PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 150, Number 9, September 2022, Pages 4037–4051
https://doi.org/10.1090/proc/15966
Article electronically published on April 15, 2022

# ASYMPTOTIC BISMUT FORMULAE FOR STOCHASTIC FUNCTIONAL DIFFERENTIAL EQUATIONS WITH INFINITE DELAY

YA WANG, FUKE WU, AND GEORGE YIN

(Communicated by Zhen-Qing Chen)

ABSTRACT. Using Malliavin calculus, this paper establishes asymptotic Bismut formulae for stochastic functional differential equations with infinite delay. Both nondegenerate and degenerate diffusion coefficients are treated. In addition, combined with the corresponding exponential ergodicity, stabilization bounds for $\nabla P_t f$ as $t \to \infty$ are derived.

## 1. INTRODUCTION

The Bismut-type formulae were first established in [1] using Malliavin calculus for stochastic differential equations (SDEs) on Remannian manifolds to obtain estimates of heat kernels and large deviations estimates. Then the formulae were extended to a larger class of diffusion semigroups in [2] by using martingale arguments. As a result, such formulae are also referred to as the Bismut-Elworthy-Li formulae. Subsequently, the approach of coupling by change of measures was introduced to derive the Bismut formulae and Harnack inequality for SDEs and stochastic partial differential equations; see [3] and references therein. Due to their wide range of applications on heat kernel estimates, functional inequalities, strong Feller property, and sensitivity analysis in finance, the Bismut formulae have been investigated under various settings; see [4, 5] for SDEs with Brownian noise, [6–8] for SDEs driven by jump-diffusion processes, [9] for the Lions derivatives of solutions to distribution dependent SDEs.

For stochastic functional differential equations (SFDEs), Bao, Wang, and Yuan established a Bismut-type formula and Harnack inequality in [10] for degenerate SFDEs using coupling by change of measures. Then they obtained a Bismut formula for semi-linear stochastic functional partial differential equations in [11] using Malliavin calculus, and derived a Bismut-type formula in [12] for the Lions derivatives of segment processes of distribution-path dependent SDEs. When the diffusion term depends on the past history, the SFDEs might have a reconstruction property [13], which causes the laws of segment processes with different initial data to be mutually singular. Thus, the strong Feller property and the classical Bismut-type

Received by the editors June 13, 2021, and, in revised form, December 9, 2021.
2020 *Mathematics Subject Classification.* Primary 60J60, 34K50.
*Key words and phrases.* Asymptotic Bismut formula, Malliavin calculus, stochastic functional differential equation, infinite delay.
The first and second authors were supported in part by the National Natural Science Foundation of China (Grant Nos. 61873320). The third author was supported in part by the National Science Foundation under grant DMS-2114649.

formulae are invalid. To circumvent the difficulty, using coupling by change of measures, Kulik and Scheutzow [14] introduced a weaker version of the formula, namely asymptotic Bismut formula, for nondegenerate SFDEs; the asymptotic Bismut formula for Lions derivatives of segment processes of nondegenerate and degenerate distribution-path-dependent SDEs was proved in [12] using Malliavin calculus.

A crucial assumption in [14] is that the first-order Fréchet derivative of the drift term is bounded and uniformly continuous on $C([-\tau,0];\mathbb{R}^d)$, where $\tau>0$ is the length of delay. In addition, it is worth pointing out that in [12, 14], only (distribution dependent) SFDEs with finite delay were considered. It is crucial to note that for SFDEs with finite delay, the classical Bismut-type formula holds only for $t>\tau$ [10–12]. For SFDEs with *infinite delay*, the classical Bismut formula is not applicable. This paper aims to relax the assumptions for the first-order Fréchet derivative of the drift term in [14] and to establish asymptotic Bismut formulae for SFDEs with infinite delay so as to extend and generalize the results in [12, 14].

Denote by $C((-\infty,0];\mathbb{R}^d)$ the family of continuous functions from $(-\infty,0]$ to $\mathbb{R}^d$, and by $\mathcal{C}_r := \big\{\phi\in C((-\infty,0];\mathbb{R}^d) : \sup_{-\infty<\theta\le 0} e^{r\theta}|\phi(\theta)| < \infty\big\}$ with norm $\|\phi\|_r = \sup_{-\infty<\theta\le 0} e^{r\theta}|\phi(\theta)|$ and $r>0$. It is known that $(\mathcal{C}_r,\|\cdot\|_r)$ is a Polish space [15]. Let $b:\mathcal{C}_r\mapsto\mathbb{R}^d$ and $\sigma:\mathcal{C}_r\mapsto\mathbb{R}^{d\times m}$ be continuous, $W(t)$ be an $m$-dimensional Wiener process, and $X_t(\theta):(-\infty,0]\ni\theta\mapsto X(t+\theta)\in\mathbb{R}^d$ be the segment process. Consider the SFDE with infinite delay

$$dX(t) = b(X_t)dt + \sigma(X_t)dW(t) \tag{1.1}$$

and initial data $X_0=\xi\in\mathcal{C}_r$.

As a preparation, we first provide some notation and definitions to be used in the rest of the paper. Let $\|\cdot\|_{\rm HS}$ denote the Hilbert-Schmidt norm, $C^1(\mathcal{C}_r,\mathbb{R}^d)$ be the family of Fréchet differentiable functions $f:\mathcal{C}_r\to\mathbb{R}^d$ with continuous derivatives, and $\nabla_\xi f(\cdot)$ or $\langle\nabla f(\cdot),\xi\rangle$ denotes the Gâteaux or Fréchet direction derivative of $f$ along the direction $\xi$. If, moreover, $\|\nabla f(\cdot)\|$ is bounded, we denote $f\in C_b^1(\mathcal{C}_r,\mathbb{R}^d)$, where $\|\cdot\|$ is the operator norm, that is, $\|\nabla f(\cdot)\| := \sup_{\|\xi\|_r\le 1}\|\langle\nabla f(\cdot),\xi\rangle\|_{\rm HS}$. Denote by $\mathcal{M}_0$ the set of probability measures on $(-\infty,0]$. For any $k>0$, further define $\mathcal{M}_k$, the subset of $\mathcal{M}_0$ by $\mathcal{M}_k := \{\mu\in\mathcal{M}_0 : \mu^{(k)} := \int_{-\infty}^0 e^{-k\theta}\mu(d\theta)<\infty\}$. Let $T>0$ be fixed and arbitrarily,

$$\mathcal{H} = \Big\{h\in C\big([0,T],\mathbb{R}^m\big) : h(0)=0,\ \dot h(t) \text{ exists a.s., } \|h\|_{\mathcal{H}}^2 := \int_0^T |\dot h(s)|^2 ds < \infty\Big\}$$

be the Cameron-Martin space, $\mathbb{P}$ be the Wiener measure, and $\Omega = C([0,T];\mathbb{R}^m)$. Then the coordinate process $W(t,\omega) := \omega(t), \omega\in\Omega$ is an $m$-dimensional Brownian motion. A function $F\in L^2(\Omega;\mathbb{P})$ is called a Malliavin differentiable along the direction $h\in\mathcal{H}$ if the following limit

$$D_hF := \lim_{\varepsilon\to 0}\frac{F(\cdot+\varepsilon h)-F(\cdot)}{\varepsilon}$$

exists in $L^2(\Omega;\mathbb{P})$ and $D_hF$ is called the Malliavin directional derivative of $F$ along the direction $h$. If the map $\mathcal{H}\ni h\mapsto D_hF\in L^2(\Omega;\mathbb{P})$ is bounded, there exists a unique $DF\in L^2(\Omega\to\mathcal{H};\mathbb{P})$ such that $\langle DF,h\rangle_{\mathcal{H}} = D_hF$ holds in $L^2(\Omega;\mathbb{P})$ for all $h\in\mathcal{H}$. In this case we denote $F\in{\rm Dom}(D)$ and call $DF$ the Malliavin derivative or gradient of $F$. For $p\ge 1$, if $h\in L^p(\Omega\to\mathcal{H};\mathbb{P})$ is an adapted stochastic process, we write $h\in L^p_a(\Omega\to\mathcal{H};\mathbb{P})$. Let $(\delta,{\rm Dom}(\delta))$ be the dual operator of $(D,{\rm Dom}(D))$, called divergence operator, which is characterized by using the following integration-by-parts formula

$$\mathbb{E}(D_hF) = \int_\Omega D_hFd\mathbb{P} = \int_\Omega F\delta(h)d\mathbb{P} = \mathbb{E}(F\delta(h)), \quad F \in \mathrm{Dom}(D),\ h \in \mathrm{Dom}(\delta).$$

According to [16, Proposition 1.3.11], all adapted $h \in L^2(\Omega \to \mathcal{H};\mathbb{P})$ belong to $\mathrm{Dom}(\delta)$ and $\delta(h) = \int_0^T \dot{h}(s)dW(s)$.

The remainder of the paper is organized as follows. Section 2 establishes asymptotic Bismut formulae and stabilization bounds of $\nabla P_tf$ as $t \to \infty$ for nondegenerate SFDEs with infinite delay. Section 3 investigates those of the degenerate SFDEs to close the paper.

## 2. Nondegenerate SFDEs with infinite delay

To ensure the existence and uniqueness of the solution and to establish the asymptotic Bismut formula, we make Assumption 2.1.

