Nonlinear Functional Analysis and Applications
Vol. 20, No. 4 (2015), pp. 659-674

http://nfaa.kyungnam.ac.kr/jour-nfaa.htm

LARGE DEVIATIONS FOR STOCHASTIC DIFFERENTIAL EQUATIONS WITH DEVIATING ARGUMENTS

M. Suvinthra[1], K. Balachandran[2] and J. K. Kim[3]

[1]Department of Mathematics, Bharathiar University
Coimbatore 641 046, India
e-mail: suvinthra@gmail.com

[2]Department of Mathematics, Bharathiar University
Coimbatore 641 046, India
e-mail: kb.maths.bu@gmail.com

[3]Department of Mathematics Education, Kyungnam University
Changwon, Gyeongnam, 51767, Korea
e-mail: jongkyuk@kyungnam.ac.kr

Abstract. In this paper, we study the large deviations for a class of stochastic differential equations with deviating arguments. The randomness is assumed to be Gaussian, and both additive and multiplicative noise types are considered. We adopt the contraction principle argument and weak convergence approach to establish the Freidlin-Wentzell type large deviation principle for the additive and multiplicative noise cases respectively.

1. Introduction

Beginning from Newton's second law of motion, almost all physical problems are modeled as differential and integral equations. The qualitative behavior of solutions of these equations is helpful to study the systematic changes occurring in the corresponding physical phenomena. The accuracy of such analysis can be improved by considering all the possible factors affecting the system during the modeling process. Among the other factors used to model the system, taking into account the occurrences of delay leads to the formation of

[0]Received April 23, 2015. Revised July 11, 2015.

[0]2010 Mathematics Subject Classification: 60H10, 60F10.

[0]Keywords: Large deviation principle, stochastic differential equations, deviating arguments.

delay differential equations, more generally identified as differential equations with deviating arguments in which case the delays are also varying. The analysis of these type of delay equations began at the end of the 18th century and are studied intensively and applied to various problems arising especially from engineering and control. To get a glimpse of a brief history and development in the study of delay equations, one can refer [10].

The construction of a systematic theory of differential equations with deviating arguments began in the 1950s and since then, the analysis of these equations became significant in many physical phenomena and has many surprising ramifications (see [7]). An existence theorem for nonlinear Volterra integral equations with deviating arguments was carried out by Banas [2] and the results were established for a more general class of nonlinear Volterra equations by Balachandran and Ilamaran [1].

In addition, when stochastic or random effects are also considered into the delay systems, they result in stochastic delay equations. A historical development of the study of stochastic delay differential equations can be found in [11] and references therein. The existence and stability of solutions of neutral hybrid stochastic infinite delay differential equations with Poisson jumps were studied by Rathinasamy and Balachandran [16]. The nonlinear stochastic systems with delays in control were analyzed for its controllability problem by Karthikeyan and Balachandran [12]. Most of the research works being carried out is regarding the existence of solutions for these equations, asymptotic stability and control theoretic problems. In recent years, attempts are made to establish the large deviations for stochastic delay differential equations.

Large deviation theory is the study of events whose probabilities of occurrence are extremely small. Those highly improbable events may have a huge impact during its occurrence and so the study of their qualitative and quantitative properties is indeed essential and indispensable. The theory is mainly concentrated in estimating the rate at which the occurrence probability becomes negligible. The origin of the study happened during the 1930s when there was a necessity to tackle the risk or ruin in an insurance company. A general framework for the theory was formulated by Varadhan [20] via a large deviation principle (LDP).

Since then, the theory was tremendously applied to problems from many areas ranging from physics to biology. For instance, it is helpful to calculate the entropy in statistical mechanics, for both equilibrium and non-equilibrium systems. The systems are usually modeled by stochastic differential equations (SDEs) and large deviations are studied for the corresponding solution distributions. The study of large deviation results to SDEs was initiated by

Varadhan [20] for diffusion processes, and it was carried over to a general class of diffusion processes by Freidlin and Wentzell in [9].

A Freidlin-Wentzell type LDP is discussed in [5] for the following stochastic differential equation:

$$\begin{cases} dX(t) &= b\big(t, X(t)\big)\mathrm{d}t + \sqrt{\epsilon}\sigma(t, X(t))\mathrm{d}W(t), \qquad t \in (0, T], \\ X(0) &= X_0. \end{cases}$$

The method developed by Freidlin and Wentzell for studying differential equations with small stochastic perturbation involves time discretization of the original problem, then using Varadhan's contraction principle to establish the LDP for the discretized problem and finally showing that LDP holds in the limit. Following their theory, several authors have studied the large deviations for stochastic differential equations.

Later, Fleming [8] introduced a stochastic control approach to study the large deviation theory. Then Dupuis and Ellis [6] formulated a weak convergence technique by combining with the stochastic control approach developed by Fleming. The major idea behind their method is that the LDP and the Laplace principle are equivalent when the underlying space is Polish (see [5]). In [4], Budhiraja and Dupuis established a variational representation for positive functionals of Brownian motion using which large deviations can be studied for a variety of differential equations. Because of the different nature of non-linearities affecting the system, each equation has to be studied individually for large deviations. The LDP for the two dimensional Navier-Stokes equation was established by Sritharan and Sundar [19], whilst the three dimensional Navier-Stokes equation with tameness was considered by Rockner *et al.* [17]. The Freidlin-Wentzell type large deviations for stochastic evolution equations was analyzed by Liu [13].

The problem of large deviations for delay equations was first studied by Scheutzow [18] for additive Gaussian noise. In [15], Mohammed and Zhang considered the stochastic delay differential equations (SDDEs) with multiplicative type noise given by

$$\begin{aligned} &dX(t) = b\big(t, X(t), X(t-\tau)\big)\mathrm{d}t + \sqrt{\epsilon}\sigma(t, X(t), X(t-\tau))\mathrm{d}W(t),\ t \in (0, T], \\ &X(0) = X_0 \end{aligned}$$

and established the LDP using the approximation method. In [14], Mo and Luo also considered the same SDDEs with both additive and multiplicative noise cases and improved the established LDP results using Varadhan's contraction principle and the weak convergence approach.

