An Averaging Principle for Two-time-Scale Functional Diffusions

Fuke Wu

School of Mathematics and Statistics
Huazhong University of Science and Technology

Joint work with Professor George Yin

Workshop for Modeling, Stochastic Control, Optimization, and Related Applications

IMA Minneapolis

May 2018

Outline

Problem

Considers the functional diffusion system with two-time scales as follows:

$$
\begin{cases}
d\xi^{\varepsilon}(t) = \frac{1}{\varepsilon}h(\xi^{\varepsilon}(t))dt + \frac{1}{\sqrt{\varepsilon}}\phi(\xi^{\varepsilon}(t))dw_1(t), \\
dx^{\varepsilon}(t) = b(x_t^{\varepsilon}, \xi^{\varepsilon}(t))dt + \psi(x_t^{\varepsilon}, \xi^{\varepsilon}(t))dw_2(t)
\end{cases}
\tag{1.1}
$$

where $x_t^{\varepsilon} := \{x^{\varepsilon}(u \wedge t) : 0 \leq u \leq T\}$ is a stopped process,
$h = (h_1, h_2, \cdots, h_m)'$: $\mathbb{R}^m \mapsto \mathbb{R}^m$, $\phi = [\phi_{ij}]_{m \times l_1} : \mathbb{R}^m \mapsto \mathbb{R}^{m \times l_1}$,
$b = (b_1, b_2, \cdots, b_n)' : C([0, T]; \mathbb{R}^n) \times \mathbb{R}^m \mapsto \mathbb{R}^n$,
$\psi = [\psi_{ij}]_{n \times l_2} : C([0, T]; \mathbb{R}^n) \times \mathbb{R}^m \mapsto \mathbb{R}^{n \times l_2}$

Problem: Averaging principle of this system as $\varepsilon \to 0$.

History 1: Two-time-scale diffusions without delay

The Two-time-scale diffusion:

$$\begin{cases} dx^\varepsilon(t) = b(x^\varepsilon(t), \xi^\varepsilon(t))dt + \sigma(x^\varepsilon(t), \xi^\varepsilon(t))dw(t), \\ d\xi^\varepsilon(t) = \frac{1}{\varepsilon} h(x^\varepsilon(t), \xi^\varepsilon(t))dt + \frac{1}{\sqrt{\varepsilon}} \zeta(x^\varepsilon(t), \xi^\varepsilon(t))dw_1(t) \end{cases}$$

Methods:

1. Perturbed Fokker-Planck Equation [Ref: Khasminskii and Yin (96, 04, 05)];
2. Martingale method + Weak Convergence
   - (i) Probability method [Ref: Skorohod 89; Kushner(84, 90)];
   - (ii) PDE method [Ref: Pardoux and Yu (01, 03, 05)]

Result: $x^\varepsilon(t) \Rightarrow x(t)$, where $x(t)$ satisfies $dx(t) = \bar{b}(x(t))dt + \bar{\sigma}(x(t))dw(t)$,

where $\bar{b}(x) = \int b(x,\xi)\mu_x(d\xi)$, $\bar{\sigma}'(x)\bar{\sigma}(x) = \int \sigma'(x,\xi)\sigma(x,\xi)\mu_x(d\xi)$, in which $\mu_x(\cdot)$ is the invariant measure of the fixed-x process

$$d\xi(t) = h(x, \xi(t))dt + \zeta(x, \xi(t))dB(t) \qquad [B(t) = w(\varepsilon t)/\sqrt{\varepsilon}].$$

History

No-delay system:

- R.Z. Khasminskii, On an averaging principle for Ito stochastic differential equations, Kybernetika 4 (1968) 260-279.
- R.Z.Khasminskii,G.Yin,Asymptotic series for singularly perturbed Kolmogorov-Fokker-Planck equations, SIAM J. Appl. Math. 56 (1996) 1766-1793.
- R.Z. Khasminskii, G. Yin, On transition densities of singularly perturbed diffusions with fast and slow components, SIAM J. Appl. Math. 56 (1996) 1794-1819.
- R.Z. Khasminskii, G. Yin, On averaging principles: an asymptotic expansion approach, SIAM J. Math. Anal. 35 (2004) 1534-1560.
- R.Z. Khasminskii, G. Yin, Limit behavior of two-time-scale diffusions revisited, J. Differential Equations, 212 (2005), 85-113.

History

- A.V. Skorohod, Asymptotic Methods of the Theory of Stochastic Differential Equations, Transactions on Mathematical Monographs, vol. 78, American Mathematical Society, Providence, RI, 1989.
- H.J.Kushner, Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems, Birkhäuser, Boston, MA, 1990.
- E. Pardoux, A.Yu. Veretennikov, On Poisson equation and diffusion approximation 1, Ann. Probab. 29 (2001) 1061-1085.
- E. Pardoux, A.Yu. Veretennikov, On Poisson equation and diffusion approximation 2, Ann. Probab. 31 (2003) 1166-1192.
- E. Pardoux, A. Yu. Veretennikov, On Poisson equation and diffusion approximation. 3, Ann. Probab. 33 (2005), 1111-1133.

History 2: Delay system

Equation: $\dot{x}^\varepsilon(t) = a(x^\varepsilon(t), x^\varepsilon(t-\tau), \xi^\varepsilon(t)) + \varepsilon^{-1} b(x^\varepsilon(t), x^\varepsilon(t-\tau), \xi^\varepsilon(t)),$

where $\xi^\varepsilon(t) = \xi(t/\varepsilon^2)$ and $\xi(\cdot)$ is a right continuous stationary ϕ-mixing process.

Method: The martingale method+Weak convergence and approximation (The perturbed test function approach) (Kushner 1984).

Result: $x^\varepsilon(t) \Rightarrow x(t)$, where $x(t)$ satisfies

$$dx(t) = (\bar{a}(x(t), x(t-\tau)) + \bar{b}(x(t), x(t-\tau)))dt + \Phi(x(t), x(t-\tau))dw(t)$$

and $\bar{a}(x,y) = \lim_{t\to\infty} \mathbb{E}a(x, y, \xi(t))$, $\int_{T_1}^{T_2} \mathbb{E}b_x(x, y, \xi(u))b(x, y, \xi(T_1))du \to \bar{b}(x, y)$, $\int_{T_1}^{T_2} \mathbb{E}b_i(x, y, \xi(u))b_j(x, y, \xi(T_1))du \to 1/2S_{1,ij}(x, y)$ as T_1, T_2, $T_2 - T_1 \to \infty$ for each x, y, and $S(x, y) = \Phi\Phi'$.

♣ **The averaging principle needs to be reestablished.**

History

Delay system:

- G. Yin, K. M. Ramachandran, A differential delay equation with wideband noise perturbations, Stochastic Processes and their Applications, 35 (1990), 231–249.
- K. M. Ramachandran, A singularly perturbed stochastic delay system with small parameter, Stochastic Analysis and Applications, 11 (1993), 209–230.
- K. M. Ramachandran, Stability of stochastic delay differential equation with a small parameter, Stochastic Analysis and Applications, 26 (2008), 710–723.

Question (Averaging Principle): When $\varepsilon \to 0$, under appropriate conditions, what is the limit system of the functional two-time-scale diffusion system (1.1).

Difficulty:

(1) $(x^{\varepsilon}(t), \xi^{\varepsilon}(t))'$ is not Markov. The perturbed Fokker–Planck equation method cannot be used.

(2) If we see $\xi(t)$ as a noise and use the Weak convergence and the martingale method, It is necessary to examine the differential of the functional $V(t, x(t), x_t)$. This method doesn't work if we cannot find the derivative of the delay term x_t.

Method: Dupire recently developed a functional Itô formula, which can be used to examine this problem.

