ITÔ'S FORMULA FOR FLOWS OF CONDITIONAL MEASURES ON SEMIMARTINGALES

XIN GUO AND JIACHENG ZHANG

Abstract. Motivated by recent development of mean-field systems with common noise, this paper establishes Itô's formula for flows of conditional probability measures under a common filtration associated with general semimartingales. This generalizes existing works on flows of conditional measures on Itô processes and flows of deterministic measure on general semimartingales. The key technical components involve constructing conditional independent copies and establishing the equivalence between stochastic integrals with respect to the conditional law of semimartingales and the conditional expectation of stochastic integrals with respect to copies of semimartingales. Itô's formula is then established for cylindrical functions through conditional independent copies, and extended to the general case through function approximations.

1. Introduction

Itô's formula. The classical Itô's formula for semimartingales is generally viewed as the stochastic counterpart of the chain rule in calculus. It stands as one of the key cornerstones of stochastic analysis [29], establishing an intrinsic connection between partial differential equations (PDEs) and diffusion processes. This formula plays a critical role in the theory of stochastic control.

With the recent development of theory of mean-field games and mean-field controls [6, 8, 11], Itô's formula has been extended to the flows of probability measures. Initially applied to diffusions [6], this extension was further developed for (discontinuous) semimartingales [24]. The generalized Itô's formula has been instrumental for deriving the master equations in mean-field games and the Bellman dynamic programming equation for McKean–Vlasov control problems, as discussed in [11] and the references therein.

One of the most recent advances in mean-field theory have been concerning mean-field controls and games with common noise (see for instance [1, 2, 3, 5, 9, 10, 13, 15, 18, 23, 27, 30, 31, 32, 33, 35, 40]) or with stochastic processes involving jumps ([7, 22, 25, 26, 28, 36]). In either case, Itô's formula has been developed respectively in [11] and in [24]. In order to analyze rigorously the McKean–Vlasov dynamics with *both semimartingales and idiosyncratic and common noises*, it is necessary to establish the Itô's formula for flows of *conditional laws* with semimartingales, which does not seem to exist to the best of our knowledge. This is the primary focus of this paper.

Our work. This paper establishes Itô's formula (Theorem 2.7) for flow of conditional laws $\mu_t = \mathrm{Law}(X_t|\mathcal{G}_t)$ driven by (possible discontinuous) semimartingales $\{X_t\}_{t\in[0,T]}$ with a general form of common noise represented by a sub filtration $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]}$ and $\mathcal{G} = \mathcal{G}_T$.

The primary task to analyze the Itô integral $\int_0^t \eta_s \mathrm{d}X_s^{\mathcal{G}}$ with $\{X_t^{\mathcal{G}} := \mathbb{E}[X_t|\mathcal{G}]\}_{t\in[0,T]}$ over some appropriate adapted process $\{\eta_t\}_{t\in[0,T]}$. This is crucial as the integral of the conditional expectation such as $\int_0^t \eta_s \mathrm{d}X_s^{\mathcal{G}}$ is not necessarily equivalent to the conditional expectation of the integral of the form $\mathbb{E}\big[\int_0^t \eta_s \mathrm{d}X_s|\mathcal{G}\big]$ (See discussions and examples in Section 3). Our analysis is built on several key technical components: constructing conditional independent copies of general stochastic processes (Theorem 3.1), establishing the equivalence between the conditional expectation of stochastic integrals with respect to seminmartingales and stochastic integrals with respect to the conditional law of semimartingales, and building the equivalence between the quadratic variation of the conditional expectation of a semimartingale and the conditional expectation of the quadratic variation of its two conditional independent copies (Theorem 3.3).

With these key technical components, Itô's formula for flows of conditional laws on semimartingales is established (Theorem 2.7): it is first derived for cylindrical functions (i.e., smooth mean-field functions with integrable forms 4.1)), which are then shown to be dense in the desired function space; finally, localization argument along with proper form of dominating convergence theorem finishes the task.

Related works. Conditional independent copied has been treated informally and in certain forms in the existing literature, for instance, when defining a semimartingale by stochastic differential equations involving both idiosyncratic and common noise in [11]. In this paper, we have generalized the construction of conditional independent copies for a broad class of semimartingales and with an arbitrary sub σ-algebra. This concept may be of independent theoretical interest, besides its crucial role in our derivation of Itô's formula for flow of conditional laws on semimartingales.

There are several methods for deriving Itô's formula and its variants, including the time discretization approach in [6] and [36] for mean-field jump diffusions, the density approach in [8] using the Fokker–Planck equation, and the particle approximation approach [11, 14], [19, 38] to approximate flows of measures by flows of empirical measures. The cylindrical function approach has proved appropriate and powerful for analysis of general semimartingales. They are initially explored in Fleming–Viot processes [21], later adopted in the analysis of polynomial diffusions [17], and most recently used in Itô's formula for flow of measures with (discontinuous) semimartingales [24].

Meanwhile, linear derivatives on the space of probability measures are known to be appropriate for characterizing the infinitesimal changes in the functional of controlled McKean–Vlasov processes when jumps are added; this has been noted previously in [7] and further explored in [24] and [39].

Notation. Throughout the paper, we will adopt the following notations, unless otherwise specified.

- $\mathbb{N}^+$ denotes the the set of all possible natural numbers, and we fix $d \in \mathbb{N}^+$.
- $\mathcal{P}(\mathbb{R}^d)$ denotes the space of probability measures on $\mathbb{R}^d$ with the topology of weak convergence. $\mathcal{P}_p(\mathbb{R}^d)$ for $p \in [1, \infty)$ denotes the probability measures with finite p-th moment, equipped with the Wasserstein-p metric.
- $C^k(\mathbb{R}^d)$ is the set of all k-th differentiable functions on $\mathbb{R}^d$ with continuous derivatives up to the k-th order, and $C_b^k(\mathbb{R}^d) \subset C^k(\mathbb{R}^d)$ is the set of all k-th differentiable functions on $\mathbb{R}^d$ with bounded and continuous derivatives up to the k-th order. And we use the convention $C(\mathbb{R}^d) = C^0(\mathbb{R}^d)$ and $C_b(\mathbb{R}^d) = C_b^0(\mathbb{R}^d)$.
- For a random variable X, $\|X\|_{L^p}$ represents its p-th moment, i.e., $\|X\|_{L^p} = \mathbb{E}[|X|^p]^{\frac{1}{p}}$.
- For vectors $\boldsymbol{a} = (a_i)_{i=1,..,d}, \boldsymbol{b} = (b_i)_{i=1,..,d} \in \mathbb{R}^d$, we denote $\boldsymbol{a}\cdot\boldsymbol{b} = \sum_{i=1}^d a_i b_i$. For matrices $\boldsymbol{C} = (C_{ij})_{i,j=1,..,d}, \boldsymbol{D} = (D_{ij})_{i,j=1,..,d} \in \mathbb{R}^{d\times d}$, we denote $\boldsymbol{C} : \boldsymbol{D} = \sum_{i,j=1}^d C_{ij} D_{ij}$.
- For $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $\varphi : \mathbb{R}^d \to \mathbb{R}$ such that $\int_{\mathbb{R}^d} |\varphi(x)| \mathrm{d}\mu(x) < \infty$, we set

$$\langle \mu, \varphi \rangle := \int_{\mathbb{R}^d} \varphi(x) \mathrm{d}\mu(x).$$

2. Assumptions and main result

Given two $\mathbb{R}^d$-valued $\mathbb{F}$- semimartingales $\boldsymbol{X} = (\boldsymbol{X}_t)_{t\in[0,T]}$ and $\boldsymbol{Y} = (\boldsymbol{Y}_t)_{t\in[0,T]}$ for $T > 0$ and $d \in \mathbb{N}^+$ on a completed filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, and $\mu = \big(\mu_t := \mathrm{Law}(\boldsymbol{X}_t|\mathcal{G}_t)\big)_{t\in[0,T]}$ with a sub filtration $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]} \subset \mathbb{F}$. Our focus is to develop an Itô's formula for the functional $\Phi(\mu_t, \boldsymbol{Y}_t)$ with the function $\Phi : \mathcal{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$, under suitable regularity conditions.

2.1. Assumptions.

Throughout the paper, we will make the following assumptions on the semimartingales $\boldsymbol{X}$ and $\boldsymbol{Y}$, the filtrations $\mathbb{G}$ and $\mathbb{F}$, and the function Φ. These are standard assumptions to ensure that stochastic integrals are properly defined (see [37]) with the presence of common noise (see [34]) and for flow of measures on semimartingales (see [24, 16]).

Assumptions on semimartingales.

Assumption 2.1 (Semimartingale).

- We assume $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{H}^p$ for $1 \le p \le \infty$ [37, Section V.2], with

$$\begin{aligned} \big\|\boldsymbol{X}\big\|_{\mathcal{H}^p} &:= \inf_{\boldsymbol{M},\boldsymbol{V}} \left\| |\boldsymbol{X}_0| + \sqrt{[\boldsymbol{M},\boldsymbol{M}]_T} + \int_0^T \big|\mathrm{d}\boldsymbol{V}_s\big| \right\|_{L^p} < \infty, \\ \big\|\boldsymbol{Y}\big\|_{\mathcal{H}^p} &:= \inf_{\boldsymbol{N},\boldsymbol{U}} \left\| |\boldsymbol{Y}_0| + \sqrt{[\boldsymbol{N},\boldsymbol{N}]_T} + \int_0^T \big|\mathrm{d}\boldsymbol{U}_s\big| \right\|_{L^p} < \infty, \end{aligned} \tag{2.1}$$

  where the infimum is taken over all possible decompositions

$$\boldsymbol{X}_t = \boldsymbol{X}_0 + \boldsymbol{M}_t + \boldsymbol{V}_t \text{ and } \boldsymbol{Y}_t = \boldsymbol{Y}_0 + \boldsymbol{N}_t + \boldsymbol{U}_t.$$

  Here $\boldsymbol{V} = (\boldsymbol{V}_t)_{t\in[0,T]}$ and $\boldsymbol{U} = (\boldsymbol{U}_t)_{t\in[0,T]}$ are adapted càdlàg processes of finite variation with $\boldsymbol{V}_0 = \boldsymbol{U}_0 = \boldsymbol{0}$, and $\boldsymbol{M} = (\boldsymbol{M}_t)_{t\in[0,T]}$ and $\boldsymbol{N} = (\boldsymbol{N}_t)_{t\in[0,T]}$ are $\mathbb{F}$-local martingales such that $\boldsymbol{M}_0 = \boldsymbol{N}_0 = \boldsymbol{0}$.
- We assume also

$$\mathbb{E}\Big[\Big(\sum_{0\le s\le T} |\Delta\boldsymbol{X}_s|\Big)^p\Big] + \mathbb{E}\Big[\Big(\sum_{0\le s\le T} |\Delta\boldsymbol{Y}_s|\Big)^p\Big] < \infty. \tag{2.2}$$

Under these assumptions, a proper form of dominated convergence theorem has been established ([37, Page 273, Lemma]) and the following two propositions hold.

Proposition 2.2 ([37, Section V, Theorem 2]). *For any $1 \le p \le \infty$, consider a norm on the space of $\mathbb{R}$-valued adapted càdlàg processes (or the space of $\mathbb{R}$-valued adapted càglàd processes) defined by*

$$\|H\|_{\mathcal{S}^p} := \Big\| \sup_{0\le t\le T} |H_t| \Big\|_{L^p}.$$

Then, for $1 \le p < \infty$, there exists a constant c_p depending only on p such that

$$\|H\|_{\mathcal{S}^p} \le c_p \|H\|_{\mathcal{H}^p},$$

for any $\mathbb{R}$-valued semimartingale H.

Proposition 2.3 (Emery inequality, [37, Section V, Theorem 3]). *Let Z be an $\mathbb{R}$-valued semimartingale, H is an $\mathbb{R}$-valued adapted càglàd process, and $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ $(1 \le p, q \le \infty)$, then*

$$\Big\| \int_0^{\cdot} H_s \mathrm{d}Z_s \Big\|_{\mathcal{H}^r} \le \|H\|_{\mathcal{S}^p} \|Z\|_{\mathcal{H}^q}.$$

Remark 2.4. Propositions 2.2 and 2.3 also hold for $\mathbb{R}^d$-valued semimartingales with appropriate forms in $d \in \mathbb{N}^+$.

These propositions are important for the subsequent analysis, for instance in the localization argument for establishing Itô's lemma for the flow of conditional measure on semimartinagles, as will be clear in Section 4.2.

Assumptions on filtration.

Assumption 2.5 (Filtration)**.** Sub-filtration $\mathbb{G} \subset \mathbb{F}$ satisfies the compatibility assumption or the conditional independence condition. That is, for each $0 \le t \le T$, $\mathcal{F}_t$ and $\mathcal{G}_T$ are independent, given $\mathcal{G}_t$, written as $\mathcal{F}_t \perp\!\!\!\perp \mathcal{G}_T \big| \mathcal{G}_t$.

In the context of mean field games, this assumption ensures a weak closure of adapted processes, as first noted in [12, Lemma 3.11], and then stated more explicitly in [4, Theorem 5.4].

One can deduce from is assumption that $\mathrm{Law}(\boldsymbol{X}_t|\mathcal{G}_t) = \mathrm{Law}(\boldsymbol{X}_t|\mathcal{G}_T)$, for any $\mathbb{F}$-adapted process $\boldsymbol{X} = (\boldsymbol{X}_t)_{t\in[0,T]}$ almost surely for each $t \le T$, according to a special form of the **(H)** hypothesis by freezing $\mathcal{G}_u = \mathcal{G}_T$ when $u \ge T$ ([20]). This allows us to define a cádlág version of $\mu_t := \mathrm{Law}(\boldsymbol{X}_t|\mathcal{G}_t)$ by taking $\mathrm{Law}(\boldsymbol{X}_t|\mathcal{G}_T)$, which will be assumed throughout the paper.

Assumptions on function Φ. We will adopt the the linear derivative adapted from [16], which has shown to be appropriate for the analyzing the flow of measures on semimartingales [24, 16].

Definition 2.6 ($C^{2,2}(\mathcal{P}_p(\mathbb{R}^d)\times\mathbb{R}^d)$ functions)**.** For $p \ge 2$, a function $\Phi(\mu, \boldsymbol{y})$ is a $C^{2,2}(\mathcal{P}_p(\mathbb{R}^d)\times\mathbb{R}^d)$ if for any $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ and $\boldsymbol{y} \in \mathbb{R}^d$, there exists a continuous mapping $(\mu, \boldsymbol{y}, \boldsymbol{x}_1) \to \frac{\delta\Phi}{\delta\mu}(\mu, \boldsymbol{y}, \boldsymbol{x}_1)$, and there exists a continuous mapping $(\mu, \boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2) \to \frac{\delta^2\Phi}{(\delta\mu)^2}(\mu, \boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2)$ that is symmetric in its two arguments and with the following properties,

- continuously differentiable: $\nabla_{\boldsymbol{y}}\Phi(\mu,\boldsymbol{y})$, $\nabla^2_{\boldsymbol{y}}\Phi(\mu,\boldsymbol{y})$, $\frac{\delta\Phi}{\delta\mu}(\mu,\boldsymbol{y},\boldsymbol{x}_1)$, $\nabla_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}(\mu,\boldsymbol{y},\boldsymbol{x}_1)$, $\nabla_{\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}(\mu,\boldsymbol{y},\boldsymbol{x}_1)$ $\nabla^2_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}(\mu,\boldsymbol{y},\boldsymbol{x}_1)$, $\nabla_{\boldsymbol{x}_1}\nabla_{\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}(\mu,\boldsymbol{y},\boldsymbol{x}_1)$, $\nabla_{\boldsymbol{x}_1}\nabla_{\boldsymbol{x}_2}\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu,\boldsymbol{y},\boldsymbol{x}_1,\boldsymbol{x}_2)$ all exist and are continuous functions for all $\boldsymbol{y},\boldsymbol{x},\boldsymbol{x}' \in \mathbb{R}^d$, $\mu \in \mathcal{P}_p(\mathbb{R}^d)$,
- uniform polynomial-growth : there exists a constant $c > 0$ such that for all $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y} \in \mathbb{R}^d$, $\mu \in \mathcal{P}(\mathbb{R}^d)$, we have

$$\Big|\nabla_{\boldsymbol{y}}\Phi\Big| \le c\big(1+|\boldsymbol{y}|^{p-1}\big), \quad \Big|\nabla^2_{\boldsymbol{y}}\Phi\Big| \le c\big(1+|\boldsymbol{y}|^{p-2}\big),$$
$$\left|\nabla_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}\right| + \left|\nabla_{\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}\right| \le c\big(1+|\boldsymbol{x}_1|^{p-1}+|\boldsymbol{y}|^{p-1}\big),$$
$$\left|\nabla^2_{\boldsymbol{x}}\frac{\delta\Phi}{\delta\mu}\right| + \left|\nabla_{\boldsymbol{x}_1\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}\right| \le c\big(1+|\boldsymbol{x}_1|^{p-2}+|\boldsymbol{y}|^{p-2}\big),$$
$$\left|\nabla_{\boldsymbol{x}_1,\boldsymbol{x}_2}\frac{\delta^2\Phi}{(\delta\mu)^2}\right| \le c\big(1+|\boldsymbol{x}_1|^{p-2}+|\boldsymbol{x}_2|^{p-2}+|\boldsymbol{y}|^{p-2}\big),$$

- fundamental theorem of calculus: for everything $\mu,\nu \in \mathcal{P}_2(\mathbb{R}^d)$, we have

$$\Phi(\mu,\boldsymbol{y}) - \Phi(\nu,\boldsymbol{y}) = \int_0^1\int_{\mathbb{R}^d}\frac{\delta\Phi}{\delta\mu}\big(\lambda\mu+(1-\lambda)\nu,\boldsymbol{y},\boldsymbol{x}_1\big)(\mu-\nu)(\mathrm{d}\boldsymbol{x}_1)\mathrm{d}\lambda,$$
$$\begin{aligned}\Phi(\boldsymbol{x}_0,\mu) - \Phi(\boldsymbol{x}_0,\nu) = &\int_0^1\int_{\mathbb{R}^d}\frac{\delta\Phi}{\delta\mu}\big(\nu,\boldsymbol{y},\boldsymbol{x}_1\big)(\mu-\nu)(\mathrm{d}\boldsymbol{x}_1)\mathrm{d}\lambda \\ &+\int_0^1\int_0^t\int_{\mathbb{R}^d}\frac{\delta^2\Phi}{(\delta\mu)^2}\big(s\mu+(1-s)\nu,\boldsymbol{y},\boldsymbol{x}_1,\boldsymbol{x}_2\big)(\mu-\nu)(\mathrm{d}\boldsymbol{x}_1)(\mu-\nu)(\mathrm{d}\boldsymbol{x}_2)\mathrm{d}s\mathrm{d}t.\end{aligned}$$

Such $\frac{\delta\Phi}{\delta\mu}$ and $\frac{\delta^2\Phi}{(\delta\mu)^2}$ are called (a version of) the *linear derivative* and the *second derivative* of Φ.