In this paper, we study the large deviation principle for stochastic differential equations with deviating arguments (SDEDA) with randomess of Gaussian type. For the additive noise case, we implement the Varadhan's contraction principle to establish the LDP. In the case of multiplicative type noise, we adopt the variational representation formulated by Budhiraja and Dupuis [4] for Brownian motion.

2. Preliminaries

Consider the following differential equation with deviating arguments in the Euclidean space $\mathbb{R}^d$:

$$\begin{cases} \mathrm{d}X(t) &= b\Big(t, X(t), X(\alpha_1(t)), \ldots, X(\alpha_m(t))\Big)\mathrm{d}t, \quad t \in (0,T], \\ X(0) &= X_0, \end{cases} \tag{2.1}$$

where $X_0 \in \mathbb{R}^d$ and for $J = [0,T]$, $b : J \times \mathbb{R}^d \times (\mathbb{R}^d)^m \to \mathbb{R}^d$, with $(\mathbb{R}^d)^m$ denoting the space of all m-tuples formed by elements from $\mathbb{R}^d$. Assume that b satisfies the Lipschitz condition and linear growth property. Then there exists some positive constants L and K such that for all $x, x_1, x_2, \ldots, x_m, y, y_1, y_2, \ldots y_m \in \mathbb{R}^d$ and $t \in J$,

$$\begin{aligned} &\|b(t, x, x_1, x_2, \ldots, x_m) - b(t, y, y_1, y_2, \ldots y_m)\| \\ &\le L\left[\|x - y\| + \sum_{i=1}^{m} \|x_i - y_i\|\right], \end{aligned} \tag{2.2}$$

$$\big\|b(t, x, x_1, x_2, \ldots, x_m)\big\| \le K\left[1 + \|x\| + \sum_{i=1}^{m} \|x_i\|\right], \tag{2.3}$$

where $\|\cdot\|$ denotes the usual Euclidean norm in $\mathbb{R}^d$. We intend to establish the LDP for (2.1) with Gaussian randomness. Let $\{\Omega, \mathcal{F}, \mathbf{P}\}$ be a complete filtered probability space equipped with a complete family of right continuous increasing sub σ-algebras $\{\mathcal{F}_t, t \in J\}$ satisfying $\{\mathcal{F}_t \subset \mathcal{F}\}$, and let $\{X^\epsilon\}$ be a family of random variables defined on this space and taking values in a Polish space E (i.e., a complete separable metric space E). Initially, we quote some definitions and results from large deviation theory:

Definition 2.1. (Rate Function) A function $I : E \to [0, \infty]$ is called a rate function if I is lower semicontinuous. A rate function I is called a good rate function if for each $N < \infty$, the level set $K_N = \{f \in E : I(f) \le N\}$ is compact in E.

Definition 2.2. (Large Deviation Principle) Let I be a rate function on E. We say the sequence $\{X^\epsilon, \epsilon > 0\}$ satisfies the large deviation principle with rate function I if the following two conditions hold:

1. *L*arge deviation upper bound. For each closed subset F of E,

$$\limsup_{\epsilon\to 0} \epsilon \log \mathbf{P}(X^\epsilon \in F) \leq -I(F).$$

2. *L*arge deviation lower bound. For each open subset G of E,

$$\liminf_{\epsilon\to 0} \epsilon \log \mathbf{P}(X^\epsilon \in G) \geq -I(G).$$

As an example, the following result by Schilder enhances a large deviation principle for a family of probability measures induced by standard Brownian motions.

Theorem 2.3. (Schilder's Thoerem) *Let* $\{X_\epsilon, \epsilon > 0\}$ *be the probability measure induced by* $W_\epsilon(\cdot) = \sqrt{\epsilon}W(\cdot)$ *on* $C_0([0,1])$ *equipped with the supremum norm, where* $W(\cdot)$ *denotes the standard Brownian motion on* $\mathbb{R}^d$*. Then* $\{X_\epsilon\}$ *satisfies in* $C_0([0,1])$*, an LDP with good rate function*

$$\tilde{I}(\phi) = \begin{cases} \frac{1}{2}\int_0^1 |\dot{\phi}(t)|^2 \mathrm{d}t, & \text{if } \phi \in H_1, \\ \\ \infty, & \text{otherwise}, \end{cases} \tag{2.4}$$

where H_1 *denotes the space of all absolutely continuous functions with square integrable derivatives equipped with the norm* $\|\phi\|_{H_1} = \left(\int_0^1 |\dot{\phi}(t)|^2 \mathrm{d}t\right)^{1/2}$.

Theorem 2.4. (Contraction Principle) *Let* E *and* $\tilde{E}$ *be Polish spaces and* $f : E \to \tilde{E}$ *be a continuous function. Suppose that* $\{X_\epsilon, \epsilon > 0\}$ *satisfies the LDP with a rate function* $I(\cdot)$*. Then* $\{\tilde{X}_\epsilon\} := \{X_\epsilon f^{-1}\}$ *also satisfies the LDP with the rate function*

$$\tilde{I}(y) = \begin{cases} \inf\{I(x) : x \in f^{-1}(y)\}, & \text{if } f^{-1}(y) \neq \phi, \\ \\ \infty, & \text{otherwise}. \end{cases} \tag{2.5}$$

The contraction principle enables to move the LDP result from one family of random variables in a probability space to another via continuous mappings.

3. The LDP with additive noise

In this section, we consider the large deviation principle for (2.1) subjected to small additive noise:

$$\begin{cases} dX(t) = b\Big(t, X(t), X(\alpha_1(t)), \dots, X(\alpha_m(t))\Big)\mathrm{d}t + \sqrt{\epsilon}\mathrm{d}W(t), \quad t \in (0,T], \\ X(0) = X_0, \end{cases}$$

with solution $X^\epsilon(t), \epsilon > 0$ and where $W(t)$ is a standard d-dimensional Brownian motion.