**Assumption 2.1.** $b \in C^1(\mathcal{C}_r, \mathbb{R}^d)$ is bounded on bounded subsets of $\mathcal{C}_r$. There exists a positive constant $K > 0$ such that for any $\phi, \psi \in \mathcal{C}_r$

$$2(\phi(0) - \psi(0), b(\phi) - b(\psi))_+ + \|\sigma(\phi) - \sigma(\psi)\|_{\mathrm{HS}}^2 \le K\|\phi - \psi\|_r^2, \tag{2.1}$$

where $(\cdot,\cdot)$ denotes the scalar product in $\mathbb{R}^d$ and $a_+ := \max\{0, a\}$ for any $a \in \mathbb{R}$.

*Remark* 2.2. Note that (1.1) has a unique strong solution under Assumption 2.1 (see e.g., [17]). By using standard arguments [18, p.160], for any $T > 0$ and $p > 0$, there exist a constant $C_{p,\xi} > 0$ and an increasing function $A(t) : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\mathbb{E}\big[\sup_{0\le t\le T} \|X_t\|_r^p\big] \le C_{p,\xi}A(T). \tag{2.2}$$

**Assumption 2.3.** For any $\phi \in \mathcal{C}_r$, the matrix $\sigma(\phi)$ admits a right inverse $\sigma^{-1}(\phi)$ and $\|\sigma^{-1}\|_\infty := \sup_{\phi\in\mathcal{C}_r} \|\sigma^{-1}(\phi)\|_{\mathrm{HS}} < \infty$.

**Assumption 2.4.** There exist constants $L$ and $\gamma \ge 1$ such that for any $\phi \in \mathcal{C}_r$, $\|\nabla b(\phi)\| \le L(1 + \|\phi\|_r^\gamma)$.

**Theorem 2.5.** *Assume Assumptions 2.1, 2.3, and 2.4 hold. Then for any $T > 0$, $f \in C_b^1(\mathcal{C}_r, \mathbb{R}), \lambda \ge 0$, and $\xi, \eta \in \mathcal{C}_r$, the following representation formula holds,*

$$\nabla_\eta P_T f(\xi) = \mathbb{E}\big\langle(\nabla f)(X_T), Z_T\big\rangle + \lambda\mathbb{E}\Big(f(X_T)\int_0^T \sigma(X_s)^{-1}Z_s dW(s)\Big), \tag{2.3}$$

*where $Z_t$ denotes the unique segment process to the SFDE with infinity delay*

$$dZ(t) = \{-\lambda Z(t) + \langle\nabla b(X_t), Z_t\rangle\}dt + \langle\nabla\sigma(X_t), Z_t\rangle dW(t) \tag{2.4}$$

*and $Z_0 = \eta \in \mathcal{C}_r$. Furthermore, there exist constants $A$, $\alpha$, and $\lambda_0 > 0$ such that for $\lambda \ge \lambda_0$,*

$$\Big|\nabla_\eta P_T f(\xi) - \lambda\mathbb{E}\Big(f(X_T)\int_0^T \sigma(X_s)^{-1}Z_s dW(s)\Big)\Big| \le A\sup_{\phi\in\mathcal{C}_r}\|\nabla f(\phi)\|\|\eta\|_r e^{-\alpha T}. \tag{2.5}$$

To prove this theorem, we need the following four lemmas. We first show that for sufficiently large $\lambda$, (2.4) is exponentially stable.

**Lemma 2.6.** *Under Assumption 2.1, for any $\lambda \geq 0$, (2.4) has a unique solution and for any $p \geq 2$, there exist $A_1$, $\alpha_1 > 0$, and sufficiently large $\lambda_p > 0$ such that for $\lambda \geq \lambda_p$,*

$$\mathbb{E}\|Z_t\|_r^p \leq A_1\|\eta\|_r^p e^{-\alpha_1 t}, \qquad t \geq 0. \tag{2.6}$$

*Proof.* By (2.1), for any $\phi, \psi \in \mathcal{C}_r$, we have

$$2\varepsilon(\phi(0), b(\psi+\varepsilon\phi) - b(\psi))_+ + \|\sigma(\psi+\varepsilon\phi) - \sigma(\psi)\|_{\text{HS}}^2 \leq K\varepsilon^2\|\phi\|_r^2,$$

which implies that

$$2(\phi(0), \langle \nabla b(\psi), \phi\rangle)_+ + \|\langle \nabla\sigma(\psi), \phi\rangle\|_{\text{HS}}^2 \leq K\|\phi\|_r^2. \tag{2.7}$$

Therefore, combined with the linearity of the inner product, (2.7) implies that (2.4) has a unique solution. Next, we proceed to prove (2.6). By the Lyapunov inequality, it suffices to prove this inequality for $p > 4$. By virtue of the Itô formula and (2.7), we obtain

$$e^{2\lambda t}|Z(t)|^2 \leq |\eta(0)|^2 + K\int_0^t e^{2\lambda s}\|Z_s\|_r^2 ds + M^\lambda(t),$$

where $M^\lambda(t) := 2\int_0^t e^{2\lambda s}\big(Z(s), \langle \nabla\sigma(X_s), Z_s\rangle dW(s)\big)$. Letting $\kappa = 2(\lambda - r)$, we have

$$e^{2rt}|Z(t)|^2 \leq |\eta(0)|^2 e^{-\kappa t} + K\int_0^t e^{-\kappa(t-s)} e^{2rs}\|Z_s\|_r^2 ds + e^{-\kappa t}M^\lambda(t). \tag{2.8}$$

Note that $\|Z_t\|_r \leq e^{-2rt}\|\eta\|_r^2 + e^{-2rt}\sup_{0\leq s\leq t} e^{2rs}|Z(s)|^2$ and $e^{2rt}\|Z_t\|_r$ is nondecreasing, which implies $\int_0^t e^{-\kappa(t-u)}e^{2ru}\|Z_u\|_r^2 du$ is nondecreasing with respective to $t$. Then from (2.8), we see that

$$e^{2rt}\|Z_t\|_r^2 \leq 2\|\eta\|_r^2 + K\int_0^t e^{-\kappa(t-u)}e^{2ru}\|Z_u\|_r^2 du + \sup_{0\leq s\leq t} e^{-\kappa s}M^\lambda(s). \tag{2.9}$$

The Hölder inequality yields that

$$\Big(\int_0^t e^{-\kappa(t-u)}e^{2ru}\|Z_u\|_r^2 du\Big)^{\frac{p}{2}} \leq \Big(\frac{p-2}{\kappa p}\Big)^{\frac{p-2}{2}} \int_0^t e^{pru}\|Z_u\|_r^p du. \tag{2.10}$$

Note that $M^\lambda(t\wedge\tau_n)$ is a square integrable martingale for any $n > \|\eta\|_r$, where $\tau_n := \inf\{t \geq 0 : \|Z_s\|_r \geq n\}$. Then by (2.7) and [19, Lemma 2.2] or [12, Lemma 7.2], there exists a positive constant $a_{p,\kappa}$ satisfying $\lim_{\kappa\to\infty} a_{p,\kappa} = 0$ such that

$$\mathbb{E}\Big[\sup_{0\leq s\leq t\wedge\tau_n} e^{-\kappa s}M^\lambda(s)\Big]^{\frac{p}{2}} \leq K^{\frac{p}{4}}a_{p,\kappa}\mathbb{E}\int_0^{t\wedge\tau_n} e^{pru}\|Z_u\|_r^p du. \tag{2.11}$$

Hence it follows from (2.9), (2.10), and (2.11) that

$$\mathbb{E}e^{pr(t\wedge\tau_n)}\|Z_{t\wedge\tau_n}\|_r^p \leq 3^{\frac{p-2}{2}}2^{\frac{p}{2}}\|\eta\|_r^p + A_{p,\kappa,K}\int_0^t e^{pr(u\wedge\tau_n)}\|Z_{u\wedge\tau_n}\|_r^p du, \tag{2.12}$$

where $A_{p,\kappa,K} := 3^{\frac{p-2}{2}}\big(\big(\frac{p-2}{\kappa p}\big)^{\frac{p-2}{2}} + K^{\frac{p}{4}}a_{p,\kappa}\big)$. The Gronwall inequality gives

$$\mathbb{E}e^{pr(t\wedge\tau_n)}\|Z_{t\wedge\tau_n}\|_r^p \leq 3^{\frac{p-2}{2}}2^{\frac{p}{2}}\|\eta\|_r^p \exp\{A_{p,\kappa,K}t\}.$$

Recall that $\lim_{n\to\infty}\tau_n = \infty$. By Fatou's lemma, we obtain

$$\mathbb{E}\|Z_t\|_r^p \leq 3^{\frac{p-2}{2}}2^{\frac{p}{2}}\|\eta\|_r^p \exp\{(A_{p,\kappa,K} - pr)t\}.$$

Since $A_{p,\kappa,K} \to 0$ as $\kappa \to \infty$ (that is $\lambda \to \infty$), there exists a positive constant $\alpha_1$ such that $\mathbb{E}\|Z_t\|_r^p \leq 3^{\frac{p-2}{2}} 2^{\frac{p}{2}} \|\eta\|_r^p e^{-\alpha_1 t}$. Then (2.6) holds. The proof is complete. $\square$

By the pathwise uniqueness of (1.1), there exists a measurable functional $\Psi : \mathcal{C}_r \times C([0,\infty);\mathbb{R}^d) \to C([0,\infty);\mathbb{R}^d)$ such that $X(\cdot,\xi) = \Psi(\xi, W(\cdot))$, $\mathbb{P}$-a.e., where $\xi \in \mathcal{C}_r$ denotes the initial data. Let $h(\cdot) \in L^\infty(\Omega \to \mathcal{H};\mathbb{P})$ be an adapted process. Consider the following SFDE with infinite delay

$$dX^{h,\varepsilon}(t) = [b(X_t^{h,\varepsilon}) + \varepsilon\sigma(X_t^{h,\varepsilon})\dot{h}(t)]dt + \sigma(X_t^{h,\varepsilon})dW(t), \quad t \in [0,T], \tag{2.13}$$

with $X_0^{h,\varepsilon} = \xi \in \mathcal{C}_r$. Let $\varepsilon \in (0,1]$ without loss of generality. Under Assumption 2.1, (2.13) has a unique strong solution. The Malliavin directional derivative of $X(t)$ along $h$ is defined as $D_h X(t,\omega) := \lim_{\varepsilon\to 0} \frac{X^{h,\varepsilon}(t)-X(t)}{\varepsilon}$, provided the limit exists in $L^2(\Omega \to C([0,T],\mathbb{R}^d);\mathbb{P})$. To establish the existence of the above limit, we provide Lemma 2.7.