2.2. **Main result.**

Theorem 2.7 (Itô's formula for flows of conditional measure on semimartingales)**.** *Given a completed filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, which supports two $\mathbb{R}^d$-valued $\mathbb{F}$-semimartingales $\boldsymbol{X} = (\boldsymbol{X}_t)_{t\in[0,T]}$ and $\boldsymbol{Y} = (\boldsymbol{Y}_t)_{t\in[0,T]}$ satisfying Assumption 2.1. Let μ_t =Law$(X_t|\mathcal{G}_t)$, with the subfiltration $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]} \subset \mathbb{F}$ satisfying Assumption 2.5. Then, for $p \geq 2$ and any $\Phi \in C^{2,2}(\mathcal{P}_p(\mathbb{R}^d) \times \mathbb{R}^d)$, we have*

$$
\begin{aligned}
&\Phi(\mu_t, \boldsymbol{Y}_t) - \Phi(\mu_0, \boldsymbol{Y}_0)\\
&= \overline{\mathbb{E}}\Bigg[\int_{0+}^t \Bigg(\nabla_{\boldsymbol{x}_1} \frac{\delta \Phi}{\delta \mu}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}) \cdot \mathrm{d}(\boldsymbol{X}')^c_s + \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1} \frac{\delta \Phi}{\delta \mu}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}']^c_s \\
&\qquad + \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_2} \frac{\delta^2 \Phi}{(\delta \mu)^2}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}, \boldsymbol{X}''_{s-}) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}'']^c_t + \nabla^2_{\boldsymbol{x}_1\boldsymbol{y}} \frac{\delta \Phi}{\delta \mu}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{Y}]^c_s \Bigg)\\
&\quad + \sum_{0<s\leq t} \Bigg(\frac{1}{2}\Bigg(\frac{\delta^2 \Phi}{(\delta \mu)^2}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_s, \boldsymbol{X}''_s) - \frac{\delta^2 \Phi}{(\delta \mu)^2}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}, \boldsymbol{X}''_s)\\
&\qquad\qquad - \frac{\delta^2 \Phi}{(\delta \mu)^2}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_s, \boldsymbol{X}''_{s-}) + \frac{\delta^2 \Phi}{(\delta \mu)^2}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}, \boldsymbol{X}''_{s-}) \Bigg)\\
&\qquad\qquad + \Bigg(\nabla_{\boldsymbol{y}} \frac{\delta \Phi}{\delta \mu}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_s) - \nabla_{\boldsymbol{y}} \frac{\delta \Phi}{\delta \mu}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}) \Bigg) \cdot \Delta \boldsymbol{Y}_s\\
&\qquad\qquad + \frac{\delta \Phi}{\delta \mu}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_s) - \frac{\delta \Phi}{\delta \mu}(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}) \Bigg) \mathbb{1}_{\{\mu_s = \mu_{s-}\}} \Bigg| \mathcal{F} \Bigg]\\
&\quad + \int_{0+}^t \nabla_{\boldsymbol{y}} \Phi(\mu_{s-}, \boldsymbol{Y}_{s-}) \cdot \mathrm{d}(\boldsymbol{Y})^c_s + \frac{1}{2} \int_{0+}^t \nabla^2_{\boldsymbol{y}\boldsymbol{y}} \Phi(\mu_{s-}, \boldsymbol{Y}_{s-}) : \mathrm{d}[\boldsymbol{Y}, \boldsymbol{Y}]^c_s\\
&\quad + \sum_{0<s\leq t} \Big(\Phi(\mu_s, \boldsymbol{Y}_s) - \Phi(\mu_{s-}, \boldsymbol{Y}_{s-}) \Big),
\end{aligned}
\tag{2.3}
$$

for all $t \in [0, T]$, where $\boldsymbol{X}', \boldsymbol{X}''$ are the conditional independent copies of $\boldsymbol{X}$ given the sub σ-algebra $\mathcal{G}_T$ defined in an enlarged probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ and $\overline{\mathbb{E}}[\cdot|\mathcal{F}]$ is the conditional expectation given the extension $\mathcal{F} \subset \overline{\mathcal{F}}$ in the enlarged probability space.

Remark 2.8. To ensure that the enlarged space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{F}}, \overline{\mathbb{P}})$ is properly defined such that $\mu_t, \boldsymbol{X}_t, \boldsymbol{Y}_t$ and $\mathcal{F}$ are naturally extended in this enlarged space, one needs the notion of independent copies of stochastic processes, as will be developed in the next section and discussed after Corollary 3.2.

3. Conditional independent copy

3.1. **Why conditional independent copy?** One of the key components in Theorem 2.7 is the concept of the conditional independent copy of stochastic processes. Here in this section, we illustrate through several examples its properties in the context of flow of conditional laws for semimartingales and its role in the derivation of Itô's formula.

Let us first recall a mean-field game with common noise where X_t is of the form

$$\mathrm{d}X_t = b_t \mathrm{d}t + \sigma_t \mathrm{d}W_t + \sigma^0_t \mathrm{d}W^0_t,$$

where W, W^0 are two independent Brownian motions and $\mathbb{G} = \{\mathcal{G}_t\}_{t\in[0,T]}$ is the sub filtration generated by W^0 and $\mathcal{G} = \mathcal{G}_T = \sigma(W^0_u, u \leq T)$. By the classical Itô's formula, the dynamic of

$g(X_t)$ is given by

$$\mathrm{d}g(X_t) = b_t g'(X_t) + \frac{1}{2}(\sigma_t^2 + (\sigma_t^0)^2)g''(X_t)\mathrm{d}t + \sigma_t g'(X_t)\mathrm{d}W_t + \sigma_t^0 g'(X_t)\mathrm{d}W_t^0,$$

and consequently the dynamic of $Z_t = \mathbb{E}[g(X_t)|\mathcal{G}]$ is given by

$$\mathrm{d}Z_t = \mathbb{E}\left[b_t g'(X_t) + \frac{1}{2}(\sigma_t^2 + (\sigma_t^0)^2)g''(X_t)\Big|\mathcal{G}\right]\mathrm{d}t + \mathbb{E}\big[\sigma_t^0 g'(X_t)|\mathcal{G}\big]\mathrm{d}W_t^0. \tag{3.1}$$

The above equation follows essentially from the fact that

$$\mathbb{E}\left[\int_0^t \eta_s \mathrm{d}W_s\Big|\mathcal{G}\right] = 0, \qquad \mathbb{E}\left[\int_0^t \eta_s \mathrm{d}W_s^0\Big|\mathcal{G}\right] = \int_0^t \mathbb{E}[\eta_s|\mathcal{G}]\mathrm{d}W_s^0, \text{ for all } t \in [0,T] \tag{3.2}$$

see [34, Lemma B.1] for details. Indeed, from (3.1) one can calculate the quadratic variation of Z_t to derive the Itô's formula for $f(Z_t)$:

$$\mathrm{d}\langle Z\rangle_t = \mathbb{E}\big[\sigma_t^0 g'(X_t)|\mathcal{G}\big]^2\mathrm{d}t = \mathbb{E}[\sigma_t^0 g'(X_t)\widetilde{\sigma}_t^0 g'(\widetilde{X}_t)|\mathcal{G}]\mathrm{d}t,$$

where $\widetilde{\sigma}_t^0, \widetilde{X}_t$ are conditional respective independent copy of σ_t^0 and X_t given $\mathcal{G}$, which is formally introduced in [11]. Under these expressions, one can derive Itô's formula for $f(Z_t)$, which depends on the precise expression of

$$\begin{aligned}\int_0^t f'(Z_s)\mathrm{d}Z_s &= \int_0^t f'(Y_s)\Big(\mathbb{E}\Big[b_s g'(X_s) + \frac{1}{2}(\sigma_s^2 + (\sigma_s^0)^2)g''(X_s)\Big|\mathcal{G}\Big]\mathrm{d}s + \mathbb{E}\big[\sigma_t^0 g'(X_s)|\mathcal{G}\big]\mathrm{d}W_s^0\Big)\\ &= \int_0^t \Big(\mathbb{E}\Big[f'(Z_s)\Big(b_s g'(X_s) + \frac{1}{2}(\sigma_s^2 + (\sigma_s^0)^2)g''(X_s)\Big)\Big|\mathcal{G}\Big]\mathrm{d}s + \mathbb{E}\big[f'(Z_s)\sigma_t^0 g'(X_s)|\mathcal{G}\big]\mathrm{d}W_s^0\Big)\end{aligned}$$

and

$$\begin{aligned}\int_0^t f''(Z_s)\mathrm{d}\langle Z\rangle_s &= \int_0^t f''(Z_s)\cdot\frac{1}{2}\mathbb{E}\big[\sigma_s^0 g'(X_s)\widetilde{\sigma}_s^0 g'(\widetilde{X}_s)\big|\mathcal{G}_T\big]\mathrm{d}s\\ &= \frac{1}{2}\int_0^t \mathbb{E}\big[f''(Z_s)\sigma_s^0 g'(X_s)\widetilde{\sigma}_s^0 g'(\widetilde{X}_s)\big|\mathcal{G}_T\big]\mathrm{d}s.\end{aligned}$$

Now, the issue arises when one considers a general continuous semimartingale X and a general sub σ-algebra $\mathcal{G}$, where the Itô's formula takes the form:

$$\mathrm{d}g(X_t) = g'(X_t)\mathrm{d}X_t + \frac{1}{2}g''(X_t)\mathrm{d}\langle X\rangle_t. \tag{3.3}$$

In this case, in order to characterize the dynamics of $Z_t = \mathbb{E}[g(X_t)|\mathcal{G}]$, one needs an appropriate expression for $\mathrm{d}Z_t$, $\mathrm{d}\langle Z\rangle_t$ and more importantly for

$$\int_0^t f'(Z_s)\mathrm{d}Z_s, \text{ and } \int_0^t f''(Z_s)\mathrm{d}\langle Z\rangle_s. \tag{3.4}$$

One naive guess of the extension of (3.2) to this general case is

$$\mathbb{E}\left[\int_0^t \eta_s \mathrm{d}X_s\Big|\mathcal{G}\right] = \int_0^t \mathbb{E}[\eta_s|\mathcal{G}]\mathrm{d}X_s^{\mathcal{G}}, \tag{3.5}$$

where $X_t^{\mathcal{G}} := \mathbb{E}[X_t|\mathcal{G}]$.

Unfortunately this does not hold in general. For instance, take X_t from (3.1) with $b_t = 0$. By (3.2), the left hand side of (3.5) becomes

$$\mathbb{E}\left[\int_0^t \eta_s \mathrm{d}X_s\Big|\mathcal{G}\right] = \mathbb{E}\left[\int_0^t \eta_s\big(\sigma_s\mathrm{d}W_s + \sigma_s^0\mathrm{d}W_s^0\big)\Big|\mathcal{G}\right] = \int_0^t \mathbb{E}\big[\eta_s\sigma_s^0\big|\mathcal{G}\big]\mathrm{d}W_s^0,$$

while the right hand side is

$$\int_0^t \mathbb{E}[\eta_s|\mathcal{G}]\mathrm{d}X_s^{\mathcal{G}} = \int_0^t \mathbb{E}[\eta_s|\mathcal{G}]\mathbb{E}[\sigma_s^0|\mathcal{G}]\mathrm{d}W_s^0,$$

and (3.5) holds if and only if η_s and σ_s^0 are conditionally uncorrelated of $\mathcal{G}$.

This is exactly why the concept of conditional independent copies of stochastic processes: instead of (3.2), let us consider the expression of the following form

$$\int_0^t f'(Z_s)\mathrm{d}Z_s = \int_0^t f'(Z_s)\mathrm{d}\mathbb{E}\big[g(X_s)|\mathcal{G}\big] = \int_0^t f'(Z_s)\mathrm{d}\mathbb{E}\big[g(X_s^{(1)})|\mathcal{F}\big] = \mathbb{E}\Big[\int_0^t f'(Z_s)\mathrm{d}g(X_s^{(1)})\Big|\mathcal{F}\Big]$$
$$= \mathbb{E}\Big[\int_0^t f'(Z_s)\Big(g'(X_s^{(1)})\mathrm{d}X_s^{(1)} + \frac{1}{2}g''(X_s^{(1)})\mathrm{d}\langle X^{(1)}\rangle_s\Big)\Big|\mathcal{F}\Big],$$

and

$$\int_0^t f''(Z_s)\mathrm{d}\langle Z\rangle_s = \int_0^t f'(Z_s)\mathrm{d}\langle\mathbb{E}\big[g(X_\cdot)|\mathcal{G}\big]\rangle_s = \mathbb{E}\Big[\int_0^t f'(Y_s)\mathrm{d}\langle g(X_\cdot^{(1)}), g(X_\cdot^{(2)})\rangle_s|\mathcal{F}\Big]$$
$$= \mathbb{E}\Big[\int_0^t f''(Z_s)g'(X_s^{(1)})g'(X_s^{(2)})\mathrm{d}\langle X^{(1)}, X^{(2)}\rangle_s\Big|\mathcal{F}\Big],$$

where $X^{(1)}$, $X^{(2)}$ are two conditional independent copies of X give $\mathcal{G}$. With these expressions, one can analyze (3.4) for a general semimartingale X, as will be detailed in (2c) and (2d) of Theorem 3.3.

3.2. **Examples.** To get some intuition for this notion of conditional independent copies of stochastic processes, let us see some examples when $\mathcal{F}$ and the sub σ-algebra $\mathcal{G}$ have some special structure.

Example 1. Let A, B are two independent random variables with $\mathcal{F} = \sigma(A, B)$ and $\mathcal{G} = \sigma(A)$. Then one can enlarge the probability space to include $A^{(i)}$ such that $A^{(i)}$ are independent and with the same distribution of A, and are independent of B. Then for any $X = f(A, B)$, the conditional independent copies of X are $X^{(i)} = f(A^{(i)}, B)$.

Example 2. Let X be the unique strong solution of the SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t + \sigma^{(0)}(X_t)\mathrm{d}W_t^{(0)},$$

where $W, W^{(0)}$ are two independent Brownian motions, and b, σ, and σ^0 satisfy appropriate continuity conditions. Let $\mathcal{F}_t = \sigma\big(\{W_s, W_s^{(0}\}_{s\le t}\big)$ and sub σ-algebra $\mathcal{G} = \sigma(W^{(0)})$, then one can enlarge the probability space to include independent Brownian motions $W^{(1)}$ and $X^{(1)}$, with $X^{(1)}$ the unique strong solution of

$$\mathrm{d}X_t^{(1)} = b(X_t^{(1)})\mathrm{d}t + \sigma(X_t^{(1)})\mathrm{d}W_t^{(1)} + \sigma^{(0)}(X_t^{(1)})\mathrm{d}W_t^{(0)},$$

and $X^{(1)}$ is the conditional independent copies of X given $\mathcal{G}$.

There are more examples of conditional independent copies of stochastic processes in the literature of mean-field games. For instance, when defining a semimartingale by stochastic differential equations involving both idiosyncratic and common noise, conditional copies are formally formulated in [11] as $\Omega = \Omega_0 \times \widetilde{\Omega}_1$, where the original probability space follows the structure $\Omega = \Omega_0 \times \Omega_1$, and $\widetilde{\Omega}_1$ represents the independent copy of the probability space Ω_1. [11, Theorem 4.14, Volume II]

In the next section, we will provide the construction of conditional independent copies for the general case of a semimartingale $\boldsymbol{X}$ and an *arbitrary* sub σ-algebra $\mathcal{G}$ which is not necessarily associated with any probability space. As emphasized earlier, this construction is necessary for establishing Itô's formula with the flow of conditional measures on semimartigales.

3.3. **Construction of conditional independent copies: general cases.** Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let us consider the extended probability space $\overline{\Omega} = \Omega^{n+1}$ defined by

$$\overline{\Omega} = \Omega^{n+1} = \{(\omega_0, \omega_1, \cdots, \omega_n) | \omega_0, \omega_1, \cdots, \omega_n \in \Omega\}, \tag{3.6}$$

and the extended σ-algebra $\overline{\mathcal{F}}$

$$\overline{\mathcal{F}} = \sigma\{A_0 \times A_1 \times \cdots \times A_n : A_0, A_1, \cdots, A_n \in \mathcal{F}\}. \tag{3.7}$$

To ensure that these copies are conditional independent, let us define the probability measure $\overline{\mathbb{P}}$ as follows: for $A_0, A_1, \cdots, A_n \in \mathcal{F}$, define

$$\overline{\mathbb{P}}\big(A_0 \times A_1 \times \cdots \times A_n\big) := \mathbb{E}\Big[\mathbb{1}_{A_0} \prod_{i=1}^{n} \mathbb{P}(A_i|\mathcal{G})\Big]. \tag{3.8}$$

This measure is properly defined since the set $\{A_0 \times A_1 \times \cdots \times A_n\}$ is a π-system. Finally, define $\widetilde{\mathcal{F}}, \widetilde{\mathcal{G}}$ for the extension of $\mathcal{F}$ and $\mathcal{G}$ into this enlarged probability space as

$$\widetilde{\mathcal{F}} := \{A \times \Omega^n : A \in \mathcal{F}\}, \quad \widetilde{\mathcal{G}} := \{A \times \Omega^n : A \in \mathcal{G}\}. \tag{3.9}$$

Then we have:

Theorem 3.1. *Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a sub σ-algebra $\mathcal{G} \subset \mathcal{F}$. Then there exists uniquely an enlarged probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \mathbb{P})$ satisfying* (3.6), (3.7) *and* (3.8)*; and for any random variable R in $(\Omega, \mathbb{F}, \mathbb{P})$, there exist an $\widetilde{R}$ and n conditional independent copies $\{R^{(i)}\}_{i=1,2,..n}$ by*

$$\widetilde{R}(\omega_0, \omega_1, \cdots, \omega_n) = R(\omega_0), \; R^{(i)}(\omega_0, \omega_1, \cdots, \omega_n) = R(\omega_i), \tag{3.10}$$

for $\overline{\omega} = (\omega_0, , \omega_1, \cdots, \omega_n) \in \overline{\Omega}$ and $\widetilde{R}$, $\{R^{(i)}\}_{i=1,2,..n}$ satisfying

- *$\widetilde{R}, \big\{R^{(i)}\big\}_{i=1,2,..n}$ are independent given $\widetilde{\mathcal{G}}$,*
- *$\widetilde{R}, \big\{R^{(i)}\big\}_{i=1,2,..n}$ are independent given $\widetilde{\mathcal{F}}$,*
- *$Law\big(R\big|\mathcal{G}\big)(\omega_0) = Law\big(\widetilde{R}\big|\widetilde{\mathcal{G}}\big)(\overline{\omega}) = Law\big(R^{(i)}\big|\widetilde{\mathcal{G}}\big)(\overline{\omega}) = Law\big(R^{(i)}\big|\widetilde{\mathcal{F}}\big)(\overline{\omega})$, for all $i = 1, 2, .., n$, a.s.*

with $\overline{\omega} = (\omega_0, \cdots, \omega_n) \in \overline{\Omega}$ and $\widetilde{\mathcal{F}}$, $\widetilde{\mathcal{G}}$ defined by (3.9).

