Supplementary material for “Continuous time reinforcement learning: A random measure approach”

Christian Bender* Nguyen Tran Thuan†‡

November 19, 2024

Abstract

This document contains supplementary materials for the main article [1]. All notations used here are in accordance with that in [1].

1 SDEs driven by orthogonal martingale measures

1.1 Martingale measures and integration

We briefly recall the notion of martingale measure initiated by Walsh [8]. We consider here the finite time interval $[0,T]$ but note that the discussion below can be readily extended for $[0,\infty)$. Let (E,d_E) be a complete and separable metric space equipped with its Borel σ-field $\mathcal{B}(E)$. A mapping $M\colon \Omega\times[0,T]\times\mathcal{B}(E)\to\mathbb{R}$ is called an $(\mathbb{F},\mathbb{P})$-*martingale measure* on $[0,T]\times\mathcal{B}(E)$ if:

1. For $A\in\mathcal{B}(E)$, $(M(t,A))_{t\in[0,T]}$ is an $\mathbf{L}^2(\mathbb{P})$-martingale adapted with $\mathbb{F}$ and $M(0,A)=0$;
2. For $t\in[0,T]$ and disjoint $A,B\in\mathcal{B}(E)$, one has $M(t,A\cup B)=M(t,A)+M(t,B)$ a.s.;
3. There exists a non-decreasing sequence $(E_n)_{n\in\mathbb{N}}\subseteq\mathcal{B}(E)$ with $\cup_{n\in\mathbb{N}}E_n=E$ such that
   (a) For any $n\in\mathbb{N}$, $\sup_{A\in\mathcal{B}(E_n)}\|M(T,A)\|_{\mathbf{L}^2(\mathbb{P})}<\infty$;
   (b) For any $n\in\mathbb{N}$, one has $\|M(T,A_k)\|_{\mathbf{L}^2(\mathbb{P})}\to 0$ for all decreasing sequence $(A_k)_{k\in\mathbb{N}}\subseteq\mathcal{B}(E_n)$ with $\cap_{k\in\mathbb{N}}A_k=\emptyset$.

An $(\mathbb{F},\mathbb{P})$-martingale measure M is said to be *continuous* if $[0,T]\ni t\mapsto M(t,A)$ is continuous for all $A\in\mathcal{B}(E)$. Note that, due to the usual conditions, we always choose the càdlàg version of the martingale $M(\cdot,A)$ for any $A\in\mathcal{B}(E)$.

An $(\mathbb{F},\mathbb{P})$-martingale measure M is *orthogonal* if $M(\cdot,A)M(\cdot,B)$ is an $(\mathbb{F},\mathbb{P})$-martingale for any disjoint $A,B\in\mathcal{B}(E)$. It is indicated by Walsh [8] (see also [5, Theorem I-4]) that if an $(\mathbb{F},\mathbb{P})$-martingale measure M is orthogonal, then there is a random positive finite measure μ_M on $\mathcal{B}([0,T]\times E)$, which is $\mathbb{F}$-predictable (i.e. $(\mu_M((0,t]\times A))_{t\in[0,T]}$ is $\mathbb{F}$-predictable for all $A\in\mathcal{B}(E)$), such that

$$\mu_M((0,t]\times A)=\langle M(\cdot,A)\rangle_t \quad \mathbb{P}\text{-a.s.},\quad \forall(t,A)\in[0,T]\times\mathcal{B}(E).$$

The measure μ_M is then called the *intensity measure* of M. Moreover, for $t\in[0,T]$, $A,B\in\mathcal{B}(E)$,

$$\langle M(\cdot,A),M(\cdot,B)\rangle_t=\langle M(\cdot,A\cap B)\rangle_t=\mu_M((0,t]\times(A\cap B))\quad \mathbb{P}\text{-a.s.}$$

The stochastic integrals driven by an orthogonal martingale measure M can be constructed via the Itô’s approach (see [5, 8]) as follows. Let $\mathbf{L}^2(\mathbb{F},\mu_M)$ be the collection of all $\mathbb{F}$-predictable H

*Department of Mathematics, Saarland University, Germany. Email: bender@math.uni-saarland.de
†Department of Mathematics, Saarland University, Germany. Email: nguyen@math.uni-saarland.de
‡Department of Mathematics, Vinh University, Vinh, Nghe An, Viet Nam. Email: thuannt@vinhuni.edu.vn

(i.e., H is $\mathcal{P}_{\mathbb{F}} \otimes \mathcal{B}(E)/\mathcal{B}(\mathbb{R})$-measurable) with $\mathbb{E}\big[\int_0^T \int_E H(t,x)^2 \mu_M(\mathrm{d}t, \mathrm{d}x)\big] < \infty$. For a simple function $H(\omega,t,x) = \sum_{i=1}^n h_{i-1}(\omega)\mathbb{1}_{(t_{i-1},t_i]}(t)\mathbb{1}_{A_i}(x)$ where $A_i \in \mathcal{B}(E)$, $0 \le t_0 < t_1 < \cdots < t_n \le T$, h_{i-1} is bounded and $\mathcal{F}_{t_{i-1}}$-measurable, $n \in \mathbb{N}$, we define

$$H \bullet M(t,A) := \sum_{i=1}^n h_{i-1}[M(t_i \wedge t, A \cap A_i) - M(t_{i-1} \wedge t, A \cap A_i)], \quad (t,A) \in [0,T] \times \mathcal{B}(E).$$

Then, it is clear that $H \bullet M$ is an $(\mathbb{F},\mathbb{P})$-martingale measure and satisfies the isometry

$$\mathbb{E}[|H \bullet M(t,A)|^2] = \mathbb{E}\left[\int_0^T \int_A H(t,x)^2 \mu_M(\mathrm{d}t, \mathrm{d}x)\right]. \tag{1.1}$$

As the family of simple functions is dense in $\mathbf{L}^2(\mathbb{F}, \mu_M)$, one can extend $H \bullet M$ for $H \in \mathbf{L}^2(\mathbb{F}, \mu_M)$ as usual to obtain a martingale measure which is also orthogonal with intensity $\mu_{H \bullet M}(\mathrm{d}t, \mathrm{d}x) = H(t,x)^2 \mu_M(\mathrm{d}t, \mathrm{d}x)$, see [5, Theorem I-6]. Moreover, (1.1) then also holds for $H \in \mathbf{L}^2(\mathbb{F}, \mu_M)$. In the sequel, we apply the integral notation

$$\int_{(0,t] \times E} H(s,x) M(\mathrm{d}s, \mathrm{d}x) := H \bullet M(t,E), \quad t \in [0,T].$$

Assume that the intensity μ_M of an orthogonal martingale measure M satisfies $\mu_M(\{t\} \times E) = 0$ for all $t \in [0,T]$ a.s. Then, by a localization argument, one can extend the stochastic integrals driven by M for $H \in \mathbf{L}^2_{\mathrm{loc}}(\mathbb{F}, \mu_M)$, where $\mathbf{L}^2_{\mathrm{loc}}(\mathbb{F}, \mu_M)$ consists of all $\mathbb{F}$-predictable H with $\int_0^T \int_E H(t,x)^2 \mu_M(\mathrm{d}t, \mathrm{d}x) < \infty$ a.s. (see, e.g., [6, Chapter 13] for the case of continuous M). Namely, we let $(\tau_n)_n$ be the localizing sequence given by $\tau_0 := 0$ and $\tau_n := \inf\big\{t > \tau_{n-1} : \int_0^t \int_E H(s,x)^2 \mu_M(\mathrm{d}s, \mathrm{d}x) > n\big\} \wedge T$, which are non-decreasing $\mathbb{F}$-stopping times and eventually constant T a.s., and define

$$\int_{(0,T] \times E} H(t,x) M(\mathrm{d}t, \mathrm{d}x) := \lim_{n \to \infty} \int_{(0,T] \times E} H(t,x) \mathbb{1}_{(0,\tau_n]}(t) M(\mathrm{d}t, \mathrm{d}x),$$

where the limit is taken in probability. One notes that the limit does not depend on the choice of the localizing sequence $(\tau_n)_n$.