**Lemma 2.7.** *Under Assumption 2.1, let $T > 0$ and $h(\cdot) \in L^\infty(\Omega \to \mathcal{H};\mathbb{P})$ be an adapted process. Then for any $p \geq 2$ and $t \in [0,T]$, there exists an increasing and continuous function $A_{p,h}(t)$ such that*

$$\mathbb{E}\Big[\sup_{0\leq s\leq t} X_s^{h,\varepsilon} - X_s\|_r^p\Big] \leq A_{p,h}(t)\varepsilon^p. \tag{2.14}$$

*Proof.* Let $Y_t^{h,\varepsilon} = \big(X_t^{h,\varepsilon} - X_t\big)/\varepsilon$ and $Y^{h,\varepsilon}(t) = Y_t^{h,\varepsilon}(0)$ for $t \geq 0$. Then we have

$$dY^{h,\varepsilon}(t) = \Big\{\frac{b(X_t^{h,\varepsilon}) - b(X_t)}{\varepsilon} + \sigma(X_t^{h,\varepsilon})\dot{h}(t)\Big\}dt + \frac{\sigma(X_t^{h,\varepsilon}) - \sigma(X_t)}{\varepsilon}dW(t), \tag{2.15}$$

with $Y_0^{h,\varepsilon} = \mathbf{0} \in \mathcal{C}_r$. By the Itô formula and Assumption 2.1, for any $p > 2$, we have

$$\begin{aligned}|Y^{h,\varepsilon}(t)|^p \leq& \frac{Kp^2}{2}\int_0^t |Y^{h,\varepsilon}(s)|^{p-2}\|Y_s^{h,\varepsilon}\|_r^2 ds \\ &+ p\int_0^t |Y^{h,\varepsilon}(s)|^{p-1}\|\sigma(X_s^{h,\varepsilon})\|_{\mathrm{HS}}|\dot{h}(s)|ds + M(t),\end{aligned} \tag{2.16}$$

where $M(t) := \frac{p}{\varepsilon}\int_0^t |Y^{h,\varepsilon}(s)|^{p-2}\big(Y^{h,\varepsilon}(s), (\sigma(X_s^{h,\varepsilon}) - \sigma(X_s))dW(s)\big)$. Since $h(\cdot) \in L^\infty(\Omega \to \mathcal{H};\mathbb{P})$, there exists a positive constant $A$ such that $\int_0^T |\dot{h}(s)|^2 ds \leq A$, $\mathbb{P}$-a.s. By Hölder's inequality, Young's inequality, and Assumption 2.1, we have

$$\begin{aligned}&p\mathbb{E}\Big[\sup_{0\leq u\leq t}\int_0^u |Y^{h,\varepsilon}(s)|^{p-1}\|\sigma(X_s^{h,\varepsilon})\|_{\mathrm{HS}}|\dot{h}(s)|ds\Big] \\ &\leq p\mathbb{E}\Big(\int_0^t |\dot{h}(s)|^2 ds\Big)^{1/2}\Big(\int_0^t |Y^{h,\varepsilon}(s)|^{2p-2}\|\sigma(X_s^{h,\varepsilon})\|_{\mathrm{HS}}^2 ds\Big)^{1/2} \\ &\leq \sqrt{A}p\mathbb{E}\Big(\sup_{0\leq s\leq t}|Y^{h,\varepsilon}(s)|^{2p-2}\int_0^t \|\sigma(X_s^{h,\varepsilon})\|_{\mathrm{HS}}^2 ds\Big)^{1/2} \\ &\leq \frac{1}{4}\mathbb{E}\sup_{0\leq s\leq t}|Y^{h,\varepsilon}(s)|^p + 2^{p-1}A^{p/2}C_1\bar{K}^p t^{(p-2)/2}\mathbb{E}\int_0^t (1 + \|X_s^{h,\varepsilon}\|_r^p)ds,\end{aligned} \tag{2.17}$$

where $C_1 := \big(4(p-1)\big)^{p-1}$ and $\bar{K} := \sqrt{K} + \|\sigma(0)\|_{\rm HS}$. By using the Burkholder-Davis-Gundy inequality, the Young inequality, and (2.1), we have

$$\mathbb{E}\Big[\sup_{0\le s\le t} M(s)\Big] \le \frac{1}{4}\mathbb{E}\Big[\sup_{0\le s\le t}|Y^{h,\varepsilon}(s)|^p\Big] + 32Kp^2\mathbb{E}\int_0^t \|Y_s^{h,\varepsilon}\|_r^p ds. \tag{2.18}$$

Note that $Y_0^{h,\varepsilon} = \mathbf{0} \in \mathcal{C}_r$ and $\|Y_s^{h,\varepsilon}\|_r \le \sup_{0\le u\le s}|Y^{h,\varepsilon}(u)|$. Substituting (2.17) and (2.18) into (2.16) gives

$$\mathbb{E}\Big[\sup_{0\le s\le t}|Y^{h,\varepsilon}(s)|^p\Big] \le C_2 t^{\frac{p-2}{2}}\int_0^t (1+\mathbb{E}\|X_s^{h,\varepsilon}\|_r^p)ds + C_3\int_0^t \mathbb{E}\Big[\sup_{0\le u\le s}|Y^{h,\varepsilon}(u)|^p\Big]ds,$$

where $C_2 := 2^p A^{p/2} C_1 \bar{K}^p$ and $C_3 := 65Kp^2$. By using a similar argument as in the derivations of (2.17) and (2.18), for any $q \ge 2$ and $h(\cdot) \in L^\infty\big(\Omega \to \mathcal{H};\mathbb{P}\big)$, there exists a continuous and increasing function $A^{q,h}(\cdot)$ such that

$$\sup_{0<\varepsilon\le 1}\mathbb{E}\|X_t^{h,\varepsilon}\|_r^q \le A^{q,h}(t), \quad \forall\, t\in[0,T]. \tag{2.19}$$

Then applying the Gronwall inequality yields that

$$\mathbb{E}\sup_{0\le s\le t}|Y^{h,\varepsilon}(s)|^p \le \bar{A}_{p,h}(t)e^{C_3 t}, \tag{2.20}$$

where $\bar{A}_{p,h}(t) := C_2 t^{p/2}(1 + A^{p,h}(t))$. By the definition of norm $\|\cdot\|_r$, we have

$$\sup_{0\le s\le t}\|Y_s^{h,\varepsilon}\|_r = \sup_{0\le s\le t}\sup_{\theta\le 0} e^{r\theta}|Y^{h,\varepsilon}(s+\theta)| \le \sup_{0\le s\le t}|Y^{h,\varepsilon}(s)|. \tag{2.21}$$

Thus we arrive at $\mathbb{E}\Big[\sup_{0\le s\le t}\|Y_s^{h,\varepsilon}\|_r^p\Big] \le \bar{A}_{p,h}(t)e^{C_3 t}$. This implies that (2.14) holds for $A_{p,h}(t) := \bar{A}_{p,h}(t)e^{C_3 t}$. The proof is complete. □

**Lemma 2.8.** *Let Assumptions 2.1 and 2.4 hold. Then for any $T > 0$, $h(\cdot) \in L^\infty(\Omega \to \mathcal{H};\mathbb{P})$, the limit*

$$D_h X_t := \lim_{\varepsilon\to 0}\frac{X_t^{h,\varepsilon} - X_t}{\varepsilon}, \quad t\in[0,T] \tag{2.22}$$

*exists in $L^2(\Omega \to C([0,T];\mathcal{C}_r);\mathbb{P})$. Moreover, the segment process $\{D_h X_t\}_{t\in[0,T]}$ uniquely solves the following SFDE with infinite delay*

$$dU^h(t) = \{\langle\nabla b(X_t), U_t^h\rangle + \sigma(X_t)\dot{h}(t)\}dt + \langle\nabla\sigma(X_t), U_t^h\rangle dW(t) \tag{2.23}$$

*with $U_0^h = \mathbf{0} \in \mathcal{C}_r$, where $U_t^h = D_h X_t$ denotes the segment process of $U^h(t)$.*

*Remark* 2.9. For any given $h(\cdot) \in L^\infty(\Omega \to \mathcal{H};\mathbb{P})$, existence of the limit in (2.22) implies that $X_t$ is Malliavin differentiable along $h$, denoted by $D_h X_t$. Moreover, the solution process $X(t)$ is also Malliavin differentiable along $h$, denoted by $D_h X(t)$, and solves uniquely (2.23) on $[0,T]$. In fact, for any $\varepsilon > 0$, by Remark 2.2, for any $h \in L_a^{2+\varepsilon}(\Omega \to \mathcal{H};\mathbb{P})$, (2.23) has a unique solution under Assumption 2.1.