Note that when $\mathcal{G}$ is a trivial σ-algebra $\{\emptyset, \Omega\}$, then the above theorem reduces to the standard construction of independent copies. Its proof is given in Section 5.1.

Now, in the context of semimartingales, there is a corresponding version of the conditional independent copy where the probability space is a filtered space and sub σ-algebra $\mathcal{G} = \mathcal{G}_T$ comes from the sub filtration $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]}$. Now define $\overline{\mathbb{F}} := (\overline{\mathcal{F}}_t)_{t\in(0,T)}$ by

$$\overline{\mathcal{F}}_t = \sigma\{A_0 \times A_1 \times \cdots \times A_n : A_0, A_1, \cdots, A_n \in \mathcal{F}_t\}, \tag{3.11}$$

and further define $\widetilde{\mathbb{F}} = (\widetilde{\mathcal{F}}_t)_{t\in(0,T)}$ and $\widetilde{\mathbb{G}} = (\widetilde{\mathcal{G}}_t)_{t\in(0,T)}$ respectively for the extension of $\mathbb{F}$ and $\mathbb{G}$ into this enlarged probability space as

$$\widetilde{\mathcal{F}}_t := \{A \times \Omega^n : A \in \mathcal{F}_t\}, \quad \widetilde{\mathcal{G}}_t := \{A \times \Omega^n : A \in \mathcal{G}_t\}. \tag{3.12}$$

With this construction, we have

Corollary 3.2. *Given two $\mathbb{R}^d$-valued $\mathbb{F}$-semimartingales $\boldsymbol{X} = (\boldsymbol{X}_t)_{t\in[0,T]}$ and $\boldsymbol{Y} = (\boldsymbol{X}_t)_{t\in[0,T]}$ on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, and a sub σ-algebra $\mathcal{G} \subset \mathcal{F}$, there exists uniquely an enlarged probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{F}} = (\mathcal{F}_t)_{t\in[0,T]}, \overline{\mathbb{P}})$ satisfying* (3.6), (3.7), (3.8) *and*

(3.11). *Moreover, one can define two conditional independent copies* $(\boldsymbol{X}', \boldsymbol{Y}')$, $(\boldsymbol{X}'', \boldsymbol{Y}'')$ *of the random processes* $(\boldsymbol{X}, \boldsymbol{Y})$ *by*

$$(\boldsymbol{X}', \boldsymbol{Y}')(\omega_0, \omega_1, \cdots, \omega_n) = (\boldsymbol{X}, \boldsymbol{Y})(\omega_1),\ (\boldsymbol{X}'', \boldsymbol{Y}'')(\omega_0, \omega_1, \cdots, \omega_n) = (\boldsymbol{X}, \boldsymbol{Y})(\omega_2),$$

and $(\boldsymbol{X}, \boldsymbol{Y})$, $(\boldsymbol{X}', \boldsymbol{Y}')$ *and* $(\boldsymbol{X}'', \boldsymbol{Y}'')$ *satisfying*

$$\begin{cases} Law\big(\boldsymbol{X}, \boldsymbol{Y}\big|\mathcal{G}\big) = Law\big(\boldsymbol{X}', \boldsymbol{Y}'\big|\mathcal{G}\big) = Law\big(\boldsymbol{X}', \boldsymbol{Y}'\big|\mathcal{F}\big) \\ \qquad = Law\big(\boldsymbol{X}'', \boldsymbol{Y}''\big|\mathcal{G}\big) = Law\big(\boldsymbol{X}'', \boldsymbol{Y}''\big|\mathcal{F}\big) \ a.s. \\ (\boldsymbol{X}, \boldsymbol{Y}), (\boldsymbol{X}', \boldsymbol{Y}'), (\boldsymbol{X}'', \boldsymbol{Y}'') \text{ are independent given } \mathcal{G}, \text{and are also independent given } \mathcal{F}. \end{cases} \tag{3.13}$$

Furthermore, if the sub σ*-algebra* $\mathcal{G} = \mathcal{G}_T$ *from a sub filtration* $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]} \subset \mathbb{F}$ *satisfies the compatibility conditions as in Assumption 2.5, then this compatibility condition holds for the natural extension of* $\overline{\mathbb{F}}$ *and* $\overline{\mathbb{G}}$ *defined by* (3.12) *in the extended probability space. That is, for each* $0 \le t \le T$, $\widetilde{\mathcal{F}}_t$ *and* $\widetilde{\mathcal{G}}_T$ *are independent given* $\overline{\mathcal{G}}_t$, *written as* $\widetilde{\mathcal{F}}_t \perp\!\!\!\perp \widetilde{\mathcal{G}}_T\big|\widetilde{\mathcal{G}}_t$.

Note that Theorem 3.1 allows us to use the notation $(\boldsymbol{X}, \boldsymbol{Y}), \mathcal{F}$, and $\mathcal{G}$ to respectively represent $(\widetilde{\boldsymbol{X}}, \widetilde{\boldsymbol{Y}}), \widetilde{\mathcal{F}}$, and $\widetilde{\mathcal{G}}$ without any ambiguity, as is adopted in the above corollary. Furthermore, Corollary 3.2 enables simplifying the notation $\widetilde{\mathbb{F}}$ and $\widetilde{\mathbb{G}}$ respectively by $\mathbb{F}$ and $\mathbb{G}$ without any ambiguity. Moreover, if we define the conditional law in the enlarged probability space as $\widetilde{\mu}_t := \mathrm{Law}\big(\widetilde{\boldsymbol{X}}\big|\widetilde{\mathcal{G}}_t\big)$, then by the compatibility condition and Theorem 3.3, we have

$$\widetilde{\mu}_t(\omega_0, \omega_1, \cdots, \omega_n) = \mathrm{Law}(\widetilde{\boldsymbol{X}}_t\big|\widetilde{\mathcal{G}}_T)(\omega_0, \omega_1, \cdots, \omega_n) = \mathrm{Law}(\boldsymbol{X}_t\big|\widetilde{\mathcal{G}}_T)(\omega_0) = \mu_t(\omega_0),$$

which in turn allows for using μ_t for $\widetilde{\mu}_t$ without any ambiguity.

3.4. **Properties of the conditional independent copies of semimartingales.** Next, we study properties of these conditional independent copies of semimartingales and explore the relation between the conditional expectation of their stochastic integrals and stochastic integrals with respect to the conditional law of semimartingales. For the sake of simplifying the notations, we will present the results in the one-dimensional case and their corresponding multi-dimensional cases can be adapted accordingly.

Theorem 3.3. *Give two* $\mathbb{R}$*-valued* $\mathbb{F}$*-semimartingales* $\{X_t\}_{t\in[0,T]}$ *and* $\{Y_t\}_{t\in[0,T]}$ *on a filtered probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. *Let* (X', Y') *and* (X'', Y'') *be their conditional independent copies defined in Corollary 3.2 in the enlarged filtered probability space* $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{F}} = (\mathcal{F}_t)_{t\in[0,T]}, \overline{\mathbb{P}})$ *satisfying* (3.6), (3.8) *and* (3.11). *Then,*

(1) X, Y, X', Y', X'', Y'' *are* $\overline{\mathbb{F}}$*-semimartingales.*

(2) *Suppose* $\mathbb{E}[|X_t|] < \infty$ *and* $\mathbb{E}[|Y_t|] < \infty$ *for all* $t \in [0, T]$, *and define*

$$X_t^{\mathcal{G}} := \overline{\mathbb{E}}\big[X_t\big|\mathcal{G}_T\big] = \overline{\mathbb{E}}\big[X_t\big|\mathcal{G}_t\big],\ \ Y_t^{\mathcal{G}} := \overline{\mathbb{E}}\big[Y_t\big|\mathcal{G}_T\big] = \overline{\mathbb{E}}\big[Y_t\big|\mathcal{G}_t\big].$$

(a) *If* $\{X_t\}_{t\in[0,T]}$ *is a finite variation process and* $\mathbb{E}\big[\big(\int_0^T |\mathrm{d}X_s|\big)^p\big] < \infty$, *for* $p \ge 1$, *then* $\{X_t^{\mathcal{G}}\}_{t\in[0,T]}$ *is also a finite variation process and*

$$\overline{\mathbb{E}}\bigg[\bigg(\int_0^T \big|\mathrm{d}X_s^{\mathcal{G}}\big|\bigg)^p\bigg] < \infty.$$

(b) $\{X_t^{\mathcal{G}}\}_{t\in[0,T]}$ *and* $\{Y_t^{\mathcal{G}}\}_{t\in[0,T]}$ *are semimartingales with respect to* $\overline{\mathbb{F}}$, $\mathbb{F}$ *and* $\mathbb{G}$.

(c) *If* $\{Z_t\}_{t\in[0,T]}$ *is an* $\mathbb{R}$*-valued* $\mathbb{F}$*-adapted càdlàg process in the filtered probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *and* $\|X\|_{\mathcal{H}_p} + \|Z\|_{\mathcal{H}_q} < \infty$ *with* $\frac{1}{p} + \frac{1}{q} = 1$, *then*

$$\int_0^t Z_{s-}\mathrm{d}X_s^{\mathcal{G}} = \overline{\mathbb{E}}\bigg[\int_0^t Z_{s-}\mathrm{d}X'_s\bigg|\mathcal{F}\bigg].$$

(d) If $\{Z_t\}_{t\in[0,T]}$ is an $\mathbb{R}$-valued $\mathbb{F}$-adapted càdlàg process in a filtered probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ and $\|X\|_{\mathcal{H}_p}+\|Y\|_{\mathcal{H}_q}+\|Z\|_{\mathcal{H}_r}<\infty$ with $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$, then

$$\int_0^t Z_{s-}\mathrm{d}\big[X^{\mathcal{G}},Y^{\mathcal{G}}\big]_s=\overline{\mathbb{E}}\bigg[\int_0^t Z_{s-}\mathrm{d}\big[X'',Y'\big]_s\bigg|\mathcal{F}\bigg],$$

$$\int_0^t Z_s\mathrm{d}\big[X^{\mathcal{G}},Y\big]_s=\overline{\mathbb{E}}\bigg[\int_0^t Z_s\mathrm{d}\big[X',Y\big]_s\bigg|\mathcal{F}\bigg].$$

The first statement in Theorem 3.3 shows that semi-martingales remain semimartingales under the enlarged filtration $\overline{F}$. (2a) and (2b) suggest that under the conditional expectation operation, the finite variation property, the semimartingale property with respect to $\overline{\mathcal{F}}$ and $\mathcal{F}$ are preserved and the martingale property of $\mathcal{G}$ is unchanged because of the compatibility condition in Assumption 2.5. In particular, (2c) and (2d) are the key equations for the derivation of the Itô's formula for cylindrical function; they ensure crucially the equivalence between the stochastic integral of the conditional expectation of a semimartingale and the conditional expectation of the Itô's integral of its conditional independent copy, as well as the equivalence of the quadratic variation of the conditional expectation of a semimartingale and the conditional expectation of the quadratic variation of its two conditional independent copies.

The proof of the theorem is presented in Section 5.2.

4. Proof of Theorem 2.7

Having established the properties of conditional independent copy of semimartingales, we are now ready to establish Itô's formula for flow of conditional laws on semimratingales, through several steps. First, we will derive the formula for cylindrical functions; we will then show that cylindrical functions are dense in $C^{2,2}(\mathbb{R}^d\times\mathcal{P}_p(\mathbb{R}^d))$. Finally, appropriate localization techniques and applications of appropriate dominance convergence theorem finish the proof.

4.1. **Step 1: Ito's formula for cylindrical functions.** First, let us recall the cylindrical functions (see for instance ([17]).

Definition 4.1 ($C^{2,2}$ cylindrical functions)**.** A function $\Phi:\mathcal{P}(\mathbb{R}^d)\times\mathbb{R}^d\to\mathbb{R}$ is a $C^{2,2}$ cylindrical function if

$$\Phi(\mu,\boldsymbol{y})=f\big(\langle\mu,g^{(1)}\rangle,..\langle\mu,g^{(n)}\rangle,\boldsymbol{y}\big),$$

where $f\in C^2(\mathbb{R}^{n+d})$ and $g\in C_b^2(\mathbb{R}^d)$.

Next, we show that when restricted to the smaller set of functions Φ, Theorem 2.7 holds. For semimartingales $\{\boldsymbol{X}_t\}_{t\in[0,T]}$ and $g^{(k)}\in C_b^2(\mathbb{R}^d)$, for $k=1,2,..,n$, by the classical Ito's formula [37, Theorem 33, Chapter II], we have

$$\begin{aligned}g^{(i)}(\boldsymbol{X}_t)-g^{(i)}(\boldsymbol{X}_0)=&\int_{0+}^t\nabla g^{(i)}(\boldsymbol{X}_{s-})\cdot\mathrm{d}\boldsymbol{X}_s^c+\frac{1}{2}\int_{0+}^t\nabla^2 g^{(i)}(\boldsymbol{X}_{s-}):\mathrm{d}[\boldsymbol{X},\boldsymbol{X}]_s^c\\&+\sum_{0<s\le t}\Big\{g^{(i)}(\boldsymbol{X}_s)-g^{(i)}(\boldsymbol{X}_{s-})\Big\},\ \text{for } t\in[0,T],\end{aligned}$$

and $\{g^{(i)}(\boldsymbol{X}_t)\}_{t\in[0,T]}$ for $i=1,..,N$ are semimartingales. Taking conditional expectation given $\mathcal{G}_T$, and recall that $\mu_t=\mathrm{Law}(\boldsymbol{X}_t|\mathcal{G}_t)=\mathrm{Law}(\boldsymbol{X}_t|\mathcal{G}_T)$, we have

$$\begin{aligned}\big\langle\mu_t,g^{(i)}\big\rangle-\big\langle\mu_0,g^{(i)}\big\rangle=\mathbb{E}\bigg[&\int_{0+}^t\nabla g^{(i)}(\boldsymbol{X}_{s-})\cdot\mathrm{d}\boldsymbol{X}_s^c+\frac{1}{2}\int_{0+}^t\nabla^2 g^{(i)}(\boldsymbol{X}_{s-}):\mathrm{d}[\boldsymbol{X},\boldsymbol{X}]_s^c\\&+\sum_{0<s\le t}\Big\{g^{(i)}(\boldsymbol{X}_s)-g^{(i)}(\boldsymbol{X}_{s-})\Big\}\bigg|\mathcal{G}_T\bigg].\end{aligned}\tag{4.1}$$

Define $\boldsymbol{Z}$ by

$$\boldsymbol{Z} := \Big\{ \boldsymbol{Z}_t = \Big(\langle \mu_t, g^{(1)} \rangle, \langle \mu_t, g^{(2)} \rangle, .., \langle \mu_t, g^{(n)} \rangle \Big) \Big\}_{t \in [0,T]},$$

and by (2b) of Theorem 3.3, $\boldsymbol{Z}$ is an $\overline{\mathbb{F}}$-semimartingale. Using the localization argument (which is detailed in Step 2), let us assume that $\boldsymbol{Y}_t$ is bounded. Then applying Itô's formula to $f\big(\boldsymbol{Z}, \boldsymbol{Y}\big)$ yields

$$\begin{aligned}
&f(\boldsymbol{Z}_t, \boldsymbol{Y}_t) - f(\boldsymbol{Z}_0, \boldsymbol{Y}_0) \\
&= \underbrace{\int_{0+}^t \nabla_{\boldsymbol{z}} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-}) \cdot \mathrm{d}\boldsymbol{Z}_s}_{J_1} + \underbrace{\frac{1}{2} \int_{0+}^t \nabla^2_{\boldsymbol{zz}} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-}) : \mathrm{d}[\boldsymbol{Z}, \boldsymbol{Z}]_s + \int_{0+}^t \nabla^2_{\boldsymbol{zy}} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-}) : \mathrm{d}[\boldsymbol{Z}, \boldsymbol{Y}]_s}_{J_2} \\
&\underbrace{- \sum_{0<s\le t} \Big\{ \nabla_{\boldsymbol{z}} f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-}) \cdot \Delta \boldsymbol{Z}_s + \frac{1}{2} \nabla^2_{\boldsymbol{zz}} f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-}) : \big(\Delta \boldsymbol{Z}_s \Delta \boldsymbol{Z}_s^\top \big) + \nabla^2_{\boldsymbol{zy}} f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-}) : \big(\Delta \boldsymbol{Z}_s \Delta \boldsymbol{Y}_s^\top \big) \Big\}}_{J_3} \\
&\quad + \int_{0+}^t \nabla_{\boldsymbol{y}} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-}) \cdot \mathrm{d}\boldsymbol{Y}^c_s + \frac{1}{2} \int_{0+}^t \nabla^2_{\boldsymbol{yy}} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-}) : \mathrm{d}[\boldsymbol{Y}, \boldsymbol{Y}]^c_s \\
&\quad + \sum_{0<s\le t} \Big\{ f(\boldsymbol{Y}_s, \boldsymbol{Z}_s) - f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-}) \Big\}.
\end{aligned} \tag{4.2}$$