Lemma 1.1. *Let M be an orthogonal $(\mathbb{F},\mathbb{P})$-martingale measure with intensity μ_M. Then, for any $\mathbb{F}$-stopping times $\sigma, \tau \colon \Omega \to [0,T]$ with $\sigma \le \tau$, $A \in \mathcal{B}(E)$, and any bounded $\mathcal{F}_\sigma$-measurable $h \colon \Omega \to \mathbb{R}$, one has, a.s.,*

$$\int_{(0,T] \times E} h \mathbb{1}_{(\sigma,\tau]}(s) \mathbb{1}_A(e) M(\mathrm{d}s, \mathrm{d}e) = h[M(\tau, A) - M(\sigma, A)].$$

Proof. It suffices to prove the assertion when $\tau = T$. We note that the stochastic integral above is defined in $\mathbf{L}^2(\mathbb{P})$ as the integrand is $\mathbb{F}$-predictable and bounded which obviously belongs to $\mathbf{L}^2(\mathbb{F}, \mu_M)$. Let $(\sigma_n)_{n \in \mathbb{N}}$ be a decreasing sequence of $\mathbb{F}$-stopping times taking finitely many values in $[0,T]$ such that $\sigma_n \to \sigma$ when $n \to \infty$. Assume $\sigma_n(\Omega) = \{s_1^n, \ldots, s_{k_n}^n\} \subset [0,T]$ with $s_{j-1}^n < s_j^n$. Since $h\mathbb{1}_{(\sigma_n,T]}\mathbb{1}_A \to h\mathbb{1}_{(\sigma,T]}\mathbb{1}_A$ in $\mathbf{L}^2(\mathbb{F}, \mu_M)$, we only need to show the assertion for σ_n in place of σ. Indeed, one has, a.s.,

$$\begin{aligned}
\int_{(0,T] \times E} h \mathbb{1}_{(\sigma_n,T]}(s) \mathbb{1}_A(e) M(\mathrm{d}s, \mathrm{d}e) &= \sum_{j=1}^{k_n} \int_{(0,T] \times E} h \mathbb{1}_{\{\sigma_n = s_j^n\}} \mathbb{1}_{(s_j^n,T]}(s) \mathbb{1}_A(e) M(\mathrm{d}s, \mathrm{d}e) \\
&= \sum_{j=1}^{k_n} h \mathbb{1}_{\{\sigma_n = s_j^n\}} [M(T,A) - M(s_j^n, A)] \\
&= h[M(T,A) - M(\sigma_n, A)]
\end{aligned}$$

where we note that $h\mathbb{1}_{\{\sigma_n = s_j^n\}}$ is $\mathcal{F}_{s_j^n}$-measurable and bounded. □

Lemma 1.2. *Let M be an orthogonal $(\mathbb{F},\mathbb{P})$-martingale measure with intensity $\mu_M(\mathrm{d}s,\mathrm{d}e) = \mu_s(\mathrm{d}e)\mathrm{d}s$ for some transition kernel $\{(\omega,s,A)\mapsto \mu_s(\omega,A), (\omega,s)\in\Omega\times[0,T], A\in\mathcal{B}(E)\}$. For an $\mathbb{F}$-stopping time $\tau\colon\Omega\to[0,T]$, we define $\mathbb{F}^\tau=(\mathcal{F}_t^\tau)_{t\in[0,T]}$ with $\mathcal{F}_t^\tau:=\mathcal{F}_{(\tau+t)\wedge T}$ and*

$$M_\tau(t,A) := M((\tau+t)\wedge T, A) - M(\tau,A), \quad (t,A)\in[0,T]\times\mathcal{B}(E).$$

Then, M_τ is an orthogonal $(\mathbb{F}^\tau,\mathbb{P})$-martingale measure with ($\mathbb{F}^\tau$-predictable) intensity $\mu_{M_\tau}(\mathrm{d}s,\mathrm{d}e) = \mathbb{1}_{(0,T-\tau]}(s)\mu_{\tau+s}(\mathrm{d}e)\mathrm{d}s$. Moreover, for $g\colon\Omega\times[0,T]\times E\to\mathbb{R}$ with $\{(\omega,s,e)\mapsto \mathbb{1}_{(\tau(\omega),T]}(s)g(\omega,s,e)\}\in\mathbf{L}^2_{\mathrm{loc}}(\mathbb{F},\mu_M)$, one has $\{(\omega,s,e)\mapsto \mathbb{1}_{(0,T-\tau(\omega)]}(s)g(\omega,\tau(\omega)+s,e)\}\in\mathbf{L}^2_{\mathrm{loc}}(\mathbb{F}^\tau,\mu_{M_\tau})$ and, a.s.,

$$\int_{(0,T]\times E} g(s,e)\mathbb{1}_{(\tau,T]}(s)M(\mathrm{d}s,\mathrm{d}e) = \int_{(0,T]\times E} g(\tau+s,e)\mathbb{1}_{(0,T-\tau]}(s)M_\tau(\mathrm{d}s,\mathrm{d}e), \tag{1.2}$$

where the stochastic integrals in the left-hand side and the right-hand side are constructed in relation to $\mathbb{F}$ and $\mathbb{F}^\tau$, respectively.

Proof. It is clear that $\mathbb{F}^\tau$ satisfies the usual conditions. According to the optional stopping theorem, one can readily show that M_τ is an orthogonal $(\mathbb{F}^\tau,\mathbb{P})$-martingale measure with ($\mathbb{F}^\tau$-predictable) intensity

$$\begin{aligned}\mu_{M_\tau}((0,t]\times A) &= \langle M_\tau(\cdot,A)\rangle_t = \langle M(\cdot,A)\rangle_{(\tau+t)\wedge T} - \langle M(\cdot,A)\rangle_\tau \\ &= \int_{(0,T]\times A}\mathbb{1}_{(\tau,(\tau+t)\wedge T]}(s)\mu_M(\mathrm{d}s,\mathrm{d}e) = \int_{(0,t]\times A}\mathbb{1}_{(0,T-\tau]}(s)\mu_{\tau+s}(\mathrm{d}e)\mathrm{d}s\end{aligned}$$

where the last equality is due to $((\tau+t)\wedge T)-\tau = (T-\tau)\wedge t$. In other words,

$$\mu_{M_\tau}(\mathrm{d}s,\mathrm{d}e) = \mathbb{1}_{(0,T-\tau]}(s)\mu_{\tau+s}(\mathrm{d}e)\mathrm{d}s.$$

For the "Moreover" part, we note that $\mathbb{1}_{(0,T-\tau]}g_\tau$ is $\mathbb{F}^\tau$-predictable, where $g_\tau(\omega,s,e):=g(\omega,\tau(\omega)+s,e)$. It readily follows from a change of variables that

$$\int_0^T\int_E \mathbb{1}_{(\tau,T]}(s)|g(s,e)|^2\mu_s(\mathrm{d}e)\mathrm{d}s = \int_0^T\int_E \mathbb{1}_{(0,T-\tau]}(s)|g(\tau+s,e)|^2\mu_{\tau+s}(\mathrm{d}e)\mathrm{d}s,$$

which yields $\mathbb{1}_{(0,T-\tau]}g_\tau\in\mathbf{L}^2_{\mathrm{loc}}(\mathbb{F}^\tau,\mu_{M_\tau})$. For (1.2), let us first consider $g(\omega,s,e)=h_a(\omega)\mathbb{1}_{(a,b]}(s)\mathbb{1}_A(e)$ for any $0\le a<b\le T$, $A\in\mathcal{B}(E)$, and any bounded and $\mathcal{F}_a$-measurable h_a. Then, Lemma 1.1 implies that, a.s.,

$$\mathrm{LHS}(1.2) = \int_{(0,T]\times E} h_a\mathbb{1}_{(\tau\vee a,\,\tau\vee b]}(s)\mathbb{1}_A(e)M(\mathrm{d}s,\mathrm{d}e) = h_a[M(\tau\vee b,A)-M(\tau\vee a,A)].$$

Remark that $(a-\tau)\vee 0$ is an $\mathbb{F}^\tau$-stopping time and h_a is $\mathcal{F}^\tau_{(a-\tau)\vee 0}$-measurable as for any Borel set B and any $r\in[0,T]$,

$$\{h_a\in B\}\cap\{(a-\tau)\vee 0\le r\} = \underbrace{(\{h_a\in B\}\cap\{a\le\tau\})}_{\in\mathcal{F}_{a\wedge\tau}\subseteq\mathcal{F}^\tau_\tau}\cup\underbrace{(\{h_a\in B\}\cap\{\tau<a\le(\tau+r)\wedge T\})}_{\in\mathcal{F}^\tau_r}.$$