B. Dupire, Functional Itô Calculus, Portfolio Research paper 2009-04, Bloomberg, 2009.

Martingale Methods

Define an infinitesimal operator $\hat{\mathcal{L}}^\varepsilon$: we say $f \in \mathcal{D}(\hat{\mathcal{L}}^\varepsilon)$, the domain of $\hat{\mathcal{L}}^\varepsilon$, and $\hat{\mathcal{L}}^\varepsilon f = g$ if

$$\mathrm{p}-\lim_{\delta\to 0}\left[\frac{\mathbb{E}_t^\varepsilon f(t+\delta)-f(t)}{\delta}-g(t)\right]=0,$$

where p-lim is defined as follows: $f =$p-lim$_\delta f^\delta$ if and only if

$$\begin{cases} \sup_{t,\delta}\mathbb{E}|f^\delta(t)|<\infty \\ \lim_{\delta\to 0}\mathbb{E}|f^\delta(t)-f(t)|=0 \quad \text{for each} \quad t. \end{cases}$$

Lemma (Kurtz, 1975)

If $f \in \mathcal{D}(\hat{\mathcal{L}}^\varepsilon)$. Then

$$f(t)-\int_0^t \hat{\mathcal{L}}^\varepsilon f(u)du =: M_\varepsilon^f(t)$$

is a martingale, and also with probability 1,

$$\mathbb{E}_t^\varepsilon f(t+s)-f(s)=\int_t^{t+s}\mathbb{E}_t^\varepsilon\hat{\mathcal{L}}^\varepsilon f(u)du.$$

Martingale Methods and Weak Convergence

Method:

(S1) For $f^\varepsilon \in \mathcal{D}(\hat{\mathcal{L}}^\varepsilon)$, then $M_f^\varepsilon(t) := f^\varepsilon(t) - \int_0^t \hat{\mathcal{L}}^\varepsilon f^\varepsilon(u)du$ is a martingale.

(S2) Suppose that L is the infinitesimal operator (the Itô operator) for a diffusion process. For bounded continuous twice differentiable function f with compact support, $M_f(t) = f(x(t)) - \int_0^t Lf(x(s))ds$ is a martingale.

(S3) For $x^\varepsilon(t)$, if $\mathbb{E}|f^\varepsilon(t) - f(x^\varepsilon(t))| \to 0$ and $\mathbb{E}|\hat{\mathcal{L}}^\varepsilon f^\varepsilon(t) - Lf(x^\varepsilon(t))| \to 0$ as $\varepsilon \to 0$ for each t and that $\sup_{t\le T,\varepsilon>0} \mathbb{E}|\hat{\mathcal{L}}^\varepsilon f^\varepsilon(t)| < \infty$ for each $T < \infty$.

(S4) Then as $\varepsilon \to 0$, $f(x^\varepsilon(t)) - \int_0^t Lf(x^\varepsilon(s))ds$ is also a martingale.

(S5) If $\{x^\varepsilon(\cdot)\}$ is tight, the limit of any weakly convergent subsequence also satisfies the above martingale representation, which implies that $x^\varepsilon(t) \Rightarrow x(t)$.

Important points:

(1) Define $f^\varepsilon(t)$;

(2) Determine $\hat{\mathcal{L}}^\varepsilon$ and $x(t)$;

(3) The tightness of $x^\varepsilon(\cdot)$.

Outline

1. Problem and Research History
2. Invariant measure of the fast-varying process
3. Functional Itô formula and the martingale representation
4. Weak convergence to the averaging system

Let $\xi^\varepsilon(t) = \xi(t/\varepsilon)$. Then $\xi(t)$ satisfies the SDE

$$d\xi(t) = h(\xi(t))dt + \phi(\xi(t))d\tilde{w}(t), \tag{2.1}$$

where $\tilde{w}(t) = w_1(\varepsilon t)/\sqrt{\varepsilon}$ is a standard Brownian motion. Let us impose following assumptions on the coefficients $h(\cdot)$ and $\phi(\cdot)$ (The dissipative condition:).

(A1) $h(\cdot)$ is locally Lipschitz continuous and there exists λ_1 such that for any y_1, $y_2 \in \mathbb{R}^m$,

$$\langle y_1 - y_2, h(y_1) - h(y_2)\rangle \le -\lambda_1 |y_1 - y_2|^2$$

and $\phi(\cdot)$ is global Lipschitz continuous, i.e., there exists λ_2 such that

$$|\phi(y_1) - \phi(y_2)|^2 \le \lambda_2 |y_1 - y_2|^2.$$

Theorem

Under **(A1)**, *Eq. (2.1) admits a unique strong solution $\xi(t;\xi_0)$, which is $\mathcal{F}_{1\varepsilon t}$-adapted and this solution holds the following properties:*

(i) *this solution is a strong homogeneous Markov process;*

(ii) $\mathbb{E}\Big(\sup_{t\in[0,T]}|\xi(t)|^2\Big) \le C_T$, *where C_T is a constant dependent on T;*

(iii) *If $2\lambda_1 > \lambda_2$, then (2.1) has a unique invariant measure $\mu(\cdot)$, which is exponentially ergodic.*

Assumption (A1) is only used to guarantee the existence and uniqueness of Eq. (2.1), and the boundedness of this solution as well as the exponential ergodicity. If other conditions can guarantee these results, these conditions can be used.

Outline

Horizontal derivative

Let us give the definitions of the *horizontal derivative* and the *vertical derivative*. Denote $D([0,T];\mathbb{R}^n)$ the space of function on $[0,T]$ with values in $\mathbb{R}^n$ which are right continuous functions with left limits.

Definition

For a stopped process $x_t = \{x(u \wedge t) : 0 \leq u \leq T\}$, a non-anticipative functional $F : [0,T] \times D([0,T];\mathbb{R}^n) \to \mathbb{R}$ is said to be horizontally differentiable at $(t,x) \in [0,T] \times D([0,T];\mathbb{R}^n)$ if the limit

$$\mathcal{D}F(t,x) = \lim_{\delta \to 0^+} \frac{F(t+\delta, x_t) - F(t, x_t)}{\delta}$$

exists. $\mathcal{D}F(t,x)$ is called as the horizontal derivative of F at (t,x)

Definition

For a stopped process $x_t = \{x(u \wedge t) : 0 \le u \le T\}$, a non-anticipative functional $F : [0, T] \times D([0, T]; \mathbb{R}^n) \to \mathbb{R}$ is said to be vertically differentiable at $(t, x) \in [0, T] \times D([0, T]; \mathbb{R}^n)$ if the functional map

$$e \mapsto F(t, x_t + e\mathbf{1}_{[t,T]})$$

is differentiable at 0, where $e = (e_1, e_2, \cdots, e_n)'$ represents the canonical basis on $\mathbb{R}^n$. Its gradient at 0 is called as the vertical derivative of F at (t, x):

$$\nabla_x F(t, x) = (\partial_1 F(t, x), \partial_2 F(t, x), \cdots, \partial_n F(t, x)),$$

where

$$\partial_i F(t, x) = \lim_{\delta \to 0^+} \frac{F(t, x_t + \delta e_i \mathbf{1}_{[t,T]}) - F(t, x_t)}{\delta}.$$

If F is vertically differentiable at all $(t, x) \in [0, T] \times D([0, T]; \mathbb{R}^n)$, then $\nabla_x F$ is a non-anticipative functional called the vertical derivative of F. Repeating this operation may give the definition of $\nabla_x^2 F(t, x) = [\partial_{ij}^2 F(t, x)]_{n \times n}$ as the gradient at 0 (if it exists) of the map

$$e \mapsto \nabla_x F(t, x_t + e\mathbf{1}_{[t,T]}).$$

Define $\mathcal{C}^{1,2}([0,T]\times C([0,T];\mathbb{R}^n);\mathbb{R})$ as the family of the non-anticipative functional which have continuously horizontal derivative, and continuous twice vertical derivative. If $X(t)$ is a semimartingale processes, the functional Itô formula can be given (R. Cont, D. A. Fournié, Functional Itô calculus and stochastic integral representation of martingales, The Annals of Probability, 41(1) (2013), 109–133.).