For arbitrarily given $g \in C\big(J;\mathbb{R}^d\big)$, we define $F(g) \in C\big(J;\mathbb{R}^d\big)$ as the unique solution to the following equation:

$$\begin{aligned} &F(g)(t) \\ &= F(g)(0) + \int_0^t b\Big(s, F(g)(s), F(g)(\alpha_1(s)), \ldots, F(g)(\alpha_m(s))\Big)ds + g(t) \end{aligned} \tag{3.1}$$

for $t \in J$ and where $C\big(J;\mathbb{R}^d\big)$ is the space of all $\mathbb{R}^d$-valued continuous functions on J.

Theorem 3.1. *$\big\{X^\epsilon(t)\big\}$ satisfies the large deviation principle in $C\big(J;\mathbb{R}^d\big)$ with the rate function*

$$I(f) := \inf\left\{\frac{1}{2}\int_0^T \|\dot{g}(t)\|^2 \mathrm{d}t; F(g) = f\right\},$$

where $g \in C\big(J;\mathbb{R}^d\big)$ is absolutely continuous, otherwise, $I(f) = \infty$.

Proof. From (3.1), we observe that for $g_1, g_2 \in C\big(J;\mathbb{R}^d\big)$ with $g_1(0) = 0$ and $g_2(0) = 0$,

$$\begin{aligned} &F(g_1)(t) - F(g_2)(t) \\ &= \int_0^t \Big[b\Big(s, F(g_1)(s), F(g_1)(\alpha_1(s)), \ldots, F(g_1)(\alpha_m(s))\Big) \\ &\quad - b\Big(s, F(g_2)(s), F(g_2)(\alpha_1(s)), \ldots, F(g_2)(\alpha_m(s))\Big)\Big]\mathrm{d}s + g_1(t) - g_2(t). \end{aligned}$$

Denote $\kappa(t) = \sup\limits_{0\le s\le t} \|F(g_1)(s) - F(g_2)(s)\|$. Since b satisfies Lipschitz condition, we have

$$\kappa(t) \le (m+1)L\int_0^t \kappa(s)\mathrm{d}s + \|g_1(t) - g_2(t)\|.$$

If we now apply Gronwall's lemma to the above obtained inequality, we get

$$\kappa(t) \le C\|g_1(t) - g_2(t)\|(1 + e^{(m+1)Lt}),$$

where C is a positive constant depending on L, m and T. Let

$$\|g_1 - g_2\|_{C\left(J;\mathbb{R}^d\right)} = \sup_{t\in J}\|g_1(t) - g_2(t)\| \le \delta.$$

Then we observe that

$$\begin{aligned} \|F(g_1) - F(g_2)\|_{C\left(J;\mathbb{R}^d\right)} &= \sup_{t\in J}\|F(g_1)(t) - F(g_2)(t)\| \\ &\le \delta C(1 + e^{(m+1)LT}), \end{aligned}$$

hence F is continuous. Noting that $X^\epsilon(t) = F(\sqrt{\epsilon}W)(t)$ and the continuity of F, the theorem follows at once from the contraction principle Theorem 2.4 and Schilder's Theorem 2.3. $\square$

4. The LDP with Multiplicative Noise

In this section, we consider the stochastic differential equation (2.1) with small multiplicative noise:

$$\begin{cases} \mathrm{d}X(t) = b\Big(t, X(t), X(\alpha_1(t)), \ldots, X(\alpha_m(t))\Big)\mathrm{d}t \\ \qquad\qquad + \sqrt{\epsilon}\sigma\Big(t, X(t), X(\alpha_1(t)), \ldots, X(\alpha_m(t))\Big)\mathrm{d}W(t), \quad t \in (0, T], \\ X(0) = X_0, \end{cases} \tag{4.1}$$

with solution $X^\epsilon, \epsilon > 0$, where $X_0 \in \mathbb{R}^d$ and $W(t)$ denotes a standard d-dimensional Brownian motion. Assume that $\sigma : J \times \mathbb{R}^d \times (\mathbb{R}^d)^m \to \mathbb{R}^d \otimes \mathbb{R}^d$ satisfies the following conditions:

$$\begin{aligned} &\|\sigma(t, x, x_1, \ldots, x_m) - \sigma(t, y, y_1, \ldots, y_m)\| \\ &\le L_2 \left(\|x - y\| + \sum_{i=1}^m \|x_i - y_i\| \right), \end{aligned} \tag{4.2}$$

$$\Big\|\sigma(t, x, x_1, x_2, \ldots, x_m)\Big\| \le K_2 \left[1 + \|x\| + \sum_{i=1}^m \|x_i\| \right] \tag{4.3}$$

for all $x, x_1, \ldots, x_m, y, y_1, \ldots, y_m \in \mathbb{R}^d$, $t \in J$ and L_2, K_2 are positive constants. Observe that (4.1) has a unique strong solution due to the conditions (2.2),(2.3),(4.2),(4.3) and the applications of Banach contraction mapping principle. Since X^ϵ is a strong solution to (4.1), it follows from the Yamada-Watanabe theorem [21] (see also [14]) that there exists a Borel-measurable function $G^\epsilon : C\big(J; \mathbb{R}^d\big) \to C\big(J; \mathbb{R}^d\big)$ such that $X^\epsilon(\cdot) = G^\epsilon\big(W(\cdot)\big)$ a.s. Let

$$\mathcal{A} = \left\{ v : v \in \mathbb{R}^d, v \text{ is} \mathcal{F}_t \text{ –predictable and } \int_0^T \|\nu(s, \omega)\|^2 \mathrm{d}s < \infty \quad \text{a.s.} \right\},$$

$$S_N = \left\{ v \in L^2\big(0, T; \mathbb{R}^d\big) : \int_0^T \|v(s)\|^2 \mathrm{d}s \le N \right\},$$

where $L^2\big(0, T; \mathbb{R}^d\big)$ is the space of all $\mathbb{R}^d$ -valued square integrable functions on J. Then S_N endowed with the weak topology is a compact Polish space. Also define

$$\mathcal{A}_N = \left\{ \nu \in \mathcal{A} : \nu(\omega) \in S_N \ \mathbf{P} - a.s \right\}.$$

We implement the theory developed by Budhiraja and Dupuis to establish the Laplace principle for the family $\{X^\epsilon : \epsilon > 0\}$ (see [14]). Indeed the Laplace principle and the large deviation principle are equivalent when the underlying space is Polish.