*Proof.* For $h(\cdot) \in L^\infty(\Omega \to \mathcal{H};\mathbb{P})$, by (2.7) and the Lipschitz continuity of $\sigma(\cdot)$, (2.23) has a unique solution. Now it remains to show that the limit $D_h X_t$ exists in $L^2(\Omega \to C([0,T];\mathcal{C}_r),\mathbb{P})$ and is the segment process $U_t^h$ of SFDE (2.23). In view of (2.21), it suffices to prove that

$$\lim_{\varepsilon\to 0}\mathbb{E}\Big[\sup_{0\le t\le T}|Y^{h,\varepsilon}(t) - U^h(t)|^2\Big] = 0.$$

By $b \in C^1(\mathcal{C}_r, \mathbb{R}^d), \sigma \in C^1(\mathcal{C}_r, \mathbb{R}^{d\times m})$, the Mean Value Theorem [20, Theorem 3.2.6] implies

$$b(X_t^{h,\varepsilon}) - b(X_t) = \int_0^1 \langle \nabla b(X_t + \theta(X_t^{h,\varepsilon} - X_t)), X_t^{h,\varepsilon} - X_t\rangle d\theta,$$
$$\sigma(X_t^{h,\varepsilon}) - \sigma(X_t) = \int_0^1 \langle \nabla \sigma(X_t + \theta(X_t^{h,\varepsilon} - X_t)), X_t^{h,\varepsilon} - X_t\rangle d\theta.$$

Let $Z^{h,\varepsilon}(t) = Y^{h,\varepsilon}(t) - U^h(t)$. It follows from (2.15) and (2.23) that

$$dZ^{h,\varepsilon}(t) = \{\langle \nabla b(X_t), Z_t^{h,\varepsilon}\rangle + I_1^\varepsilon(t)\}dt + \{\langle \nabla\sigma(X_t), Z_t^{h,\varepsilon}\rangle + I_2^\varepsilon(t)\}dW(t),$$

where

$$I_1^\varepsilon(t) := (\sigma(X_t^{h,\varepsilon}) - \sigma(X_t))\dot{h}(t) + \int_0^1 \langle \nabla b(X_t + \theta(X_t^{h,\varepsilon} - X_t)) - \nabla b(X_t), Y_t^{h,\varepsilon}\rangle d\theta,$$
$$I_2^\varepsilon(t) := \int_0^1 \langle \nabla \sigma(X_t + \theta(X_t^{h,\varepsilon} - X_t)) - \nabla\sigma(X_t), Y_t^{h,\varepsilon}\rangle d\theta.$$

Applying the Itô formula and (2.7) gives that

$$|Z^{h,\varepsilon}(t)|^2 \le \int_0^t 2K\|Z_s^{h,\varepsilon}\|_r^2 + |Z_s^{h,\varepsilon}|^2 + |I_1^\varepsilon(s)|^2 + 2\|I_2^\varepsilon(s)\|_{\mathrm{HS}}^2 ds + \widetilde{M}(t), \tag{2.24}$$

where $\widetilde{M}(t) := 2\int_0^t \big(Z^{h,\varepsilon}(s), \langle\nabla\sigma(X_s), Z_s^{h,\varepsilon}\rangle + I_2^\varepsilon(s)\big)dW(s)$. By a similar argument as in the derivation of (2.18), we obtain

$$\mathbb{E}\Big[\sup_{0\le s\le t} \widetilde{M}(s)\Big] \le \frac{1}{2}\mathbb{E}\Big[\sup_{0\le s\le t}|Z^{h,\varepsilon}(s)|^2\Big] + 128\mathbb{E}\int_0^t K\|Z_s^{h,\varepsilon}\|_r^2 + \|I_2^\varepsilon(s)\|_{\mathrm{HS}}^2 ds. \tag{2.25}$$

From (2.24) and (2.25), there exist constants $C_4$ and $C_5 > 0$ such that

$$\mathbb{E}\Big[\sup_{0\le s\le t}|Z^{h,\varepsilon}(s)|^2\Big] \le C_4\mathbb{E}\int_0^t |I_1^\varepsilon(s)|^2 + \|I_2^\varepsilon(s)\|_{\mathrm{HS}}^2 ds + C_5\int_0^t \mathbb{E}\Big[\sup_{0\le u\le s}|Z^{h,\varepsilon}(u)|^2\Big]ds.$$

Then the Gronwall inequality implies that

$$\mathbb{E}\Big[\sup_{0\le s\le t}|Z^{h,\varepsilon}(s)|^2\Big] \le C_4 e^{C_5 t}\mathbb{E}\int_0^t |I_1^\varepsilon(s)|^2 + \|I_2^\varepsilon(s)\|_{\mathrm{HS}}^2 ds. \tag{2.26}$$

By virtue of the definition of $I_1^\varepsilon(s)$ and $I_2^\varepsilon(s)$, we have

$$|I_1^\varepsilon(s)|^2 + \|I_2^\varepsilon(s)\|_{\mathrm{HS}}^2 \le 2\|\sigma(X_s^{h,\varepsilon}) - \sigma(X_s)\|_{\mathrm{HS}}^2|\dot{h}(s)|^2 + 2J^\varepsilon(s)\|Y_s^{h,\varepsilon}\|_r^2, \tag{2.27}$$

where $J^\varepsilon(s) := \int_0^1 \|\nabla b(X_s + \theta(X_s^{h,\varepsilon} - X_s)) - \nabla b(X_s)\|^2 + \|\nabla\sigma(X_s + \theta(X_s^{h,\varepsilon} - X_s)) - \nabla\sigma(X_s)\|^2 d\theta$. By virtue of Assumption 2.4 and (2.7), there exists an increasing function $C_6 > 0$ such that for all $s \in [0, T]$

$$J^\varepsilon(s) \le C_6\big(1 + \|X_s\|_r^{2\gamma} + \|X_s^{h,\varepsilon} - X_s\|_r^{2\gamma}\big). \tag{2.28}$$

In view of $h \in L^\infty(\Omega \to \mathcal{H}, \mathbb{P})$ and Lemma 2.7, we obtain

$$\begin{aligned}&\limsup_{\varepsilon\to 0}\int_0^T \mathbb{E}\|\sigma(X_s^{h,\varepsilon}) - \sigma(X_s)\|_{\mathrm{HS}}^2|\dot{h}(s)|^2 ds\\ &\le K\|h\|_{L^\infty(\Omega\to\mathcal{H},\mathbb{P})}^2 \limsup_{\varepsilon\to 0}\mathbb{E}\Big[\sup_{0\le s\le T}\|X_s^{h,\varepsilon} - X_s\|_r^2\Big] = 0.\end{aligned} \tag{2.29}$$

Since $\nabla b$ and $\nabla\sigma$ are continuous, Lemma 2.7 implies that $J^\varepsilon(s)$ and $J^\varepsilon(s)\|Y_s^{h,\varepsilon}\|_r^2$ converge to 0 in probability as $\varepsilon \to 0$. In addition, it follows from (2.19) and

(2.14) that $\{(1+\|X_s\|_r^{2\gamma})\|Y_s^{h,\varepsilon}\|_r^2\}_{\varepsilon\in(0,1]}$ and $\|X_s^{h,\varepsilon}-X_s\|_r^{2\gamma}\|Y_s^{h,\varepsilon}\|_r^2$ are uniformly integrable for any fixed $s\in[0,T]$. Hence by the dominated convergence theorem, $\lim_{\varepsilon\to 0}\mathbb{E}J^\varepsilon(s)\|Y_s^{h,\varepsilon}\|_r^2=0$. In light of Lemma 2.7, (2.19), and (2.28), $\lim_{\varepsilon\to 0}\int_0^T\mathbb{E}J^\varepsilon(s)\|Y_s^{h,\varepsilon}\|_r^2ds=0$. Then it follows from (2.26), (2.27), and (2.29) that

$$\lim_{\varepsilon\to 0}\mathbb{E}\Big[\sup_{0\le s\le T}|Z^{h,\varepsilon}(s)|^2\Big]\le C_4e^{C_5T}\lim_{\varepsilon\to 0}\int_0^T\mathbb{E}\{|I_1^\varepsilon(s)|^2+\|I_2^\varepsilon(s)\|_{\mathrm{HS}}^2\}ds=0.$$

Hence this proof is completed. □

**Lemma 2.10.** *Let Assumptions 2.1 and 2.4 hold. Then for any $T>0$, the limit*

$$\nabla_\eta X_t:=\lim_{\varepsilon\to 0}\frac{X_t(\xi+\varepsilon\eta)-X_t(\xi)}{\varepsilon},\quad t\in[0,T],\ \xi,\eta\in\mathcal{C}_r \tag{2.30}$$

*exists in $L^2(\Omega\to C([0,T];\mathcal{C}_r);\mathbb{P})$ and is the unique functional solution to the following SFDE*

$$dV(t)=\langle\nabla b(X_t),V_t\rangle dt+\langle\nabla\sigma(X_t),V_t\rangle dW(t),\qquad V_0=\eta\in\mathcal{C}_r. \tag{2.31}$$

*Proof.* It is easy to see from (2.7) that (2.31) has a unique solution, and the proof of existence of the limit in (2.30) is similar to that of Lemma 2.8. We omit it here. □

*Proof of Theorem 2.5.* Let $h(t)=\lambda\int_0^t\sigma^{-1}(X_s)Z(s)ds$, which is an adapted stochastic process. By Lemma 2.6 and Assumption 2.3, Hölder's inequality leads to

$$\mathbb{E}\Big(\int_0^T|\dot h(s)|^2ds\Big)^2\le\lambda^4\|\sigma^{-1}\|_\infty^4T\int_0^T\mathbb{E}|Z(s)|^4ds<\infty.$$