Let us compute J_1, J_2, J_3 separately. First, we note that $g^{(i)}$ are bounded, $\boldsymbol{Y}_t$ is bounded, $f \in C^2(\mathbb{R}^{n+d})$, we have $\nabla_{\boldsymbol{z}} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-})$, $\nabla^2_{\boldsymbol{zy}} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-})$ and $\nabla^2_{\boldsymbol{z}} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-})$ are bounded and we confirm the conditions for applying (2c),(2d) of Theorem 3.3. Note that for the conditional copy $\boldsymbol{X}'$, we have

$$\begin{aligned}
g^{(i)}(\boldsymbol{X}'_t) - g^{(i)}(\boldsymbol{X}'_0) &= \int_{0+}^t \nabla g^{(i)}(\boldsymbol{X}'_{s-}) \cdot \mathrm{d}(\boldsymbol{X}')^c_s + \frac{1}{2} \int_{0+}^t \nabla^2 g^{(i)}(\boldsymbol{X}_{s-}) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}']^c_s \\
&\quad + \sum_{0<s\le t} \Big\{ g^{(i)}(\boldsymbol{X}'_s) - g^{(i)}(\boldsymbol{X}'_{s-}) \Big\}, \text{ for } t \in [0,T],
\end{aligned}$$

therefore, by (2c) of Theorem 3.3, $\mathrm{d}\boldsymbol{Z}_s = \mathrm{d}\big(g^{(1)}(\boldsymbol{X}_s), \cdots, g^{(n)}(\boldsymbol{X}_s)\big)^{\mathcal{G}})$, we have

$$\begin{aligned}
J_1 = \overline{\mathbb{E}}\Bigg[&\int_{0+}^t \sum_{i=1}^n \partial_{z_i} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-}) \Big(\nabla g^{(i)}(\boldsymbol{X}'_{s-}) \cdot \mathrm{d}(\boldsymbol{X}')^c_s + \frac{1}{2} \nabla^2 g^{(i)}(\boldsymbol{X}'_{s-}) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}']^c_s \Big) \\
&+ \sum_{0<s\le t} \sum_{i=1}^n \partial_{z_i} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-}) \Big\{ g^{(i)}(\boldsymbol{X}'_s) - g^{(i)}(\boldsymbol{X}'_{s-}) \Big\} \Big| \mathcal{F} \Bigg]
\end{aligned} \tag{4.3}$$

For J_2, computing $[g^{(i)}(\boldsymbol{X}'), g^{(j)}(\boldsymbol{X}'')]_t$ for $1 \le i, j \le n$, we have

$$\begin{aligned}
\big[g^{(i)}(\boldsymbol{X}'_t), g^{(j)}(\boldsymbol{X}''_t) \big]_t &= \big[(g^{(i)}(\boldsymbol{X}'))^c, (g^j(\boldsymbol{X}''))^c \big]_t + \sum_{0<s\le t} \Delta g^{(i)}(\boldsymbol{X}'_s) \Delta g^{(j)}(\boldsymbol{X}''_s) \\
&= \int_{0+}^t \Big(\nabla g^{(i)}(\boldsymbol{X}'_{s-})^\top \nabla g^{(j)}(\boldsymbol{X}''_{s-}) \Big) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}'']^c_s \\
&\quad + \sum_{0<s\le t} \Big\{ \big(g^{(i)}(\boldsymbol{X}'_s) - g^{(i)}(\boldsymbol{X}'_{s-}) \big) \big(g^{(j)}(\boldsymbol{X}''_s) - g^{(j)}(\boldsymbol{X}''_{s-}) \big) \Big\},
\end{aligned}$$

and computing $[g^{(i)}(\boldsymbol{X'}), Y^{(k)}]_t$ for $1 \le i \le n$, $1 \le k \le d$, we have

$$\begin{aligned}\big[g^{(i)}(\boldsymbol{X}'_t), Y^{(j)}\big]_t &= \big[(g^{(i)}(\boldsymbol{X'}))^c, (Y^j)^c\big]_t + \sum_{0<s\le t} \Delta g^{(i)}(\boldsymbol{X}'_s)\Delta Y^{(j)}_s \\ &= \int_{0+}^t \nabla g^{(i)}(\boldsymbol{X}'_{s-}) \cdot \mathrm{d}[\boldsymbol{X'}, Y^{(j)}]^c_s + \sum_{0<s\le t} \Big\{\big(g^{(i)}(\boldsymbol{X}'_s) - g^{(i)}(\boldsymbol{X}'_{s-})\big)\big(Y^{(k)}_s - Y^{(k)}_{s-}\big)\Big\}.\end{aligned}$$

Consequently, using (2d) of Theorem 3.3, we have

$$\begin{aligned}J_2 = \overline{\mathbb{E}}\Bigg[&\int_{0+}^t \Bigg(\frac{1}{2}\sum_{i,j=1}^n \partial^2_{z_i z_j} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-})\Big(\big(\nabla g^{(i)}(\boldsymbol{X}'_{s-})\nabla g^{(j)}(\boldsymbol{X}''_{s-})^\top\big) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}'']^c_s\Big) \\ &\qquad + \sum_{i=1}^n\sum_{k=1}^d \partial^2_{z_i y_k} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-})\Big(\nabla g^{(i)}(\boldsymbol{X}'_{s-}) \cdot \mathrm{d}[\boldsymbol{X}', Y^{(k)}]^c_s\Big)\Bigg) \\ &+ \sum_{0<s\le t}\Bigg(\frac{1}{2}\sum_{i,j=1}^n \partial^2_{z_i z_j} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-})\Big\{\big(g^{(i)}(\boldsymbol{X}'_s) - g^{(i)}(\boldsymbol{X}'_{s-})\big)\big(g^{(j)}(\boldsymbol{X}'_s) - g^{(j)}(\boldsymbol{X}'_{s-})\big)\Big\} \\ &\qquad + \sum_{i=1}^n\sum_{k=1}^d \partial^2_{z_i y_k} f(\boldsymbol{Z}_{s-}, \boldsymbol{Y}_{s-})\Big\{\big(g^{(i)}(\boldsymbol{X}'_s) - g^{(i)}(\boldsymbol{X}'_{s-})\big)\big(Y^{(k)}_s - Y^{(k)}_{s-}\big)\Big\}\Bigg)\Bigg|\mathcal{F}\Bigg].\end{aligned} \tag{4.4}$$

Now, to deal with J_3, note that

$$\begin{aligned}\Delta Z^{(i)}_t &= \big\langle \mu_t, g^{(i)}\big\rangle - \lim_{s\nearrow t}\big\langle \mu_s, g^{(i)}\big\rangle = \overline{\mathbb{E}}\big[g^{(i)}(\boldsymbol{X}_t)\big|\mathcal{G}_T\big] - \lim_{s\nearrow t}\overline{\mathbb{E}}\big[g^{(i)}(\boldsymbol{X}_s)\big|\mathcal{G}_T\big] \\ &= \overline{\mathbb{E}}\big[g^{(i)}(\boldsymbol{X}_t)\big|\mathcal{G}_T\big] - \overline{\mathbb{E}}\big[\lim_{s\nearrow t} g^{(i)}(\boldsymbol{X}_s)\big|\mathcal{G}_T\big] = \overline{\mathbb{E}}\big[g^{(i)}(\boldsymbol{X}_t) - g^{(i)}(\boldsymbol{X}_{t-})\big|\mathcal{G}_T\big] \\ &= \overline{\mathbb{E}}\big[g^{(i)}(\boldsymbol{X}'_t) - g^{(i)}(\boldsymbol{X}'_{t-})\big|\mathcal{F}\big],\end{aligned}$$

where the third equation holds by the conditional dominated convergence theorem. Therefore, we can rewrite J_3 as

$$\begin{aligned}J_3 = &\sum_{0<s\le t}\Bigg\{\overline{\mathbb{E}}\Bigg[\frac{1}{2}\sum_{i,j=1}^n \partial^2_{z_i z_j} f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-})\big(g^{(i)}(\boldsymbol{X}'_s) - g^{(i)}(\boldsymbol{X}'_{s-})\big)\big(g^{(j)}(\boldsymbol{X}''_s) - g^{(j)}(\boldsymbol{X}''_{s-})\big) \\ &\qquad\qquad + \sum_{i=1}^n \partial_{z_i} f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-})\big(g^{(i)}(\boldsymbol{X}'_s) - g^{(i)}(\boldsymbol{X}'_{s-})\big)\Bigg|\mathcal{F}\Bigg]\Bigg\} \\ &+ \sum_{0<s\le t}\overline{\mathbb{E}}\Bigg[\sum_{i=1}^n\sum_{k=1}^d \partial^2_{z_i y_k} f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-})\big(g^{(i)}(\boldsymbol{X}'_s) - g^{(i)}(\boldsymbol{X}'_{s-})\big)\big(Y^{(k)}_s - Y^{(k)}_{s-}\big)\Bigg|\mathcal{F}\Bigg],\end{aligned} \tag{4.5}$$

since $\partial^2_{z_i z_j} f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-})$, $\partial_{z_i} f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-})$, $\partial^2_{z_i y_k} f(\boldsymbol{Y}_{s-}, \boldsymbol{Z}_{s-})$, $Y^{(k)}_s$ and $Y^{(k)}_{s-}$ are $\mathcal{F}$-measurable random variables for $s \in (0, t]$, $1 \le i, j \le n$ and $1 \le k \le d$. Recall that $\Phi(\mu, \boldsymbol{y}) = f\big(\langle \mu, g^{(1)}\rangle, ..\langle \mu, g^{(n)}\rangle, \boldsymbol{y}\big)$, simple calculation of linear derivatives gives us a version of the derivatives by

$$\nabla_{\boldsymbol{y}}\Phi(\mu, \boldsymbol{y}) = \nabla_{\boldsymbol{y}} f\big(\langle \mu, g^{(1)}\rangle, ..\langle \mu, g^{(n)}, \boldsymbol{y}\rangle\big), \quad \nabla^2_0\Phi(\mu, \boldsymbol{y}) = \nabla^2_{\boldsymbol{yy}} f\big(\langle \mu, g^{(1)}\rangle, ..\langle \mu, g^{(n)}\rangle, \boldsymbol{y}\big),$$

$$\frac{\delta\Phi}{\delta\mu}(\mu, \boldsymbol{y}, \boldsymbol{x}_1) = \sum_{i=1}^n \partial_{z_i} f\big(\langle \mu, g^{(1)}\rangle, ..\langle \mu, g^{(n)}\rangle, \boldsymbol{y}\big) g^{(i)}(\boldsymbol{x}_1),$$

$$\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu, \boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{i,j=1}^n \partial_{z_i z_j} f\big(\langle \mu, g^{(1)}\rangle, ..\langle \mu, g^{(n)}\rangle, \boldsymbol{y}\big) g^{(i)}(\boldsymbol{x}_1) g^{(j)}(\boldsymbol{x}_2).$$

Therefore, we can rewrite J_1 by

$$\begin{aligned}&\overline{\mathbb{E}}\Big[\int_{0+}^{t}\Big(\nabla_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-})\cdot\mathrm{d}(\boldsymbol{X}')^c_s+\nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_s,\boldsymbol{X}'_s):\mathrm{d}[\boldsymbol{X}',\boldsymbol{X}']^c_t\Big)\\&+\sum_{0<s\le t}\Big(\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s)-\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-})\Big)\Big|\mathcal{F}\Big],\end{aligned}\tag{4.6}$$

J_2 by

$$\begin{aligned}&\overline{\mathbb{E}}\Big[\int_{0+}^{t}\Big(\nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_2}\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu_{s-},\boldsymbol{Y}_s,\boldsymbol{X}'_s,\boldsymbol{X}''_s):\mathrm{d}[\boldsymbol{X}',\boldsymbol{X}'']^c_t+\nabla^2_{\boldsymbol{x}_1\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}):\mathrm{d}[\boldsymbol{X}',\boldsymbol{Y}]^c_s\Big)\\&+\sum_{0<s\le t}\Big(\frac{1}{2}\Big(\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s,\boldsymbol{X}''_s)-\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-},\boldsymbol{X}''_s)\\&\qquad-\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s,\boldsymbol{X}''_{s-})+\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-},\boldsymbol{X}''_{s-})\Big)\\&\qquad+\Big(\nabla_{\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s)-\nabla_{\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-})\Big)\cdot\Delta\boldsymbol{Y}_s\Big)\Big|\mathcal{F}\Big],\end{aligned}\tag{4.7}$$

and J_3 by

$$\begin{aligned}\sum_{0<s\le t}\Big(&\frac{1}{2}\overline{\mathbb{E}}\Big[\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s,\boldsymbol{X}''_s)-\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-},\boldsymbol{X}''_s)\\&\quad-\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s,\boldsymbol{X}''_{s-})+\frac{\delta^2\Phi}{(\delta\mu)^2}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-},\boldsymbol{X}''_{s-})\Big|\mathcal{F}\Big]\\&+\overline{\mathbb{E}}\Big[\Big(\nabla_{\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s)-\nabla_{\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-})\Big)\cdot\Delta\boldsymbol{Y}_s\Big|\mathcal{F}\Big]\\&+\overline{\mathbb{E}}\Big[\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s)-\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-})\Big|\mathcal{F}\Big]\Big).\end{aligned}\tag{4.8}$$

Using the argument at the end of [24, Section 3.2.1], one can cancel out redundant terms in J_2 and J_3 to get the first summation in the right hand side of (2.3). Therefore, combining (4.6), (4.7), and (4.8) and plugging them back into (4.2) complete the proof.

4.2. **Step 2: localization argument.** We claim that for a sequence of $\mathbb{F}$ stopping time $\{\tau_n\}_{n\in\mathbb{N}}$ such that $\tau_n\to T$, in order to prove that (2.3) holds for $\Phi\in C^{2,2}(\mathcal{P}_p(\mathbb{R}^d)\times\mathbb{R}^d)$ with the process $(\mu,\boldsymbol{Y})=(\mu_{t\wedge\tau_n},\boldsymbol{Y}_{t\wedge\tau_n})_{t\in[0,T]}$, it suffices to prove that (2.3) holds for Φ with the truncated process $(\mu^{(\tau_n)},\boldsymbol{Y}^{(\tau_n)}):=(\mu_{t\wedge\tau_n},\boldsymbol{Y}_{t\wedge\tau_n})_{t\in[0,T]}$.

Indeed, suppose (2.3) holds for the truncated process, for $t<T$. Let $n\to\infty$, it is not hard to see that the left hand side of (2.3) converges and last three terms of the right hand side of (2.3) converge to the desired equation. Now, let us split the conditional expectation term of the right hand side of (2.3) into several terms. First, in order to prove that

$$\overline{\mathbb{E}}\Big[\int_{0+}^{t\wedge\tau_n}\nabla_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-})\cdot\mathrm{d}(\boldsymbol{X}')^c_s\Big|\mathcal{F}\Big]\to\overline{\mathbb{E}}\Big[\int_{0+}^{t}\nabla_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-})\cdot\mathrm{d}(\boldsymbol{X}')^c_s\Big|\mathcal{F}\Big],\tag{4.9}$$

almost surely as $n \to \infty$, we separate it by,

$$\begin{aligned}\int_{0+}^{t\wedge\tau_n} \nabla_{\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}\big) \cdot \mathrm{d}(\boldsymbol{X}')^c_s = \\ \int_{0+}^{t\wedge\tau_n} \nabla_{\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}\big) \cdot \mathrm{d}\boldsymbol{X}'_s - \sum_{0<s\le t\wedge\tau_n} \nabla_{\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}\big) \cdot \Delta\boldsymbol{X}'_s\end{aligned} \tag{4.10}$$

Note that the first term is uniformly bounded by

$$\sup_{0\le s\le T} \left| \int_{0+}^{s} \nabla_{\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{u-}, \boldsymbol{Y}_{u-}, \boldsymbol{X}'_{u-}\big) \cdot \mathrm{d}\boldsymbol{X}'_u \right|,$$

and

$$\begin{aligned}\overline{\mathbb{E}}\Big[\sup_{0\le s\le T} \Big| \int_{0+}^{s} \nabla_{\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{u-}, \boldsymbol{Y}_{u-}, \boldsymbol{X}'_{u-}\big) \cdot \mathrm{d}\boldsymbol{X}'_u \Big|\Big] &= \Big\| \int_{0+}^{\cdot} \nabla_{\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{u-}, \boldsymbol{Y}_{u-}, \boldsymbol{X}'_{u-}\big) \cdot \mathrm{d}\boldsymbol{X}'_u \Big\|_{\mathcal{S}_1} \\ &\le c_1 \Big\| \int_{0+}^{\cdot} \nabla_{\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{u-}, \boldsymbol{Y}_{u-}, \boldsymbol{X}'_{u-}\big) \cdot \mathrm{d}\boldsymbol{X}'_u \Big\|_{\mathcal{H}_1} \\ &\le c_1 \Big\| \nabla_{\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{\cdot}, \boldsymbol{Y}_{\cdot}, \boldsymbol{X}'_{\cdot}\big) \Big\|_{\mathcal{S}_{p'}} \big\|\boldsymbol{X}'\big\|_{\mathcal{H}_p} \le c_1 C \big\| |\boldsymbol{X}|^{p-1} + |\boldsymbol{Y}|^{p-1} + 1 \big\|_{\mathcal{S}_{p'}} \big\|\boldsymbol{X}\big\|_{\mathcal{H}_p} \\ &\le c_p c_1 C \big(\big\|\boldsymbol{X}\big\|^p_{\mathcal{H}_p} + \big\|\boldsymbol{Y}\big\|^p_{\mathcal{H}_p} + 1 \big) \big\|\boldsymbol{X}\big\|_{\mathcal{H}_p} < \infty,\end{aligned}$$

where $p' = \frac{p}{p-1}$, c_1, c_p are constants in Proposition 2.2 and C is a generic constant that may vary line by line. The first inequality is due to Proposition 2.2, the second inequality holds because of Proposition 2.3, the third inequality holds by $\Phi \in C^{2,2}(\mathcal{P}_p(\mathbb{R}^d) \times \mathbb{R}^d)$, and the last one holds because $\frac{1}{p} + \frac{1}{p'} = 1$. Meanwhile, the second term of the right hand side of (4.10) is bounded by