Theorem 4.1. *Suppose there exists a measurable map* $G^0 : C\big(J;\mathbb{R}^d\big) \to C\big(J;\mathbb{R}^d\big)$ *such that the following two conditions hold:*

(i) *Let* $\{\nu^\epsilon : \epsilon > 0\} \subset \mathcal{A}_N$ *for some* $N < \infty$. *If* ν^ϵ *converge to* ν *in distribution as* S_N*-valued random elements, then*
$$G^\epsilon\left(W(\cdot) + \frac{1}{\sqrt{\epsilon}}\int_0^\cdot \nu^\epsilon(s)\mathrm{d}s\right) \to G^0\left(\int_0^\cdot \nu(s)\mathrm{d}s\right)$$
in distribution as $\epsilon \to 0$.

(ii) *For each* $N < \infty$, *the set* $K_N = \left\{G^0\left(\int_0^\cdot v(s)\mathrm{d}s\right) : v \in S_N\right\}$ *is a compact subset of* $C\big(J;\mathbb{R}^d\big)$.

Then the family $\{X^\epsilon, \epsilon > 0\}$ *satisfies the Laplace principle in* $C(J;\mathbb{R}^d)$ *with the rate function* I *given by*
$$I(h) = \inf\left\{\frac{1}{2}\int_0^T \|v(t)\|^2\mathrm{d}t; X_v = h \text{ and } v \in L^2(0,T;\mathbb{R}^d)\right\} \tag{4.4}$$
for each $h \in C(J;\mathbb{R}^d)$ *with the convention that the infimum of an empty set is infinity.*

Hence establishing a Laplace principle is now simplified to satisfing the assumptions (i) and (ii) for our system. We first introduce the skeleton equation associated to (4.1) with a control term $v \in L^2\big(0,T;\mathbb{R}^d\big)$:
$$\begin{cases} dz_v(t) = b\Big(t, z_v(t), z_v(\alpha_1(t)), \ldots, z_v(\alpha_m(t))\Big)\mathrm{d}t \\ \qquad\qquad + \sigma\Big(t, z_v(t), z_v(\alpha_1(t)), \ldots, z_v(\alpha_m(t))\Big)v(t)\mathrm{d}t, \quad t \in (0,T]; \\ z_v(0) = X_0 \end{cases} \tag{4.5}$$
with solution z_v. The main theorem in this section is the following.

Theorem 4.2. *The family* $\{X^\epsilon(t)\}$ *of* (4.1) *satisfies the large deviation principle* (*equivalently, Laplace principle*) *in* $C\big(J;\mathbb{R}^d\big)$ *with good rate function*
$$I(f) := \inf\left\{\frac{1}{2}\int_0^T \|v(t)\|^2\mathrm{d}t; z_v = f\right\}, \tag{4.6}$$
where $v \in L^2\big(0,T;\mathbb{R}^d\big)$, *otherwise,* $I(f) = \infty$.

For a proof, it suffices to show that (i) and (ii) are true to our system and is done in the following two lemmas:

Lemma 4.3. (Compactness) *Define* $G^0 : C\big(J;\mathbb{R}^d\big) \to C\big(J;\mathbb{R}^d\big)$ *by*

$$G^0(g) := \begin{cases} z_v, & \text{if } g = \int_0^{\cdot} v(s)\mathrm{d}s \quad \text{for some} \quad v \in L^2\big(0,T;\mathbb{R}^d\big), \\ 0\ , & \text{otherwise.} \end{cases}$$

Then, for each $N < \infty$*, the set*

$$K_N = \left\{ G^0\left(\int_0^{\cdot} v(s)\mathrm{d}s \right) : v \in S_N \right\}$$

is a compact subset of $C\big(J;\mathbb{R}^d\big)$.

Proof. We first prove that the map

$$\begin{aligned} z_v(t) = X_0 &+ \int_0^t b\Big(s, z_v(s), z_v(\alpha_1(s)), \ldots, z_v(\alpha_m(s))\Big)\mathrm{d}s \\ &+ \int_0^t \sigma\Big(s, z_v(s), z_v(\alpha_1(s)), \ldots, z_v(\alpha_m(s))\Big)v(s)\mathrm{d}s \end{aligned} \tag{4.7}$$

is continuous from S_N to $C\big(J;\mathbb{R}^d\big)$. Consider a sequence $\{v_n\} \in S_N$ such that $v^n \to v$ weakly in S_N as $n \to \infty$. From equation (4.7), we have

$$\begin{aligned} &z_{v_n}(t) - z_v(t) \\ &= \int_0^t \Big[b\Big(s, z_{v_n}(s), z_{v_n}(\alpha_1(s)), \ldots, z_{v_n}(\alpha_m(s))\Big) \\ &\qquad\qquad - b\Big(s, z_v(s), z_v(\alpha_1(s)), \ldots, z_v(\alpha_m(s))\Big)\Big]\,\mathrm{d}s \\ &\quad + \int_0^t \Big[\sigma\Big(s, z_{v_n}(s), z_{v_n}(\alpha_1(s)), \ldots, z_{v_n}(\alpha_m(s))\Big) \\ &\qquad\qquad - \sigma\Big(s, z_v(s), z_v(\alpha_1(s)), \ldots, z_v(\alpha_m(s))\Big)\Big]\, v_n(s)\mathrm{d}s \\ &\quad + \int_0^t \sigma\Big(s, z_v(s), z_v(\alpha_1(s)), \ldots, z_v(\alpha_m(s))\Big)\Big(v_n(s) - v(s)\Big)\mathrm{d}s. \end{aligned}$$