This shows that $h(\cdot)\in L_a^4(\Omega\to\mathcal{H},\mathbb{P})\in\mathrm{Dom}(\delta)$. Hence, (2.23) has a unique solution for this $h(\cdot)$ under Assumption 2.1. By Lemmas 2.8 and 2.10, $u_t:=\nabla_\eta X_t-U_t^h$ satisfies the following SFDE

$$du(t)=\{\langle\nabla b(X_t),u_t\rangle-\lambda Z(t)\}dt+\langle\nabla\sigma(X_t),u_t\rangle dW(t)$$

with $u_0=\eta\in\mathcal{C}_r$. By the strong uniqueness of solution to (2.4), $Z_t=u_t=\nabla_\eta X_t-U_t^h$ for $t\in[0,T]$. Then for any $f\in C_b^1(\mathcal{C}_r,\mathbb{R})$ and $\lambda>0$, the chain rule yields

$$\nabla_\eta P_tf(\xi)=\mathbb{E}\langle(\nabla f)(X_t),\nabla_\eta X_t\rangle=\mathbb{E}\langle(\nabla f)(X_t),Z_t\rangle+\mathbb{E}\langle(\nabla f)(X_t),U_t^h\rangle. \tag{2.32}$$

Let

$$h_n(t)=\int_0^t\dot h(s)1_{\{|\dot h(s)|\le n\}}ds=\lambda\int_0^t\sigma^{-1}(X_s)Z(s)1_{\{\lambda|\sigma^{-1}(X_s)Z(s)|\le n\}}ds,\ t\in[0,T].$$

Obviously, $h_n(t)\in L^\infty(\Omega\to\mathcal{H};\mathbb{P})$. By Lemma 2.8, $U_t^{h_n}=D_{h_n}X_t$ for any $t\in[0,T]$. Furthermore, the chain rule and integration-by-parts formula yields

$$\mathbb{E}\langle(\nabla f)(X_T),U_T^{h_n}\rangle=\mathbb{E}D_{h_n}f(X_T)=\mathbb{E}f(X_T)\delta(h_n). \tag{2.33}$$

By (2.7), (2.23), Assumption 2.3, and a similar approach as in the derivations of (2.17) and (2.18), we have

$$\begin{aligned}\mathbb{E}\|U_T^h-U_T^{h_n}\|_r^2&\le C(T)\mathbb{E}\Big(\int_0^T|\sigma(X_s)(\dot h(s)-\dot h_n(s))|ds\Big)^2\\&\le\lambda^2C(T)T\int_0^T\mathbb{E}|Z(s)|^21_{\{\lambda|\sigma^{-1}(X_s)Z(s)|>n\}}ds\ \to 0\ \text{ as }\ n\to\infty.\end{aligned} \tag{2.34}$$

By virtue of Assumption 2.3 and Lemma 2.6, as $n \to \infty$, we have

$$\mathbb{E}|\delta(h) - \delta(h_n)|^2 = \lambda^2 \mathbb{E} \int_0^T |\sigma^{-1}(X_s)Z(s)|^2 1_{\{\lambda|\sigma^{-1}(X_s)Z(s)|>n\}} ds \to 0. \tag{2.35}$$

Since $f \in C_b^1(\mathcal{C}_r, \mathbb{R})$, it follows from (2.33), (2.34), and (2.35) that

$$\mathbb{E}\langle(\nabla f)(X_T), U_T^h\rangle = \mathbb{E} f(X_T)\delta(h) = \lambda \mathbb{E}\Big(f(X_T) \int_0^T \sigma^{-1}(X_s)Z(s)dW(s)\Big).$$

Hence, from (2.32),

$$\nabla_\eta P_T f(\xi) = \mathbb{E}\langle(\nabla f)(X_T), Z_T\rangle + \lambda \mathbb{E}\Big(f(X_T) \int_0^T \sigma^{-1}(X_s)Z(s)dW(s)\Big),$$

which is the desired assertion (2.3). This together with (2.6) yields (2.5). □

Proposition 2.11 presents the stabilization bound for $\nabla P_t f$ as $t \to \infty$. To proceed, we first introduce the following notation. Let $\mathcal{P}(\mathcal{C}_r)$ denote the family of probability measures on $\mathcal{C}_r$. Given a nonnegative and continuous functional $F$ on $\mathcal{C}_r$ and a constant $\gamma \in (0, 1]$, define a distance-like function by $d_{F,\gamma}(\xi, \eta) := \sqrt{(1 \wedge \|\xi - \eta\|_r^\gamma)(1 + F(\xi) + F(\eta))}$. For $d_{F,\gamma}$, the associated Wasserstein distance between two probability measures $\mu, \nu \in \mathcal{P}(\mathcal{C}_r)$ is defined as follows:

$$\mathbb{W}_{d_{F,\gamma}}(\mu, \nu) = \inf_{\Pi \in \mathscr{C}(\mu,\nu)} \int_{\mathcal{C}_r \times \mathcal{C}_r} d_{F,\gamma}(\xi, \eta)\Pi(d\xi, d\eta),$$

where $\mathscr{C}(\mu, \nu)$ denotes the collection of all couplings of $\mu$ and $\nu$. Let $\mathbb{L}(F)$ denote the family of Lipschitz continuous function with respective to $d_{F,\gamma}$ on $\mathcal{C}_r$, that is,

$$\mathbb{L}(F) := \Big\{f : \|f\|_{F,\gamma} := \sup_{\xi \neq \eta, \xi, \eta \in \mathcal{C}_r} \frac{|f(\xi) - f(\eta)|}{d_{F,\gamma}(\xi, \eta)} < \infty\Big\}.$$

**Proposition 2.11.** *Assume Assumptions 2.1, 2.3, and 2.4 hold. Assume further that there exist a continuous functional $V : \mathcal{C}_r \to [0, \infty)$ with $\lim_{\|\xi\|_r \to \infty} V(\xi) = \infty$ and constants $C_V, \theta > 0$ such that*

$$P_t V(\xi) := \int_{\mathcal{C}_r} V(\eta)P_t(\xi, d\eta) \le C_V e^{-\theta t} V(\xi) + C_V \tag{2.36}$$

*holds for all $\xi \in \mathcal{C}_r$ and $t \ge 0$. Then for any $\gamma \in (0, 1]$, there exist positive constants $\bar{\alpha}$ and $C$ such that for any $f \in C_b^1(\mathcal{C}_r) \bigcap \mathbb{L}(V)$ and $\xi \in \mathcal{C}_r$, $t > 0$,*

$$\|\nabla P_t f(\xi)\| \le A \sup_{\phi \in \mathcal{C}_r} \|\nabla f(\phi)\| e^{-\alpha t} + C\|f\|_{V,\gamma} e^{-\bar{\alpha} t}. \tag{2.37}$$

*Proof.* Under assumptions of this proposition, it follows from [17] that $P_t$ has a unique invariant probability measure $\mu$ and there exist constants $C_0$ and $\rho > 0$ such that

$$\mathbb{W}_{d_{V,\gamma}}(P_t(\xi, \cdot), \mu) \le C_0 e^{-\rho t} \sqrt{1 + V(\xi)}, \ t \ge 0. \tag{2.38}$$

For any $t_0 \in [0, t]$, we have

$$\begin{aligned}\Big|\mathbb{E}\Big(f(X_t) \int_0^t \sigma(X_s)^{-1} Z_s dW(s)\Big)\Big| &\le \Big|\mathbb{E}\Big(f(X_t) \int_{t_0}^t \sigma(X_s)^{-1} Z_s dW(s)\Big)\Big| \\ &\quad + \Big|\mathbb{E}\Big(f(X_t) \int_0^{t_0} \sigma(X_s)^{-1} Z_s dW(s)\Big)\Big| =: I_1 + I_2.\end{aligned} \tag{2.39}$$

By virtue of (2.6), (2.36), and the Hölder inequality, we obtain

$$\begin{aligned} I_1 &= \Big|\mathbb{E}\Big((f(X_t)-f(0))\int_{t_0}^t \sigma(X_s)^{-1}Z_s dW(s)\Big)\Big| \\ &\le \|f\|_{V,\gamma}\sqrt{1+\mathbb{E}V(X_t)+V(0)}\Big(\mathbb{E}\int_{t_0}^t |\sigma(X_s)^{-1}Z_s|^2 ds\Big)^{1/2} \\ &\le \frac{\|\eta\|_r}{\alpha_1}\|f\|_{V,\gamma}\|\sigma^{-1}\|_\infty\sqrt{(1+C_V+C_Ve^{-\theta t}V(\xi)+V(0))A_1}e^{-\frac{\alpha_1 t_0}{2}}. \end{aligned} \tag{2.40}$$

Note that $|P_{t-t_0}f(X_{t_0})-\mu(f)| \le C_0\|f\|_{V,\gamma}\sqrt{1+V(X_{t_0})}e^{-\rho(t-t_0)}$, where $\mu(f) := \int_{\mathcal{C}_r} f(\zeta)\mu(d\zeta)$, which follows from (2.38). By Markov property and (2.6), we have

$$\begin{aligned} I_2 &= \Big|\mathbb{E}\Big((P_{t-t_0}f(X_{t_0})-\mu(f))\int_0^{t_0}\sigma(X_s)Z_s dW(s)\Big)\Big| \\ &\le C_0\|f\|_{V,\gamma}e^{-\rho(t-t_0)}\mathbb{E}\Big(\sqrt{1+V(X_{t_0})}\Big|\int_0^{t_0}\sigma(X_s)^{-1}Z_s dW(s)\Big|\Big) \\ &\le \frac{C_0}{\alpha_1}\|f\|_{V,\gamma}e^{-\rho(t-t_0)}\sqrt{1+C_Ve^{-\theta t_0}V(\xi)+C_V}\|\sigma^{-1}\|_\infty\sqrt{A_1}\|\eta\|_r. \end{aligned} \tag{2.41}$$