$$\begin{aligned}\sup_{0\le s\le T} \Big| \nabla_{\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}\big) \Big| \cdot \sum_{0<s\le T} |\Delta\boldsymbol{X}'_s| &\le C \sup_{0\le s\le T} \big(1 + |\boldsymbol{X}_s|^{p-1} + |\boldsymbol{Y}_s|^{p-1}\big) \cdot \sum_{0<s\le T} |\Delta\boldsymbol{X}'_s| \\ &\le C \bigg(1 + \sup_{0\le s\le T} |\boldsymbol{X}_s|^p + \sup_{0\le s\le T} |\boldsymbol{Y}_s|^p + \Big(\sum_{0<s\le T} |\Delta\boldsymbol{X}'_s| \Big)^p \bigg),\end{aligned} \tag{4.11}$$

where C is a generic constant that may vary line by line, and the right hand side of the above inequality is integrable due to the fact that $\|\boldsymbol{X}\|_{\mathcal{S}_p} \le c_p \|\boldsymbol{X}\|_{\mathcal{H}_p} < \infty$, $\|\boldsymbol{Y}\|_{\mathcal{S}_p} \le c_p \|\boldsymbol{Y}\|_{\mathcal{H}_p} < \infty$ and the assumption on $\boldsymbol{X}$. Therefore, using the conditional dominated theorem, we conclude (4.9). Next, to prove that

$$\begin{aligned}\overline{\mathbb{E}}\Big[\int_{0+}^{t\wedge\tau_n} \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}\big) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}']^c_s \Big| \mathcal{F} \Big] \\ \to \overline{\mathbb{E}}\Big[\int_{0+}^{t} \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}\big) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}']^c_s \Big| \mathcal{F} \Big],\end{aligned} \tag{4.12}$$

almost surely as $n \to \infty$, we note that

$$\int_{0+}^{t\wedge\tau_n} \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}\big) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}']^c_s$$

is bounded by

$$\sup_{0\le s\le T} \Big| \int_{0+}^{s} \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{u-}, \boldsymbol{Y}_{u-}, \boldsymbol{X}'_{u-}\big) : \mathrm{d}[\boldsymbol{X}', \boldsymbol{X}']^c_u \Big| \le \sup_{0\le s\le T} \Big| \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1} \frac{\delta\Phi}{\delta\mu}\big(\mu_{s-}, \boldsymbol{Y}_{s-}, \boldsymbol{X}'_{s-}\big) \Big| \cdot [\boldsymbol{X}', \boldsymbol{X}']^c_T$$

and

$$\overline{\mathbb{E}}\Big[\sup_{0\le s\le T}\Big|\nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\Big|\cdot[\boldsymbol{X}',\boldsymbol{X}']^c_T\Big]\le\Big\|\nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}\big(\mu_\cdot,\boldsymbol{Y}_\cdot,\boldsymbol{X}'_\cdot\big)\Big\|_{\mathcal{S}_{p''}}\big\|[\boldsymbol{X}',\boldsymbol{X}']^c_T\big\|_{L_{p/2}}$$
$$\le C\Big\|1+|\boldsymbol{X}'|^{p-2}+|\boldsymbol{Y}|^{p-2}\Big\|_{\mathcal{S}_{p''}}\big\|[\boldsymbol{X}',\boldsymbol{X}']^c_T\big\|_{L_{p/2}}$$
$$\le c_{p''}C\Big(1+\|\boldsymbol{X}'\|^p_{\mathcal{H}_p}+\|\boldsymbol{Y}\|^p_{\mathcal{H}_p}\Big)\big\|[\boldsymbol{X}',\boldsymbol{X}']^c_T\big\|_{L_{p/2}}<\infty,$$

where $c_{p''}$ is the constant in Proposition 2.2, C is a generic constant that may vary from line to line and $p''=\frac{p-2}{2p}$. Thus, (4.12) holds due to the conditional dominated theorem, we conclude . Similar arguments holds also for the convergence of $\int_{0+}^{t\wedge\tau_n}\nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_2}\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-},\boldsymbol{X}''_{s-}\big):\mathrm{d}[\boldsymbol{X}',\boldsymbol{X}'']_s$ and $\int_{0+}^{t\wedge\tau_n}\nabla^2_{\boldsymbol{x}_1\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big):\mathrm{d}[\boldsymbol{X}',\boldsymbol{Y}]_s$.

Next, let us deal with the convergence of

$$\sum_{0<s\le t\wedge\tau_n}\bigg(\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s,\boldsymbol{X}''_s\big)-\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-},\boldsymbol{X}''_s\big)$$
$$-\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s,\boldsymbol{X}''_{s-}\big)+\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-},\boldsymbol{X}''_{s-}\big)\bigg)\mathbb{1}_{\{\mu_s=\mu_{s-}\}}.\tag{4.13}$$

With simple application of the fundamental theorem of calculus for the linear derivative, we have

$$\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s,\boldsymbol{X}''_s\big)-\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-},\boldsymbol{X}''_s\big)$$
$$-\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s,\boldsymbol{X}''_{s-}\big)+\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-},\boldsymbol{X}''_{s-}\big)$$
$$=\int_0^1\int_0^1\nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_2}\frac{\delta^2\Phi}{(\delta\mu)^2}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}+\lambda(\boldsymbol{X}'_s-\boldsymbol{X}'_{s-}),\boldsymbol{X}''_{s-}+\gamma(\boldsymbol{X}''_s-\boldsymbol{X}''_{s-})\big):\Big((\Delta\boldsymbol{X}'_s)^\top\Delta\boldsymbol{X}''_s\Big)\mathrm{d}\lambda\mathrm{d}\gamma,$$

and the right hand side of the above equation is bounded by

$$C\cdot|\Delta\boldsymbol{X}'_s|\cdot|\Delta\boldsymbol{X}''_s|\cdot\Big(1+|\boldsymbol{X}'_s|^{p-2}+|\boldsymbol{X}''_s|^{p-2}\Big).$$

Therefore, (4.13) is bounded by

$$\sup_{0\le s\le T}\Big(1+|\boldsymbol{X}'_s|^{p-2}+|\boldsymbol{X}''_s|^{p-2}\Big)\big(\sum_{0<s\le T}|\Delta\boldsymbol{X}'_s|\cdot|\Delta\boldsymbol{X}''_s|\big)$$
$$\le C\bigg(\sup_{0\le t\le T}\Big(1+|\boldsymbol{X}'_t|^{p-2}+|\boldsymbol{X}''_t|^{p-2}\Big)^{\frac{p}{p-2}}+\Big(\sum_{0<s\le T}|\Delta\boldsymbol{X}'_s|^2+\sum_{0<s\le T}|\Delta\boldsymbol{X}''_s|^2\Big)^{\frac{p}{2}}\bigg)$$
$$\le C\bigg(\sup_{0\le t\le T}\Big(1+|\boldsymbol{X}'_t|^p+|\boldsymbol{X}''_t|^p\Big)+\sum_{0<s\le T}|\Delta\boldsymbol{X}'_s|^p+\sum_{0<s\le T}|\Delta\boldsymbol{X}''_s|^p\bigg),$$

where C is a generic constant that may vary from line by line. Therefore, the right hand side of the above inequality is integrable and the convergence of the conditional expectation of (4.13) is ensured by conditional dominated convergence theorem. Similar arguments can be applied to deduce the convergence of the conditional expectation of $\big(\nabla_{\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s\big)-\nabla_{\boldsymbol{y}}\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\big)\cdot\Delta\boldsymbol{Y}_s$ and $\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_s\big)-\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)$, s and thus we conclude that (2.3) holds for the original process $(\mu,\boldsymbol{Y})$ for $0\le t<T$.

For the case of $t = T$, taking $n \to \infty$, left hand side equals $\Phi(\mu_{T-}, \boldsymbol{Y}_{T-}) - \Phi(\mu_0, \boldsymbol{Y}_0)$ while all terms at the right hand sides is evaluated at T except the last term being

$$\sum_{0<s<t} \Big(\Phi(\mu_s, \boldsymbol{Y}_s) - \Phi(\mu_{s-}, \boldsymbol{Y}_{s-}) \Big),$$

adding $\Phi(\mu_T, \boldsymbol{Y}_T) - \Phi(\mu_{T-}, \boldsymbol{Y}_{T-})$ to both sides, we finish the proof of the claim.

4.3. **Step 3: cylindrical functions are dense in $C^{2,2}(\mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d))$.** The following result slightly generalizes that in [16, Theorem 4.4, Theorem 4.10, Section 4] as it includes the convergences of the derivatives of the linear derivative.

Proposition 4.2. *Let $\Phi \in C^{2,2}(\mathcal{P}_p(\mathbb{R}^d) \times \mathbb{R}^d)$, then there exist a sequence of $C^{2,2}$ cylinder functions $\{f_n\}_{n\in\mathbb{N}^+}$ such that one has the point-wise convergence*

$$\begin{aligned} &\Big(f_n, \nabla_{\boldsymbol{y}} f_n, \nabla^2_{\boldsymbol{y}} f_n, \frac{\delta f_n}{\delta\mu}, \nabla_{\boldsymbol{x}_1} \frac{\delta f_n}{\delta\mu}, \nabla^2_{\boldsymbol{x}_1} \frac{\delta f_n}{\delta\mu}, \nabla_{\boldsymbol{x}_1\boldsymbol{y}} \frac{\delta f_n}{\delta\mu}, \nabla_{\boldsymbol{x}_1\boldsymbol{x}_2} \frac{\delta^2 f_n}{(\delta\mu)^2} \Big) \\ &\qquad \to \Big(\Phi, \nabla_{\boldsymbol{y}} \Phi, \nabla^2_{\boldsymbol{y}} \Phi, \frac{\delta \Phi}{\delta\mu}, \nabla_{\boldsymbol{x}_1} \frac{\delta \Phi}{\delta\mu}, \nabla^2_{\boldsymbol{x}_1} \frac{\delta \Phi}{\delta\mu}, \nabla_{\boldsymbol{x}_1\boldsymbol{y}} \frac{\delta \Phi}{\delta\mu}, \nabla_{\boldsymbol{x}_1\boldsymbol{x}_2} \frac{\delta^2 \Phi}{(\delta\mu)^2} \Big) \end{aligned} \tag{4.14}$$

as $n \to \infty$, for all $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}, \in \mathbb{R}^d$, $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. Moreover, there exists a constant $c > 0$ such that

$$\begin{aligned} |\nabla_{\boldsymbol{y}} f_n|(\mu, \boldsymbol{y}) &\le c(1 + |\boldsymbol{y}|^{p-1}), & |\nabla^2_{\boldsymbol{y}} f_n|(\mu, \boldsymbol{y}) &\le c(1 + |\boldsymbol{y}|^{p-2}), \\ \left| \frac{\delta f_n}{\delta\mu} \right| &\le c\big(1 + |\boldsymbol{x}_1|^p + |\boldsymbol{y}|^p\big), & \left| \nabla_{\boldsymbol{x}_1} \frac{\delta f_n}{\delta\mu} \right| &\le c\big(1 + |\boldsymbol{x}_1|^{p-1} + |\boldsymbol{y}|^{p-1}\big), \\ \left| \nabla^2_{\boldsymbol{x}_1} \frac{\delta f_n}{\delta\mu} \right| &\le c_K\big(1 + |\boldsymbol{x}_1|^{p-2} + |\boldsymbol{y}|^{p-2}\big), & \left| \nabla^2_{\boldsymbol{x}_1\boldsymbol{y}} \frac{\delta f_n}{\delta\mu} \right| &\le c\big(1 + |\boldsymbol{x}_1|^{p-2} + |\boldsymbol{y}|^{p-2}\big), \\ & & \left| \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_2} \frac{\delta^2 f_n}{(\delta\mu)^2} \right| &\le c\big(1 + |\boldsymbol{x}_1|^{p-2} + |\boldsymbol{x}_2|^{p-2} + |\boldsymbol{y}|^{p-2}\big), \end{aligned} \tag{4.15}$$

for all $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y} \in \mathbb{R}^d$, $\mu \in \mathcal{P}_p(\mathbb{R}^d)$.

The proof is similar to that of [16, Section 4]. The main step of their proof consists of constructing an operator $T_n : C_b(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ and the corresponding adjoint $T_n^* : \mathcal{P}_p(\mathbb{R}^d) \to \mathcal{P}_p(\mathbb{R}^d)$ corresponding square $T_n^{\otimes 2} : C_b(\mathbb{R}^{2d}) \to C_b(\mathbb{R}^{2d})$ such that $T_n\varphi_n \to \varphi$, when $\varphi_n \to \varphi$ and $T_n^*\mu \to \mu$. Define $f_n(\mu, \boldsymbol{y}) := \Phi(T_n^*\mu, \boldsymbol{y})$, then the key identities to be established are

$$\frac{\delta f_n}{\delta\mu}(\mu, \boldsymbol{y}, \boldsymbol{x}_1) = T_n\Big(\frac{\delta\Phi}{\delta\mu}(T_n^*\mu, \boldsymbol{y}, \cdot) \Big)(\boldsymbol{x}_1), \quad \frac{\delta^2 f_n}{(\delta\mu)^2}(\mu, \boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2) = T_n^{\otimes 2} \frac{\delta^2\Phi}{(\delta\mu)^2}(T_n^*\mu, \boldsymbol{y}, \cdot, \cdot)(\boldsymbol{x}_1, \boldsymbol{x}_2).$$

for all $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y} \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. To this end, it suffices to note that the construction of T_n ensures

$$T_n \nabla_{\boldsymbol{x}_1} \varphi(\boldsymbol{x}_1) = \nabla_{\boldsymbol{x}_1} T_n \varphi(\boldsymbol{x}_1), \quad T_n^{\otimes 2} \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_2} \psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_2} T_n \psi(\boldsymbol{x}_1, \boldsymbol{x}_2),$$

for all $\varphi \in C_b^1(\mathbb{R}^d)$ and $\psi \in C_b^2(\mathbb{R}^{2d})$.

4.4. **Final step.** Define $\tau_k := \inf\{t \in [0, T] : |\boldsymbol{Y}_t| > k\}$, then by Step 1, we know that Itô's formula (2.3) holds for the truncated process $(\mu^{(\tau_n)}, \boldsymbol{Y}^{(\tau_n)}) := (\mu_{t\wedge\tau_n}, \boldsymbol{Y}_{t\wedge\tau_n})_{t\in[0,T]}$ with the sequences of cylindrical functions $\{f_n\}_{n\in\mathbb{N}^+}$. By Step 2, the localization argument shows that Itô's formula (2.3) holds for the sequences of cylindrical functions $\{f_n\}_{n\in\mathbb{N}^+}$.

In order to pass the limit from f_n to Φ, let us consider the first term in the right hand side of (2.3) and the remaining terms follow similarly. For the first term, we need to prove that there exists a subsequence $\{n_k\}_{k\in\mathbb{N}^+}$ such that

$$\mathbb{E}\Big[\int_{0+}^{t}\nabla_{\boldsymbol{x}_1}\frac{\delta f_{n_k}}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\cdot\mathrm{d}(\boldsymbol{X}')^c_s\Big|\mathcal{F}\Big]\to\mathbb{E}\Big[\int_{0+}^{t}\nabla_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\cdot\mathrm{d}(\boldsymbol{X}')^c_s\Big|\mathcal{F}\Big],\tag{4.16}$$

as $k\to\infty$. To this end, we first prove that

$$\int_{0+}^{t}\nabla_{\boldsymbol{x}_1}\frac{\delta f_n}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\cdot\mathrm{d}(\boldsymbol{X}')^c_s\xrightarrow{L^1}\int_{0+}^{t}\nabla_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\cdot\mathrm{d}(\boldsymbol{X}')^c_s,\tag{4.17}$$

as $n\to\infty$. By Proposition 4.2, we have

$$\Big||\nabla_{\boldsymbol{x}_1}\frac{\delta f_n}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\Big|\le c\big(1+|\boldsymbol{X}_{s-}|^{p-1}+|\boldsymbol{Y}_{s-}|^{p-1}\big),$$

where the $\mathcal{S}_{p'}$ norm of the right hand side is finite with $p'=\frac{p}{p-1}$. Now (4.17) follows from $\boldsymbol{X}\in\mathcal{H}^p$ and the dominated convergence theorem of stochastic integral ([37, Page 273, Lemma]). Next, since

$$\begin{aligned}&\mathbb{E}\Big[\Big|\mathbb{E}\Big[\int_{0+}^{t}\nabla_{\boldsymbol{x}_1}\frac{\delta f_n}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\cdot\mathrm{d}(\boldsymbol{X}')^c_s\Big|\mathcal{F}\Big]-\mathbb{E}\Big[\int_{0+}^{t}\nabla_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\cdot\mathrm{d}(\boldsymbol{X}')^c_s\Big|\mathcal{F}\Big]\Big|\Big]\\&\quad\le\mathbb{E}\Big[\Big|\int_{0+}^{t}\nabla_{\boldsymbol{x}_1}\frac{\delta f_n}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\cdot\mathrm{d}(\boldsymbol{X}')^c_s-\int_{0+}^{t}\nabla_{\boldsymbol{x}_1}\frac{\delta\Phi}{\delta\mu}\big(\mu_{s-},\boldsymbol{Y}_{s-},\boldsymbol{X}'_{s-}\big)\cdot\mathrm{d}(\boldsymbol{X}')^c_s\Big|\Big]\to0,\end{aligned}$$

and L_1 convergences implies convergence in subsequence, (4.16) holds. The remaining terms follow similarly due to the uniform bound (4.15), and thus the proof of Theorem 2.7 is complete.

5. Proof of Section 3

5.1. **Construction of the conditional independent copy.** This subsection is devoted to the proof of the precise construction of conditional independent copy as in Theorem 3.1. For ease of exposition, we will show how to enlarge the probability space with two independent copies, and any finite number of copies can be easily adapted.

Recall the definition in Section 3, let us consider the probability space $\overline{\Omega}$ and the σ-algebra $\overline{\mathcal{F}}$ defined by

$$\overline{\Omega}=\Omega^3=\{(\omega_0,\omega_1,\omega_2)|\omega_0,\omega_1,\omega_2\in\Omega\},\qquad\overline{\mathcal{F}}=\sigma\{A_0\times A_1\times A_2:A_0,A_1,A_2\in\mathcal{F}\}\tag{5.1}$$

and probability measure $\overline{\mathbb{P}}$ is defined as follows: for $A_0,A_1,A_2\in\mathcal{F}$,

$$\overline{\mathbb{P}}\big(A_0\times A_1\times A_2\big):=\mathbb{E}\Big[\mathbb{1}_{A_0}\mathbb{P}(A_1|\mathcal{G})\mathbb{P}(A_2|\mathcal{G})\Big],\tag{5.2}$$

Since the set $\{A_0\times A_1\times A_2:A_0,A_1,A_2\in\mathcal{F}\}$ is a π-system, the measure on $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}})$ is properly defined. Next, use the notation $\widetilde{\mathcal{F}},\widetilde{\mathcal{G}}$ for the nature extension into this enlarged probability space as

$$\widetilde{\mathcal{F}}:=\{A\times\Omega\times\Omega:A\in\mathcal{F}\},\quad\widetilde{\mathcal{G}}:=\{A\times\Omega\times\Omega:A\in\mathcal{G}\},\tag{5.3}$$

and define $\mathcal{F}'$, $\mathcal{F}''$ as

$$\mathcal{F}':=\{\Omega\times A\times\Omega:A\in\mathcal{F}\},\quad\mathcal{F}'':=\{\Omega\times\Omega\times A:A\in\mathcal{F}\}.\tag{5.4}$$

With these constructions, we have the following propositions.