Theorem (Functional Itô formula)

For any $V\in\mathcal{C}^{1,2}([0,T]\times C([0,T];\mathbb{R}^n);\mathbb{R})$,

$$\begin{aligned} V(t,X_t) &= V(0,X_0)+\int_0^t \mathcal{D}V(u,X_u)du+\int_0^t \nabla_x V(u,X_u)dX(u)\\ &\quad+\frac{1}{2}\int_0^t \mathrm{trace}(\nabla_x^2 V(u,X_u)d[X](u)),\quad a.s. \end{aligned} \tag{3.1}$$

where $[X](u)$ *is the quadratic variation process of the solution process* $X(t)$*. In particular,* $Y(t)=V(t,X_t)$ *is a continuous semimartingale.*

let us consider the following stochastic functional differential equation

$$dX(t) = B(t, X_t)dt + \Psi(t, X_t)dW(t), \tag{3.2}$$

with the deterministic initial value $X(0) \in \mathbb{R}^n$. Substituting the solution process $X(t)$ and the corresponding quadratic variation process $[X](u)$ into Eq. (3.1) gives that

$$V(t, X_t) - V(0, X_0) = \int_0^t \mathcal{L}V(u, X_u)du + \int_0^t \nabla_x V(u, X_u)\Psi(u, X_u)dW(u), \quad a.s. \tag{3.3}$$

where

$$\mathcal{L} = \mathcal{D} + \sum_{i=1}^n B_i(t, x)\partial_i + \frac{1}{2}\sum_{i,j=1}^n \Psi_i(t, x)\Psi_j(t, x)\partial_{ij}^2 \tag{3.4}$$

is an infinitesimal generator for any $x \in C([0, T]; \mathbb{R}^n)$. Under suitable conditions,

$$V(t, X_t) - V(0, X_0) - \int_0^t \mathcal{L}V(u, X_u)du \tag{3.5}$$

is a martingale with respect to the σ-algebra filtration $\mathcal{F}_t^W$.

If $V(t, X_t) = v(t, X(t)) \in C^{1,2}([0,T] \times \mathbb{R}^n; \mathbb{R})$, the generator $\mathcal{L}$ becomes L according to the standard Itô formula.

$$L(t,x)\cdot = \frac{\partial \cdot}{\partial t} + \sum_{i=1}^{n} B_i(t,x)\frac{\partial \cdot}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^{n} \Psi_i(t,x)\Psi_j(t,x)\frac{\partial^2 \cdot}{\partial y_i \partial y_j} \quad (3.6)$$

for $x \in C([0,T]; \mathbb{R}^n)$ and $y = (y_1, y_2, \ldots, y_n)' \in \mathbb{R}^n$, and

$$v(t, X(t)) - v(0, X(0)) - \int_0^t L(u, x_u) v(u, X(u)) du$$

is a martingale with respect to the filtration $\mathcal{F}_t^W$. This also implies that for any $\delta > 0$,

$$\mathbb{E}_t^W v(t+\delta, X(t+\delta)) - v(t, X(t)) = \int_t^{t+\delta} L(u, x_u) v(u, X(u)) du \ \text{ a.s.} \quad (3.7)$$

Similarly, let us define $\mathbb{C}^{1,2,2}([0,T]\times\mathbb{R}^n\times C([0,T];\mathbb{R});\mathbb{R})$ and give the following mixed functional Itô formula.

Theorem (Mixed functional Itô formula)

Let $X(t)$ be the solution of (3.2). For any $V\in\mathbb{C}^{1,2,2}([0,T]\times\mathbb{R}^n\times C([0,T];\mathbb{R}^n);\mathbb{R})$,

$$\begin{aligned}
&V(t,X(t),X_t)\\
&=V(0,X(0),X_0)+\int_0^t\mathcal{D}V(u,X(u),X_u)du\\
&\quad+\int_0^t[V_y(u,X(u),X_u)+\nabla_xV(u,X(u),X_u)]dX(u)\\
&\quad+\frac{1}{2}\int_0^t|[V_{yy}(u,X(u),X_u)+\nabla_x^2V(u,X(u),X_u)+2\nabla_xV_y(u,X(u),X_u)]d[X](u)|\\
&=V(0,X(0),X_0)+\int_0^t\mathbb{L}(u,X_u)V(u,X(u),X_u)du\\
&\quad+\int_0^t[V_y(u,X(u),X_u)+\nabla_xV(u,X(u),X_u)]\Psi(u,X_u)dW(u)\quad a.s.,
\end{aligned}\tag{3.8}$$

where $\mathcal{D}$, ∇_x and ∇_x^2 are defined as above, and

Theorem (continue)

$$V_y(t,y,x) = \Big(\frac{\partial V(t,y,x)}{\partial y_1}, \frac{\partial V(t,y,x)}{\partial y_2}, \dots, \frac{\partial V(t,y,x)}{\partial y_2}\Big),$$

$$V_{yy} = \Big[\frac{\partial^2 V(t,y,x)}{\partial y_i \partial y_j}\Big]_{n\times n} \text{ and } \nabla_x V_y(t,y,x) = \Big[\partial_i\Big(\frac{\partial V(t,y,x)}{\partial y_j}\Big)\Big]_{n\times n},$$

and

$$\mathbb{L}(t,x)\cdot = \mathcal{D}\cdot + \sum_{i=1}^n B_i(t,x)\Big[\frac{\partial}{\partial y_i} + \partial_i\Big]\cdot + \frac{1}{2}\sum_{i,j=1}^n \Psi_i(t,x)\Psi_j(u,x)\Big[\frac{\partial^2}{\partial y_i \partial y_j} + 2\partial_i\Big(\frac{\partial}{\partial y_j}\Big) + \partial_{ij}^2\Big]\cdot$$

is an infinitesimal generator for any $(t,x) \in [0,T] \times C([0,T];\mathbb{R}^n)$*. In particular,* $Y(t) = V(t, X(t), X_t)$ *is a continuous semimartingale and*

$$V(t,X(t),X_t) - V(0,X(0),X_0) - \int_0^t \mathbb{L}(u,X_u)V(u,X(u),X_u)du$$

is a martingale with respect to the filtration $\mathcal{F}_t^W$*.*

Outline

Tightness and truncation technique

In order to obtain the desired weak convergence, we need to prove the tightness first. We need to verify

$$\lim_{N_0\to\infty}\limsup_{\varepsilon\to 0}\mathbb{P}\Big(\sup_{t\le T}|x^{\varepsilon}(t)|\ge N_0\Big)=0 \quad \text{for each} \quad T<\infty.$$

The verification is usually quite involved, and requires complicated calculations. To circumvent the difficulties, we use the truncation technique as follows. For each $N>0$ sufficient large such that $|x(0)|\le N$, let us introduce the following truncated equation

$$dx^{\varepsilon,N}(t)=b^N(x_t^{\varepsilon,N},\xi^{\varepsilon}(t))dt+\psi^N(x_t^{\varepsilon,N},\xi^{\varepsilon}(t))dw_2(t), \tag{4.1}$$

where $x_t^{\varepsilon,N}=\{x^{\varepsilon,N}(t\wedge u):0\le u\le T\}$, $b^N(x_t^{\varepsilon,N},\xi^{\varepsilon}(t))=b(x_t^{\varepsilon,N},\xi^{\varepsilon}(t))q(x_t^{\varepsilon,N})$, $\psi^N(x_t^{\varepsilon,N},\xi^{\varepsilon}(t))=\psi(x_t^{\varepsilon,N},\xi^{\varepsilon}(t))q(x_t^{\varepsilon,N})$, and

$$q(x)=\begin{cases}1, & \text{when} \quad x\in C([0,T];S_N),\\ 0, & \text{when} \quad x\in C([0,T];\mathbb{R}^n-S_{N+1}),\\ \text{smooth}, & \text{otherwise.}\end{cases}$$

It is obvious that $x^{\varepsilon,N}(t)=x^{\varepsilon}(t)$ up until the first exit from $S_N=\{x:|x|\le N\}$.

Relationship among those σ-algebra filtration

Let $\mathcal{F}_t^{\varepsilon,N} = \sigma\{x^{\varepsilon,N}(s), s \leq t\}$. We can also give the corresponding definitions for $\overline{\mathcal{M}}^{\varepsilon,N}$ and $\hat{\mathcal{L}}^{\varepsilon,N}$. It is obvious that $\mathcal{F}_t^{\varepsilon,N} \subset \sigma(\mathcal{F}_t^{w_2} \vee \mathcal{F}_t^{\xi^\varepsilon})$, where $\mathcal{F}_t^{w_2}$ and $\mathcal{F}_t^{\xi^\varepsilon}$ are generated by the Brownian motion $w_2(t)$ and the stochastic process $\xi^\varepsilon(t) = \xi(t/\varepsilon)$, respectively. Since $\xi(\cdot)$ is the solution of (2.1), $\mathcal{F}_t^{\xi^\varepsilon} \subset \mathcal{F}_{t/\varepsilon}^{\tilde{w}}$, where $\mathcal{F}_t^{\tilde{w}}$ is the σ-algebra filtration generated by the Brownian motion $\tilde{w}(t)$. These show that $\mathcal{F}_t^{\varepsilon,N} \subset \sigma\left\{\mathcal{F}_t^{w_2} \vee \mathcal{F}_{\frac{t}{\varepsilon}}^{\tilde{w}}\right\}$.