Since $(a-\tau,b-\tau]\cap(0,T-\tau] = ((a-\tau)\vee 0,(b-\tau)\vee 0]$, Lemma 1.1 implies that, a.s.,

$$\begin{aligned}\mathrm{RHS}(1.2) &= \int_{(0,T]\times E} h_a\mathbb{1}_{((a-\tau)\vee 0,\,(b-\tau)\vee 0]}(s)\mathbb{1}_A(e)M_\tau(\mathrm{d}s,\mathrm{d}e) \\ &= h_a[M_\tau((b-\tau)\vee 0,A)-M_\tau((a-\tau)\vee 0,A)] \\ &= h_a[M(\tau+(b-\tau)\vee 0,A)-M(\tau+(a-\tau)\vee 0,A)] \\ &= h_a[M(\tau\vee b,A)-M(\tau\vee a,A)] = \mathrm{LHS}(1.2).\end{aligned}$$

By the linearity, (1.2) holds for all linear combinations of such functions g. Due to the denseness we obtain (1.2) for $g\in\mathbf{L}^2(\mathbb{F},\mu_M)$, and finally, by a localizing argument we infer that (1.2) holds for $g\in\mathbf{L}^2_{\mathrm{loc}}(\mathbb{F},\mu_M)$. □

1.2 SDEs driven by orthogonal martingale measures

Let $\{M^{(1)}, \dots, M^{(\ell)}\}$ be a collection of (càdlàg) $(\mathbb{F}, \mathbb{P})$-martingale measures on $[0,T] \times \mathcal{B}(E)$. Assume that each $M^{(j)}$ is an orthogonal martingale measure with (random) intensity measure $\mu^{(j)}$ which satisfies

$$\mu^{(j)}(\omega, \mathrm{d}s, \mathrm{d}e) = \mu_s^{(j)}(\omega, \mathrm{d}e)\mathrm{d}s \quad \mathbb{P}\text{-a.s. } \omega \in \Omega$$

for some transition kernel $\{(\omega, s, A) \mapsto \mu_s^{(j)}(\omega, A), (\omega, s) \in \Omega \times [0,T], A \in \mathcal{B}(E)\}$, $j = 1, \dots, \ell$.

Let $\beta \colon \Omega \times [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ be $\mathcal{P}_{\mathbb{F}} \otimes \mathcal{B}(\mathbb{R}^m)/\mathcal{B}(\mathbb{R}^m)$-measurable, $\alpha \colon \Omega \times [0,T] \times \mathbb{R}^m \times E \to \mathbb{R}^{m \times \ell}$ be $\mathcal{P}_{\mathbb{F}} \otimes \mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(E)/\mathcal{B}(\mathbb{R}^{m \times \ell})$-measurable and consider the following m-dimensional SDE

$$Y_t = \eta + \int_0^t \beta(s, Y_{s-})\mathrm{d}s + \int_{(0,t] \times E} \alpha(s, Y_{s-}, e) M(\mathrm{d}s, \mathrm{d}e), \quad t \in [0,T], \tag{1.3}$$

for some $\mathcal{F}_0$-measurable $\mathbb{R}^m$-valued random variable η, and for $M := (M^{(1)}, \dots, M^{(\ell)})^{\mathsf{T}}$.

Definition 1.3. A process $Y \colon \Omega \times [0,T] \to \mathbb{R}^m$ is a *strong solution* to the SDE (1.3) with initial condition η if:

(i) $Y_0 = \eta$ a.s.;

(ii) $Y = (Y_t)_{t \in [0,T]}$ is a càdlàg and $\mathbb{F}$-adapted process;

(iii) $\int_0^T |\beta(s, Y_{s-})|\mathrm{d}s + \sum_{i=1}^m \sum_{j=1}^\ell \int_{(0,T] \times E} |\alpha^{(i,j)}(s, Y_{s-}, e)|^2 \, \mu_s^{(j)}(\mathrm{d}e)\mathrm{d}s < \infty$ a.s.;

(iv) The SDE (1.3) is satisfied for all $t \in [0,T]$ a.s.

Proposition 1.4. *Assume that there exist constants $K_\beta, K_\alpha \geq 0$ not depending on (ω, s, y_1, y_2) such that, for $\mathbb{P}$-a.s. $\omega \in \Omega$ and for all $s \in [0,T]$, $y_1, y_2 \in \mathbb{R}^m$,*

$$\begin{aligned} &|\beta(\omega, s, y_1) - \beta(\omega, s, y_2)| \leq K_\beta |y_1 - y_2|, \\ &4\ell \sum_{i=1}^m \sum_{j=1}^\ell \int_E |\alpha^{(i,j)}(\omega, s, y_1, e) - \alpha^{(i,j)}(\omega, s, y_2, e)|^2 \, \mu_s^{(j)}(\omega, \mathrm{d}e) \leq K_\alpha^2 |y_1 - y_2|^2, \end{aligned} \tag{1.4}$$

and that

$$K_0^2 := \mathbb{E}\left[T \int_0^T |\beta(s,0)|^2 \mathrm{d}s + 4\ell \sum_{i=1}^m \sum_{j=1}^\ell \int_0^T \int_E |\alpha^{(i,j)}(s, 0, e)|^2 \, \mu_s^{(j)}(\mathrm{d}e)\mathrm{d}s\right] < \infty. \tag{1.5}$$

Then, for any $\mathcal{F}_0$-measurable initial condition η, the SDE (1.3) has a unique (up to an indistinguishability) strong solution Y.

Proof. **Existence.** Let us fix an $\mathcal{F}_0$-measurable η.

Case 1: $\eta \in \mathbf{L}^2(\mathbb{P})$. We use the usual Picard iterations. Let $Y^0 = (Y_t^0)_{t \in [0,T]}$ with $Y_t^0 := \eta$ for all $t \in [0,T]$, and inductively define the sequence of process $(Y^n)_{n \in \mathbb{N}}$ via

$$Y_t^n := \eta + \int_0^t \beta(s, Y_{s-}^{n-1})\mathrm{d}s + \int_{(0,t] \times E} \alpha(s, Y_{s-}^{n-1}, e) M(\mathrm{d}s, \mathrm{d}e), \quad t \in [0,T].$$

To show that Y^n is well-defined, we consider

$$\Theta_t := \int_0^t \beta(s, 0)\mathrm{d}s + \int_{(0,t] \times E} \alpha(s, 0, e) M(\mathrm{d}s, \mathrm{d}e), \quad t \in [0,T].$$

Combining Doob's maximal inequality, the inequality $(x_1 + \cdots + x_\ell)^2 \le \ell(x_1^2 + \cdots + x_\ell^2)$, Itô's isometry with using (1.5) we infer that Θ is an adapted and $\mathbb{R}^m$-valued càdlàg process with

$$\mathbb{E}\bigg[\sup_{0\le t\le T}|\Theta_t|^2\bigg] \le 2\mathbb{E}\bigg[T\int_0^T|\beta(s,0)|^2\mathrm{d}s + 4\ell\sum_{i=1}^m\sum_{j=1}^\ell\int_0^T\int_E|\alpha^{(i,j)}(s,0,e)|^2\,\mu_s^{(j)}(\mathrm{d}e)\mathrm{d}s\bigg] = 2K_0^2.$$

It then follows from the square integrability of η and Fubini's theorem that

$$\begin{aligned}\mathbb{E}\bigg[\sup_{0\le t\le T}|Y_t^1 - Y_t^0|^2\bigg] &\le 2\mathbb{E}\bigg[\sup_{0\le t\le T}|Y_t^1 - Y_t^0 - \Theta_t|^2\bigg] + 2\mathbb{E}\bigg[\sup_{0\le t\le T}|\Theta_t|^2\bigg]\\ &\le 4T^2K_\beta^2\,\mathbb{E}[|\eta|^2] + 4TK_\alpha^2\,\mathbb{E}[|\eta|^2] + 4K_0^2\\ &\le K_{(1i)}^2(1 + \mathbb{E}[|\eta|^2]) \qquad (1i)\end{aligned}$$

for

$$K_{(1i)}^2 := 4\max\{K_0^2, T^2K_\beta^2 + TK_\alpha^2\}.$$

We deduce by induction using the same arguments as above that Y^n is well-defined and square integrable for all $n \in \mathbb{N}$. For any $n \ge 1$ and $t \in [0,T]$,

