)=h(t,X_t,\mu_t)$. Moreover, based on $h\in C_b^{1,2,2}([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d))$, it holds that $\bar{h}$ is C^2 in x. However, we don't know the differentiability of $\bar{h}$ in t. Note that the differentiability of $\bar{h}$ in t comes from two parts-$h(t,x,\mu)$ in t for fixed x,μ and $h(s,x,\mu_t)$ in t for fixed s,x. Therefore, to apply the classical Itô formula to $\bar{h}(t,X_t)$, we only need to consider the second part by $h\in C_b^{1,2,2}([0,T]\times\mathbb{R}^d\times\mathcal{M}_2(\mathbb{R}^d))$.

Step 1. We study the differentiability of $h(s,x,\mu_t)$ in t.

Here, we follow the method in [3] to deal with it. For the convenience to our expression, we take $H(\mu_t):=h(s,x,\mu_t)$. For any positive integer K, set

$$x^1,x^2,\cdots,x^K\in\mathbb{R}^d, H^K(x^1,x^2,\cdots,x^K):=H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{x^l}\Big), \tag{17}$$

and then $H^K(x^1,x^2,\cdots,x^K)$ is a function on $\mathbb{R}^{d\times K}$. Moreover, by [3, Proposition 3.1, Page 17], it holds that H^K is C^2 on $\mathbb{R}^{d\times K}$ and

$$\partial_{x^i}H^K(x^1,x^2,\cdots,x^K) \;=\; \frac{1}{K}\partial_\mu H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{x^l}\Big)(x^i),$$

$$\partial^2_{x^ix^j}H^K(x^1,x^2,\cdots,x^K) = \frac{1}{K}\partial_y\partial_\mu H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{x^l}\Big)(x^i)\delta_{i,j}+\frac{1}{K^2}\partial^2_\mu H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{x^l}\Big)(x^i,x^j). \tag{18}$$

Besides, we take K independent copies $X^l_t, l=1,2,\cdots,K$ of X_t. That is,

$$\mathrm{d}X^l_t=b(t,X^l_t,\mathcal{L}_{X^l_t})\mathrm{d}t+\sigma(t,X^l_t,\mathcal{L}_{X^l_t})\mathrm{d}B^l_t+\int_{\mathbb{U}_0}f(t,X^l_{t-},\mathcal{L}_{X^l_t},u)\tilde{N}^l(\mathrm{d}t,\mathrm{d}u), \quad l=1,2,\cdots,K$$

where $B^l,N^l,l=1,2,\cdots,K$ are mutually independent and have the same distribution to that of B,N, respectively. And then applying the Itô formula to $H^K(X^1_t,X^2_t,\cdots,X^K_t)$ and taking the expectation on two sides, we obtain that for $0\leqslant t<t+h\leqslant T$

$$\begin{aligned}
\mathbb{E}H^K(X^1_{t+h},X^2_{t+h},\cdots,X^K_{t+h}) &= \mathbb{E}H^K(X^1_t,X^2_t,\cdots,X^K_t)\\
&\quad+\int_t^{t+h}\mathbb{E}\partial_{x^i}H^K(X^1_s,X^2_s,\cdots,X^K_s)b(s,X^i_s,\mathcal{L}_{X^i_s})\mathrm{d}s\\
&\quad+\frac{1}{2}\int_t^{t+h}\mathbb{E}\partial^2_{x^ix^i}H^K(X^1_s,X^2_s,\cdots,X^K_s)\sigma\sigma^*(s,X^i_s,\mathcal{L}_{X^i_s})\mathrm{d}s\\
&\quad+\int_t^{t+h}\int_{\mathbb{U}_0}\mathbb{E}\Big[H^K\Big(X^1_s+f(s,X^1_s,\mathcal{L}_{X^1_s},u),X^2_s,\cdots,X^K_s\Big)\\
&\qquad -H^K(X^1_s,X^2_s,\cdots,X^K_s)\\
&\qquad -\partial_{x^1}H^K(X^1_s,X^2_s,\cdots,X^K_s)f(s,X^1_s,\mathcal{L}_{X^1_s},u)\Big]\nu(\mathrm{d}u)\mathrm{d}s\\
&\quad+\cdots\\
&\quad+\int_t^{t+h}\int_{\mathbb{U}_0}\mathbb{E}\Big[H^K\Big(X^1_s,X^2_s,\cdots,X^K_s+f(s,X^K_s,\mathcal{L}_{X^K_s},u)\Big)\\
&\qquad -H^K(X^1_s,X^2_s,\cdots,X^K_s)\\
&\qquad -\partial_{x^K}H^K(X^1_s,X^2_s,\cdots,X^K_s)f(s,X^K_s,\mathcal{L}_{X^K_s},u)\Big]\nu(\mathrm{d}u)\mathrm{d}s\\
&= \mathbb{E}H^K(X^1_t,X^2_t,\cdots,X^K_t)\\
&\quad+K\int_t^{t+h}\mathbb{E}\partial_{x^1}H^K(X^1_s,X^2_s,\cdots,X^K_s)b(s,X^1_s,\mathcal{L}_{X^1_s})\mathrm{d}s\\
&\quad+\frac{K}{2}\int_t^{t+h}\mathbb{E}\partial^2_{x^1x^1}H^K(X^1_s,X^2_s,\cdots,X^K_s)\sigma\sigma^*(s,X^1_s,\mathcal{L}_{X^1_s})\mathrm{d}s\\
&\quad+K\int_t^{t+h}\int_{\mathbb{U}_0}\int_0^1\mathbb{E}\Big[\Big(\partial_{x^1}H^K(X^1_s+\eta f(s,X^1_s,\mathcal{L}_{X^1_s},u),X^2_s,\cdots,X^K_s)\\
&\qquad -\partial_{x^1}H^K(X^1_s,X^2_s,\cdots,X^K_s)\Big)f(s,X^1_s,\mathcal{L}_{X^1_s},u)\Big]\mathrm{d}\eta\nu(\mathrm{d}u)\mathrm{d}s,
\end{aligned}$$