Let

$$\zeta^n(t) = \int_0^t \sigma\big(s, z_v(s), z_v(\alpha_1(s)), \ldots, z_v(\alpha_m(s))\big)\big(v_n(s) - v(s)\big)\mathrm{d}s.$$

Since σ satisfies the linear growth property, we have

$$\begin{aligned}
&\sup_{t\in J}\|\zeta^n(t)\| \\
&\leq \int_0^T \Big\|\sigma\Big(s, z_v(s), z_v(\alpha_1(s)), \ldots, z_v(\alpha_m(s))\Big)\Big(v_n(s)-v(s)\Big)\Big\|\,\mathrm{d}s \\
&\leq \left(\int_0^T \Big\|\sigma\Big(s, z_v(s), z_v(\alpha_1(s)), \ldots, z_v(\alpha_m(s))\Big)\Big\|^2 \mathrm{d}s\right)^{\frac{1}{2}} \left(\int_0^T \|v_n(s)-v(s)\|^2\,\mathrm{d}s\right)^{\frac{1}{2}} \\
&\leq C < \infty.
\end{aligned}$$

Since $v_n \to v$ weakly in $L^2\big(0,T;\mathbb{R}^d\big)$, by Arzéla-Ascoli theorem we could conclude that $\zeta^n \to 0$ in $C\big(J;\mathbb{R}^d\big)$. This implies

$$\lim_{n\to\infty}\sup_{t\in J}\|\zeta^n(t)\| = 0. \tag{4.8}$$

Set $\kappa^n(t) = \sup_{0\leq s\leq t}\|z_{v_n}(s) - z_v(s)\|$, using the Lipschitz continuity of b and σ, we arrive at

$$\kappa^n(t) \leq \sup_{0\leq s\leq t}\|\zeta^n(s)\| + (m+1)\int_0^t \kappa^n(s)\big(L + L_2\|v_n(s)\|\big)\mathrm{d}s.$$

Now making use of Gronwall's inequality,

$$\kappa^n(t) \leq C_1 \sup_{0\leq s\leq t}\|\zeta^n(s)\|\,(1+e^{Ct}),$$

where C, C_1 are constants depending on L, L_2, m, N and T. Hence

$$\begin{aligned}
\|z_{v_n}(t) - z_v(t)\|_{C\left(J;\mathbb{R}^d\right)} &= \sup_{0\leq t\leq T}\|z_{v_n}(t) - z_v(t)\| \\
&\leq C_1 \sup_{t\in J}\|\zeta^n(t)\|\,(1+e^{CT})
\end{aligned}$$

and so $z_{v_n} \to z_v$ in $C\big(J;\mathbb{R}^d\big)$ by virtue of (4.8). Also since the space S_N is compact, it follows that the set $K_N = \big\{G^0\big(\int_0^{\cdot} v(s)\mathrm{d}s\big) :\ v \in S_N\big\}$ for $N < \infty$ is compact. $\square$

Next we intend to verify the weak convergence condition (i) of Theorem 4.1. For this we first consider the stochastic differential equation

$$\begin{aligned}
&X^\epsilon_{\nu^\epsilon}(t) \\
&= X_0 + \int_0^t \Big[b\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big) \\
&\quad + \sigma\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big)\nu^\epsilon(s)\Big]\,\mathrm{d}s \\
&\quad + \sqrt{\epsilon}\int_0^t \sigma\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big)\mathrm{d}W(s), \quad t \in J.
\end{aligned} \tag{4.9}$$

The existence of solutions for the above equation (4.9) follows easily from the Girsanov's theorem (see [19] for a proof of similar kind). We now move on to the weak convergence result.

Lemma 4.4. (Weak Convergence) *Let $\{\nu^\epsilon : \epsilon > 0\} \subset \mathcal{A}_N$ for some $N < \infty$. Assume ν^ϵ converges to ν in distribution as S_N-valued random elements, then*

$$G^\epsilon\left(W(\cdot) + \frac{1}{\sqrt{\epsilon}}\int_0^\cdot \nu^\epsilon(s)\mathrm{d}s\right) \to G^0\left(\int_0^\cdot \nu(s)\mathrm{d}s\right)$$

in distribution as $\epsilon \to 0$.

Proof. Without loss of generality, assume that $\epsilon \leq \frac{1}{4}$. Applying Itö's formula, we have

$$\begin{aligned}
&d\,\|X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\|^2 \\
&= 2\Big(X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\Big)\Big[b\Big(t, X^\epsilon_{\nu^\epsilon}(t), X^\epsilon_{\nu^\epsilon}(\alpha_1(t)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(t))\Big) \\
&\quad - b\Big(t, z^\nu(t), z_\nu(\alpha_1(t)), \ldots, z_\nu(\alpha_m(t))\Big)\Big]\mathrm{d}t \\
&\quad + 2\Big(X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\Big)\Big[\sigma\Big(t, X^\epsilon_{\nu^\epsilon}(t), X^\epsilon_{\nu^\epsilon}(\alpha_1(t)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(t))\Big)\nu^\epsilon(t) \\
&\quad - \sigma\Big(t, z^\nu(t), z_\nu(\alpha_1(t)), \ldots, z_\nu(\alpha_m(t))\Big)\nu(t)\Big]\mathrm{d}t \\
&\quad + \epsilon\|\sigma\Big(t, X^\epsilon_{\nu^\epsilon}(t), X^\epsilon_{\nu^\epsilon}(\alpha_1(t)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(t))\Big)\|^2\mathrm{d}t \\
&\quad + 2\sqrt{\epsilon}\Big(X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\Big)\sigma\Big(t, X^\epsilon_{\nu^\epsilon}(t), X^\epsilon_{\nu^\epsilon}(\alpha_1(t)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(t))\Big)\mathrm{d}W(t).
\end{aligned}$$