$$\begin{aligned}\frac{1}{2}\mathbb{E}\bigg[\sup_{0\le s\le t}|Y_s^{n+1} - Y_s^n|^2\bigg] &\le \mathbb{E}\bigg[\sup_{0\le s\le t}\bigg|\int_0^s[\beta(r,Y_{r-}^n) - \beta(r,Y_{r-}^{n-1})]\mathrm{d}r\bigg|^2\bigg]\\ &\quad + \mathbb{E}\bigg[\sup_{0\le s\le t}\bigg|\int_{(0,s]\times E}[\alpha(r,Y_{r-}^n,e) - \alpha(r,Y_{r-}^{n-1},e)]M(\mathrm{d}r,\mathrm{d}e)\bigg|^2\bigg]\\ &\le t\mathbb{E}\bigg[\int_0^t|\beta(s,Y_s^n) - \beta(s,Y_s^{n-1})|^2\mathrm{d}s\bigg]\\ &\quad + 4\ell\sum_{i=1}^m\sum_{j=1}^\ell\mathbb{E}\bigg[\int_0^t\int_E|\alpha^{(i,j)}(s,Y_s^n,e) - \alpha^{(i,j)}(s,Y_s^{n-1},e)|^2\,\mu_s^{(j)}(\mathrm{d}e)\mathrm{d}s\bigg]\\ &\le (tK_\beta^2 + K_\alpha^2)\,\mathbb{E}\bigg[\int_0^t|Y_s^n - Y_s^{n-1}|^2\mathrm{d}s\bigg], \qquad (1.6)\end{aligned}$$

where we use Doob's maximal inequality and Itô's isometry in the second inequality and use Fubini's theorem, together with (1.4), in the third inequality. Then, for

$$K_{(2i)}^2 := 2(TK_\beta^2 + K_\alpha^2), \qquad (2i)$$

and for any $t \in [0,T]$, $n \ge 1$, one has

$$\mathbb{E}\bigg[\sup_{0\le s\le t}|Y_s^{n+1} - Y_s^n|^2\bigg] \le K_{(2i)}^2\int_0^t\mathbb{E}\bigg[\sup_{0\le r\le s}|Y_r^n - Y_r^{n-1}|^2\bigg]\mathrm{d}s.$$

Iterating the estimate above, we get for any $n \in \mathbb{N}$,

$$\mathbb{E}\bigg[\sup_{0\le s\le T}|Y_s^{n+1} - Y_s^n|^2\bigg] \le K_{(2i)}^{2n}\frac{T^n}{n!}\mathbb{E}\bigg[\sup_{0\le r\le T}|Y_r^1 - Y_r^0|^2\bigg]. \qquad (1.7)$$

Combining Markov's inequality with (1.7) yields

$$\sum_{n=0}^\infty\mathbb{P}\bigg(\bigg\{\sup_{0\le s\le T}|Y_s^{n+1} - Y_s^n| \ge \frac{1}{2^n}\bigg\}\bigg) \le \mathbb{E}\bigg[\sup_{0\le r\le T}|Y_r^1 - Y_r^0|^2\bigg]\sum_{n=0}^\infty\frac{\big(4TK_{(2i)}^2\big)^n}{n!} < \infty.$$

By the Borel–Cantelli lemma, there is an event Ω_0 with probability one such that for any $\omega \in \Omega_0$, there exists $n_\omega \in \mathbb{N}$ such that $\sup_{0\le s\le T} |Y_s^{n+1}(\omega) - Y_s^n(\omega)| < 2^{-n}$ for all $n \ge n_\omega$. We then deduce that $Y^n(\omega)$ converges uniformly on $[0,T]$ for $\omega \in \Omega_0$. For all $t \in [0,T]$, define

$$Y_t(\omega) := \begin{cases} \lim_{n\to\infty} Y_t^n(\omega) & \text{if } \omega \in \Omega_0 \\ 0 & \text{if } \omega \notin \Omega_0. \end{cases}$$

By the uniform convergence and the completeness of the underlying filtration, Y has càdlàg paths and is adapted. Now, by the triangle inequality, it follows from (1.7) that

$$\begin{aligned} \Big\| \sup_{0\le s\le T} |Y_s - Y_s^0| \Big\|_{\mathbf{L}^2(\mathbb{P})} &= \Big\| \lim_{n\to\infty} \sup_{0\le s\le T} |Y_s^n - Y_s^0| \Big\|_{\mathbf{L}^2(\mathbb{P})} \\ &\le \sum_{j=0}^{\infty} K_{(2i)}^j \sqrt{\frac{T^j}{j!}} \Big\| \sup_{0\le r\le T} |Y_r^1 - Y_r^0| \Big\|_{\mathbf{L}^2(\mathbb{P})} \\ &\le \sqrt{2\mathrm{e}^{2K_{(2i)}^2 T}} K_{(1i)} \sqrt{1 + \mathbb{E}[|\eta|^2]}, \end{aligned}$$

which then yields

$$\mathbb{E}\Big[\sup_{0\le s\le T} |Y_s|^2 \Big] \le 2\mathbb{E}[|\eta|^2] + 4\mathrm{e}^{2K_{(2i)}^2 T} K_{(1i)}^2 (1 + \mathbb{E}[|\eta|^2]) \le K_{(3i)}^2 (1 + \mathbb{E}[|\eta|^2]), \tag{1.8}$$

where

$$K_{(3i)}^2 := 2 + 4\mathrm{e}^{2K_{(2i)}^2 T} K_{(1i)}^2. \tag{3i}$$

By (1.4), (1.5) and (1.8), the following process Z is well-defined in $\mathbf{L}^2(\mathbb{P})$,

$$Z_t := Y_0 + \int_0^t \beta(s, Y_{s-}) \mathrm{d}s + \int_{(0,t]\times E} \alpha(s, Y_{s-}, e) M(\mathrm{d}s, \mathrm{d}e), \quad t \in [0,T].$$

We now show that $Z = Y$. Indeed, proceeding as in (1.6) with Z in place of Y^n, we get

$$\begin{aligned} \Big\| \sup_{0\le s\le T} |Z_s - Y_s^{n+1}| \Big\|_{\mathbf{L}^2(\mathbb{P})} &\le \sqrt{T} K_{(2i)} \Big\| \sup_{0\le s\le T} |Y_s - Y_s^n| \Big\|_{\mathbf{L}^2(\mathbb{P})} \\ &\le \sqrt{T} K_{(2i)} \sum_{j=n}^{\infty} K_{(2i)}^j \sqrt{\frac{T^j}{j!}} \Big\| \sup_{0\le r\le T} |Y_r^1 - Y_r^0| \Big\|_{\mathbf{L}^2(\mathbb{P})} \\ &\xrightarrow{n\to\infty} 0, \end{aligned}$$

which completes the proof for the existence when the initial condition is square integrable.

Case 2: $\eta \notin \mathbf{L}^2(\mathbb{P})$. For $k \in \mathbb{N}$, following the construction in **Case 1**, we let $Y(k) = (Y_t(k))_{t\in[0,T]}$ be the strong solution of (1.3) with initial condition $\eta_k := \eta \mathbb{1}_{\{|\eta|\le k\}} \in \mathbf{L}^2(\mathbb{P})$. It follows from (1.8) that $\mathbb{E}\big[\sup_{0\le s\le T} |Y_s(k)|^2\big] < \infty$. We prove that

$$Y(k)\mathbb{1}_{\{|\eta|\le l\}} = Y(l)\mathbb{1}_{\{|\eta|\le l\}}, \quad \forall k \ge l \ge 1 \tag{1.9}$$

by showing

$$\mathbb{E}\Big[\sup_{0\le t\le T} |Y_t(k)\mathbb{1}_{\{|\eta|\le l\}} - Y_t(l)\mathbb{1}_{\{|\eta|\le l\}}|^2 \Big] = 0. \tag{1.10}$$

Indeed, since $\mathbb{1}_{\{|\eta|\le l\}}$ is bounded and $\mathcal{F}_0$-measurable, we may move it inside the stochastic integrals to get, a.s.,

$$Y_t(k)\mathbb{1}_{\{|\eta|\le l\}} = \eta_l + \int_0^t \mathbb{1}_{\{|\eta|\le l\}}\beta(s, Y_{s-}(k))\mathrm{d}s + \int_{(0,t]\times E} \mathbb{1}_{\{|\eta|\le l\}}\alpha(s, Y_{s-}(k), e)M(\mathrm{d}s, \mathrm{d}e),$$

and it in particular holds when k is replaced by l. Noting that

$$\mathbb{1}_{\{|\eta|\le l\}}F(x) - \mathbb{1}_{\{|\eta|\le l\}}F(y) = F(x\mathbb{1}_{\{|\eta|\le l\}}) - F(y\mathbb{1}_{\{|\eta|\le l\}}),$$

we then derive from the same lines as in (1.6) that, for any $t \in [0, T]$,

$$\mathbb{E}\Big[\sup_{0\le s\le t} |Y_s(k)\mathbb{1}_{\{|\eta|\le l\}} - Y_s(l)\mathbb{1}_{\{|\eta|\le l\}}|^2\Big] \le K^2_{(2i)} \int_0^t \mathbb{E}\Big[\sup_{0\le r\le s} |Y_r(k)\mathbb{1}_{\{|\eta|\le l\}} - Y_r(l)\mathbb{1}_{\{|\eta|\le l\}}|^2\Big]\mathrm{d}s$$