where the convention that the repeated indices stand for the summation is used in the first equality, and the property of the same distributions for $X^l_t, l=1,2,\cdots,K$ is used in the second equality. Inserting (17) (18) in the above equality, we get that

$$\mathbb{E}H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{X^l_{t+h}}\Big) = \mathbb{E}H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{X^l_t}\Big)+\int_t^{t+h}\mathbb{E}\partial_\mu H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{X^l_s}\Big)(X^1_s)b(s,X^1_s,\mathcal{L}_{X^1_s})\mathrm{d}s$$

$$
\begin{aligned}
&+\frac{1}{2}\int_t^{t+h}\mathbb{E}\partial_y\partial_\mu H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{X_s^l}\Big)(X_s^1)\sigma\sigma^*(s,X_s^1,\mathcal{L}_{X_s^1})\mathrm{d}s\\
&+\frac{1}{2K}\int_t^{t+h}\mathbb{E}\partial_\mu^2 H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{X_s^l}\Big)(X_s^1,X_s^1)\sigma\sigma^*(s,X_s^1,\mathcal{L}_{X_s^1})\mathrm{d}s\\
&+\int_t^{t+h}\int_{\mathbb{U}_0}\int_0^1\mathbb{E}\Big[\Big(\partial_\mu H\Big(\frac{1}{K}\delta_{X_s^1+\eta f(s,X_s^1,\mathcal{L}_{X_s^1},u)}+\frac{1}{K}\sum_{l=2}^K\delta_{X_s^l}\Big)\\
&\qquad\qquad\circ(X_s^1+\eta f(s,X_s^1,\mathcal{L}_{X_s^1},u))\\
&\qquad\quad-\partial_\mu H\Big(\frac{1}{K}\sum_{l=1}^K\delta_{X_s^l}\Big)(X_s^1)\Big)f(s,X_s^1,\mathcal{L}_{X_s^1},u)\Big]\mathrm{d}\eta\nu(\mathrm{d}u)\mathrm{d}s.
\end{aligned}
$$

Next, we take the limit on two sides of the above equality. Note that

$$
\lim_{K\to\infty}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\rho^2\left(\frac{1}{K}\sum_{l=1}^K\delta_{X_t^l},\mu_t\right)\right]=0,
$$

which comes from [6, Section 5]. And then as $K\to\infty$, by continuity and boundedness of $H,\partial_\mu H,\partial_y\partial_\mu H$, and boundedness of $\partial_\mu^2 H,b,\sigma$, it follows from the dominated convergence theorem that

$$
\begin{aligned}
H(\mu_{t+h}) &= H(\mu_t)+\int_t^{t+h}\mathbb{E}\partial_\mu H\Big(\mu_s\Big)(X_s^1)b(s,X_s^1,\mathcal{L}_{X_s^1})\mathrm{d}s\\
&\quad+\frac{1}{2}\int_t^{t+h}\mathbb{E}\partial_y\partial_\mu H\Big(\mu_s\Big)(X_s^1)\sigma\sigma^*(s,X_s^1,\mathcal{L}_{X_s^1})\mathrm{d}s\\
&\quad+\int_t^{t+h}\int_{\mathbb{U}_0}\int_0^1\mathbb{E}\Big[\Big(\partial_\mu H(\mu_s)\Big(X_s^1+\eta f(s,X_s^1,\mathcal{L}_{X_s^1},u)\Big)\\
&\qquad\qquad-\partial_\mu H(\mu_s)(X_s^1)\Big)f(s,X_s^1,\mathcal{L}_{X_s^1},u)\Big]\mathrm{d}\eta\nu(\mathrm{d}u)\mathrm{d}s.
\end{aligned}
$$