Proposition 5.1. *Given $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}})$ defined by (5.1) and (5.2) and the sub-algebras $\widetilde{\mathcal{F}},\widetilde{\mathcal{G}}$, $\mathcal{F}'$, and $\mathcal{F}''$ defined by (5.3) and (5.4). Suppose $A,B\in\mathcal{F}$, let $\widetilde{A}=A\times\Omega\times\Omega$, $A'=\Omega\times A\times\Omega$, $A''=\Omega\times\Omega\times A$, $B'=\Omega\times B\times\Omega$ and $C''=\Omega\times\Omega\times C$. Then*

(1) For $\overline{\omega} = (\omega_0, \omega_1, \omega_2) \in \overline{\Omega}$

$$\mathbb{P}(A|\mathcal{G})(\omega_0) = \overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}})(\omega_0, \omega_1, \omega_2), \ \textit{a.s. under } \overline{\mathbb{P}}.$$

(2) Given $\widetilde{\mathcal{G}}$, then $\widetilde{A}, A'$, and A'' have the same conditional law. That is,

$$\overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}}) = \overline{\mathbb{P}}(A'|\widetilde{\mathcal{G}}) = \overline{\mathbb{P}}(A''|\widetilde{\mathcal{G}}), \ \textit{a.s. under } \overline{\mathbb{P}}.$$

Moreover, this conditional probability $\overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}})$ coincides with the conditional probability of A' and A'' given $\widetilde{\mathcal{F}}$, that is

$$\overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}}) = \overline{\mathbb{P}}(A'|\widetilde{\mathcal{F}}) = \overline{\mathbb{P}}(A''|\widetilde{\mathcal{F}}), \ \textit{a.s. under } \overline{\mathbb{P}}.$$

(3) Given $\widetilde{\mathcal{G}}$, then $\widetilde{\mathcal{F}}, \mathcal{F}'$, and $\mathcal{F}''$ are conditionally independent. That is,

$$\overline{\mathbb{P}}\big(\widetilde{A} \cap B' \cap C''|\widetilde{\mathcal{G}}\big) = \overline{\mathbb{P}}\big(\widetilde{A}|\widetilde{\mathcal{G}}\big)\overline{\mathbb{P}}\big(B'|\widetilde{\mathcal{G}}\big)\overline{\mathbb{P}}\big(C''|\widetilde{\mathcal{G}}\big), \ \textit{a.s. under } \overline{\mathbb{P}}.$$

Moreover, given $\widetilde{\mathcal{F}}$, then $\mathcal{F}'$ and $\mathcal{F}''$ are conditionally independent, That is

$$\overline{\mathbb{P}}\big(B' \cap C''|\widetilde{\mathcal{F}}\big) = \overline{\mathbb{P}}\big(B'|\widetilde{\mathcal{F}}\big)\overline{\mathbb{P}}\big(C''|\widetilde{\mathcal{F}}\big), \ \textit{a.s. under } \overline{\mathbb{P}}.$$

Proof. To prove (1), note that $\mathbb{P}[A|\mathcal{G}]$ is $\mathcal{G}$ measurable in $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, then $\mathbb{P}[A|\mathcal{G}](\omega_1)$ as a random variable in $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{F}}, \overline{\mathbb{P}})$ is $\widetilde{\mathcal{G}}$ measurable . For any $\widetilde{G} \in \widetilde{\mathcal{G}}$, there exists a $G \in \mathcal{G}$ such that $\widetilde{G} = G \times \Omega \times \Omega \in \widetilde{\mathcal{G}}$. Recall the definition of $\overline{\mathbb{P}}$ at (5.2) and the property of conditional expectation, we have

$$\overline{\mathbb{P}}\big[\widetilde{A} \cap \widetilde{G}\big] = \overline{\mathbb{P}}\big[(A \cap G) \times \Omega \times \Omega\big] = \mathbb{P}[A \cap G] = \mathbb{E}\Big[\mathbb{P}[A|\mathcal{G}]\mathbb{1}_G\Big] = \int_\Omega \mathbb{P}[A|\mathcal{G}](\omega)\mathbb{1}_G(\omega)\mathbb{P}(\mathrm{d}\omega).$$

Again by (5.2), we obtain $\overline{\mathbb{P}}(\mathrm{d}\omega_1 \times \Omega \times \Omega) = \mathbb{P}(\mathrm{d}\omega_1)$, therefore,

$$\begin{aligned} \overline{\mathbb{P}}\big[\widetilde{A} \cap \widetilde{G}\big] &= \int_{\Omega^3} \mathbb{P}[A|\mathcal{G}](\omega_0)\mathbb{1}_B(\omega_0)\overline{\mathbb{P}}(\mathrm{d}\omega_0\mathrm{d}\omega_1\mathrm{d}\omega_2) \\ &= \int_{\Omega^3} \mathbb{P}[A|\mathcal{G}](\omega_0)\mathbb{1}_{\widetilde{G}}(\omega_0, \omega_1, \omega_2)\overline{\mathbb{P}}(\mathrm{d}\omega_0\mathrm{d}\omega_1\mathrm{d}\omega_2) = \overline{\mathbb{E}}\Big[\mathbb{P}[A|\mathcal{G}]\mathbb{1}_{\widetilde{G}}\Big], \end{aligned} \tag{5.5}$$

which leads to $\mathbb{P}[A|\mathcal{G}](\omega_0) = \overline{\mathbb{P}}[\widetilde{A}|\widetilde{\mathcal{G}}](\omega_0, \omega_1, \omega_2)$ a.s. under $\overline{P}$.

To prove (2), for any $\widetilde{G} \in \widetilde{\mathcal{G}}$, there exists a $G \in \mathcal{G}$ such that $\widetilde{G} = G \times \Omega \times \Omega \in \widetilde{\mathcal{G}}$. Then,

$$\overline{\mathbb{P}}\big[A' \cap \widetilde{G}\big] = \overline{\mathbb{P}}\big[G \times A \times \Omega\big] = \mathbb{E}\big[\mathbb{1}_G\mathbb{P}(A|\mathcal{G})\big] = \overline{\mathbb{E}}\Big[\mathbb{P}[A|\mathcal{G}]\mathbb{1}_{\widetilde{G}}\Big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{P}}[\widetilde{A}|\widetilde{\mathcal{G}}]\mathbb{1}_{\widetilde{G}}\Big],$$

where the second equation holds because of (5.2), the third equation holds due to the same calculation as (5.5) and the last equation holds by (1). This leads to $\overline{\mathbb{P}}(A'|\widetilde{\mathcal{G}}) = \overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}})$ and similarly, we have $\overline{\mathbb{P}}(A''|\widetilde{\mathcal{G}}) = \overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}})$.

As for the second statement of (2), note that $\overline{\mathbb{P}}(A'|\widetilde{\mathcal{G}})$ is $\widetilde{\mathcal{G}}$ measurable, thus $\widetilde{\mathcal{F}}$ measurable. For any $\widetilde{F} \in \widetilde{\mathcal{F}}$, there exists a $F \in \mathcal{F}$ such that $\widetilde{F} = F \times \Omega \times \Omega \in \widetilde{\mathcal{F}}$. Then,

$$\overline{\mathbb{P}}(A' \cap \widetilde{F}) = \overline{\mathbb{P}}\big[F \times A \times \Omega\big] = \mathbb{E}\big[\mathbb{1}_F\mathbb{P}(A|\mathcal{G})\big] = \overline{\mathbb{E}}\Big[\mathbb{P}[A|\mathcal{G}]\mathbb{1}_{\widetilde{F}}\Big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{P}}[\widetilde{A}|\widetilde{\mathcal{G}}]\mathbb{1}_{\widetilde{F}}\Big],$$

where the second equation holds because of (5.2), the third equation holds due to the same calculation as (5.5) and the last equation holds by (1). This leads to $\overline{\mathbb{P}}(A'|\widetilde{\mathcal{F}}) = \overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}})$ and similarly, we have $\overline{\mathbb{P}}(A''|\widetilde{\mathcal{F}}) = \overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}})$.

To prove (3), for any $\widetilde{G} \in \widetilde{\mathcal{G}}$, there exists a $G \in \mathcal{G}$ such that $\widetilde{G} = G \times \Omega \times \Omega \in \widetilde{\mathcal{G}}$. Then,

$$\begin{aligned} \overline{\mathbb{P}}(\widetilde{A} \cap B' \cap C'' \cap \widetilde{G}) &= \overline{\mathbb{P}}((A \cap G) \times B \times C) = \mathbb{E}\Big[\mathbb{1}_{A\cap G}\mathbb{P}(B|\mathcal{G})\mathbb{P}(C|\mathcal{G})\Big] \\ &= \mathbb{E}\Big[\mathbb{E}[\mathbb{1}_{A\cap G}|\mathcal{G}]\mathbb{P}(B|\mathcal{G})\mathbb{P}(C|\mathcal{G})\Big] = \mathbb{E}\Big[\mathbb{1}_G\mathbb{P}(A|\mathcal{G})\mathbb{P}(B|\mathcal{G})\mathbb{P}(C|\mathcal{G})\Big] \end{aligned}$$

Using the same calculation as in (5.5) and note that $\mathbb{P}(A|\mathcal{G}) = \overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}})$, $\mathbb{P}(B|\mathcal{G}) = \overline{\mathbb{P}}(B'|\widetilde{\mathcal{G}})$ and $\mathbb{P}(C|\mathcal{G}) = \overline{\mathbb{P}}(C''|\widetilde{\mathcal{G}})$, we have

$$\overline{\mathbb{P}}(\widetilde{A} \cap B' \cap C'' \cap \widetilde{G}) = \overline{\mathbb{E}}\Big[\mathbb{1}_{\widetilde{G}}\mathbb{P}(A|\mathcal{G})\mathbb{P}(B|\mathcal{G})\mathbb{P}(C|\mathcal{G})\Big] = \overline{\mathbb{E}}\Big[\mathbb{1}_{\widetilde{G}}\overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}})\overline{\mathbb{P}}(B'|\widetilde{\mathcal{G}})\overline{\mathbb{P}}(C''|\widetilde{\mathcal{G}})\Big].$$

This leads to

$$\overline{\mathbb{P}}\big(\widetilde{A} \cap B' \cap C''|\widetilde{\mathcal{G}}\big) = \overline{\mathbb{P}}(\widetilde{A}|\widetilde{\mathcal{G}})\overline{\mathbb{P}}(B'|\widetilde{\mathcal{G}})\overline{\mathbb{P}}(C''|\widetilde{\mathcal{G}}), \text{ a.s. under } \overline{\mathbb{P}}.$$

As for the second statement of (3), for any $\widetilde{F} \in \widetilde{\mathcal{F}}$, there exists a $F \in \mathcal{F}$ such that $\widetilde{F} = F \times \Omega \times \Omega \in \widetilde{\mathcal{F}}$. Then,

$$\begin{aligned}\overline{\mathbb{P}}(B' \cap C'' \cap \widetilde{F}) &= \overline{\mathbb{P}}(F \times B \times C) = \mathbb{E}\big[\mathbb{1}_F\mathbb{P}(B|\mathcal{G})\mathbb{P}(C|\mathcal{G})\big] \\ &= \overline{\mathbb{E}}\big[\mathbb{1}_{\widetilde{F}}\mathbb{P}(B|\mathcal{G})\mathbb{P}(C|\mathcal{G})\big] = \overline{\mathbb{E}}\big[\mathbb{1}_{\widetilde{F}}\overline{\mathbb{P}}(B'|\widetilde{\mathcal{F}})\overline{\mathbb{P}}(C''|\overline{\mathcal{F}})\big],\end{aligned}$$

where the second equation holds because of (5.2), the third equation holds due to the same calculation as (5.5) and the last equation holds by (1) and (2). Therefore,

$$\overline{\mathbb{P}}\big(B' \cap C''|\widetilde{\mathcal{F}}\big) = \overline{\mathbb{P}}\big(B'|\widetilde{\mathcal{F}}\big)\overline{\mathbb{P}}\big(C''|\widetilde{\mathcal{F}}\big), \text{ a.s. under } \overline{\mathbb{P}},$$

which completes the whole proof. $\square$

Proof of Theorem 3.1. For any random variable R in $(\Omega, \mathcal{F}, \mathbb{P})$, recall that the natural extension $\widetilde{R}$ and the two conditional independent copies R', R'' of R are defined by

$$\widetilde{R}(\omega_0, \omega_1, \omega_2) = R(\omega_0),\ R'(\omega_0, \omega_1, \omega_2) = R(\omega_1),\ R''(\omega_0, \omega_1, \omega_2) = R(\omega_2).$$

Then

$$\operatorname{Law}\big(R|\mathcal{G}\big)(\omega_0) = \operatorname{Law}\big(\widetilde{R}|\widetilde{\mathcal{G}}\big)(\overline{\omega}) = \operatorname{Law}\big(R'|\widetilde{\mathcal{G}}\big)(\overline{\omega}) = \operatorname{Law}\big(R''\big|\widetilde{\mathcal{F}}\big)(\overline{\omega}) \text{ a.s. under } \overline{\mathbb{P}},$$

is a direct consequence of (2) of Proposition 5.1 where $\overline{\omega} = (\omega_0, \omega_1, \omega_2)$. Meanwhile, (3) leads to

$$R, R' \text{ and } R'' \text{ are independent given } \widetilde{\mathcal{G}} \text{ and } R, R' \text{ and } R'' \text{ are independent given } \mathcal{F},$$

which completes the proof of Theorem 3.1. $\square$

Proof of Corollary 3.2. It suffices to show that extension of the sub-filtration $\widetilde{\mathbb{G}} \subset \widetilde{\mathbb{F}}$ satisfies the compatibility assumption or the conditional independence condition, that is

$$\widetilde{\mathcal{F}}_t \perp\!\!\!\perp \widetilde{\mathcal{G}}_T\big|\widetilde{\mathcal{G}}_t, \text{ for all } 0 \le t \le T.$$

For any $\widetilde{F}_t \in \widetilde{\mathcal{F}}_t$, $\widetilde{G}_t \in \widetilde{\mathcal{G}}_t$ and $\widetilde{G}_T \in \widetilde{\mathcal{G}}_T$, there exists $F_t \in \mathcal{F}_t$, $G_t \in \mathcal{G}_t$ and $G_T \in \mathcal{G}_T$ such that $\widetilde{F}_t = F_t \times \Omega \times \Omega$, $\widetilde{G}_t = G_t \times \Omega \times \Omega$ and $\widetilde{G}_T = G_T \times \Omega \times \Omega$. Then,

$$\begin{aligned}\overline{\mathbb{P}}(\widetilde{F}_t \cap \widetilde{G}_t \cap \widetilde{G}_T) &= \overline{\mathbb{P}}((F_t \cap G_t \cap G_T) \times \Omega \times \Omega) = \mathbb{P}(F_t \cap G_t \cap G_T) \\ &= \mathbb{E}\big[\mathbb{P}(F_t \cap G_T|\mathcal{G}_t)\mathbb{1}_{G_t}\big] = \mathbb{E}\big[\mathbb{P}(F_t|\mathcal{G}_t)\mathbb{P}(G_T|\mathcal{G}_t)\mathbb{1}_{G_t}\big] \\ &= \overline{\mathbb{E}}\big[\mathbb{P}(F_t|\mathcal{G}_t)\mathbb{P}(G_T|\mathcal{G}_t)\mathbb{1}_{\widetilde{G}_t}\big] = \overline{\mathbb{E}}\big[\overline{\mathbb{P}}(\widetilde{F}_t|\widetilde{\mathcal{G}}_t)\overline{\mathbb{P}}(\widetilde{G}_T|\widetilde{\mathcal{G}}_t)\mathbb{1}_{\widetilde{G}_t}\big],\end{aligned}$$

where the fourth equation holds since $\mathcal{F}_t \perp\!\!\!\perp \mathcal{G}_T|\mathcal{G}_t$, the fifth equation holds because of the same calculation as (5.5) and the last equation holds because of (2) in Proposition 5.1. Therefore,

$$\overline{\mathbb{P}}(\widetilde{F}_t \cap \widetilde{G}_T|\widetilde{\mathcal{G}}_t) = \overline{\mathbb{P}}(\widetilde{F}_t|\widetilde{\mathcal{G}}_t)\overline{\mathbb{P}}(\widetilde{G}_T|\widetilde{\mathcal{G}}_t),$$

and thus completes the proof. $\square$

5.2. **Proof of Theorem 3.3.** For (1), due to the decomposablity (see [37, Theorem 1, Chapter 4]) let us assume that

$$X_t = M_t + V_t,$$

where M_t is an $\mathbb{F}$-local martingale and V_t is finite variation process. To prove that X is a $\overline{\mathbb{F}}$-semimartingale, it suffices to show that M_t is a $\overline{\mathbb{F}}$-local martingale. Let $\{\tau_k\}_{k\in\mathbb{N}^+}$ be a sequence of $\mathbb{F}$-stopping times such that $M^{\tau_k} := (M_{t\wedge\tau_k})_{t\in[0,T]}$ is an $\mathbb{F}$-martingale for all $k \in \mathbb{N}^+$ and as $k \to \infty$, we have $\tau_k \to \infty$. We claim that

$$\overline{\mathbb{E}}[M_{t\wedge\tau_k}|\overline{\mathcal{F}}_s] = M_{s\wedge\tau_k}.$$