which then yields (1.10) with the aid of Gronwall's lemma. Note that $Y(\cdot)$ is uniformly Cauchy in probability as

$$\mathbb{P}\bigg(\Big\{\sup_{0\le t\le T} |Y_t(k) - Y_t(l)| > \varepsilon\Big\}\bigg) \le \mathbb{P}(\{|\eta| > l\}) \xrightarrow{k,l\to\infty} 0, \quad \forall \varepsilon > 0,$$

which then implies the existence of $\mathcal{Y}$ such that $Y(k) \xrightarrow{k\to\infty} \mathcal{Y}$ uniformly on $[0, T]$ in probability. Consequently, $\mathcal{Y}$ is adapted and càdlàg, which ensures that

$$\int_0^T |\beta(s, \mathcal{Y}_{s-})|\mathrm{d}s + \sum_{i=1}^m \sum_{j=1}^\ell \int_0^T \int_E |\alpha^{(i,j)}(s, \mathcal{Y}_{s-}, e)|^2 \mu_s^{(j)}(\mathrm{d}e)\mathrm{d}s < \infty \quad \mathbb{P}\text{-a.s.}$$

under (1.4), (1.5) and the càdlàg property of $\mathcal{Y}$. Now, letting $k \to \infty$ in (1.9) yields $\mathcal{Y}\mathbb{1}_{\{|\eta|\le l\}} = Y(l)\mathbb{1}_{\{|\eta|\le l\}}$ for any $l \in \mathbb{N}$. We define

$$\mathcal{Z}_t := \eta + \int_0^t \beta(s, \mathcal{Y}_{s-})\mathrm{d}s + \int_{(0,t]\times E} \alpha(s, \mathcal{Y}_{s-}, e)M(\mathrm{d}s, \mathrm{d}e), \quad t \in [0, T]$$

so that, a.s,

$$\mathcal{Z}_t\mathbb{1}_{\{|\eta|\le l\}} = \eta_l + \int_0^t \mathbb{1}_{\{|\eta|\le l\}}\beta(s, \mathcal{Y}_{s-})\mathrm{d}s + \int_{(0,t]\times E} \mathbb{1}_{\{|\eta|\le l\}}\alpha(s, \mathcal{Y}_{s-}, e)M(\mathrm{d}s, \mathrm{d}e), \quad t \in [0, T].$$

Then,

$$\mathbb{E}\Big[\sup_{0\le t\le T} |\mathcal{Z}_t\mathbb{1}_{\{|\eta|\le l\}} - Y_t(l)\mathbb{1}_{\{|\eta|\le l\}}|^2\Big] \le TK^2_{(2i)}\mathbb{E}\Big[\sup_{0\le s\le T} |\mathcal{Y}_s\mathbb{1}_{\{|\eta|\le l\}} - Y_s(l)\mathbb{1}_{\{|\eta|\le l\}}|^2\Big] = 0$$

which shows that $\mathcal{Z}\mathbb{1}_{\{|\eta|\le l\}} = Y(l)\mathbb{1}_{\{|\eta|\le l\}} = \mathcal{Y}\mathbb{1}_{\{|\eta|\le l\}}$ for any $l \in \mathbb{N}$. Letting $l \to \infty$ we conclude that $\mathcal{Z} = \mathcal{Y}$, and thus, $\mathcal{Y}$ solves (1.3).

Uniqueness. Assume that $\mathcal{Y}$ and $\tilde{\mathcal{Y}}$ solve (1.3) with an initial condition η. Define $T_0 := 0$ and

$$\begin{aligned} T_n := T \wedge \inf\bigg\{t > T_{n-1} : \int_0^t |\beta(s, \mathcal{Y}_s)|^2\mathrm{d}s + 4\ell \sum_{i=1}^m \sum_{j=1}^\ell \int_0^t \int_E |\alpha^{(i,j)}(s, \mathcal{Y}_s, e)|^2 \mu_s^{(j)}(\mathrm{d}e)\mathrm{d}s > n\bigg\} \\ \wedge \inf\bigg\{t > T_{n-1} : \int_0^t |\beta(s, \tilde{\mathcal{Y}}_s)|^2\mathrm{d}s + 4\ell \sum_{i=1}^m \sum_{j=1}^\ell \int_0^t \int_E |\alpha^{(i,j)}(s, \tilde{\mathcal{Y}}_s, e)|^2 \mu_s^{(j)}(\mathrm{d}e)\mathrm{d}s > n\bigg\}. \end{aligned}$$

Then, $(T_n)_n$ is a sequence of non-decreasing stopping times which are eventually constant T a.s. Note that, for any $n \in \mathbb{N}$, one has $\mathbb{E}\big[\sup_{0\le t\le T} |\mathcal{Y}_{t\wedge T_n} - \eta|^2\big] < \infty$, as well as for $\tilde{\mathcal{Y}}$. Using the same arguments as for (1.6), we infer that $\mathbb{E}\big[\sup_{0\le t\le T} |\mathcal{Y}_{t\wedge T_n} - \tilde{\mathcal{Y}}_{t\wedge T_n}|^2\big] = 0$, and consequently, $\mathcal{Y}_{\cdot\wedge T_n} = \tilde{\mathcal{Y}}_{\cdot\wedge T_n}$ for all n. Letting $n \to \infty$ and using the càdlàg property we derive $\mathcal{Y} = \tilde{\mathcal{Y}}$. $\square$

Remark 1.5. The proof of Proposition 1.4 reveals that, if in addition $\eta \in \mathbf{L}^2(\mathbb{P})$ then the strong solution of (1.3) satisfies

$$\mathbb{E}\Big[\sup_{0\le t\le T}|Y_t|^2\Big] \le K(1+\mathbb{E}[|\eta|^2])$$

for some constant $K \ge 0$ depending only on $K_\alpha, K_\beta, K_0, T$.

2 Miscellaneous

2.1 Proof of [1, Lemma 6.1]

The assumption $\int_{[0,1]^d}\eta_s^{(k)}(u)\eta_s^{(k')}(u)\mathrm{d}u = \mathbb{1}_{\{k=k'\}}$ for $\mathbb{P}\otimes\boldsymbol{\lambda}_{[0,T]}$-a.e. $(\omega,s)\in\Omega\times[0,T]$ particularly implies that $\mathbb{E}\big[\int_0^T\int_{[0,1]^d}|\eta_s^{(k)}(u)|^2\mathrm{d}u\mathrm{d}s\big] = T$. Hence, for any (k,l), $(\eta^{(k)}\cdot M_{B^{(l)}})$ is a square integrable $(\mathbb{F},\mathbb{P})$-martingale null at 0. Since $M_{B^{(l)}}$ is a continuous martingale measure (see [5, Section II(3)]), the process $(\eta^{(k)}\cdot M_{B^{(l)}})$ is also continuous as indicated in [5, Propisition I-6(1)]. As $M_{B^{(l)}}$ and $M_{B^{(l')}}$ are independent for $l\neq l'$ by assumption, it is straightforward to prove that the product $(\eta^{(k)}\cdot M_{B^{(l)}})(\eta^{(k')}\cdot M_{B^{(l')}})$ is also a continuous $(\mathbb{F},\mathbb{P})$-martingale, which thus implies that $\langle(\eta^{(k)}\cdot M_{B^{(l)}}),(\eta^{(k')}\cdot M_{B^{(l')}})\rangle = 0$. We compute the quadratic covariation using [5, Proposition I-6(2)], a.s.,

$$\big\langle(\eta^{(k)}\cdot M_{B^{(l)}}),(\eta^{(k')}\cdot M_{B^{(l')}})\big\rangle_t = \mathbb{1}_{\{l=l'\}}\int_0^t\int_{[0,1]^d}\eta_s^{(k)}(u)\eta_s^{(k')}(u)\mathrm{d}u\mathrm{d}s = \mathbb{1}_{\{(k,l)=(k',l')\}}t.$$

Thus, the desired conclusion follows from the Lévy characterization for Brownian motion. □