To proceed, the following assumptions are needed.

(A2) $b(x,\xi)$, $\psi(x,\xi)$, $\nabla_x b_i(x,\xi)$, $\nabla_x \psi_{ij}(x,\xi)$, $\nabla_x^2 b_i(x,\xi)$ and $\nabla_x^2 \psi_{ij}(x,\xi), i,j = 1,\cdots,n$ are boundedness preserving with respect to $x \in C([0,T];\mathbb{R}^n)$, and continuous and bounded with respect to $\xi \in \mathbb{R}^m$ for $x \in G$, where $G \subset \mathbb{R}^n$ is a compact set;

(A3) For $G \subset \mathbb{R}^n$ being a compact set, $x \in C([0,T];G)$, $b(x,\cdot)$ and $\psi'(x,\cdot)\psi(x,\cdot) = A(x,\cdot) = [a_{ij}(x,\cdot)]_{n\times n}$ are integrable functionals with respect to the measure μ, and assume that

$$\begin{cases} \int_{\mathbb{R}^n} b(x,\xi)\mu(d\xi) = \bar{b}(x), \\ \int_{\mathbb{R}^n} a_{ij}(x,\xi)\mu(d\xi) = \bar{a}_{ij}(x), \end{cases}$$

that is, $\bar{b}(x) = \mu b(x,\xi)$ and $\bar{a}_{ij}(x) = \mu a_{ij}(x,\xi)$. Moreover, $A(\cdot,\cdot)$ is nonnegative definite (this shows that $\bar{A}(\cdot) = [\bar{a}_{ij}(\cdot)]_{n\times n}$ is also nonnegative definite) and set $\bar{\psi}'(\cdot)\bar{\psi}(\cdot) = A(\cdot)$.

(A4) The following equation

$$dx(t) = \bar{b}(x_t)dt + \bar{\psi}(x_t)dB(t) \tag{4.2}$$

has a unique weak solution (here uniqueness in the sense of distribution) on $[0,T]$ for each continuous deterministic initial value $x(0)$, where $B(t)$ is a standard Brownian motion, $\bar{b} = (\bar{b}_1, \bar{b}_2, \cdots, \bar{b}_n)'$, $\bar{\sigma} = [\bar{\sigma}_{ij}]_{n\times n}$.

Remark: According to the Markov property of $\xi(t)$, **(A3)** implies that for any bounded continuous functional $b(x,\cdot)$ and $a_{ij}(x,\cdot)$, and deterministic initial value $\xi(0)$,

$$\begin{cases} \lim_{t\to\infty} \mathcal{P}_t^{\xi} b(x,\xi(0)) = \lim_{t\to\infty} \mathbb{E}_t^{\xi} b(x,\xi(0)) = \lim_{t\to\infty} \mathbb{E} b(x,\xi(t)) = \bar{b}(x), \\ \lim_{t\to\infty} \mathcal{P}_t^{\xi} a_{ij}(x,\xi(0)) = \lim_{t\to\infty} \mathbb{E}_t^{\xi} a_{ij}(x,\xi(0)) = \lim_{t\to\infty} \mathbb{E} a_{ij}(x,\xi(t)) = \bar{a}_{ij}(x), \end{cases}$$

where $\mathbb{E}_t^{\xi}$ represent the conditional mathematical expectation with respect to the σ-algebra filtration $\mathcal{F}_t^{\xi}$ generated by the solution process $\xi(s)$ of Eq.(2.1).

According to Theorem 5 and the martingale representation formula (3.6), for solution $x(t)$ of (4.2), applying the operator L to the function $v \in C^{1,2}([0,T]\times\mathbb{R}^n;\mathbb{R})$ yields

$$M_f(t) = v(t,x(t)) - v(0,x(0)) - \int_0^t L(u,x_u)v(u,x(u))du \qquad (4.3)$$

is a martingale, where for $x \in C([0,T];\mathbb{R}^n)$,

$$L(t,x)\cdot = \frac{\partial\cdot}{\partial t} + \sum_{i=1}^n \bar{b}_i(x)\frac{\partial\cdot}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^n \bar{a}_{ij}(x)\frac{\partial^2\cdot}{\partial y_i\partial y_j}. \qquad (4.4)$$

We also say that $x(\cdot)$ solves the martingale problem for operator L if (4.3) holds.

We also use the operators L^N, $\mathcal{L}^N$ and $\mathbb{L}^N$, which means that the operators L, $\mathcal{L}$ and $\mathbb{L}$ with $x, y, \bar{b}, \bar{\psi}$ and $\bar{A}$ replaced by $x^N, y^N, \bar{b}^N, \bar{\psi}^N$ and $\bar{A}^N$, respectively.

Under these assumptions, we first present the following tightness of $\{x^{\varepsilon,N}(\cdot)\}$ in the space $C([0,T];\mathbb{R}^n)$.

Theorem

Under assumptions (A2), there exists a unique strong solution $x^{\varepsilon,N}(t)$ for the stochastic functional differential truncated equation (4.1) for any initial value $|x(0)| \leq N$, and this solution is continuous and $\mathcal{F}_t^{\varepsilon,N}$-adapted. Moreover, for each $N > 0$, this solution $\{x^{\varepsilon,N}(\cdot)\}$ is tight in $C([0,T];\mathbb{R}^n)$.

To prove this theorem, we need the following lemma (also see Kushner1984).

Lemma

Let $X^\varepsilon(\cdot)$ denote a sequence of $\mathcal{F}_t^\varepsilon$-measurable processes with paths in $C([0,T];\mathbb{R}^n)$ satisfying

$$\lim_{K\to\infty}\limsup_{\varepsilon\to 0}\mathbb{P}\{\sup_{t\le T}|X^\varepsilon(t)|\ge K\}=0 \tag{4.5}$$

for each $T<\infty$, $\delta>0$, and $\varepsilon>0$. Let there be a random variable $\gamma_\varepsilon(\delta)$ such that

$$\begin{cases} \mathbb{E}_t^\varepsilon\gamma_\varepsilon(\delta)\ge\mathbb{E}_t^\varepsilon\min\{1,|X^\varepsilon(t+v)-X^\varepsilon(t)|^2\}, \quad \textit{all}\ \ 0\le v\le\delta, \quad t\le T, \\ \lim\limits_{\delta\to 0}\limsup\limits_{\varepsilon\to 0}\mathbb{E}\gamma_\varepsilon(\delta)=0, \end{cases} \tag{4.6}$$

where $\mathbb{E}_t^\varepsilon$ is the conditional mathematical expectation with respect to $\mathcal{F}_t^\varepsilon$. Then $X^\varepsilon(\cdot)$ is tight in $C([0,\infty);\mathbb{R}^n)$. Instead (4.6), suppose that for each $T<\infty$,

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\sup_{\tau\le T}\mathbb{E}\min\{1,|X^\varepsilon(\rho+\delta)-X^\varepsilon(\rho)|^2\}=0, \tag{4.7}$$

where ρ is an $\mathcal{F}_t^\varepsilon$-stopping time. Then $X^\varepsilon(\cdot)$ is tight in $C([0,T];\mathbb{R}^n)$.

Theorem

If (A1) – (A3) hold, then $\{x^\varepsilon(\cdot)\}$ *is tight in* $C([0,T];\mathbb{R}^n)$*, and the limit of any weakly convergent subsequence satisfies equation (4.2).*

To do this, we shall apply the following lemma (Ref. Yin 90 or Kushner 84).