Recall that

$$\overline{\mathcal{F}}_s = \sigma\{A_0 \times A_1 \times A_2 : A_0, A_1, A_2 \in \mathcal{F}_s\},$$

it suffices to show that

$$\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_0\times A_1\times A_2}\big] = \overline{\mathbb{E}}\big[M_{s\wedge\tau_k}\mathbb{1}_{A_0\times A_1\times A_2}\big] \tag{5.6}$$

for all $A_0, A_1, A_2 \in \mathcal{F}_s$. Note that $M_{t\wedge\tau_k}\mathbb{1}_{A_0\times\Omega\times\Omega}$, $\mathbb{1}_{\Omega\times A_1\times\Omega}$ and $\mathbb{1}_{\Omega\times\Omega\times A_2}$ are conditional independent of given $\mathcal{G}_T$ by (3) in Proposition 5.1, we have

$$\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_0\times A_1\times A_2}\big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_0\times\Omega\times\Omega}\big|\mathcal{G}_T\big]\overline{\mathbb{E}}\big[\mathbb{1}_{\Omega\times A_1\times\Omega}\big|\mathcal{G}_T\big]\overline{\mathbb{E}}\big[\mathbb{1}_{\Omega\times\Omega\times A_2}\big|\mathcal{G}_T\big]\Big]. \tag{5.7}$$

By (2) in Proposition 5.1 and $\mathcal{F}_s \perp\!\!\!\perp \mathcal{G}_T|\mathcal{G}_s$, we know that

$$\overline{\mathbb{E}}\big[\mathbb{1}_{\Omega\times A_1\times\Omega}\big|\mathcal{G}_T\big] = \overline{\mathbb{P}}(A_1|\mathcal{G}_T) = \overline{\mathbb{P}}(A_1|\mathcal{G}_s), \text{ and similarly } \overline{\mathbb{E}}\big[\mathbb{1}_{\Omega\times\Omega\times A_2}\big|\mathcal{G}_T\big] = \overline{\mathbb{P}}(A_2|\mathcal{G}_s). \tag{5.8}$$

Plugging back in (5.7), we have

$$\begin{aligned}\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_0\times A_1\times A_2}\big] &= \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_0\times\Omega\times\Omega}\big|\mathcal{G}_T\big]\overline{\mathbb{P}}\big(A_1|\mathcal{G}_s\big)\overline{\mathbb{P}}\big(A_2|\mathcal{G}_s\big)\Big]\\ &= \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_0\times\Omega\times\Omega}\big|\mathcal{G}_s\big]\overline{\mathbb{P}}\big(A_1|\mathcal{G}_s\big)\overline{\mathbb{P}}\big(A_2|\mathcal{G}_s\big)\Big]\end{aligned} \tag{5.9}$$

where the second inequality holds by the tower rule. Since M^{τ_k} is a $\mathcal{F}$-martingale, we get

$$\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_0\times\Omega\times\Omega}\big|\mathcal{G}_s\big] = \overline{\mathbb{E}}\big[\mathbb{1}_{A_0\times\Omega\times\Omega}\mathbb{E}[M_{t\wedge\tau_k}|\mathcal{F}_s]\big|\mathcal{G}_s\big] = \overline{\mathbb{E}}\big[M_{s\wedge\tau_k}\mathbb{1}_{A_0\times\Omega\times\Omega}\big|\mathcal{G}_s\big].$$

Therefore,

$$\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_0\times A_1\times A_2}\big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M_{s\wedge\tau_k}\mathbb{1}_{A_0\times\Omega\times\Omega}\big|\mathcal{G}_s\big]\overline{\mathbb{P}}\big(A_1|\mathcal{G}_s\big)\overline{\mathbb{P}}\big(A_2|\mathcal{G}_s\big)\Big].$$

Repeating the calculation as in (5.7), (5.8) and (5.9) we achieve

$$\overline{\mathbb{E}}\big[M_{s\wedge\tau_k}\mathbb{1}_{A_0\times A_1\times A_2}\big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M_{s\wedge\tau_k}\mathbb{1}_{A_0\times\Omega\times\Omega}\big|\mathcal{G}_s\big]\overline{\mathbb{P}}\big(A_1|\mathcal{G}_s\big)\overline{\mathbb{P}}\big(A_2|\mathcal{G}_s\big)\Big],$$

which leads to (5.6).

To show that X' is an $\overline{\mathbb{F}}$-semimartingale, by the construction of the extended probability space, we can define M', V', and τ_k' to be respectively the conditional independent copy of M, V, and τ_k. Then it suffices to show that it suffices to show that

$$\mathbb{E}\big[M'_{t\wedge\tau_k'}\mathbb{1}_{A_0\times A_1\times A_2}\big] = \mathbb{E}\big[M'_{s\wedge\tau_k'}\mathbb{1}_{A_0\times A_1\times A_2}\big].$$

Similarly as the procedure of showing (5.6), we have

$$\begin{aligned}&\overline{\mathbb{E}}\big[M'_{t\wedge\tau_k'}\mathbb{1}_{A_0\times A_1\times A_2}\big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M'_{t\wedge\tau_k'}\mathbb{1}_{\Omega\times A_2\times\Omega}\big|\mathcal{G}_T\big]\overline{\mathbb{P}}(A_0|\mathcal{G}_s)\overline{\mathbb{P}}(A_2|\mathcal{G}_s)\Big]\\ &= \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_1\times\Omega\times\Omega}\big|\mathcal{G}_T\big]\overline{\mathbb{P}}(A_0|\mathcal{G}_s)\overline{\mathbb{P}}(A_2|\mathcal{G}_s)\Big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M_{t\wedge\tau_k}\mathbb{1}_{A_1\times\Omega\times\Omega}\big|\mathcal{G}_s\big]\overline{\mathbb{P}}(A_0|\mathcal{G}_s)\overline{\mathbb{P}}(A_2|\mathcal{G}_s)\Big]\\ &= \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M_{s\wedge\tau_k}\mathbb{1}_{A_1\times\Omega\times\Omega}\big|\mathcal{G}_s\big]\overline{\mathbb{P}}(A_0|\mathcal{G}_s)\overline{\mathbb{P}}(A_2|\mathcal{G}_s)\Big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M_{s\wedge\tau_k}\mathbb{1}_{A_1\times\Omega\times\Omega}\big|\mathcal{G}_T\big]\overline{\mathbb{P}}(A_0|\mathcal{G}_s)\overline{\mathbb{P}}(A_2|\mathcal{G}_s)\Big]\\ &= \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[M'_{s\wedge\tau_k'}\mathbb{1}_{\Omega\times A_2\times\Omega}\big|\mathcal{G}_T\big]\overline{\mathbb{P}}(A_0|\mathcal{G}_s)\overline{\mathbb{P}}(A_2|\mathcal{G}_s)\Big] = \mathbb{E}\big[M'_{s\wedge\tau_k'}\mathbb{1}_{A_0\times A_1\times A_2}\big],\end{aligned}$$

where the first and last equation mimic the argument in (5.7), (5.8) and (5.9), the second and the sixth equation hold since (2) in Proposition 5.1, the third and the fifth equation is because of the tower rule and the fourth equation follows from the martingale property. This completes the proof of (1).

For (2a), it suffices to prove that

$$\overline{\mathbb{E}}\Big[\Big(\sup_{\sigma_m}\sum_{i=1}^{k_m}\big|X^{\mathcal{G}}_{t^m_{i+1}}-X^{\mathcal{G}}_{t^m_i}\big|\Big)^p\Big]<\infty,$$

where the supreme is taken over all partition $\sigma_m : \{0 = t^m_0 \leq t^m_1 \leq \cdots \leq t^m_{k_m} = T\}$ of $[0, T]$. Since

$$\sum_{i=1}^{k_m}\big|X^{\mathcal{G}}_{t^m_{i+1}}-X^{\mathcal{G}}_{t^m_i}\big|=\sum_{i=1}^{k_m}\Big|\overline{\mathbb{E}}\big[X_{t^m_{i+1}}-X_{t^m_i}\big|\mathcal{G}_T\big]\Big|\leq\overline{\mathbb{E}}\Big[\sum_{i=1}^{k_m}\big|X_{t^m_{i+1}}-X_{t^m_i}\big|\Big|\mathcal{G}_T\Big],$$

we have,

$$\begin{aligned}
&\overline{\mathbb{E}}\Big[\Big(\sup_{\sigma_m}\sum_{i=1}^{k_m}\big|X^{\mathcal{G}}_{t^m_{i+1}}-X^{\mathcal{G}}_{t^m_i}\big|\Big)^p\Big]\\
&=\overline{\mathbb{E}}\Big[\sup_{\sigma_m}\Big(\sum_{i=1}^{k_m}\big|X^{\mathcal{G}}_{t^m_{i+1}}-X^{\mathcal{G}}_{t^m_i}\big|\Big)^p\Big]\leq\overline{\mathbb{E}}\Big[\sup_{\sigma_m}\Big(\overline{\mathbb{E}}\Big[\sum_{i=1}^{k_m}\big|X_{t^m_{i+1}}-X_{t^m_i}\big|\Big|\mathcal{G}_T\Big]\Big)^p\Big]\\
&\leq\overline{\mathbb{E}}\Big[\Big(\mathbb{E}\Big[\sup_{\sigma_m}\sum_{i=1}^{k_m}\big|X_{t^m_{i+1}}-X_{t^m_i}\big|\Big|\mathcal{G}_T\Big]\Big)^p\Big]\leq\overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\Big[\Big(\sup_{\sigma_m}\sum_{i=1}^{k_m}\big|X_{t^m_{i+1}}-X_{t^m_i}\big|\Big)^p\Big|\mathcal{G}_T\Big]\Big]\\
&=\overline{\mathbb{E}}\Big[\Big(\sup_{\sigma_m}\sum_{i=1}^{k_m}\big|X_{t^m_{i+1}}-X_{t^m_i}\big|\Big)^p\Big]=\overline{\mathbb{E}}\Big[\Big(\int_0^T\big|\mathrm{d}X_s\big|\Big)^p\Big]<\infty,
\end{aligned}$$

which leads to (2a).

For (2b), due to (2a) and the decomposablity (see [37, Theorem 1, Chapter 4]), it suffices to show that if X is a $\mathbb{F}$-local martingale, then $X^{\mathcal{G}}$ is a local martingale with respect to $\overline{\mathbb{F}}$, $\mathbb{F}$ and $\mathbb{G}$ and thanks to (1), it can be further reduced to prove that $X^{\mathcal{G}}$ is a local martingale with respect to $\mathbb{F}$ and $\mathbb{G}$. Let τ_k be the stopping time $\tau_k := \inf\{t \in [0, T] : [X, X]_t > k\}$ and consider the truncated process $X^{\tau_k} := \{X_{t\wedge\tau_k}\}_{t\in[0,T]}$. By Burkholder-Davis-Gundy inequality,

$$\overline{\mathbb{E}}\big[(X^{\tau_k})^*_t\big]^2\leq\overline{\mathbb{E}}\big[\big((X^{\tau_k})^*_t\big)^2\big]\leq C\overline{\mathbb{E}}\big[\big[X^{\tau_k},X^{\tau_k}\big]_t\big]\leq k, \tag{5.10}$$

for some constant C and $(X^{\tau_k})^*_t = \sup_{0\leq s\leq t}|X_{s\wedge\tau_k}|$. Let $\{\widetilde{\tau}_\ell\}_{\ell\in\mathbb{N}^+}$ be the sequence of the $\mathbb{F}$-stopping times such that $\tau_\ell \to T$ and $X^{\tilde{\tau}_\ell} = \{X_{t\wedge\tilde{\tau}_\ell}\}_{t\in[0,T]}$ are $\mathbb{F}$-martingales for all $\ell \in \mathbb{N}^+$. Then $X^{\tilde{\tau}_\ell\wedge\tau_k} = \{X_{t\wedge\tilde{\tau}_\ell\wedge\tau_k}\}_{t\in[0,T]}$ are $\mathbb{F}$-martingales for all $\ell \in \mathbb{N}^+$. Using conditional dominate convergence theorem, $|X_{t\wedge\tau_k\wedge\tilde{\tau}_\ell}| \leq (X^{\tau_k})^*_t$ and (5.10), we have

$$X_{s\wedge\tau_k}=\lim_{\ell\to\infty}X_{s\wedge\tau_k\wedge\tilde{\tau}_\ell}=\lim_{\ell\to\infty}\overline{\mathbb{E}}\big[X_{t\wedge\tau_k\wedge\tilde{\tau}_\ell}\big|\mathcal{F}_s\big]=\overline{\mathbb{E}}\Big[\lim_{\ell\to\infty}X_{t\wedge\tau_k\wedge\tilde{\tau}_\ell}\big|\mathcal{F}_s\Big]=\overline{\mathbb{E}}\big[X_{t\wedge\tau_k}\big|\mathcal{F}_s\big].$$

Combining with (5.10), we have that $X^{\tau_k} = \{X_{t\wedge\tau_k}\}_{t\in[0,T]}$ is indeed a square integrable $\mathbb{F}$-martingale.

Now we claim that $(X^{\tau_k})^{\mathcal{G}}$ is a $\mathbb{G}$-martingale, and using the property of **H** hyphothesis, and the compatiability condition of $\mathbb{F}$ and $\mathbb{G}$ (i.e. $\mathcal{F}_t \perp\!\!\!\perp \mathcal{G}_T|\mathcal{G}_t$ for all $t \in [0, T]$), it is also an $\mathbb{F}$-martingale (see [20]). The integrability holds trivially since X_t is integrable. And the martingale property holds since

$$\overline{\mathbb{E}}\Big[X^{\mathcal{G}}_{t\wedge\tau_k}\big|\mathcal{G}_s\Big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[X_{t\wedge\tau_k}|\mathcal{G}_T\big]\big|\mathcal{G}_s\Big] = \overline{\mathbb{E}}\big[X_{t\wedge\tau_k}|\mathcal{G}_s\big] = \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[X_{t\wedge\tau_k}|\mathcal{F}_s\big]\big|\mathcal{G}_s\Big]$$
$$= \overline{\mathbb{E}}\big[X_{s\wedge\tau_k}|\mathcal{G}_s\big] = \overline{\mathbb{E}}\big[X_{s\wedge\tau_k}|\mathcal{G}_T\big] = X^{\mathcal{G}}_{s\wedge\tau_k}.$$

This completes the proof of the claim that $X^{\mathcal{G}}$ is an $\mathbb{G}$-martingale and thus leads to(2b).

For (2c), let $\sigma_m : 0 = t_0^m \leq t_1^m \leq ... \leq t_{k_m}^m = t$ be a sequence of partitions tend to identity (the definition of a sequence tending to identity can be found in the [37, Page 64, defintion] with a slight modification of the terminal time), such that

$$\int_0^t Z_{s-}\mathrm{d}X^{\mathcal{G}}_s = \lim_{m\to\infty}\sum_{k=0}^{k_m-1} Z_{t_k^m}(X^{\mathcal{G}}_{t^m_{k+1}} - X^{\mathcal{G}}_{t^m_k}), \quad \int_0^t Z_{s-}\mathrm{d}X'_s = \lim_{m\to\infty}\sum_{k=0}^{k_m-1} Z_{t_k^m}(X'_{t^m_{k+1}} - X'_{t^m_k}),$$

where the convergence is understood as almost surely convergence. Such partition can be found thanks to [37, Theorem 21, Chapter II]. Then

$$\int_0^t Z_{s-}\mathrm{d}X^{\mathcal{G}}_s = \lim_{m\to\infty}\sum_{k=0}^{k_m-1} Z_{t_k^m}(X^{\mathcal{G}}_{t^m_{k+1}} - X^{\mathcal{G}}_{t^m_k}) = \lim_{m\to\infty}\sum_{k=0}^{k_m-1} Z_{t_k^m}\mathbb{E}\Big[X_{t^m_{k+1}} - X_{t^m_k}\Big|\mathcal{G}_T\Big]$$
$$= \lim_{m\to\infty}\sum_{k=0}^{k_m-1} Z_{t_k^m}\overline{\mathbb{E}}\Big[X'_{t^m_{k+1}} - X'_{t^m_k}\Big|\mathcal{F}\Big] = \lim_{m\to\infty}\sum_{k=0}^{k_m-1}\mathbb{E}\Big[Z_{t_k^m}(X'_{t^m_{k+1}} - X'_{t^m_k})\Big|\mathcal{F}\Big].$$

Define Z^m as the process

$$Z^m_t = Z_0\mathbb{1}_{\{0\}}(t) + \sum_{k=0}^{k_m-1} Z_{t_k^m}\mathbb{1}_{(t_k^m, t_{k+1}^m]}(t),$$

then $\int_0^t Z_{s-}\mathrm{d}X^{\mathcal{G}}_s = \lim_{m\to\infty}\mathbb{E}\big[\int_0^t Z^m_s\mathrm{d}X'_s\big|\mathcal{F}\big]$, and

$$\sum_{k=0}^{k_m-1}\overline{\mathbb{E}}\Big[Z_{t_k^m}\big(X'_{t^m_{k+1}} - X'_{t^m_k}\big)\Big] - \overline{\mathbb{E}}\Big[\int_0^t Z_{s-}\mathrm{d}X'_s\Big] = \overline{\mathbb{E}}\Big[\int_0^t (Z^m_s - Z_{s-})\mathrm{d}X'_s\Big].$$

Recall that X' has the decomposition $X' = M' + V'$ where M' is a local martingale and V' being a adapted finite variation process and

$$\Big\|\sqrt{[M', M']_T} + \int_0^T \big|\mathrm{d}V'_s\big|\Big\|_{L^p} < \infty.$$

Therefore,

$$\overline{\mathbb{E}}\Big[\Big|\int_0^t (Z^m_s - Z''_{s-})\mathrm{d}X'_s\Big|\Big] \leq \overline{\mathbb{E}}\Big[\Big|\int_0^t (Z^m_s - Z_{s-})\mathrm{d}M'_s\Big|\Big] + \overline{\mathbb{E}}\Big[\Big|\int_0^t (Z^m_s - Z_{s-})\mathrm{d}V'_s\Big|\Big]. \tag{5.11}$$