Thus, by simple calculus we obtain that

$$
\begin{aligned}
\partial_t H(\mu_t) &= \int_{\mathbb{R}^d}\langle b(t,y,\mu_t),\partial_\mu H(\mu_t)(y)\rangle\mu_t(\mathrm{d}y)+\frac{1}{2}\int_{\mathbb{R}^d}\mathrm{tr}\Big((\sigma^*\sigma)(t,y,\mu_t)\partial_y\partial_\mu H(\mu_t)(y)\Big)\mu_t(\mathrm{d}y)\\
&\quad+\int_{\mathbb{U}_0}\int_0^1\int_{\mathbb{R}^d}\Big[\Big(\partial_\mu H(\mu_t)\Big(y+\eta f(t,y,\mu_t,u)\Big)-\partial_\mu H(\mu_t)(y)\Big)f(t,y,\mu_t,u)\Big]\mu_t(\mathrm{d}y)\mathrm{d}\eta\nu(\mathrm{d}u).
\end{aligned}
$$

Step 2. We prove (5).

By **Step 1.**, we know that $\bar{h}(t,x)$ is C^1 in t and C^2 in x. Therefore the classical Itô formula admits us to obtain that

$$
\begin{aligned}
\mathrm{d}h(t,X_t,\mathcal{L}_{X_t}) &= \mathrm{d}\bar{h}(t,X_t)\\
&= \partial_t\bar{h}(t,X_t)\mathrm{d}t+\langle\partial_x\bar{h}(t,X_t),b(t,X_t,\mu_t)\rangle\mathrm{d}t+\langle\partial_x\bar{h}(t,X_t),\sigma(t,X_t,\mu_t)\mathrm{d}B_t\rangle\\
&\quad+\int_{\mathbb{U}_0}\Big[\bar{h}\Big(t,X_t+f(t,X_t,\mu_t,u)\Big)-\bar{h}(t,X_t)\Big]\tilde{N}(\mathrm{d}t,\mathrm{d}u)\\
&\quad+\frac{1}{2}\mathrm{tr}\Big(\sigma\sigma^*(t,X_t,\mu_t)\partial_x^2\bar{h}(t,X_t)\Big)\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{\mathbb{U}_0}\bigg[\bar{h}\Big(t, X_t + f(t, X_t, \mu_t, u)\Big) - \bar{h}(t, X_t) - \langle f(t, X_t, \mu_t, u), \partial_x \bar{h}(t, X_t)\rangle\bigg]\nu(\mathrm{d}u)\mathrm{d}t \\
= \ & \partial_t h(t, X_t, \mu_t)\mathrm{d}t + \partial_t h(s, x, \mu_t)|_{s=t, x=X_t}\mathrm{d}t \\
&+\langle \partial_x h(t, X_t, \mu_t), b(t, X_t, \mu_t)\rangle \mathrm{d}t + \langle \partial_x h(t, X_t, \mu_t), \sigma(t, X_t, \mu_t)\mathrm{d}B_t\rangle \\
&+\int_{\mathbb{U}_0}\bigg[h\Big(t, X_t + f(t, X_t, \mu_t, u), \mu_t\Big) - h(t, X_t, \mu_t)\bigg]\tilde{N}(\mathrm{d}t, \mathrm{d}u) \\
&+\frac{1}{2}\mathrm{tr}\Big(\sigma\sigma^*(t, X_t, \mu_t)\partial_x^2 h(t, X_t, \mu_t)\Big)\mathrm{d}t \\
&+\int_{\mathbb{U}_0}\bigg[h\Big(t, X_t + f(t, X_t, \mu_t, u), \mu_t\Big) - h(t, X_t, \mu_t) \\
&\qquad\qquad -\langle f(t, X_t, \mu_t, u), \partial_x h(t, X_t, \mu_t)\rangle\bigg]\nu(\mathrm{d}u)\mathrm{d}t \\
= \ & \partial_t h(t, X_t, \mu_t)\mathrm{d}t + \int_{\mathbb{R}^d}\langle b(t, y, \mu_t), \partial_\mu h(t, X_t, \mu_t)(y)\rangle \mu_t(\mathrm{d}y)\mathrm{d}t \\
&+\frac{1}{2}\int_{\mathbb{R}^d}\mathrm{tr}\Big((\sigma^*\sigma)(t, y, \mu_t)\partial_y\partial_\mu h(t, X_t, \mu_t)(y)\Big)\mu_t(\mathrm{d}y)\mathrm{d}t \\
&+\int_{\mathbb{U}_0}\int_0^1\int_{\mathbb{R}^d}\bigg\langle \partial_\mu h(t, X_t, \mu_t)\Big(y + \eta f(t, y, \mu_t, u)\Big) \\
&\qquad\qquad -\partial_\mu h(t, X_t, \mu_t)(y), f(t, y, \mu_t, u)\bigg\rangle \mu_t(\mathrm{d}y)\mathrm{d}\eta\nu(\mathrm{d}u)\mathrm{d}t \\
&+\langle \partial_x h(t, X_t, \mu_t), b(t, X_t, \mu_t)\rangle \mathrm{d}t + \langle \partial_x h(t, X_t, \mu_t), \sigma(t, X_t, \mu_t)\mathrm{d}B_t\rangle \\
&+\int_{\mathbb{U}_0}\bigg[h\Big(t, X_t + f(t, X_t, \mu_t, u), \mu_t\Big) - h(t, X_t, \mu_t)\bigg]\tilde{N}(\mathrm{d}t, \mathrm{d}u) \\
&+\frac{1}{2}\mathrm{tr}\Big(\sigma\sigma^*(t, X_t, \mu_t)\partial_x^2 h(t, X_t, \mu_t)\Big)\mathrm{d}t \\
&+\int_{\mathbb{U}_0}\bigg[h\Big(t, X_t + f(t, X_t, \mu_t, u), \mu_t\Big) - h(t, X_t, \mu_t) \\
&\qquad\qquad -\langle f(t, X_t, \mu_t, u), \partial_x h(t, X_t, \mu_t)\rangle\bigg]\nu(\mathrm{d}u)\mathrm{d}t.
\end{aligned}
$$