Lemma

Let $X^\varepsilon(\cdot)$ *be* $\mathbb{R}^n$*-valued and defined on* $[0,T]$*, with* $X^\varepsilon(0)$ *being deterministic. Let* $\{X^\varepsilon(\cdot)\}$ *be tight on* $C([0,T];\mathbb{R}^n)$*. Suppose (A4) holds, and for each* $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$ *(or any dense subset of it), each* $T<\infty$*, there exists* $f^\varepsilon(\cdot)\in\mathcal{D}(\hat{\mathcal{L}}^\varepsilon)$*, such that*

$$\text{p-}\lim_{\varepsilon\to 0}[f^\varepsilon(\cdot)-f(X^\varepsilon(\cdot))]=0 \tag{4.8}$$

and

$$\text{p-}\lim_{\varepsilon\to 0}[\hat{\mathcal{L}}^\varepsilon f^\varepsilon(\cdot)-\mathcal{L}f(X^\varepsilon(\cdot))]=0. \tag{4.9}$$

Then, $X^\varepsilon(\cdot)\Rightarrow X(\cdot)$*, where* $X(\cdot)$ *is the weak solution of the stochastic differential equation (4.2).*

Under the truncation technique, for $x^{\varepsilon,N}(t)$ and any $f(\cdot) \in C_0^4(\mathbb{R}^n;\mathbb{R})$, we need to look for the function $f^{\varepsilon,N}(\cdot) \in \mathcal{D}(\hat{\mathcal{L}}^{\varepsilon,N})$ and test the corresponding conditions:

$$\begin{cases} \sup_{t,\varepsilon} \mathbb{E}|f^{\varepsilon,N}(t) - f(x^{\varepsilon,N}(t))| < \infty, \\ \lim_{\varepsilon\to 0} \mathbb{E}|f^{\varepsilon,N}(t) - f(x^{\varepsilon,N}(t))| = 0 \quad \text{for each} \quad t. \end{cases} \tag{4.10}$$

Similarly, to prove (4.9) for the above $x^{\varepsilon,N}(t)$ and $f(\cdot)$, we also need to test the similar condition as follows:

$$\begin{cases} \sup_{t,\varepsilon} \mathbb{E}|\hat{\mathcal{L}}^{\varepsilon,N} f^{\varepsilon,N}(t) - L^N(t, x_t^{\varepsilon,N}) f(x^{\varepsilon,N}(t))| < \infty, \\ \lim_{\varepsilon\to 0} \mathbb{E}|\hat{\mathcal{L}}^{\varepsilon,N} f^{\varepsilon,N}(t) - L^N(t, x_t^{\varepsilon,N}) f(x^{\varepsilon,N}(t))| = 0 \quad \text{for each} \quad t. \end{cases} \tag{4.11}$$

Sketch of Proof

For any $f(\cdot) \in C_0^4(\mathbb{R}^n; \mathbb{R})$, we choose

$$f_1^{\varepsilon,N}(t) := V_1(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N}) = \int_t^T f_y(x^{\varepsilon,N}(t))\mathbb{E}_t^{\xi^\varepsilon}[b^N(x_t^{\varepsilon,N}, \xi^\varepsilon(u)) - \bar{b}^N(x_t^{\varepsilon,N})]du, \tag{4.12}$$

$$f_2^{\varepsilon,N}(t) := V_2(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N}) = \sum_{i,j=1}^n \int_t^T f_{y_i y_j}(x^{\varepsilon,N}(t))\mathbb{E}_t^{\xi^\varepsilon}[a_{ij}^N(x_t^{\varepsilon,N}, \xi^\varepsilon(u)) - \bar{a}_{ij}^N(x_t^{\varepsilon,N})]du. \tag{4.13}$$

Making change of variable u/ε to u implies that

$$f_1^{\varepsilon,N}(t) = \varepsilon \int_{t/\varepsilon}^{T/\varepsilon} f_y(x^{\varepsilon,N}(t))\mathbb{E}_{t/\varepsilon}^{\xi}[b^N(x_t^{\varepsilon,N}, \xi(u)) - \bar{b}^N(x_t^{\varepsilon,N})]du, \tag{4.14}$$

$$f_2^{\varepsilon,N}(t) = \varepsilon \sum_{i,j=1}^n \int_{t/\varepsilon}^{T/\varepsilon} f_{y_i y_j}(x^{\varepsilon,N}(t))\mathbb{E}_{t/\varepsilon}^{\xi}[a_{ij}^N(x_t^{\varepsilon,N}, \xi(u)) - \bar{a}_{ij}^N(x_t^{\varepsilon,N})]du. \tag{4.15}$$

where $\mathbb{E}_t^\xi$ is the conditional expectation with respect to $\mathcal{F}_t^\xi$. Define

$$f^{\varepsilon,N}(t) = f(x^{\varepsilon,N}(t)) + f_1^{\varepsilon,N}(t) + \frac{1}{2} f_2^{\varepsilon,N}(t). \tag{4.16}$$

Assumption (A2) shows that $b(x,\xi)$ is boundedness preserving on x and $b(x,\cdot)$ is continuous and bounded, which implies that $b^N(x,\cdot)$ is bounded continuous for any $x \in C([0,T];S_N)$, i.e., $b^N(x,\cdot) \in C_b(\mathbb{R}^m;\mathbb{R}^n)$. Note that the invariant measure μ is strongly exponential ergodic. According to assumption (A3), applying homogeneity of $\xi(\cdot)$ gives

$$\begin{aligned}
\sup_{t\le T}|f_1^{\varepsilon,N}(t)| &= \varepsilon \sup_{t\le T}\Big|\int_{t/\varepsilon}^{T/\varepsilon} f_y(x^{\varepsilon,N}(t))\mathbb{E}^{\xi}_{t/\varepsilon}[b^N(x_t^{\varepsilon,N},\xi(u)) - \bar{b}^N(x_t^{\varepsilon,N})]du\Big|, \\
&\le \varepsilon|f_y(x^{\varepsilon,N}(t))|\sup_{t\le T}\int_{t/\varepsilon}^{T/\varepsilon} |\mathcal{P}_{t/\varepsilon,u}b^N(x_t^{\varepsilon,N},\xi(t/\varepsilon)) - \mu b^N(x_t^{\varepsilon,N},\xi)|du \\
&= \varepsilon|f_y(x^{\varepsilon,N}(t))|\sup_{t\le T}\int_{t/\varepsilon}^{T/\varepsilon} |\mathcal{P}_{u-t/\varepsilon}b^N(x_t^{\varepsilon,N},\xi(t/\varepsilon)) - \mu b^N(x_t^{\varepsilon,N},\xi)|du \\
&\le \varepsilon K \sup_{t\le T}\int_{t/\varepsilon}^{T/\varepsilon} e^{-\frac{\tilde{\xi}}{2}(u-t/\varepsilon)}du, \\
&= \varepsilon K(1-e^{-\frac{\tilde{\xi}(T-t)}{2\varepsilon}}) = O(\varepsilon), \qquad (4.17)
\end{aligned}$$

which implies

$$|f_1^{\varepsilon,N}(t)| \to 0 \ \text{ w.p.1}$$

as $\varepsilon \to 0$ since $f(\cdot) \in C_0^4(\mathbb{R}^n;\mathbb{R})$.

Applying the technique similar to (4.17) gives

$$\sup_{t\leq T}|f_2^{\varepsilon,N}(t)|=O(\varepsilon), \tag{4.18}$$

which implies $|f_2^{\varepsilon,N}(t)|\to 0$ w. p. 1 as $\varepsilon\to 0$ since $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$. These imply that there exist $\varepsilon_0>0$ and η, for any $\varepsilon\in(0,\varepsilon_0)$,

$$\mathbb{E}|f_1^{\varepsilon,N}(t)|\vee\mathbb{E}|f_2^{\varepsilon,N}(t)|<\eta$$

and

$$\lim_{\varepsilon\to 0}\mathbb{E}|f^{\varepsilon,N}(t)-f(x^{\varepsilon,N}(t))|=0 \quad \text{for each} \quad t. \tag{4.19}$$

These shows (4.10) holds. According to Definition of p-lim,

$$\text{p-}\lim_{\varepsilon\to 0}[f^{\varepsilon,N}(\cdot)-f(x^{\varepsilon,N}(\cdot))]=0,$$

that is, (4.8) holds.