Taking $t_0 = t/2$ it follows from (2.40), (2.41), and (2.39) that there exists a constant $C > 0$ such

$$\Big|\mathbb{E}\Big(f(X_t)\int_0^t \sigma(X_s)^{-1}Z_s dW(s)\Big)\Big| \le C(e^{-\alpha_1 t/4}+e^{-\rho t/2})\|f\|_{V,\gamma}\|\eta\|_r,$$

which together with (2.5) yields

$$\|\nabla P_t f(\xi)\| \le A\sup_{\phi\in\mathcal{C}_r}\|\nabla f(\phi)\|e^{-\alpha t}+C(e^{-\alpha_1 t/4}+e^{-\rho t/2})\|f\|_{V,\gamma}.$$

Hence, (2.37) holds for $\bar\alpha = (\alpha_1/4)\wedge(\rho/2)$. The proof is concluded. □

## 3. Degenerate SFDEs with infinite delay

This section considers the following SFDE with infinite delay and degenerate diffusion coefficients on $\mathbb{R}^{2d} := \mathbb{R}^d\times\mathbb{R}^d$:

$$\begin{cases} dX^{(1)}(t) = b^{(1)}(X_t)dt, \\ dX^{(2)}(t) = b^{(2)}(X_t)dt + \sigma(X_t)dW(t) \end{cases} \tag{3.1}$$

with $X_0 = \xi := (\xi^{(1)}, \xi^{(2)}) \in \mathcal{C}_r\times\mathcal{C}_r$, where $X_t := (X_t^{(1)}, X_t^{(2)}) \in \mathcal{C}_r\times\mathcal{C}_r$ for some $r > 0$, $\{W(t)\}_{t\ge0}$ is an $m$-dimensional Brownian motion, $b := (b^{(1)}, b^{(2)}) : \mathcal{C}_r\times\mathcal{C}_r \mapsto \mathbb{R}^{2d}$ and $\sigma : \mathcal{C}_r\times\mathcal{C}_r \mapsto \mathbb{R}^{d\times m}$ are continuous functionals.

For (3.1), the diffusion term is degenerate and it does not satisfy the assumptions of Theorem 2.5. To establish the asymptotic Bismut formula for (3.1), we make Assumptions 3.1 and 3.2.

**Assumption 3.1.** $b \in C^1(\mathcal{C}_r\times\mathcal{C}_r, \mathbb{R}^{2d})$ is bounded on bounded sets. There exist constants $K_1$, $K_2$, $K_3$, $\kappa > 0$ satisfying $\inf_{0<\varepsilon<1}\frac{2K_1\varepsilon+(\varepsilon+32)K_3}{2(1-\varepsilon)\varepsilon} < r\wedge\kappa$, such that for any $\phi := (\phi^{(1)}, \phi^{(2)}), \psi := (\psi^{(1)}, \psi^{(2)}) \in \mathcal{C}_r\times\mathcal{C}_r$, $\|\nabla_\phi\sigma(\psi)\|_{\mathrm{HS}}^2 \le K_3\|\phi\|_r^2$ and

$$\big(\phi^{(1)}(0), \nabla_\phi b^{(1)}(\psi)\big)+\big(\phi^{(2)}(0), \nabla_\phi b^{(2)}(\psi)\big) \le K_1\|\phi\|_r^2+K_2|\phi^{(2)}(0)|\|\phi\|_r-\kappa|\phi^{(1)}(0)|^2.$$

**Assumption 3.2.** $b \in C^1(\mathcal{C}_r \times \mathcal{C}_r, \mathbb{R}^{2d})$ is bounded on bounded sets. $\|\nabla\sigma\|$ and $\|\nabla b^{(2)}\|$ are bounded. There exist constants $\beta, \kappa > 0$ with $\inf_{0<\alpha<\kappa\wedge r} \frac{\beta^2}{\alpha(\kappa-\alpha)} < 4$ such that

$$\big(\phi^{(1)}(0), \nabla_\phi b^{(1)}(\psi)\big) \le \beta|\phi^{(1)}(0)|\|\phi\|_r - \kappa|\phi^{(1)}(0)|^2.$$

**Theorem 3.3.** *Assume Assumptions 2.3, 2.4, and 3.1 or 3.2 hold. Then for any* $T > 0, \lambda > 0$, $f \in C_b^1(\mathcal{C}_r \times \mathcal{C}_r, \mathbb{R})$, *and* $\xi, \eta \in \mathcal{C}_r \times \mathcal{C}_r$,

$$\nabla_\eta P_T f(\xi) = \mathbb{E}\langle(\nabla f)(X_t), Z_T\rangle + \lambda\mathbb{E}\Big(f(X_T)\int_0^T \sigma(X_s)^{-1}Z_s^{(2)}dW(s)\Big), \tag{3.2}$$

*where* $Z_t$ *is the segment process of* $Z(t) := (Z^{(1)}(t), Z^{(2)}(t)) \in \mathbb{R}^{2d}$ *which solves uniquely the following SFDE*

$$dZ(t) = \{-\lambda\big(\mathbf{0}, Z^{(2)}(t)\big) + \langle\nabla b(X_t), Z_t\rangle\}dt + \big(\mathbf{0}, \langle\nabla\sigma(X_t), Z_t\rangle dW(t)\big), \tag{3.3}$$

*with* $Z_0 = \eta \in \mathcal{C}_r \times \mathcal{C}_r$. *Moreover, there exist constants* $c$, $\alpha$, *and* $\lambda_0 > 0$ *such that for any* $\lambda > \lambda_0$

$$\Big|\nabla_\eta P_T f(\xi) - \lambda\mathbb{E}\Big(f(X_T)\int_0^T \sigma(X_s)^{-1}Z_s^{(2)}dW(s)\Big)\Big| \le c\|\eta\|_r e^{-\alpha T}. \tag{3.4}$$

We first derive Lemma 3.4 before proving Theorem 3.3.

**Lemma 3.4.** *Let Assumptions 2.3 and 3.1 or 3.2 hold. Then (3.3) has a unique solution. Moreover, there exist constants* $p_0 > 2$, $c$, $\theta > 0$, *and* $\lambda_0 > 0$ *large enough such that for any* $\lambda > \lambda_0$,

$$\mathbb{E}\|Z_t\|_r^p \le c\|\eta\|_r^p e^{-\theta t}, \ \ p \in (0, p_0], \ \ t \ge 0. \tag{3.5}$$

*Proof.* Note that (3.3) has a unique solution under Assumption 3.1 or 3.2 and for any $q > 0$ there exist a nonnegative continuous increasing function $A_2(t)$ and constant $C > 0$ such that

$$\mathbb{E}\|Z_t\|_r^q \le C\|\eta\|_r^q A_2(t), \ \ t \ge 0. \tag{3.6}$$

To proceed, we prove (3.5) under Assumptions 3.1 and 3.2.

(i) *Assumption 3.1 holds.* Let $\lambda \ge 2\kappa$. Then for some $p > 2$ and any $\alpha \le r \wedge \kappa$, applying the Itô formula, Assumption 3.1 and the Young inequality gives

$$\begin{aligned} &e^{p\alpha t}|Z(t)|^p \\ &\quad \le |Z(0)|^p + p\int_0^t e^{p\alpha s}|Z(s)|^{p-2}\Big\{\alpha|Z(s)|^2 + K_1\|Z_s\|_r^2 + K_2|Z^{(2)}(s)|\|Z_s\|_r \\ &\qquad - \kappa|Z^{(1)}(s)|^2 - \lambda|Z^{(2)}(s)|^2 + \frac{K_3(p-1)}{2}\|Z_s\|_r^2\Big\}ds + pN(t) \\ &\quad \le |Z(0)|^p + p\big(K_1 + \frac{K_3(p-1)}{2} + \frac{K_2^2}{2\lambda}\big)\int_0^t e^{p\alpha s}|Z(s)|^{p-2}\|Z_s\|_r^2 ds + pN(t), \end{aligned} \tag{3.7}$$

where $N(t) := \int_0^t e^{p\alpha s}|Z(s)|^{p-2}\big(Z^{(2)}(s), \langle\nabla\sigma(X_s), Z_s\rangle dW(s)\big)$. Similar to (2.18), by Assumption 3.1, there exists a constant $\varepsilon \in (0,1)$ such that

$$p\mathbb{E}\Big[\sup_{0\le u\le t} N(u)\Big] \le \varepsilon\mathbb{E}\Big[\sup_{0\le s\le t} e^{p\alpha s}|Z(s)|^p\Big] + \frac{8p^2K_3}{\varepsilon}\mathbb{E}\int_0^t e^{p\alpha s}|Z(s)|^{p-2}\|Z_s\|_r^2 ds. \tag{3.8}$$

Let $C_6 = 2K_1 + (p-1+16p/\varepsilon)K_3 + K_2^2/\lambda$, then by (3.7) and (3.8), we have

$$\mathbb{E}\Big[\sup_{0\le s\le t} e^{p\alpha s}|Z(s)|^p\Big] \le \frac{1}{1-\varepsilon}|Z(0)|^p + \frac{p}{2(1-\varepsilon)}C_6\mathbb{E}\int_0^t e^{p\alpha s}\|Z_s\|_r^p ds.$$

Since $\alpha \le r$, then $\|Z_t\|_r^p \le e^{-p\alpha t}\|\eta\|_r^p + e^{-p\alpha t}\sup_{0\le s\le t} e^{p\alpha s}|Z(s)|^p$. Thus we obtain

$$\mathbb{E}e^{p\alpha t}\|Z_t\|_r^p \le \frac{2-\varepsilon}{1-\varepsilon}\|\eta\|_r^p + \frac{p}{2(1-\varepsilon)}C_6\int_0^t \mathbb{E}e^{p\alpha s}\|Z_s\|_r^p ds.$$

Applying the Gronwall inequality leads to

$$\mathbb{E}\|Z_t\|_r^p \le \frac{2-\varepsilon}{1-\varepsilon}\|\eta\|_r^p \exp\Big\{\frac{p}{2(1-\varepsilon)}C_6 t - p\alpha t\Big\}.$$

Take $\alpha = r\wedge\kappa$. Noting that $\inf_{0<\varepsilon<1}\frac{2K_1\varepsilon+(\varepsilon+32)K_3}{2(1-\varepsilon)\varepsilon} < r\wedge\kappa$, we can choose constants $\varepsilon_0\in(0,1)$, $p>2$ sufficiently close to 2 and $\lambda_0$ large enough such that for any $\lambda>\lambda_0$, $\frac{1}{2(1-\varepsilon_0)}C_6 - \alpha < 0$. Thus, there exists a $\theta>0$ such that (3.5) holds.