This is just right (5). The proof is completed.

Acknowledgements:

The authors are very grateful to Professor Xicheng Zhang for valuable discussions. The first author also thanks Professor Renming Song for providing him an excellent environment to work in the University of Illinois at Urbana-Champaign.

References

[1] R. Buckdahn, J. Li, S. Peng, C. Rainer: Mean-field stochastic differential equations and associated PDEs. *Ann Probab,* 45(2) (2017), 824-878.

[2] P. Cardaliaguet: *Notes on mean field games (from P.L. Lion's lectures at College de France).* https://www.ceremade.dauphine.fr/cardalia/MFG100629.pdf, 2013.

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

[4] C. Dellacherie and P. A. Meyer: *Probabilities and Potential B: Theory of Martingales.* North-Holland, Amsterdam/New York/Oxford, 1982.

[5] T. Hao and J. Li: Mean-field SDEs with jumps and nonlocal integral-PDEs, *Nonlinear Differential Equations Appl.* 23 (2) (2016), 1-51.

[6] J. Horowitz and R. L. Karandikar: Mean rates of convergence of empirical measures in the Wasserstein metric, *Journal of Computational and Applied Mathematics.* 55(1994)261-273.

[7] N. Ikeda and S. Watanabe: *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., North-Holland/Kodanska, Amsterdam/Tokyo, 1989.

[8] J. Jacod and A.N. Shiryaev: *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin, 1987.

[9] J. Li, Mean-field forward and backward SDEs with jumps and associated nonlocal quasilinear integral-PDEs, *Stoch. Proc. Appl.* 128 (2018)3118-3180.

[10] L. Lin, F. Xu, Q. Zhang: The link between stochastic differential equations with non-Markovian coefficients and backward stochastic partial differential equations. Preprint, 2019.

[11] Jr. H.P. McKean: Propagation of chaos for a class of non-linear parabolic equations. Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967).

[12] Jr. H.P. McKean: A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA* 56 (1966), 1907-1911.

[13] K. R. Parthasarathy: *Probability measures on metric spaces.* AMS Chelsea Publishing, 2005.

[14] H. Qiao and J.-L. Wu: Characterising the path-independence of the Girsanov transformation for non-Lipschnitz SDEs with jumps, *Statistics and Probability Letters*, 119(2016)326-333.

[15] H. Qiao and J.-L. Wu: On the path-independence of the Girsanov transformation for stochastic evolution equations with jumps in Hilbert spaces, Discrete and Continuous Dynamical Systems-B, 24(2019)1449-1467.

[16] P. Ren and F.-Y. Wang: Space-Distribution PDEs for path independent additive functionals of McKean-Vlasov SDEs, to appear in *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*

[17] P. Ren and F. Yang: Path independence of additive functionals for stochastic differential equations under G-framework, *Front. Math. China* 14 (2019), no. 1, 135-148.

[18] A. Truman, F. -Y. Wang, J.-L. Wu, W. Yang: A link of stochastic differential equations to nonlinear parabolic equations, *SCIENCE CHINA Mathematics*, 55 (2012) 1971-1976.

[19] J.-L. Wu and 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.