For $(t,y,x)\in[0,T]\times\mathbb{R}^n\times C([0,T];\mathbb{R}^n)$, let us define

$$V(t,y,x)=V_1(t,y,x)+\frac{1}{2}V_2(t,y,x).$$

According to the definitions of $f^{\varepsilon,N}(\cdot)$, $\mathcal{L}^{\varepsilon,N}$, L^N and $\mathbb{L}^N$, applying L^N and $\mathbb{L}^N$ gives

$$\begin{aligned}\mathcal{L}^{\varepsilon,N}f^{\varepsilon,N}(t) &= \text{p-}\lim_{\delta\to 0}\frac{\mathbb{E}_t^{\varepsilon,N}f^{\varepsilon,N}(t+\delta)-f^{\varepsilon,N}(t)}{\delta}\\ &= \text{p-}\lim_{\delta\to 0}\frac{\mathbb{E}_t^{\varepsilon,N}f(x^{\varepsilon,N}(t+\delta))-f(x^{\varepsilon,N}(t))}{\delta}\\ &\quad+\text{p-}\lim_{\delta\to 0}\frac{\mathbb{E}_t^{\varepsilon,N}V(t+\delta,x^{\varepsilon,N}(t+\delta),x_{t+\delta}^{\varepsilon,N})-V(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})}{\delta}\\ &= L^N(t,x_t^{\varepsilon,N})f(x^{\varepsilon,N}(t))+\mathbb{L}^N(t,x_t^{\varepsilon,N})V(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N}),\end{aligned}\tag{4.20}$$

where $\mathbb{E}_t^{\varepsilon,N}$ is the conditional expectation with respect to $\mathcal{F}_t^{\varepsilon,N}=\sigma\{x^{\varepsilon,N}(s);s\le t\}$, and

$$\mathbb{L}^N V(t,y,x)=\mathbb{L}^N(t,x)V_1(t,y,x)+\frac{1}{2}\mathbb{L}^N(t,x)V_2(t,y,x).$$

Let us firstly examine $\mathbb{L}^N(t,x)V_1(t,y,x)$. Note that

$$\begin{aligned}&\mathbb{L}^N(t,x_t^{\varepsilon,N})V_1(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})\\ &=\varepsilon\sum^{n}\mathbb{L}^N(t,x_t^{\varepsilon,N})\Big[\int^{T/\varepsilon}f_{y_i}(x^{\varepsilon,N}(t))\mathbb{E}^{\xi}_{t/\varepsilon}[b_i^N(x_t^{\varepsilon,N},\xi(u))-\bar{b}_i^N(x_t^{\varepsilon,N})]du\Big].\end{aligned}\tag{4.21}$$

Note that $\mathcal{F}_t^{\varepsilon,N} \subset \sigma(\mathcal{F}_t^{w_2} \vee \mathcal{F}_t^{\xi^\varepsilon})$. This implies that when we apply $\mathbb{E}_t^{\varepsilon,N}$ to $\xi(\cdot)$, it is equivalent to using $\mathbb{E}_t^{\xi^\varepsilon}$. This, together with the definition of $\mathbb{L}^N$ yields

$$\varepsilon \mathbb{L}^N(t, x_t^{\varepsilon,N})\Big[\int_{t/\varepsilon}^{T/\varepsilon} f_{y_i}(x^{\varepsilon,N}(t))\mathbb{E}_{t/\varepsilon}^{\xi}(b_i^N(x_t^{\varepsilon,N},\xi(u)) - \bar{b}_i^N(x_t^{\varepsilon,N}))du\Big]$$
$$= I_1^\varepsilon + \varepsilon\sum_{j=1}^n I_{2j}^\varepsilon + \frac{\varepsilon}{2}\sum_{k,j=1}^n I_{3kj}^\varepsilon + \varepsilon\sum_{j=1}^n I_{4j}^\varepsilon + \varepsilon\sum_{k,j=1}^n I_{5kj}^\varepsilon + \frac{\varepsilon}{2}\sum_{k,j=1}^n I_{6kj}^\varepsilon, \qquad (4.22)$$

where $I_1^\varepsilon(y, x, \xi^\varepsilon(t)) = -f_{y_i}(y)[b_i^N(x, \xi^\varepsilon(t)) - \bar{b}_i^N(x)]$ and

$$I_{2j}^\varepsilon(t,y,x) = \int_{t/\varepsilon}^{T/\varepsilon} f_{y_iy_j}(y)\mathbb{E}_{t/\varepsilon}^{\xi}[b_i^N(x,\xi(u)) - \bar{b}_i^N(x)]du b_j^N(x,\xi^\varepsilon(t)),$$

$$I_{3kj}^\varepsilon(t,y,x) = \int_{t/\varepsilon}^{T/\varepsilon} f_{y_iy_ky_j}(y)\mathbb{E}_{t/\varepsilon}^{\xi}[b_i^N(x,\xi(u)) - \bar{b}_i^N(x)]du a_{kj}^N(x,\xi^\varepsilon(t)),$$

$$I_{4j}^\varepsilon(t,y,x) = \int_{t/\varepsilon}^{T/\varepsilon} f_{y_i}(y)\partial_j[\mathbb{E}_{t/\varepsilon}^{\xi}(b_i^N(x,\xi(u)) - \bar{b}_i^N(x))]du b_j^N(x,\xi^\varepsilon(t)),$$

$$I_{5kj}^\varepsilon(t,y,x) = \int_{t/\varepsilon}^{T/\varepsilon} f_{y_iy_k}(y)\partial_j[\mathbb{E}_{t/\varepsilon}^{\xi}(b_i^N(x,\xi(u)) - \bar{b}_i^N(x))]du a_{kj}^N(x,\xi^\varepsilon(t)),$$

$$I_{6kj}^\varepsilon(t,y,x) = \int_{t/\varepsilon}^{T/\varepsilon} f_{y_i}(y)\partial_{kj}^2[\mathbb{E}_{t/\varepsilon}^{\xi}(b_i^N(x,\xi(u)) - \bar{b}_i^N(x))]du a_{kj}^N(x,\xi^\varepsilon(t)).$$

Recall that $b^N(x,\cdot)\in C_b(\mathbb{R}^m;\mathbb{R}^n)$. This, together with boundedness preserving property of $b(x,\xi)$ on x, yields that $b^N(x,\xi)$ is bounded for any $x\in C([0,T];S_N)$ and $\xi\in\mathbb{R}^m$. Since and $f(\cdot)\in C_0^4(\mathbb{R}^n;\mathbb{R})$, similar to (4.17),

$$\begin{aligned}
&\sup_{t\le T}|I_{2j}^{\varepsilon}(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})|\\
&=\sup_{t\le T}\Big|\int_{t/\varepsilon}^{T/\varepsilon} f_{y_iy_j}(x^{\varepsilon,N}(t))\mathbb{E}^{\xi}_{t/\varepsilon}[b_i^N(x_t^{\varepsilon,N},\xi(u))-\bar{b}_i^N(x_t)]du\,b_j^N(x_t^{\varepsilon,N},\xi^{\varepsilon}(t))\Big|\\
&\le K_N\sup_{t\le T}\Big|\int_{t/\varepsilon}^{T/\varepsilon}|f_{y_iy_j}(x^{\varepsilon,N}(t))|e^{-\frac{\tilde{\varsigma}}{2}(u-t/\varepsilon)}du\Big|=\frac{2}{\tilde{\varsigma}}K_N(1-e^{-\frac{\tilde{\varsigma}(T-t)}{2\varepsilon}})<\infty,
\end{aligned}$$

which implies

$$\varepsilon\sup_{t\le T}\Big|\sum_{j=1}^n I_{2j}^{\varepsilon}(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})\Big|\le\varepsilon\sup_{t\le T}\sum_{j=1}^n|I_{2j}^{\varepsilon}(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})|=O(\varepsilon).$$

Likewise, we have

$$\frac{\varepsilon}{2}\sup_{t\le T}\Big|\sum_{k,j=1}^n I_{3kj}^{\varepsilon}(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})\Big|\le\frac{\varepsilon}{2}\sup_{t\le T}\sum_{k,j=1}^n|I_{3kj}^{\varepsilon}(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})|=O(\varepsilon)$$

since $A^N(x,\xi)$ is bounded for any $x\in C([0,T];S_N)$ and $\xi\in\mathbb{R}^m$ according to $a_{ij}(x,\cdot)\in C_b(\mathbb{R}^m;\mathbb{R})$ and the bounded preserving property of $\psi(x,\xi)$ for x.