By using Young's inequality, the above inequality can be transformed as

$$\begin{aligned}
&d\,\|X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\|^2 \\
&\leq 2\Big(X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\Big)\Big[b\Big(t, X^\epsilon_{\nu^\epsilon}(t), X^\epsilon_{\nu^\epsilon}(\alpha_1(t)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(t))\Big) \\
&\quad - b\Big(t, z^\nu(t), z_\nu(\alpha_1(t)), \ldots, z_\nu(\alpha_m(t))\Big)\Big]\mathrm{d}t \\
&\quad + 2\Big(X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\Big)\sigma\Big(t, z^\nu(t), z_\nu(\alpha_1(t)), \ldots, z_\nu(\alpha_m(t))\Big)\Big(\nu^\epsilon(t) - \nu(t)\Big)\mathrm{d}t \\
&\quad + \Big\|\sigma\Big(t, X^\epsilon_{\nu^\epsilon}(t), X^\epsilon_{\nu^\epsilon}(\alpha_1(t)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(t))\Big) \\
&\quad -\sigma\Big(t, z^\nu(t), z_\nu(\alpha_1(t)), \ldots, z_\nu(\alpha_m(t))\Big)\Big\|^2\mathrm{d}t + |X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)|^2\,\|\nu^\epsilon(t)\|^2\mathrm{d}t \\
&\quad + \epsilon\,\Big\|\sigma\Big(t, z^\nu(t), z_\nu(\alpha_1(t)), \ldots, z_\nu(\alpha_m(t))\Big)\Big\|^2\mathrm{d}t \\
&\quad + 2\sqrt{\epsilon}\Big(X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\Big)\sigma\big(t, X^\epsilon_{\nu^\epsilon}(t), X^\epsilon_{\nu^\epsilon}(\alpha_1(t)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(t))\big)\mathrm{d}W(t).
\end{aligned}$$

As before, we define

$$\zeta^\epsilon(t) = \int_0^t \sigma\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\Big(\nu^\epsilon(s) - \nu(s)\Big)\mathrm{d}s.$$

Also define

$$f(u) = \int_0^\cdot \sigma\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)u(s)\mathrm{d}s.$$

By the linear growth of σ, we see that the map $f : S_N \to C\big(J; \mathbb{R}^d\big)$ is a bounded continuous function. Note that S_N is endowed with the weak topology and ν^ϵ converge to ν in distribution as S_N-valued random elements. Then $\zeta^\epsilon \to 0$ in distribution as $\epsilon \to 0$ follows immediately by Theorem A.3.6 in [6]. By virtue of the Itô's formula again

$$\begin{aligned}
&\int_0^t \Big(X^\epsilon_{\nu^\epsilon}(s) - z^\nu(s)\Big)\sigma\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\Big(\nu^\epsilon(s) - \nu(s)\Big)\mathrm{d}s \\
&= \Big(X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\Big)\zeta^\epsilon(t) - \int_0^t \Big[b\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big) \\
&\quad - b\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\Big]\zeta^\epsilon(s)\mathrm{d}s \\
&\quad - \int_0^t \Big[\sigma\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big)\nu^\epsilon(s) \\
&\quad - \sigma\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\nu(s)\Big]\zeta^\epsilon(s)\mathrm{d}s \\
&\quad - \sqrt{\epsilon}\int_0^t \sigma\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big)\zeta^\epsilon(s)\mathrm{d}W(s) \\
&=: I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

Applying Young's inequality, we get

$$I_1 \le \frac{1}{4}\left\|X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\right\|^2 + \left\|\zeta^\epsilon(t)\right\|^2.$$

It is easy to see that

$$\begin{aligned}
I_2 \le \sup_{s\in[0,t]}\|\zeta^\epsilon(s)\| \int_0^t \Big\| & b\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big) \\
& - b\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\Big\|\,\mathrm{d}s.
\end{aligned} \tag{4.10}$$

By using the Hölder's inequality, we have

$$
\begin{aligned}
I_3 \leq & \sup_{s\in[0,t]} \|\zeta^\epsilon(s)\| \int_0^t \Big\|\sigma\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big)\nu^\epsilon(s) \\
& \qquad - \sigma\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\nu(s)\Big\| \mathrm{d}s \\
\leq & \sqrt{N} \sup_{s\in[0,t]} \|\zeta^\epsilon(s)\| \left\{ \left(\int_0^t \left\|\sigma\big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\big)\right\|^2 \mathrm{d}s \right)^{\frac{1}{2}} \right. \\
& \qquad \left. + \left(\int_0^t \left\|\sigma\big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\big)\right\|^2 \mathrm{d}s \right)^{\frac{1}{2}} \right\} \\
\leq & \sup_{s\in[0,t]} |\zeta^\epsilon(s)| \left\{ \sqrt{N} \left(\int_0^t \left\|\sigma\big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\big)\right\|^2 \mathrm{d}s \right)^{\frac{1}{2}} + c_1 \right\},
\end{aligned}
$$

where in the last step above, we use the fact that

$$
\begin{aligned}
& \int_0^t \left\|\sigma\big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\big)\right\|^2 \mathrm{d}s \\
& \leq K_2 \int_0^t \Big(1 + \|z^\nu(s)\|^2 + \sum_{i=1}^m \|z_\nu(\alpha_i(s))\|^2\Big) \mathrm{d}s < \infty.
\end{aligned}
$$