2.2 Proof of [1, Proposition 6.5]

(1) Recall from [1, Subsection 6.2] that the operators $\mathcal{L}_{\mathbf{h}}$ and $\mathcal{L}_h$ coincide, if $\mathbf{h}$ executes h. Then applying Itô's formula, we obtain as in the proof of [1, Proposition 5.4] that

$$J(t,X_t^{\mathbf{h}}) - \int_0^t\left(\frac{\partial J}{\partial t}(s,X_s^{\mathbf{h}}) + (\mathcal{L}_hJ(s,\cdot))(s,X_s^{\mathbf{h}})\right)\mathrm{d}s$$

is a local martingale. Inserting the partial differential equation, we observe that

$$J(t,X_t^{\mathbf{h}}) + \lambda\int_0^t\int_{\mathbb{R}}\dot h(s,X_s^{\mathbf{h}},y)\log\dot h(s,X_s^{\mathbf{h}},y)\mathrm{d}y\mathrm{d}s$$

is a local martingale, and hence a martingale, by the boundedness assumptions on J and on the entropy. Thus, a.s.,

$$\begin{aligned}J(t,X_t^{\mathbf{h}}) &= \mathbb{E}\left[J(T,X_T^{\mathbf{h}}) + \lambda\int_0^T\int_{\mathbb{R}}\dot h(s,X_s^{\mathbf{h}},y)\log\dot h(s,X_s^{\mathbf{h}},y)\mathrm{d}y\mathrm{d}s\,\middle|\,\mathcal{F}_t\right]\\ &\quad-\lambda\int_0^t\int_{\mathbb{R}}\dot h(s,X_s^{\mathbf{h}},y)\log\dot h(s,X_s^{\mathbf{h}},y)\mathrm{d}y\mathrm{d}s\\ &= \mathbb{E}\left[g(X_T^{\mathbf{h}}) + \lambda\int_t^T\int_{\mathbb{R}}\dot h(s,X_s^{\mathbf{h}},y)\log\dot h(s,X_s^{\mathbf{h}},y)\mathrm{d}y\mathrm{d}s\,\middle|\,\mathcal{F}_t\right]\\ &= \mathcal{J}_t^{\mathbf{h}},\end{aligned}$$

i.e., J is a value function of $\mathbf{h}$.

(2) If $\tilde J$ is a value function of $\mathbf{h}$, then $(\tilde J(t,X_t^{\mathbf{h}}))_{t\ge0}$ is a modification of $\mathcal{J}$. Hence,

$$\tilde J(t,X_t^{\mathbf{h}}) + \lambda\int_0^t\int_{\mathbb{R}}\dot h(s,X_s^{\mathbf{h}},y)\log\dot h(s,X_s^{\mathbf{h}},y)\mathrm{d}y\mathrm{d}s \tag{2.1}$$

inherits the martingale property of

$$\mathcal{J}_t^{\mathbf{h}} + \lambda \int_0^t \int_{\mathbb{R}} \dot{h}(s, X_s^{\mathbf{h}}, y) \log \dot{h}(s, X_s^{\mathbf{h}}, y) \mathrm{d}y \mathrm{d}s.$$

Conversely, if the process in (2.1) is a martingale, then the last part of the proof of (1) can be repeated with $\tilde{J}$ in place of J to conclude that $\tilde{J}$ is a value function of $\mathbf{h}$. □

2.3 Proof of [1, Lemma 7.1]

Recall the representation of $\mathcal{X}$ in [1, Theorem 5.1]. For $l = 1, \ldots, p$, [5, Section II(2)] asserts that $\int_0^\cdot \int_{[0,1]^d} f_l^{(k)}(s,u) M_{B^{(l)}}(\mathrm{d}s, \mathrm{d}u)$ is a continuous square integrable martingale with quadratic variation $\int_0^\cdot \int_{[0,1]^d} |f_l^{(k)}(s,u)|^2 \mathrm{d}u \mathrm{d}s$. The boundedness of f_{p+1}, f_{p+2} and [1, Eq. (7.1)] imply

$$\int_0^T \int_{\mathbb{R}_0^q \times [0,1]^d} [|f_{p+1}(s,z,u)|^2 |z|^2 \mathbb{1}_{\{0<|z|\le R\}} + |f_{p+2}(s,z,u)| \mathbb{1}_{\{|z|>R\}}] \mu_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) < \infty,$$

which shows that the process driven by $\tilde{M}_J$ is a square integrable martingale, and that against M_J is an a.s. finite variation process. Hence, $\mathcal{X}$ is an $\mathbb{R}^m$-valued semimartingale.

According to [5, Proposition I-6], the quadratic covariation matrix of the continuous martingale part of $\mathcal{X}$ is

$$\begin{aligned}
&\left\langle \sum_{l=1}^p \int_0^\cdot \int_{[0,1]^d} f_l^{(k)}(s,u) M_{B^{(l)}}(\mathrm{d}s, \mathrm{d}u), \sum_{l'=1}^p \int_0^\cdot \int_{[0,1]^d} f_{l'}^{(k')}(s,u) M_{B^{(l')}}(\mathrm{d}s, \mathrm{d}u) \right\rangle \\
&= \sum_{l=1}^p \int_0^\cdot \int_{[0,1]^d} (f_l^{(k)} f_l^{(k')})(s,u) \mathrm{d}u \mathrm{d}s = C^{\mathcal{X},(k,k')}.
\end{aligned}$$

For the jump part, it follows from [7, Ch.3, Theorem 1] that

$$\Delta \mathcal{X}_r = \int_{\{r\} \times \mathbb{R}_0^q \times [0,1]^d} [f_{p+1}(s,z,u)|z| \mathbb{1}_{\{0<|z|\le R\}} + f_{p+2}(s,z,u) \mathbb{1}_{\{|z|>R\}}] M_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u), \ r \in [0,T] \ \mathbb{P}\text{-a.s.}$$

Let $A \in \mathcal{B}(\mathbb{R}_0^m)$ with $A \cap B_m(\kappa) = \emptyset$ for some $\kappa > 0$ where $B_m(\kappa) = \{y \in \mathbb{R}^m : |y| < \kappa\}$. Since f_{p+1} is bounded, there exists $\varepsilon > 0$ sufficiently small such that

$$\begin{aligned}
&\left\{(r,z,u) : f_{p+1}(r,z,u)|z| \mathbb{1}_{\{0<|z|\le R\}} + f_{p+2}(r,z,u) \mathbb{1}_{\{|z|>R\}} \in A\right\} \\
&= \left\{(r,z,u) : f_{p+1}(r,z,u)|z| \mathbb{1}_{\{\varepsilon<|z|\le R\}} + f_{p+2}(r,z,u) \mathbb{1}_{\{|z|>R\}} \in A\right\}.
\end{aligned}$$

We define the process (L^Z, L^U) depending on ε via

$$(L_t^Z, L_t^U) := \int_{(0,t] \times \{|z|>\varepsilon\} \times [0,1]^d} (z,u) M_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u), \quad t \in [0,T].$$

Let $N_{\mathcal{X}}$ be the random jump measure of $\mathcal{X}$. Then

$$\begin{aligned}
N_{\mathcal{X}}((s,t] \times A) &= \sum_{s<r\le t} \mathbb{1}_{\{\Delta \mathcal{X}_r \in A\}} \\
&= \sum_{s<r\le t} \mathbb{1}_{\left\{f_{p+1}(r, \Delta L_r^Z, \Delta L_r^U)|\Delta L_r^Z| \mathbb{1}_{\{\varepsilon<|\Delta L_r^Z|\le R\}} + f_{p+2}(r, \Delta L_r^Z, \Delta L_r^U) \mathbb{1}_{\{|\Delta L_r^Z|>R\}} \in A\right\}} \\
&= \int_s^t \int_{\mathbb{R}_0^q \times [0,1]^d} \mathbb{1}_A\big(f_{p+1}(r,z,u)|z| \mathbb{1}_{\{\varepsilon<|z|\le R\}} + f_{p+2}(r,z,u) \mathbb{1}_{\{|z|>R\}}\big) M_J(\mathrm{d}r, \mathrm{d}z, \mathrm{d}u) \\
&= \int_s^t \int_{\{0<|z|\le R\} \times [0,1]^d} \mathbb{1}_A(f_{p+1}(r,z,u)|z|) M_J(\mathrm{d}r, \mathrm{d}z, \mathrm{d}u) \\
&\quad + \int_s^t \int_{\{|z|>R\} \times [0,1]^d} \mathbb{1}_A(f_{p+2}(r,z,u)) M_J(\mathrm{d}r, \mathrm{d}z, \mathrm{d}u).
\end{aligned}$$