By (A2), $\nabla_x b^N(x,\xi)$ is boundedness preserving for any $x \in C([0,T];\mathbb{R}^n)$, and continuous and bounded with respect to $\xi \in \mathbb{R}^m$. These show that for any $i,j = 1,\dots,n$, $\partial_j b_i^N(x,\cdot) \in C_b(\mathbb{R}^m;\mathbb{R})$ for any $x \in C([0,T];S_N)$, hence applying similar technique to the estimations of I_{2j}^ε and recalling the boundedness of $b^N(x,\xi)$ for any $x \in C([0,T];S_N)$ and $\xi \in \mathbb{R}^m$ give

$$\begin{aligned}
&\sup_{t\le T} |I_{4j}^\varepsilon(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N})| \\
&= \sup_{t\le T}\left| \int_{t/\varepsilon}^{T/\varepsilon} f_{y_i}(x^{\varepsilon,N}(t))\partial_j[\mathbb{E}_{t/\varepsilon}^{\xi}(b_i^N(x_t^{\varepsilon,N},\xi(u)) - \bar{b}_i^N(x_t^{\varepsilon,N}))]du b_j^N(x_t^{\varepsilon,N},\xi^\varepsilon(t))\right| \\
&= \sup_{t\le T}\left| \int_{t/\varepsilon}^{T/\varepsilon} f_{y_i}(x^{\varepsilon,N}(t))[\mathbb{E}_{t/\varepsilon}^{\xi}\partial_j b_i^N(x_t^{\varepsilon,N},\xi(t)) - \partial_j\bar{b}_i^N(x_t^{\varepsilon,N})]du b_j^N(x_t^{\varepsilon,N},\xi^\varepsilon(t))\right| \\
&\le K_N \sup_{t\le T}\int_{t/\varepsilon}^{T/\varepsilon} |f_{y_i}(x^{\varepsilon,N}(t))| e^{-\frac{\tilde{\varsigma}}{2}(u-t/\varepsilon)}du = \frac{2}{\tilde{\varsigma}}K_N(1-e^{-\frac{\tilde{\varsigma}(T-t)}{2\varepsilon}}) < \infty,
\end{aligned}$$

which shows that

$$\varepsilon \sup_{t\le T}\left|\sum_{j=1}^n I_{4j}^\varepsilon(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N})\right| \le \varepsilon\sup_{t\le T}\sum_{j=1}^n |I_{4j}^\varepsilon(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N})| = O(\varepsilon).$$

The same technique as the estimation of I_{3j}^ε gives

$$\frac{\varepsilon}{2}\sup_{t\le T}\left|\sum_{k,j=1}^n I_{5kj}^\varepsilon(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N})\right| = O(\varepsilon).$$

Similarly, (A2) also shows that for any $i, j = 1, 2, \ldots, n$, $\partial^2_{ij} b^N(x, \cdot) \in C_b(\mathbb{R}^m; \mathbb{R})$ for any $x \in C([0, T]; S_N)$. This, together with the boundedness of $a^N(x, \xi)$ for any $x \in C([0, T]; S_N)$ and $\xi \in \mathbb{R}^m$, yields

$$
\begin{aligned}
&\sup_{t\le T} |I^{\varepsilon}_{6kj}(t, x^{\varepsilon,N}(t), x^{\varepsilon,N}_t)| \\
&\le \sup_{t\le T} \left| \int_{t/\varepsilon}^{T/\varepsilon} f_{y_i}(x^{\varepsilon,N}(t)) \partial^2_{kj} [\mathbb{E}^{\xi}_{t/\varepsilon}(b^N_i(x^{\varepsilon,N}_t, \xi(u)) - \bar{b}^N_i(x^{\varepsilon,N}_t))] du a^N_{kj}(x^{\varepsilon,N}_t, \xi^{\varepsilon}(t)) \right| \\
&\le K_N \sup_{t\le T} \left| \int_{t/\varepsilon}^{T/\varepsilon} f_{y_i}(x^{\varepsilon,N}(t)) [\mathbb{E}^{\xi}_{t/\varepsilon}(\partial^2_{kj} b^N_i(x^{\varepsilon,N}_t, \xi(u)) - \partial^2_{kj} \bar{b}^N_i(x^{\varepsilon,N}_t))] du \right| \\
&\le K_N \sup_{t\le T} \int_{t/\varepsilon}^{T/\varepsilon} |f_{y_i}(x^{\varepsilon,N}(t))| e^{-\frac{\tilde{\varsigma}}{2}(u - t/\varepsilon)} du \\
&\le \frac{2K_N}{\tilde{\varsigma}} (1 - e^{-\frac{\tilde{\varsigma}(T-t)}{2\varepsilon}}) < \infty,
\end{aligned}
$$

which shows that

$$
\frac{\varepsilon}{2} \sum_{k,j=1}^{n} I^{\varepsilon}_{6kj}(t, x^{\varepsilon,N}(t), x^{\varepsilon,N}_t) \le \frac{\varepsilon}{2} \sum_{k,j=1}^{n} |I^{\varepsilon}_{6kj}(t, x^{\varepsilon,N}(t), x^{\varepsilon,N}_t)| = O(\varepsilon).
$$

Substituting these estimations into (4.22) gives

$$\begin{aligned} & \mathbb{L}^N(t, x_t^{\varepsilon,N}) V_1(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N}) \\ = \quad & O(\varepsilon) - \sum_{i=1}^{n} [f_{y_i}(x^{\varepsilon,N}(t))(b_i^N(x_t^{\varepsilon,N}, \xi^{\xi}(t)) - \bar{b}_i^N(x_t^{\varepsilon,N}))] \\ = \quad & O(\varepsilon) - f_y(x^{\varepsilon,N}(t))[b^N(x_t^{\varepsilon,N}, \xi^{\varepsilon}(t)) - \bar{b}^N(x_t^{\varepsilon,N})]. \end{aligned} \tag{4.23}$$

To proceed, let us estimate $\mathbb{L}^N(t, x_t^{\varepsilon,N}) V_2(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N})$. Note that

$$\begin{aligned} & \mathbb{L}^N(t, x_t^{\varepsilon,N}) V_2(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N}) \\ & = \varepsilon \sum_{i,j=1}^{n} \mathbb{L}^N(t, x_t^{\varepsilon,N}) \Big[\int_{t/\varepsilon}^{T/\varepsilon} f_{y_i y_j}(x^{\varepsilon,N}(t)) \mathbb{E}_{t/\varepsilon}^{\xi} (a_{ij}^N(x_t^{\varepsilon,N}, \xi(u)) - \bar{a}_{ij}^N(x_t^{\varepsilon,N})) du \Big]. \end{aligned} \tag{4.24}$$

We have

$$\varepsilon \mathbb{L}^N(t, x_t^{\varepsilon,N})\Big[\int_{t/\varepsilon}^{T/\varepsilon} f_{y_iy_j}(x^{\varepsilon,N}(t))\mathbb{E}^{\xi}_{t/\varepsilon}(a_{ij}^N(x_t^{\varepsilon,N}, \xi(u)) - \bar{a}_{ij}^N(x_t^{\varepsilon,N}))du\Big]$$
$$= J_1^\varepsilon + \varepsilon \sum_{k=1}^n J_{2k}^\varepsilon + \varepsilon \sum_{k=1}^n J_{3k}^\varepsilon + \frac{\varepsilon}{2} \sum_{k,l=1}^n J_{4kl}^\varepsilon + \frac{\varepsilon}{2} \sum_{k,l=1}^n J_{5kl}^\varepsilon + \frac{\varepsilon}{2} \sum_{k,l=1}^n J_{6kl}^\varepsilon, \qquad (4.25)$$

where $J_1^\varepsilon(y, x, \xi^\varepsilon(t)) = -f_{y_iy_j}(y)[a_{ij}^N(x, \xi^\varepsilon(t)) - \bar{a}_{ij}^N(x)]$ and

$$J_{2k}^\varepsilon(t, y, x) = \int_{t/\varepsilon}^{T/\varepsilon} f_{y_iy_jy_k}(y)\mathbb{E}^{\xi}_{t/\varepsilon}[a_{ij}^N(x, \xi(u)) - \bar{a}_{ij}^N(x)]du b_k^N(x, \xi^\varepsilon(t)),$$