(ii) *Assumption* 3.2 *holds.* In terms of (3.3), for $\lambda > r$, we have

$$e^{\lambda t}Z^{(2)}(t) = Z^{(2)}(0) + \int_0^t e^{\lambda s}\langle\nabla b^{(2)}(X_s), Z_s\rangle ds + \int_0^t e^{\lambda s}\langle\nabla\sigma(X_s), Z_s\rangle dW(t).$$

Then for any $\alpha\in(0, r\wedge\kappa)$, we have

$$\begin{aligned} e^{\alpha t}Z^{(2)}(t) =& Z^{(2)}(0)e^{-(\lambda-\alpha)t} + \int_0^t e^{-(\lambda-\alpha)(t-s)}e^{\alpha s}\langle\nabla b^{(2)}(X_s), Z_s\rangle ds \\ &+ \int_0^t e^{-(\lambda-\alpha)(t-s)}e^{\alpha s}\langle\nabla\sigma(X_s), Z_s\rangle dW(t). \end{aligned}$$

Noting that $\|\nabla b^{(2)}\|$ and $\|\nabla\sigma\|$ are bounded, then by using similar arguments to derive (2.10) and (2.11), for any $p>2$ there exist a constant $c_1>0$ and a function $a_{p,s}:\mathbb{R}_+\to\mathbb{R}_+$ with $\lim_{s\to\infty}a_{p,s}=0$ such that

$$\mathbb{E}\sup_{0\le s\le t} e^{p\alpha s}|Z^{(2)}(s)|^p \le c_1|Z^{(2)}(0)|^p + a_{p,\lambda-\alpha}\int_0^t \mathbb{E}e^{p\alpha s}\|Z_s\|_r^p ds.$$

Using the definition of norm $\|\cdot\|_r$, we obtain

$$e^{p\alpha t}\mathbb{E}\|Z^{(2)}(s)\|_r^p \le (c_1+1)|Z^{(2)}(0)|^p + a_{p,\lambda-\alpha}\int_0^t e^{p\alpha s}\mathbb{E}\|Z_s\|_r^p ds. \tag{3.9}$$

For any $\varepsilon>0$, by virtue of Assumption 3.2 and (3.3), we have

$$\begin{aligned} &e^{\kappa t}\big(|Z^{(1)}(t)|^2+\varepsilon\big)^{\frac12} - \big(|Z^{(1)}(0)|^2+\varepsilon\big)^{\frac12} \\ &\le \int_0^t e^{\kappa s}\Big\{\kappa\big(|Z^{(1)}(s)|^2+\varepsilon\big)^{\frac12} + \frac{\beta|Z^{(1)}(s)|\|Z_s\|_r - \kappa|Z^{(1)}(s)|^2}{\big(|Z^{(1)}(s)|^2+\varepsilon\big)^{\frac12}}\Big\}ds \\ &\le \int_0^t e^{\kappa s}\Big(\beta\|Z_s\|_r + \frac{\kappa\varepsilon}{\big(|Z^{(1)}(s)|^2+\varepsilon\big)^{\frac12}}\Big)ds. \end{aligned}$$

Then, letting $\varepsilon\to0$ gives

$$e^{\kappa t}|Z^{(1)}(t)| \le |Z^{(1)}(0)| + \beta\int_0^t e^{\kappa s}\|Z_s\|_r ds.$$

For any $\alpha \in (0, r \wedge \kappa)$, we have

$$e^{\alpha t}|Z^{(1)}(t)| \leq |Z^{(1)}(0)|e^{-(\kappa-\alpha)t} + \beta \int_0^t e^{-(\kappa-\alpha)(t-s)} e^{\alpha s}\|Z_s\|_r ds,$$

which further implies that there exist constants $c_2$ and $\varepsilon_1 > 0$ such that

$$\begin{aligned} e^{p\alpha t}\mathbb{E}\|Z_t^{(1)}\|_r^p &\leq \|Z_0^{(1)}\|_r^p + \mathbb{E} \sup_{0\leq s\leq t} e^{p\alpha s}|Z^{(1)}(s)|^p \\ &\leq c_2\|Z_0^{(1)}\|_r^p + (1+\varepsilon_1)\beta^p \mathbb{E} \sup_{0\leq s\leq t} \Big(\int_0^s e^{-(\kappa-\alpha)(s-u)} e^{\alpha u}\|Z_u\|_r du\Big)^p \\ &\leq c_2\|Z_0^{(1)}\|_r^p + (1+\varepsilon_1)\beta^p \Big[\frac{p-1}{p(\kappa-\alpha)}\Big]^{p-1} \int_0^t e^{p\alpha s}\mathbb{E}\|Z_s\|_r^p ds, \end{aligned} \tag{3.10}$$

where we have used the inequality $(a+b)^p \leq c_\varepsilon a^p + (1+\varepsilon)b^p$ for $a, b, \varepsilon, c_\varepsilon > 0, p > 1$ and Hölder's inequality. By (3.9) and (3.10), there exist constants $\varepsilon_2$ and $c_{\varepsilon_2,p} > 0$ satisfying $\lim_{\varepsilon\to 0} c_{\varepsilon,p} = \infty$ and $c_{\varepsilon_1,\varepsilon_2} > 0$ such that

$$e^{p\alpha t}\mathbb{E}\|Z_t\|_r^p \leq c_{\varepsilon_1,\varepsilon_2}\|Z_0\|_r^p + \Gamma_{\lambda,\varepsilon_1,\varepsilon_2}\int_0^t e^{p\alpha s}\mathbb{E}\|Z_s\|_r^p ds,$$

where $\Gamma_{\lambda,\varepsilon_1,\varepsilon_2} := \big\{(1+\varepsilon_2)(1+\varepsilon_1)\beta^p \big[\frac{p-1}{p(\kappa-\alpha)}\big]^{p-1} + (1+c_{\varepsilon_2,p})a_{p,\lambda-\alpha}\big\}$. Applying the Gronwall inequality yields that

$$\mathbb{E}\|Z_t\|_r^p \leq c_{\varepsilon_1,\varepsilon_2}\|Z_0\|_r^p \exp[(\Gamma_{\lambda,\varepsilon_1,\varepsilon_2} - p\alpha)t]. \tag{3.11}$$

It is easy to see from Assumption 3.2 that there exists $\alpha_0 \in (0, \kappa \wedge r)$ such that $\beta^2/[2(\kappa-\alpha_0)] < 2\alpha_0$. Then we can find constants $\varepsilon_1, \varepsilon_2$ small enough, $p > 2$ close to 2 enough and $\lambda_0$ large enough such that

$$\Big\{(1+\varepsilon_2)(1+\varepsilon_1)\beta^p\Big[\frac{p-1}{p(\kappa-\alpha_0)}\Big]^{p-1} + (1+c_{\varepsilon_2,p})a_{p,\lambda-\alpha_0}\Big\} < p\alpha_0,$$

which implies that (3.5) holds for some $\theta > 0$. The proof is complete. □

*Proof of Theorem* 3.3. For any $h \in L^\infty(\Omega \to \mathcal{H}, \mathbb{P})$, by Lemma 2.8, the segment process $X_t$ is Malliavin differentiable along the direction $h$ and the directional derivative $D_h X_t$ is the unique solution to the following $2d$-dimensional SFDE

$$dU^h(t) = \big(\langle\nabla b(X_s), U_t^h\rangle + (\mathbf{0}, \sigma(X_t)\dot{h}(t))\big)dt + \big(\mathbf{0}, \langle\nabla\sigma(X_t), U_t^h\rangle dW(t), \tag{3.12}$$

on $t \in [0, T]$, with $U_0^h = \mathbf{0} \in \mathcal{C}_r \times \mathcal{C}_r$, where $U^h(t) := (U^{h,(1)}(t), U^{h,(2)}(t)) \in \mathbb{R}^{2d}$. Let $h(t) = \lambda\int_0^t \sigma^{-1}(X_s)Z^{(2)}(s)ds$, which is an adapted stochastic process. From Lemma 3.4 and Assumption 2.3, we see that

$$\mathbb{E}\Big(\int_0^T |\dot{h}(s)|^2 ds\Big)^{\frac{p_0}{2}} \leq \lambda^{p_0}\|\sigma^{-1}\|_\infty^{p_0} T^{\frac{p_0-2}{2}} \int_0^T \mathbb{E}|Z^{(2)}(s)|^{p_0} ds < \infty,$$

which implies $h \in L_a^{p_0}(\Omega \to \mathcal{H}, \mathbb{P}) \in \mathrm{Dom}(\delta)$, where $p_0 > 2$ comes from Lemma 3.4. And (3.12) has a unique solution for this $h$ under assumptions of Theorem 3.3.