As $\mu_J(\mathrm{d}r, \mathrm{d}z, \mathrm{d}u) = \nu_r(\mathrm{d}z)\mathrm{d}u\mathrm{d}r$ is the predictable compensator of $M_J(\mathrm{d}r, \mathrm{d}z, \mathrm{d}u)$, it implies that

$$\nu^{\mathcal{X}}((s,t] \times A) = \int_s^t \int_{\{0<|z|\le R\}\times[0,1]^d} \mathbb{1}_A(f_{p+1}(r,z,u)|z|)\nu_r(\mathrm{d}z)\mathrm{d}u\mathrm{d}r + \int_s^t \int_{\{|z|>R\}\times[0,1]^d} \mathbb{1}_A(f_{p+2}(r,z,u))\nu_r(\mathrm{d}z)\mathrm{d}u\mathrm{d}r.$$

This result can be extended to $A \in \mathcal{B}(\mathbb{R}_0^m)$ by using the approximation sequence $(A \cap B_m(\frac{1}{n}))_{n\in\mathbb{N}}$. For the predictable finite variation part $\mathfrak{b}^{\mathcal{X}}$, one has, a.s.,

$$\begin{aligned}
\mathcal{Y}_t &:= \mathcal{X}_t - \sum_{l=1}^p \int_0^t \int_{[0,1]^d} f_l(s,u) M_{B^{(l)}}(\mathrm{d}s, \mathrm{d}u) - \int_0^t \int_{\mathbb{R}_0^m} (y - \mathfrak{h}(y)) N_{\mathcal{X}}(\mathrm{d}s, \mathrm{d}y) \\
&= \int_0^t \int_{[0,1]^d} f_0(s,u)\mathrm{d}u\mathrm{d}s + \int_0^t \int_{\{0<|z|\le R\}\times[0,1]^d} f_{p+1}(s,z,u)|z|\tilde{M}_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) \\
&\quad + \int_0^t \int_{\{|z|>R\}\times[0,1]^d} f_{p+2}(s,z,u) M_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) \\
&\quad - \int_0^t \int_{\{0<|z|\le R\}\times[0,1]^d} [f_{p+1}(s,z,u)|z| - \mathfrak{h}(f_{p+1}(s,z,u)|z|)] M_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) \\
&\quad - \int_0^t \int_{\{|z|>R\}\times[0,1]^d} [f_{p+2}(s,z,u) - \mathfrak{h}(f_{p+2}(s,z,u))] M_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) \\
&= \int_0^t \int_{[0,1]^d} f_0(s,u)\mathrm{d}u\mathrm{d}s + \int_0^t \int_{\{|z|>R\}\times[0,1]^d} \mathfrak{h}(f_{p+2}(s,z,u))\nu_s(\mathrm{d}z)\mathrm{d}u\mathrm{d}s \\
&\quad - \int_0^t \int_{\{0<|z|\le R\}\times[0,1]^d} [f_{p+1}(s,z,u)|z| - \mathfrak{h}(f_{p+1}(s,z,u)|z|)]\nu_s(\mathrm{d}z)\mathrm{d}u\mathrm{d}s \\
&\quad + \int_0^t \int_{\{0<|z|\le R\}\times[0,1]^d} f_{p+1}(s,z,u)|z|\tilde{M}_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) \\
&\quad + \int_0^t \int_{\{|z|>R\}\times[0,1]^d} \mathfrak{h}(f_{p+2}(s,z,u))\tilde{M}_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) \\
&\quad - \int_0^t \int_{\{0<|z|\le R\}\times[0,1]^d} [f_{p+1}(s,z,u)|z| - \mathfrak{h}(f_{p+1}(s,z,u)|z|)]\tilde{M}_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u),
\end{aligned}$$

where in the last equality we use the fact that $\int F\tilde{M}_J = \int FM_J - \int F\mu_J$ if F is predictable and μ_J-integrable, see [4, Proposition II.1.28]. By identifying the predictable finite variation component of $\mathcal{Y}$, we obtain the desired expression of $\mathfrak{b}^{\mathcal{X}}$. □

2.4 Proof of [1, Lemma 8.1]

(1) This is straightforward.

(2) By a localizing procedure, we only need to show the desired relation under integrability condition $\mathbb{E}\big[\int_0^T |Y_s(\xi_s^\Pi)|^2\mathrm{d}s\big] < \infty$. Then, it is sufficient to prove the relation on $(t_{i-1}, t_i]$ for any $\mathbb{F}^\Pi$-predictable Y with $\mathbb{E}\big[\int_{t_{i-1}}^{t_i} |Y_s(\xi_{t_i}^\Pi)|^2\mathrm{d}s\big] < \infty$. Assume $Y_s(u) = \sum_{j=1}^k h_{j-1}\mathbb{1}_{(r_{j-1},r_j]}(s)\mathbb{1}_{A_j}(u)$ for $k \in \mathbb{N}$, $t_{i-1} \le r_0 < r_1 < \cdots < r_k = t_i$, $A_j \in \mathcal{B}([0,1]^d)$, h_{j-1} is bounded and $\mathcal{F}_{r_{j-1}}^\Pi$-measurable. Then, by the definition of $M_{B^{(l)}}^\Pi$, one has, a.s.,

$$\int_{(t_{i-1},t_i]\times[0,1]^d} Y_s(u) M_{B^{(l)}}^\Pi(\mathrm{d}s, \mathrm{d}u) = \sum_{j=1}^k h_{j-1} \int_{r_{j-1}}^{r_j} \mathbb{1}_{A_j}(\xi_{t_i}^\Pi)\mathrm{d}B_s^{(l)} = \int_{t_{i-1}}^{t_i} Y_s(\xi_{t_i}^\Pi)\mathrm{d}B_s^{(l)}.$$

The conclusion for $Y \in \mathbf{L}^2(\mathbb{F}^\Pi, M_D^\Pi)$ can be derived from a standard approximation argument where one notes that the Itô isometry coincides for both integrals driven by $M^\Pi_{B^{(l)}}$ and $B^{(l)}$.

(3) By writing $Y = \max\{Y,0\} - \max\{-Y,0\}$, we may assume $Y \ge 0$, and then the first relation follows from the argument in [1, proof of Proposition 4.3]. For the second relation, by a localizing argument, it suffices to show the desired relation under $\mathbb{E}\big[\int_0^T \int_{\mathbb{R}_0^q} |Y_s(z,\xi_s^\Pi)|^2\, \nu_s(\mathrm{d}z)\mathrm{d}s\big] < \infty$. This can be achieved in the usual way by first proving for $(-n \vee Y \wedge n)\mathbb{1}_{\{|z|>1/n\}}$ in place of Y, and then taking the limit in $\mathbf{L}^2(\mathbb{P})$ when $n \to \infty$ with the aid of Itô's isometry. $\square$

2.5 On the Poisson random measure M_J

Assume the Lévy process L as defined in [1, Proposition 4.3]. Let $\{T_j^n\}_{n,j\ge 0}$ be the of jump times of L given by

$$\begin{aligned} &T_0^0 := 0, \quad T_j^0 := \inf\{t > T_{j-1}^0 : |\Delta L_t| > 1\}, \quad j \ge 1,\\ &T_0^n := 0, \quad T_j^n := \inf\{t > T_{j-1}^n : 1/(n+1) < |\Delta L_t| \le 1/n\}, \quad j \ge 1, \quad n \ge 1. \end{aligned}$$

Let $\{\xi_j^n\}_{n,j\ge 0}$ be i.i.d. with uniform distribution on $[0,1]^d$. Assume that $\{\xi_j^n\}_{n,j\ge 0}$ is independent of L. We define the Poisson random measure M_J on $[0,T] \times \mathbb{R}_0^q \times [0,1]^d$ by

$$M_J(\omega, \mathrm{d}t, \mathrm{d}z, \mathrm{d}u) = \sum_{n=0}^\infty \sum_{j=1}^\infty \delta_{(T_j^n(\omega), \Delta L_{T_j^n(\omega)}(\omega), \xi_j^n(\omega))}(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u).$$

We note that, in general, there is no semimartingale which possesses M_J as the associated random jump measure because $\int_0^T \int_{0<|z|^2+|u|^2\le 1} (|z|^2 + |u|^2)\mu_J(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u)$ might be infinite, except the case $\int_0^T \int_{\mathbb{R}_0^q} \nu_t(\mathrm{d}z)\mathrm{d}t < \infty$ (i.e. L is of finite activities).