Combining the preceding three inequalities yields

$$
\begin{aligned}
& \|X^\epsilon_{\nu^\epsilon}(t) - z^\nu(t)\|^2 \\
& \leq 4 \int_0^t \|X^\epsilon_{\nu^\epsilon}(s) - z^\nu(s)\| \Big\| b\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big) \\
& \qquad - b\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\Big\| \mathrm{d}s \\
& \quad + 4c_1 \sup_{s\in[0,t]} \|\zeta^\epsilon(s)\| + 4 \sup_{s\in[0,t]} \|\zeta^\epsilon(s)\| \int_0^t \Big\| b\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big) \\
& \qquad - b\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\Big\| \mathrm{d}s \\
& \quad + 4\sqrt{N} \sup_{s\in[0,t]} |\zeta^\epsilon(s)| \left\{ \left(\int_0^t \left\|\sigma\big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\big)\right\|^2 \mathrm{d}s \right)^{\frac{1}{2}} \right\} \\
& \quad + 4 \sup_{s\in[0,t]} \|\zeta^\epsilon(s)\|^2 + 2 \int_0^t \Big\| \sigma\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big) \\
& \qquad - \sigma\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\Big\|^2 \mathrm{d}s \\
& \quad + 2 \int_0^t \|X^\epsilon_{\nu^\epsilon}(s) - z^\nu(s)\|^2 \, \|\nu^\epsilon(s)\|^2 \mathrm{d}s \\
& \quad + 2\epsilon \int_0^t \Big\| \sigma\Big(s, z^\nu(s), z_\nu(\alpha_1(s)), \ldots, z_\nu(\alpha_m(s))\Big)\Big\|^2 \mathrm{d}s \\
& \quad + 4\sqrt{\epsilon} \int_0^t \big(X^\epsilon_{\nu^\epsilon}(s) - z^\nu(s) - \zeta^\epsilon(s)\big) \sigma\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)), \ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big) \mathrm{d}W(s).
\end{aligned}
$$

Similarly, set $\kappa^\epsilon(t) = \sup_{0\le s\le t} \|X^\epsilon_{\nu^\epsilon}(s) - z^\nu(s)\|^2$. Then by the assumptions on $b(\cdot)$ and $\sigma(\cdot)$, the inequality can be continued as

$$\begin{aligned}
&\kappa^\epsilon(t)\\
&\le C\int_0^t \big(1+\|\nu^\epsilon(s)\|^2\big)\kappa^\epsilon(s)\mathrm{d}s + 4c_1 \sup_{s\in[0,t]}\|\zeta^\epsilon(s)\|\\
&\quad + c_2 \sup_{s\in[0,t]}\|\zeta^\epsilon(s)\| \left\{1+\int_0^t \left(\|X^\epsilon_{\nu^\epsilon}(s)\| + \sum_{i=1}^m \|X^\epsilon_{\nu^\epsilon}(\alpha_i(s))\| \right)\mathrm{d}s\right\}\\
&\quad + c_3 \sup_{s\in[0,t]}\|\zeta^\epsilon(s)\| \left\{1+\left[\int_0^t \left(\left\|X^\epsilon_{\nu^\epsilon}(s)\right\|^2 + \sum_{i=1}^m \|X^\epsilon_{\nu^\epsilon}(\alpha_i(s))\|^2 \right)\mathrm{d}s\right]^{\frac{1}{2}}\right\}\\
&\quad + c_4\left(\epsilon + \sup_{s\in[0,t]}\|\zeta^\epsilon(s)\|^2\right) + 4\sqrt{\epsilon}\sup_{\theta\in[0,t]}\left|\int_0^\theta \big(X^\epsilon_{\nu^\epsilon}(s) - z^\nu(s) - \zeta^\epsilon(s)\big)\right.\\
&\quad \left.\times\, \sigma\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)),\ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big)\mathrm{d}W(s)\right|.
\end{aligned}$$

Now, Gronwall's Lemma implies

$$\begin{aligned}
\kappa^\epsilon(t) \le C\Bigg\{ &\sup_{s\in[0,t]}\|\zeta^\epsilon(s)\| \Bigg[1+\int_0^t \left(\|X^\epsilon_{\nu^\epsilon}(s)\| + \sum_{i=1}^m \|X^\epsilon_{\nu^\epsilon}(\alpha_i(s))\| \right)\mathrm{d}s\\
&+\left(\int_0^t \left(\|X^\epsilon_{\nu^\epsilon}(s)\|^2 + \sum_{i=1}^m \|X^\epsilon_{\nu^\epsilon}(\alpha_i(s))\|^2 \right)\mathrm{d}s\right)^{\frac{1}{2}}\Bigg] + \epsilon\\
&+ \sup_{s\in[0,t]}\|\zeta^\epsilon(s)\|^2 + \sqrt{\epsilon}\sup_{\theta\in[0,t]}\Bigg\|\int_0^\theta \big(X^\epsilon_{\nu^\epsilon}(s) - z^\nu(s) - \zeta^\epsilon(s)\big)\\
&\times\, \sigma\Big(s, X^\epsilon_{\nu^\epsilon}(s), X^\epsilon_{\nu^\epsilon}(\alpha_1(s)),\ldots, X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big)\mathrm{d}W(s)\Bigg\|\Bigg\}. \qquad (4.11)
\end{aligned}$$

The lemma is proved as soon as we show that $\kappa^\epsilon(T) \to 0$ in distribution as $\epsilon \to 0$. To this end, we first define the stopping time

$$\tau^{M,\epsilon} := \inf\left\{ t \le T : \sup_{s\in[0,t]} \|X^\epsilon_{\nu^\epsilon}(s)\|^2 > M \right\},$$

where M is some constant large enough. The Burkhölder-Davis-Gundy inequality allows us to bound the expectation of the last term on the right side of (4.11) by