For the first term in (5.11), we use Burkholder-Davis-Gundy inequality so that

$$\overline{\mathbb{E}}\Big[\Big|\int_0^t (Z^m_s - Z_{s-})\mathrm{d}M'_s\Big|\Big] \leq \overline{\mathbb{E}}\Big[\sup_{0\leq s\leq t}\Big|\int_0^s (Z^m_u - Z_{u-})\mathrm{d}M'_u\Big|^p\Big]$$
$$\leq c_1\overline{\mathbb{E}}\Big[\Big(\int_0^t (Z^m_u - Z_{u-})^2\mathrm{d}[M', M']_s\Big)^{p/2}\Big]$$

for some constant c_1. Note that $\int_0^t (Z^m_u - Z_{u-})^2\mathrm{d}[M', M']_s \leq 4\sup_{0\leq s\leq t}(Z_s)^2[M', M']_t$, and

$$\overline{\mathbb{E}}\Big[\Big(\sup_{0\leq s\leq t}(Z_s)^2[M', M']_t\Big)^{1/2}\Big] \leq \Big(\overline{\mathbb{E}}\Big[\sup_{0\leq s\leq t}(Z_s)^q\Big]\overline{\mathbb{E}}\big[[M', M']_t^{p/2}\big]\Big)^{1/2} \leq C\sqrt{\|Z\|^q_{\mathcal{H}^q}\overline{\mathbb{E}}\big[[M', M']_t^{p/2}\big]} < \infty$$

Since Z is a càdlàg function, we know that $Z^m_s \to Z_{s-}, a.s.$ and $|Z^m_s - Z_s| \le \sup_{0\le u\le s}|Z_u| \in S_q$, and we can use the dominated convergence theorem (the classical one and the one in $\mathcal{H}_p$, see [37, Page 273, Lemma]) to get

$$\overline{\mathbb{E}}\Big[\int_0^t (Z^m_s - Z_{s-})\mathrm{d}M'_s\Big] \to 0.$$

For the second term in (5.11), we have

$$\overline{\mathbb{E}}\Big[\Big|\int_0^t (Z^m_s - Z_{s-})\mathrm{d}V'_s\Big|\Big] \le \overline{\mathbb{E}}\Big[\int_0^t |Z^m_s - Z_{s-}|\big|\mathrm{d}V'_s\big|\Big].$$

Noting that $\int_0^t |Z^m_s - Z_{s-}|\big|\mathrm{d}V'_s\big| \le 2\sup_{0\le s\le t}|Z_s|\int_0^t |\mathrm{d}V_t|$, and similarly, we have

$$\overline{\mathbb{E}}\Big[\Big|\int_0^t (Z^m_s - Z_{s-})\mathrm{d}V'_s\Big|\Big] \to 0,$$

and therefore,

$$\lim_{m\to\infty}\overline{\mathbb{E}}\Big[\Big|\int_0^t Z^m_s \mathrm{d}X'_s - \int_0^t Z_{s-}\mathrm{d}X'_s\Big|\Big] = 0.$$

Now define

$$R^m := \overline{\mathbb{E}}\Big[\int_0^t Z^m_{s-}\mathrm{d}X'_s\Big|\mathcal{F}\Big],\ R := \overline{\mathbb{E}}\Big[\int_0^t Z_{s-}\mathrm{d}X'_s\Big|\mathcal{F}\Big],$$

for all $A\in\mathcal{F}$, we have

$$\overline{\mathbb{E}}[(R-R_m)\mathbb{1}_A] = \overline{\mathbb{E}}\Big[\Big(\int_0^t Z^m_s\mathrm{d}X'_s - \int_0^t Z_{s-}\mathrm{d}X'_s\Big)\mathbb{1}_A\Big] \le \overline{\mathbb{E}}\Big[\Big|\int_0^t Z^m_s\mathrm{d}X'_s - \int_0^t Z_{s-}\mathrm{d}X'_s\Big|\Big] \to 0.$$

Let $A = \{R - R_m > \epsilon\}$, we have

$$\overline{\mathbb{P}}\big(R_m - R > \epsilon\big) \le \frac{\overline{\mathbb{E}}[(R_m - R)\mathbb{1}_A]}{\epsilon} \to 0,$$

and similarly

$$\overline{\mathbb{P}}\big(R_m - R < \epsilon\big) \to 0.$$

Therefore $R_m \to R$ in probability. Meanwhile $R_m \to \int_0^t Z_{s-}\mathrm{d}X^{\mathcal{G}}_s$ almost surely. Therefore,

$$\int_0^t Z_{s-}\mathrm{d}X^{\mathcal{G}}_s = \overline{\mathbb{E}}\Big[\int_0^t Z^m_{s-}\mathrm{d}X'_s\Big|\mathcal{F}\Big].$$

For (2d), let us first prove that

$$[X^{\mathcal{G}}, Y^{\mathcal{G}}]_t = \overline{\mathbb{E}}\Big[[X'', Y']_t\Big|\mathcal{G}\Big].$$

By the definition of the quadratic variation,

$$\big[X^{\mathcal{G}}, Y^{\mathcal{G}}\big]_t = X^{\mathcal{G}}_t Y^{\mathcal{G}}_t - \int_0^t X^{\mathcal{G}}_{s-}\mathrm{d}Y^{\mathcal{G}}_s - \int_0^t Y^{\mathcal{G}}_{s-}\mathrm{d}X^{\mathcal{G}}_s,$$

where the last two terms are well defined because of (2b). The first term $X^{\mathcal{G}}_t Y^{\mathcal{G}}_t$ can be calculated by

$$X^{\mathcal{G}}_t Y^{\mathcal{G}}_t = \overline{\mathbb{E}}\big[X_t|\mathcal{G}_T\big]\overline{\mathbb{E}}\big[Y_t|\mathcal{G}_T\big] = \overline{\mathbb{E}}\big[X''_t|\mathcal{F}\big]\overline{\mathbb{E}}\big[Y'_t|\mathcal{F}\big] = \overline{\mathbb{E}}\big[X''_t Y'_t|\mathcal{F}\big],$$

where the last equation holds due to (3) in Proposition 5.1. For the second term $\int_0^t X^{\mathcal{G}}_{s-}\mathrm{d}Y^{\mathcal{G}}_s$, let $\sigma_m : 0 = t^m_0 \leq t^m_1 \leq ... \leq t^m_{k_m} = t$ be a sequence of partitions tend to identity in the sense of [37, Page 64, defintion], with a slight modification of the terminal time, such that

$$\int_0^t X^{\mathcal{G}}_{s-}\mathrm{d}Y^{\mathcal{G}}_s = \lim_{m\to\infty}\sum_{k=0}^{k_m-1} X^{\mathcal{G}}_{t^m_k}(Y^{G}_{t^m_{k+1}} - Y^{\mathcal{G}}_{t^m_k}),\quad \int_0^t X''_{s-}\mathrm{d}Y'_s = \lim_{m\to\infty}\sum_{k=0}^{k_m-1} X''_{t^m_k}(Y'_{t^m_{k+1}} - Y'_{t^m_k}),$$

where the convergence is understood as almost surely convergence. Therefore,

$$\begin{aligned}\int_0^t X^{\mathcal{G}}_{s-}\mathrm{d}Y^{\mathcal{G}}_s &= \lim_{m\to\infty}\sum_{k=0}^{k_m-1} X^{\mathcal{G}}_{t^m_k}(Y^{\mathcal{G}}_{t^m_{k+1}} - Y^{\mathcal{G}}_{t^m_k}) = \lim_{m\to\infty}\sum_{k=0}^{k_m-1}\mathbb{E}\big[X_{t^m_k}|\mathcal{G}_T\big]\overline{\mathbb{E}}\big[Y_{t^m_{k+1}} - Y_{t^m_k}|\mathcal{G}_T\big]\\ &= \lim_{m\to\infty}\sum_{k=0}^{k_m-1}\mathbb{E}\big[X''_{t^m_k}|\mathcal{F}\big]\mathbb{E}\big[Y'_{t^m_{k+1}} - Y'_{t^m_k}|\mathcal{F}\big] = \lim_{m\to\infty}\sum_{k=0}^{k_m-1}\overline{\mathbb{E}}\big[X''_{t^m_k}\big(Y'_{t^m_{k+1}} - Y'_{t^m_k}\big)|\mathcal{F}\big],\end{aligned}$$

where the last equation holds due to (3) in Proposition 5.1. Similar as the proof of (2c), we can conclude that

$$\int_0^t X^{\mathcal{G}}_{s-}\mathrm{d}Y^{\mathcal{G}}_s = \overline{\mathbb{E}}\bigg[\int_0^t X''_{s-}\mathrm{d}Y'_s\bigg|\mathcal{F}\bigg],$$

and similarly

$$\int_0^t Y^{\mathcal{G}}_{s-}\mathrm{d}X^{\mathcal{G}}_s = \overline{\mathbb{E}}\bigg[\int_0^t Y'_{s-}\mathrm{d}X''_s\bigg|\mathcal{F}\bigg].$$

Therefore

$$[X^{\mathcal{G}}, Y^{\mathcal{G}}]_t = \overline{\mathbb{E}}\Big[[X'', Y']_t\Big|\mathcal{F}\Big].$$

Now let $\sigma_m : 0 = t^m_0 \leq t^m_1 \leq ... \leq t^m_{k_m} = t$ being a sequence of partitions tend to identity in the sense of [37, Page 64, defintion] with a slight modification of the terminal time, such that

$$\int_0^t Z_{s-}\mathrm{d}[X^{\mathcal{G}}, Y^{\mathcal{G}}]_s = \lim_{m\to\infty}\sum_{k=0}^{k_m-1} Z_{t^m_k}\big([X^{\mathcal{G}}, Y^{\mathcal{G}}]_{t^m_{k+1}} - [X^{\mathcal{G}}, Y^{\mathcal{G}}]_{t^m_k}\big),$$

and

$$\int_0^t Z_{s-}\mathrm{d}[X'', Y']_s = \lim_{m\to\infty}\sum_{k=0}^{k_m-1} Z_{t^m_k}([X'', Y']_{t^m_{k+1}} - [X'', Y']_{t^m_k}),$$

To repeat the same argument as before, we need to notice that $\big|\int_0^t(Z^m_s - Z_{s-})\mathrm{d}[X'', Y']_s\big| \leq 2\sup_{0\leq s\leq t}|Z_s|[X'', Y']_t$ and

$$\begin{aligned}\overline{\mathbb{E}}\bigg[\sup_{0\leq s\leq t}|Z_s|[X'', Y']_t\bigg] &\leq \overline{\mathbb{E}}[(\sup_{0\leq s\leq t}|Z_s|)^r]^{1/r}\overline{\mathbb{E}}\Big[[X'', X'']^p_t\Big]^{1/p}\overline{\mathbb{E}}\Big[[Y', Y']^q_t\Big]^{1/q}\\ &\leq C\|Z\|_{H_r}\|X\|_{H_p}\|Y\|_{H_q}\end{aligned},$$

where the first inequality is due to Cauchy inequality and Kunita-Watanabe Inequality and C is a generic constant. Now repeating the arguments in (2c), we achieve

$$\int_0^t Z_{s-}\mathrm{d}\big[X^{\mathcal{G}}, Y^{\mathcal{G}}\big]_t = \overline{\mathbb{E}}\bigg[\int_0^t Z_{s-}\mathrm{d}\big[X'', Y'\big]_s\bigg|\mathcal{F}\bigg],$$

and similarly, we can get

$$\int_0^t Z_{s-}\big[X^{\mathcal{G}}, Y\big]_s = \overline{\mathbb{E}}\bigg[\int_0^t Z_{s-}\big[X', Y\big]_s\bigg|\mathcal{F}\bigg],$$

which completes the proof. $\square$

References

[1] Saran Ahuja. Wellposedness of mean field games with common noise under a weak monotonicity condition. *SIAM J. Control Optim.*, 54(1):30–48, 2016.

[2] Saran Ahuja, Weiluo Ren, and Tzu-Wei Yang. Forward-backward stochastic differential equations with monotone functionals and mean field games with common noise. *Stochastic Process. Appl.*, 129(10):3859–3892, 2019.

[3] Erhan Bayraktar and Yuchong Zhang. A rank-based mean field game in the strong formulation. *Electron. Commun. Probab.*, 21:Paper No. 72, 12, 2016.

[4] Mathias Beiglböck and Daniel Lacker. Denseness of adapted processes among causal couplings. *arXiv preprint arXiv:1805.03185*, 2018.

[5] Alain Bensoussan, Jens Frehse, and Sheung Chi Phillip Yam. The master equation in mean field theory. *J. Math. Pures Appl. (9)*, 103(6):1441–1474, 2015.

[6] Rainer Buckdahn, Juan Li, Shige Peng, and Catherine Rainer. Mean-field stochastic differential equations and associated pdes. *Annals of probability*, 45(2):824–878, 2017.

[7] Matteo Burzoni, Vincenzo Ignazio, A Max Reppen, and H Mete Soner. Viscosity solutions for controlled mckean–vlasov jump-diffusions. *SIAM Journal on Control and Optimization*, 58(3):1676–1699, 2020.

[8] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. *The master equation and the convergence problem in mean field games:(ams-201)*. Princeton University Press, 2019.

[9] René Carmona, François Delarue, and Daniel Lacker. Mean field games with common noise. *Ann. Probab.*, 44(6):3740–3803, 2016.

[10] René Carmona and Francois Delarue. The master equation for large population equilibriums. In *Stochastic Analysis and Applications 2014: In Honour of Terry Lyons*, pages 77–128. Springer, 2014.

[11] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications I-II*. Springer, 2018.

[12] René Carmona, François Delarue, and Daniel Lacker. Mean field games with common noise. *Annals of Probability*, 44(6):3740–3803, 2016.

[13] René Carmona, Jean-Pierre Fouque, and Li-Hsien Sun. Mean field games and systemic risk. *Commun. Math. Sci.*, 13(4):911–933, 2015.

[14] Jean-François Chassagneux, Dan Crisan, and François Delarue. *A probabilistic approach to classical solutions of the master equation for large population equilibria*, volume 280. American Mathematical Society, 2022.

[15] Michele Coghi and Franco Flandoli. Propagation of chaos for interacting particles subject to environmental noise. *Ann. Appl. Probab.*, 26(3):1407–1442, 2016.

[16] Alexander MG Cox, Sigrid Källblad, Martin Larsson, and Sara Svaluto-Ferro. Controlled measure-valued martingales: a viscosity solution approach. *arXiv preprint arXiv:2109.00064*, 2021.

[17] Christa Cuchiero, Martin Larsson, and Sara Svaluto-Ferro. Probability measure-valued polynomial diffusions. 2019.

[18] Donald Dawson and Jean Vaillancourt. Stochastic McKean-Vlasov equations. *NoDEA Nonlinear Differential Equations Appl.*, 2(2):199–229, 1995.

[19] Gonçalo Dos Reis and Vadim Platonov. Itô-Wentzell-Lions formula for measure dependent random fields under full and conditional measure flows. *Potential Analysis*, 59(3):1313–1344, 2023.

[20] Robert J Elliott, Monique Jeanblanc, and Marc Yor. On models of default risk. *Mathematical Finance*, 10(2):179–195, 2000.

[21] Wendell H Fleming and Michel Viot. Some measure-valued markov processes in population genetics theory. *Indiana University Mathematics Journal*, 28(5):817–843, 1979.

[22] Guanxing Fu and Ulrich Horst. Mean field games with singular controls. *SIAM Journal on Control and Optimization*, 55(6):3833–3868, 2017.

[23] Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions. Mean field games and applications. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 205–266. Springer, Berlin, 2011.

[24] Xin Guo, Huyên Pham, and Xiaoli Wei. Itô's formula for flows of measures on semimartingales. *Stochastic Processes and their Applications*, 159:350–390, 2023.

[25] Mokhtar Hafayed, Abdelmadjid Abba, and Syed Abbas. On mean-field stochastic maximum principle for near-optimal controls for Poisson jump diffusion with applications. *International Journal of Dynamics and Control*, 2:262–284, 2014.

[26] Mokhtar Hafayed, Shahlar Meherrem, Şaban Eren, and Deniz Hasan Guçoglu. On optimal singular control problem for general Mckean-Vlasov differential equations: Necessary and sufficient optimality conditions. *Optimal Control Applications and Methods*, 39(3):1202–1219, 2018.

[27] William R. P. Hammersley, David Siska, and Lukasz Szpruch. Weak existence and uniqueness for McKean-Vlasov SDEs with common noise. *arXiv preprint arXiv:1908.00955*, 2019.
[28] Yaozhong Hu, Bernt Øksendal, and Agnes Sulem. Singular mean-field control games. *Stochastic Analysis and Applications*, 35(5):823–851, 2017.
[29] Kiyosi Itô. On a formula concerning stochastic differentials. *Nagoya Mathematical Journal*, 3:55–65, 1951.
[30] Vassili N. Kolokoltsov and Marianna Troeva. On mean field games with common noise and Mckean-Vlasov spdes. *Stochastic Analysis and Applications*, 37(4):522–549, 2019.
[31] Peter Kotelenez. A class of quasilinear stochastic partial differential equations of McKean-Vlasov type with mass conservation. *Probab. Theory Related Fields*, 102(2):159–188, 1995.
[32] Thomas G Kurtz and Jie Xiong. Particle representations for a class of nonlinear spdes. *Stochastic Processes and their Applications*, 83(1):103–126, 1999.
[33] Daniel Lacker. A general characterization of the mean field limit for stochastic differential games. *Probab. Theory Related Fields*, 165(3-4):581–648, 2016.
[34] Daniel Lacker, Mykhaylo Shkolnikov, and Jiacheng Zhang. Superposition and mimicking theorems for conditional Mckean–Vlasov equations. *Journal of the European Mathematical Society*, 2022.
[35] Daniel Lacker and Kevin Webster. Translation invariant mean field games with common noise. *Electron. Commun. Probab.*, 20:no. 42, 13, 2015.
[36] Juan Li. Mean-field forward and backward sdes with jumps and associated nonlocal quasi-linear integral-pdes. *Stochastic Processes and their Applications*, 128(9):3118–3180, 2018.
[37] Philip E Protter. *Stochastic differential equations*. Springer, 2005.
[38] Goncalo dos Reis and Vadim Platonov. On the relation between Stratonovich and Itô integrals with functional integrands of conditional measure flows. *arXiv preprint arXiv:2111.03523*, 2021.
[39] Mehdi Talbi, Nizar Touzi, and Jianfeng Zhang. Dynamic programming equation for the mean field optimal stopping problem. *SIAM Journal on Control and Optimization*, 61(4):2140–2164, 2023.
[40] Jean Vaillancourt. On the existence of random McKean-Vlasov limits for triangular arrays of exchangeable diffusions. *Stochastic Anal. Appl.*, 6(4):431–446, 1988.

Department of Industrial Engineering and Operations Research, University of California, Berkeley, Berkeley, California, USA.
Email address: `xinguo@berkeley.edu`

Department of Industrial Engineering and Operations Research, University of California, Berkeley, Berkeley, California, USA.
Email address: `jiachengz@berkeley.edu`