2.6 On the independence of $(M_{B^{(1)}}, \dots, M_{B^{(p)}})$ and M_J

Assume that $(M_{B^{(1)}}, \dots, M_{B^{(p)}})$ and M_J define on the same probability space, then

$$\left\{ \int_{(0,T]\times[0,1]^d} g_l(s,u) M_{B^{(l)}}(\mathrm{d}s, \mathrm{d}u) \,\middle|\, g_l\colon [0,T]\times[0,1]^d \to \mathbb{R} \text{ measurable and bounded}, l = 1, \dots, p \right\}$$

is independent of

$$\left\{ \int_{(0,T]\times\mathbb{R}_0^q\times[0,1]^d} h(s,z,u) M_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) \,\middle|\, h\colon [0,T] \times \mathbb{R}_0^q \times [0,1]^d \to [0,\infty) \text{ measurable} \right\}.$$

Indeed, it is sufficient to show that

$$G = \left(\sum_{l=1}^p \int_0^t \int_{[0,1]^d} g_l(s,u) M_{B^{(l)}}(\mathrm{d}s, \mathrm{d}u) \right)_{t\in[0,T]}$$

is independent of

$$H = \left(\int_{(0,t]\times\{|z|>\kappa\}\times[0,1]^d} h(s,z,u) M_J(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) \right)_{t\in[0,T]}$$

for all (non-random) measurable and bounded g_l, $h \ge 0$ and $\kappa > 0$. It is clear that H is of finite variation and G is a continuous martingale (see [5, Section II(3)]), and both are processes with independent increments. Observe that $[G,H]_t = \sum_{0\le s\le t} \Delta G_s \Delta H_s = 0$ for $t \in [0,T]$ a.s. It then follows from [3, Theorem 11.43] that G and H are independent.

3 Weak convergence in the Skorokhod topology

3.1 Skorokhod spaces and weak convergence

Fix $T \in (0,\infty)$ and let $\mathbb{D}_T(\mathbb{R}^m)$ be the family of all càdlàg functions $f\colon [0,T] \to \mathbb{R}^m$ and Λ_T consists of all strictly increasing and continuous $\lambda\colon [0,T] \to [0,T]$ with $\lambda(0)=0$, $\lambda(T)=T$. We equip $\mathbb{D}_T(\mathbb{R}^m)$ with the Skorokhod metric

$$d_T^m(x,y) := \inf_{\lambda\in\Lambda_T} \max\left\{ \sup_{0\le s<t\le T} \left| \log \frac{\lambda(t)-\lambda(s)}{t-s} \right|, \sup_{0\le t\le T} |x(t) - y(\lambda(t))| \right\}.$$

It is well-known that $(\mathbb{D}_T(\mathbb{R}^m), d_T^m)$ is a complete and separable metric space (see [2, Section 14]), however, it is not a topological vector space. It is also convenient to work with the metric $\tilde{d}_T^m$, which defines the same topology as d_T^m does, given by

$$\tilde{d}_T^m(x,y) := \inf_{\lambda\in\Lambda_T} \max\left\{ \sup_{0\le t\le T} |\lambda(t) - t|, \sup_{0\le t\le T} |x(t) - y(\lambda(t))| \right\}.$$

However, $(\mathbb{D}_T(\mathbb{R}^m), \tilde{d}_T^m)$ is not complete.

An $\mathbb{R}^m$-valued càdlàg process $X = (X_t)_{t\in[0,T]}$ can be regarded as an $\mathcal{F}/\mathcal{B}(\mathbb{D}_T(\mathbb{R}^m))$-measurable function $X\colon \Omega \to \mathbb{D}_T(\mathbb{R}^m)$ where $\mathcal{B}(\mathbb{D}_T(\mathbb{R}^m))$ is the Borel σ-algebra induced by the Skorokhod metric d_T^m. A sequence of $\mathbb{R}^m$-valued càdlàg processes $(X^n)_{n\in\mathbb{N}}$, where X^n is defined on $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$, is said to be *weakly convergent* to a càdlàg process X defined on $(\Omega, \mathcal{F}, \mathbb{P})$ if

$$\mathbb{E}^n[f(X^n)] \xrightarrow{n\to\infty} \mathbb{E}[f(X)], \quad \forall f \in C_b(\mathbb{D}_T(\mathbb{R}^m)),$$

where $\mathbb{E}^n$ and $\mathbb{E}$ are the expectation under $\mathbb{P}^n$ and $\mathbb{P}$, respectively. We then write $X^n \xrightarrow{\mathcal{D}_T} X$.

3.2 A limit theorem of Jacod–Shiryaev for triangular arrays

For the reader's convenience, we recall (and adapt to our setting) a limit theorem establishing the weak convergence of triangular arrays which we use to prove the main result in this article.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and suppose that $\{U_i^n, \mathcal{G}_i^n : i \ge 0\}$, $n \in \mathbb{N}$, are adapted sequences of $\mathbb{R}^d$-valued random variables. For each $n \in \mathbb{N}$, we consider a change of time $\sigma_n\colon \Omega \times [0,\infty) \to [0,\infty)$ with respect to $(\mathcal{G}_i^n)_{i\ge 0}$, i.e.,

(a) $\sigma_n(\cdot, 0) = 0$;

(b) For any ω, $\sigma_n(\omega, \cdot)$ is increasing, right-continuous, with jumps equal to 1;

(c) For any $t \ge 0$, $\sigma_n(\cdot, t)$ is a $(\mathcal{G}_i^n)_{i\ge0}$-stopping time.

Theorem 3.1 ([4], Theorem VIII.2.29)**.** *Assume a sequence of d-dimensional semimartingales $(X^n)_{n\in\mathbb{N}}$ where $X_t^n = \sum_{i=1}^{\sigma_n(t)} U_i^n$, $t \ge 0$. Let X be a d-dimensional process with independent increments and without fixed time of discontinuity, having characteristics $(\mathfrak{b}, C, \nu)$ in relation to a truncation function $\mathfrak{h}$. Set $\widetilde{C}_t^{(k,l)} := C_t^{(k,l)} + \int_0^t \int_{\mathbb{R}^d} (\mathfrak{h}^{(k)}\mathfrak{h}^{(l)})(y)\nu(\mathrm{d}s, \mathrm{d}y)$ as in [4, II.5.8]. If there exists some dense subset D of $[0,\infty)$ such that, as $n \to \infty$,*

$$\sup_{0\le s\le t} \left| \sum_{i=1}^{\sigma_n(s)} \mathbb{E}[\mathfrak{h}(U_i^n)|\mathcal{G}_{i-1}^n] - \mathfrak{b}_s \right| \xrightarrow{\mathbb{P}} 0 \quad \forall t \ge 0,$$

$$\sum_{i=1}^{\sigma_n(t)} \left(\mathbb{E}[(\mathfrak{h}^{(k)}\mathfrak{h}^{(l)})(U_i^n)|\mathcal{G}_{i-1}^n] - \mathbb{E}[\mathfrak{h}^{(k)}(U_i^n)|\mathcal{G}_{i-1}^n]\,\mathbb{E}[\mathfrak{h}^{(l)}(U_i^n)|\mathcal{G}_{i-1}^n] \right) \xrightarrow{\mathbb{P}} \widetilde{C}_t^{(k,l)} \quad \forall t \in D,$$

$$\sum_{i=1}^{\sigma_n(t)} \mathbb{E}[g(U_i^n)|\mathcal{G}_{i-1}^n] \xrightarrow{\mathbb{P}} \int_0^t \int_{\mathbb{R}^d} g(y)\nu(\mathrm{d}s, \mathrm{d}y) \quad \forall t \in D, g \in C_1(\mathbb{R}^d),$$

then X^n converges weakly to X in the Skorokhod topology on the space $\mathbb{D}_\infty(\mathbb{R}^d)$ of càdlàg functions $F\colon [0,\infty) \to \mathbb{R}^d$. Here, $C_1(\mathbb{R}^d) \subset C_2(\mathbb{R}^d)$ is a particular class of test functions vanishing around zero and is introduced in [4, VII.2.7].

References

[1] Bender, C. and Thuan, N.T. (2024). Continuous time reinforcement learning: A random measure approach. Preprint.

[2] Billingsley, P. (1999). *Convergence of probability measures*, 2nd ed. John Wiley & Sons, Inc.

[3] He, S., Wang, J. and Yan, J. (1992). *Semimartingale theory and stochastic calculus.* Taylor & Francis.

[4] Jacod, J. and Shiryaev, A. (2003). *Limit theorems for stochastic processes*, 2nd ed. Springer Berlin, Heidelberg.

[5] El Karoui, N. and Méléard, S. (1990). Martingale measures and stochastic calculus. *Probab. Theory Related Fields* **84**, 83–101.

[6] Kushner, H.J. and Dupuis, P. (2001). *Numerical methods for stochastic control problems in continuous time*, 2nd ed. Springer, New York.

[7] Liptser, R. and Shiryaev, A. (1989). *Theory of martingales.* Kluwer Academic Publishers.

[8] Walsh, J. (1986). An introduction to stochastic partial differential equations. *Lecture Notes in Maths.* **1180**, 265–439.