By Lemma 2.10, the segment process to (3.1) is Fréchet differentiable and the directional derivative $\nabla_\eta X_t$ is the unique segment process to the following SFDE

$$dV(t) = \langle\nabla b(X_t), V_t\rangle dt + (\mathbf{0}, \langle\nabla\sigma(X_t), V_t\rangle)dW(t),\ t \in [0, T], V_0 = \eta. \tag{3.13}$$

Because (3.3) has a unique solution, from (3.12) and (3.13), $Z_t := \nabla_\eta X_t - U_t^h$ is the unique segment process to (3.3) for $t \in [0, T]$. Then using similar arguments as in the derivation of Theorem 2.5 yields the desired results. This proof is complete. □

To present the bound for $\nabla P_t f$ as $t \to \infty$, we need the stronger assumption:

**Assumption 3.5.** $b \in C^1(\mathcal{C}_r \times \mathcal{C}_r, \mathbb{R}^{2d})$ is bounded on bounded sets. There exist a measure $\mu \in M_{2r}$ and constants $\hat{k}_1, \hat{k}_2, \hat{\kappa} > 0$ with $\inf_{0<\varepsilon<1} \frac{2\hat{k}_1\mu^{(2r)}\varepsilon+(\varepsilon+32)\hat{k}_2\mu^{(2r)}}{2(1-\varepsilon)\varepsilon} < r \wedge \hat{\kappa}$ such that for any $\phi := (\phi^{(1)}, \phi^{(2)}), \psi := (\psi^{(1)}, \psi^{(2)}) \in \mathcal{C}_r \times \mathcal{C}_r$

$$(\phi(0) - \psi(0), b(\phi) - b(\psi)) \leq -\hat{\kappa}|\phi(0) - \psi(0)|^2 + \hat{k}_1 \int_{-\infty}^{0} |\phi(\theta) - \psi(\theta)|^2 \mu(d\theta),$$

$$\|\sigma(\phi) - \sigma(\psi)\|_{\mathrm{HS}}^2 \leq \hat{k}_2 \int_{-\infty}^{0} |\phi(\theta) - \psi(\theta)|^2 \mu(d\theta).$$

**Proposition 3.6.** *Under Assumptions 2.3, 2.4, and 3.5, there exist constants $\bar{\alpha}$, $C > 0$ such that for any $f \in C_b^1(\mathcal{C}_r \times \mathcal{C}_r, \mathbb{R})$ and $\xi \in \mathcal{C}_r \times \mathcal{C}_r$ and $t > 0$,*

$$\|\nabla P_t f(\xi)\| \leq C\|\nabla f\|_\infty (1 + \|\xi\|_r) e^{-\bar{\alpha} t}. \tag{3.14}$$

*Proof.* By [21, Theorems 5.1 and 4.3], (3.1) has a unique invariant probability measure $\mu$ under Assumption 3.5 and for any $\xi \in \mathcal{C}_r \times \mathcal{C}_r$, $f \in C_b^1(\mathcal{C}_r \times \mathcal{C}_r, \mathbb{R})$, there exist constants $C_i$ and $\tilde{\alpha}_i > 0$ for $i = 1, 2$ such that

$$\mathbb{E}\|X_t\|_r^2 \leq C_1(1 + \|\xi\|_r e^{-\tilde{\alpha}_1 t}) \quad \text{and} \quad |P_t f(\xi) - \mu(f)| \leq C_2 \|\nabla f\|_\infty (1 + \|\xi\|_r) e^{-\tilde{\alpha}_2 t}.$$

Then by using similar arguments as in the proof of Proposition 2.11, (3.14) follows. □

## References


[1] Jean-Michel Bismut, *Large deviations and the Malliavin calculus*, Progress in Mathematics, vol. 45, Birkhäuser Boston, Inc., Boston, MA, 1984. MR755001

[2] K. D. Elworthy and X.-M. Li, *Formulae for the derivatives of heat semigroups*, J. Funct. Anal. **125** (1994), no. 1, 252–286, DOI 10.1006/jfan.1994.1124. MR1297021

[3] Feng-Yu Wang, *Harnack inequalities for stochastic partial differential equations*, SpringerBriefs in Mathematics, Springer, New York, 2013, DOI 10.1007/978-1-4614-7934-5. MR3099948

[4] Xicheng Zhang, *Stochastic flows and Bismut formulas for stochastic Hamiltonian systems*, Stochastic Process. Appl. **120** (2010), no. 10, 1929–1949, DOI 10.1016/j.spa.2010.05.015. MR2673982

[5] Feng-Yu Wang and Xi-Cheng Zhang, *Derivative formula and applications for degenerate diffusion semigroups* (English, with English and French summaries), J. Math. Pures Appl. (9) **99** (2013), no. 6, 726–740, DOI 10.1016/j.matpur.2012.10.007. MR3055216

[6] Xicheng Zhang, *Derivative formulas and gradient estimates for SDEs driven by $\alpha$-stable processes*, Stochastic Process. Appl. **123** (2013), no. 4, 1213–1228, DOI 10.1016/j.spa.2012.11.012. MR3016221

[7] Feng-Yu Wang, Lihu Xu, and Xicheng Zhang, *Gradient estimates for SDEs driven by multiplicative Lévy noise*, J. Funct. Anal. **269** (2015), no. 10, 3195–3219, DOI 10.1016/j.jfa.2015.09.007. MR3401615

[8] Linlin Wang, Longjie Xie, and Xicheng Zhang, *Derivative formulae for SDEs driven by multiplicative $\alpha$-stable-like processes*, Stochastic Process. Appl. **125** (2015), no. 3, 867–885, DOI 10.1016/j.spa.2014.10.011. MR3303960

[9] Panpan Ren and Feng-Yu Wang, *Bismut formula for Lions derivative of distribution dependent SDEs and applications*, J. Differential Equations **267** (2019), no. 8, 4745–4777, DOI 10.1016/j.jde.2019.05.016. MR3983053

[10] Jianhai Bao, Feng-Yu Wang, and Chenggui Yuan, *Derivative formula and Harnack inequality for degenerate functionals SDEs*, Stoch. Dyn. **13** (2013), no. 1, 1250013, 22, DOI 10.1142/S021949371250013X. MR3007251

[11] Jianhai Bao, Feng-Yu Wang, and Chenggui Yuan, *Bismut formulae and applications for functional SPDEs*, Bull. Sci. Math. **137** (2013), no. 4, 509–522, DOI 10.1016/j.bulsci.2012.11.005. MR3054274

[12] Jianhai Bao, Panpan Ren, and Feng-Yu Wang, *Bismut formula for Lions derivative of distribution-path dependent SDEs*, J. Differential Equations **282** (2021), 285–329, DOI 10.1016/j.jde.2021.02.019. MR4219322

[13] Michael Scheutzow, *Exponential growth rates for stochastic delay differential equations*, Stoch. Dyn. **5** (2005), no. 2, 163–174, DOI 10.1142/S0219493705001468. MR2147280

[14] Alexei Kulik and Michael Scheutzow, *Well-posedness, stability and sensitivities for stochastic delay equations: a generalized coupling approach*, Ann. Probab. **48** (2020), no. 6, 3041–3076, DOI 10.1214/20-AOP1449. MR4164460

[15] Yoshiyuki Hino, Satoru Murakami, and Toshiki Naito, *Functional-differential equations with infinite delay*, Lecture Notes in Mathematics, vol. 1473, Springer-Verlag, Berlin, 1991, DOI 10.1007/BFb0084432. MR1122588

[16] David Nualart, *The Malliavin calculus and related topics*, 2nd ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR2200233

[17] Ya Wang, Fuke Wu, George Yin, and Chao Zhu, *Stochastic functional differential equations with infinite delay under non-Lipschitz coefficients: Existence and uniqueness, Markov property, ergodicity, and asymptotic log-Harnack inequality*, Stochastic Process. Appl. **149** (2022), 1–38, DOI 10.1016/j.spa.2022.03.008. MR4405475

[18] Xuerong Mao, *Stochastic differential equations and applications*, 2nd ed., Horwood Publishing Limited, Chichester, 2008, DOI 10.1533/9780857099402. MR2380366

[19] Abdelhadi Es-Sarhir, Michael Scheutzow, and Onno van Gaans, *Invariant measures for stochastic functional differential equations with superlinear drift term*, Differential Integral Equations **23** (2010), no. 1-2, 189–200. MR2588808

[20] P. Drábek and J. Milota, *Methods of nonlinear analysis: applications to differential equations*, Springer, Basel, 2013.

[21] Fuke Wu, George Yin, and Hongwei Mei, *Stochastic functional differential equations with infinite delay: existence and uniqueness of solutions, solution maps, Markov properties, and ergodicity*, J. Differential Equations **262** (2017), no. 3, 1226–1252, DOI 10.1016/j.jde.2016.10.006. MR3582192

School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, Hubei 430074, People's Republic of China

*Email address*: y_wang@hust.edu.cn

School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, Hubei 430074, People's Republic of China

*Email address*: wufuke@hust.edu.cn

Department of Mathematics, University of Connecticut, Storrs, Connecticut 06269-1009

*Email address*: gyin@uconn.edu