$$\begin{aligned}
&\mathbf{E}\sup_{\theta\in[0,\tau^{M,\epsilon}]}\left\|\int_0^\theta \big(X^\epsilon_{\nu^\epsilon}(s)-z^\nu(s)-\zeta^\epsilon(s)\big)\sigma\Big(s,X^\epsilon_{\nu^\epsilon}(s),X^\epsilon_{\nu^\epsilon}(\alpha_1(s)),\ldots,X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big)\mathrm{d}W(s)\right\| \\
&\le C\mathbf{E}\left\{\int_0^{\tau^{M,\epsilon}} \left\|X^\epsilon_{\nu^\epsilon}(s)-z^\nu(s)-\zeta^\epsilon(s)\right\|^2\left\|\sigma\Big(s,X^\epsilon_{\nu^\epsilon}(s),X^\epsilon_{\nu^\epsilon}(\alpha_1(s)),\ldots,X^\epsilon_{\nu^\epsilon}(\alpha_m(s))\Big)\right\|^2\mathrm{d}s\right\}^{\frac{1}{2}} \\
&\le C\mathbf{E}\left\{\sup_{s\in[0,\tau^{M,\epsilon}]}\left\|X^\epsilon_{\nu^\epsilon}(s)-z^\nu(s)-\zeta^\epsilon(s)\right\|^2\int_0^{\tau^{M,\epsilon}}\left(1+\|X^\epsilon_{\nu^\epsilon}(s)\|^2+\sum_{i=1}^m\|X^\epsilon_{\nu^\epsilon}(\alpha_i(s))\|^2\right)\mathrm{d}s\right\} \\
&\le C. \qquad (4.12)
\end{aligned}$$

Hence the stochastic integral term on the right hand side of (4.11) tends to 0 as $\epsilon\to 0$. The following inequality is a consequence of It$\hat{o}$'s formula, H$\ddot{o}$lder's inequality, Burkh$\ddot{o}$lder-Davis-Gundy's inequality, linear growth of b and σ, Gronwall's lemma and is standard:

$$\sup_{\epsilon\in[0,\frac{1}{4}]}\mathbf{E}\left(\sup_{s\in J}\|X^\epsilon_{\nu^\epsilon}(s)\|^2\right)<\infty.$$

Due to this inequality, it follows that the Chebycheff's inequality is applicable here, yielding that there exists a suitable constant C such that

$$\liminf_{\epsilon\to 0}\mathbf{P}\left\{\tau^{M,\epsilon}=T\right\}\ge 1-\frac{C}{M}. \qquad (4.13)$$

Combining with (4.11)-(4.13) and recall that $\zeta^\epsilon\to 0$ in distribution as $\epsilon\to 0$, we may immediately get the result $\kappa^\epsilon(T)\to 0$ in distribution as $\epsilon\to 0$, and the lemma is established. □

Acknowledgments: The first author(M. Suvinthra) would like to thank the Department of Science and Technology, Government of India, New Delhi for their financial support under the INSPIRE Fellowship Scheme.

References

[1] K. Balachandran and S. Ilamaran, *An existence theorem for a Volterra integral equation with deviating arguments*, J. Appl. Math. Stoch. Anal., **3** (1990), 155–162.

[2] J. Banas, *An existence theorem for nonlinear Volterra integral equation with deviating argument*, Rend. Circ. Mat. Palermo, **35** (1986), 82–89.

[3] M. Boue and P. Dupuis, *A variational representation for certain functionals of Brownian motion*, Ann. Probab., **26** (1998), 1641–1659.

[4] A. Budhiraja and P. Dupuis, *A variational representation for positive functionals of infinite dimensional Brownian motion*, Probab. Math. Statist., **20** (2000), 39–61.

[5] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Springer, New York (2007).

[6] P. Dupuis and R.S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*, Wiley-Interscience, New York (1997).

[7] L.E. Elsgolts and S.B. Norkin, *Introduction to the Theory and Application of Differential Equations with Deviating Arguments*, Academic Press, New York (1973).

[8] W.H. Fleming, *A stochastic control approach to some large deviations problems*, In: Dolcetta, C., Fleming, W.H., Zolezzi, T. (eds.) Recent Mathematical Methods in Dynamic Programming, in: Springer Lecture Notes in Math., **1119** (1985), 52–66.

[9] M.I. Freidlin and A.D. Wentzell, *On small random perturbations of dynamical systems*, Russian Math. Surveys, **25** (1970), 1–55.

[10] J.K. Hale, *History of delay equations*, In Arino, O., Hbid, M.L., Dads, E.A. (eds.) Delay Differential Equations and Applications, pp. 1-28. Springer, Netherlands (2006).

[11] A.F. Ivanov, Y.I. Kazmerchuk and A.V. Swishchuk, *Theory, stochastic stability and applications of stochastic delay differential equations: A survey of results*, Differential Equations and Dynamical Systems, **11** (2003), 55–115.

[12] S. Karthikeyan and K. Balachandran, *On controllability for a class of stochastic impulsive systems with delays in control*, Int. J. Syst. Sci., **44** (2013), 67–76.

[13] W. Liu, *Large deviations for stochastic evolution equations with small multiplicative noise*, Appl. Math. Optim., **61** (2010), 27–56.

[14] C. Mo and J. Luo, *Large deviations for stochastic differential delay equations*, Nonlinear Anal., **80** (2013), 202–210.

[15] S.A. Mohammed and T.S. Zhang, *Large deviations for stochastic systems with memory*, Disc. Contin. Dyn. Syst. Ser. B, **6** (2006), 881–893.

[16] A. Rathinasamy and K. Balachandran, *Existence and stability of solutions of neutral hybrid stochastic infinite delay differential equations with Poisson jumps*, Nonlinear Funct. Anal. Appl., **16** (2011), 123–139.

[17] M. Rockner, T. Zhang and X. Zhang, *Large deviations for stochastic tamed 3D Navier-Stokes equations*, Appl. Math. Optim., **61** (2010), 267–285.

[18] M. Scheutzow, *Qualitative behavior of stochastic delay equations with a bounded memory*, Stochastics, **12** (1984), 41–80.

[19] S.S. Sritharan and P. Sundar, *Large deviations for two dimensional Navier-Stokes equations with multiplicative noise*, Stoch. Proc. Appl., **116** (2006), 1636–1659.

[20] S.R.S. Varadhan, *Asymptotic probabilities and differential equations*, Comm. Pure Appl. Math., **19** (1966), 261–286.

[21] T. Yamada and S. Watanabe, *On the uniqueness of solutions of stochastic differential equations*, J. Math. Kyoto Univ., **11** (1971), 155–167.