$$J_{3k}^\varepsilon(t, y, x) = \int_{t/\varepsilon}^{T/\varepsilon} f_{y_iy_j}(y)\partial_k[\mathbb{E}^{\xi}_{t/\varepsilon}(a_{ij}^N(x, \xi(u)) - \bar{a}_{ij}^N(x))]du b_k^N(x, \xi^\varepsilon(t)),$$

$$J_{4kl}^\varepsilon(t, y, x) = \int_{\varepsilon/t}^{T/\varepsilon} f_{y_iy_jy_ky_l}(y)\mathbb{E}^{\xi}_{t/\varepsilon}[a_{ij}^N(x, \xi(u)) - \bar{a}_{ij}^N(x)]du a_{kl}^N(x, \xi^\varepsilon(t)),$$

$$J_{5kl}^\varepsilon(t, y, x) = \int_{\varepsilon/t}^{T/\varepsilon} f_{y_iy_jy_k}(y)\partial_l[\mathbb{E}^{\xi}_{t/\varepsilon}(a_{ij}^N(x, \xi(u)) - \bar{a}_{ij}^N(x))]du a_{kl}^N(x, \xi^\varepsilon(t)),$$

$$J_{6kl}^\varepsilon(t, y, x) = \int_{\varepsilon/t}^{T/\varepsilon} f_{y_iy_j}(y)\partial_{kl}^2[\mathbb{E}^{\xi}_{t/\varepsilon}(a_{ij}^N(x, \xi(u)) - \bar{a}_{ij}^N(x))]du a_{kl}^N(x, \xi(t)).$$

Recall that $a_{ij}(x,\xi)=\sum_{l=1}^{l_2}\psi_{il}(x,\xi)\psi_{lj}(x,\xi)$. Condition (A2) shows that $\psi(x,\xi)$, $\nabla_x\psi_{ij}(x,\xi)$, and $\nabla_x^2\psi_{ij}(x,\xi)$, $i,j=1,2,\dots,n$ are boundedness preserving with respect to $x\in C([0,T];\mathbb{R}^n)$, and continuous and bounded with respect to $\xi\in\mathbb{R}^m$ for $x\in C([0,T];S_N)$. These imply that for any $x\in C([0,T];S_N)$, $A(x,\cdot)\in C_b(\mathbb{R}^m;\mathbb{R}^{n\times n})$, $\partial_k a_{ij}(x,\cdot)$ and $\partial_{kl}^2 a_{ij}(x,\cdot)\in C_b(\mathbb{R}^m;\mathbb{R})$ for $i,j=1,2,\dots,n$. Applying the same technique as estimate of $I_{2j}^\varepsilon(t,x(t),x_t)$ yields

$$\varepsilon\sup_{t\le T}\left|\sum_{k=1}^n J_{2k}^\varepsilon(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})\right|=O(\varepsilon)\quad\text{and}\quad\frac{\varepsilon}{2}\sup_{t\le T}\left|\sum_{k,l=1}^n J_{4kl}^\varepsilon(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})\right|=O(\varepsilon).$$

Similar to the estimations of I_{4j}^ε, I_{5kj}^ε, and I_{6kj}^ε, we can obtain

$$\varepsilon\sup_{t\le T}\left|\sum_{k=1}^n J_{3k}^\varepsilon(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})\right|=O(\varepsilon)\quad\text{and}$$
$$\frac{\varepsilon}{2}\sup_{t\le T}\left|\sum_{k,l=1}^n J_{vkl}^\varepsilon(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})\right|=O(\varepsilon)\quad(v=5,6).$$

These estimate shows

$$\mathbb{L}^N(t,x_t^{\varepsilon,N})V_2(t,x^{\varepsilon,N}(t),x_t^{\varepsilon,N})=O(\varepsilon)-\frac{1}{2}\sum_{i,j=1}^n f_{y_iy_j}(x^{\varepsilon,N}(t))[a_{ij}^N(x_t^{\varepsilon,N},\xi^\varepsilon(t))-\bar{a}_{ij}^N(x_t^{\varepsilon,N})].\tag{4.26}$$

The estimates of $\mathbb{L}^N(t, x_t^{\varepsilon,N})V_1(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N})$ and $\mathbb{L}^N(t, x_t^{\varepsilon,N})V_2(t, x^{\varepsilon,N}(t), x_t^{\varepsilon,N})$, together with (4.20), yield

$$\begin{aligned}\hat{\mathcal{L}}^{\varepsilon,N} f^{\varepsilon,N}(t) &= O(\varepsilon) + L^N(t, x_t^{\varepsilon,N}) f(x^{\varepsilon,N}(t)) - f_y(x^{\varepsilon,N}(t))[b^N(x_t^{\varepsilon,N}, \xi^\varepsilon(t)) - \bar{b}^N(x_t^{\varepsilon,N})] \\ &\quad - \frac{1}{2}\sum_{i,j=1}^n f_{y_i y_j}(x^{\varepsilon,N}(t))[a_{ij}^N(x_t^{\varepsilon,N}, \xi^\varepsilon(t)) - \bar{a}_{ij}^N(x_t^{\varepsilon,N})]. \qquad (4.27)\end{aligned}$$

Applying the generator L^N defined by (3.6) to the solution process $x^\varepsilon(t)$ in the stochastic functional differential equation (4.1) gives

$$L^N(t, x_t^{\varepsilon,N}) f(x^{\varepsilon,N}(t)) = f_y(x^{\varepsilon,N}(t)) b^N(x_t^{\varepsilon,N}, \xi^\varepsilon(t)) + \frac{1}{2}\sum_{i,j=1}^n f_{y_i y_j}(x^{\varepsilon,N}(t)) a_{ij}^N(x_t^{\varepsilon,N}, \xi^\varepsilon(t)).$$

Substituting this into (4.27) yields

$$\begin{aligned}\hat{\mathcal{L}}^{\varepsilon,N} f^{\varepsilon,N}(t) &= O(\varepsilon) + f_y(x^{\varepsilon,N}(t))\bar{b}^N(x_t^{\varepsilon,N}) + \frac{1}{2}\sum_{i,j=1}^n f_{y_i y_j}(x^{\varepsilon,N}(t))\bar{a}_{ij}^N(x_t^{\varepsilon,N}) \\ &= O(\varepsilon) + L^N(t, x_t^{\varepsilon,N}) f(x^{\varepsilon,N}(t)), \qquad (4.28)\end{aligned}$$

which implies (4.9) as $\varepsilon \to 0$. This, together with (4.19) yields $x^{\varepsilon,N}(\cdot) \Rightarrow x^N(\cdot)$ as $\varepsilon \to 0$ by virtue of Lemma 11, where $x^N(\cdot)$ solves the martingale problem with operator L^N.

Moving from the truncated processes to that of un-truncated processes, the argument is similar to that of (Kushner, 1984). For any continuous deterministic initial value $x(0)$, let $\mathbb{P}(\cdot)$ and $\mathbb{P}^N(\cdot)$ denote the probabilities induced by $x(\cdot)$ and $x^N(\cdot)$, respectively, on the Borel sets of $C([0,T];\mathbb{R}^n)$. By (A4), the martingale problem has a unique solution for each $x(0)$, so $\mathbb{P}(\cdot)$ is unique. For each $T<\infty$, the uniqueness of $\mathbb{P}(\cdot)$ implies that $\mathbb{P}(\cdot)$ agrees with $\mathbb{P}^N(\cdot)$ on all Borel sets of the set of paths in $C([0,T];S_N)$ for each $t\le T$. However, $\mathbb{P}\{\sup_{t\le T}|x(t)|\le N\}\to 1$ as $N\to\infty$. This together with the weak convergence of $x^{\varepsilon,N}(\cdot)$ implies that $x^{\varepsilon}(\cdot)\Rightarrow x(\cdot)$. Moreover, the uniqueness implies that the limit does not depend on the chosen subsequences.

The proof is thus completed.

Future works

(i) If the noise $\xi^\varepsilon(t)$ depends on $x^\varepsilon(t)$ or x_t^ε;

(ii) Other functional or delay forms.

Thank you for your attention