

**A DIFFERENTIAL DELAY EQUATION WITH  
WIDEBAND NOISE PERTURBATIONS**

By

**G. Yin**

and

**K.M. Ramachandran**

**IMA Preprint Series # 444**

August 1988

# A DIFFERENTIAL DELAY EQUATION WITH WIDEBAND NOISE PERTURBATIONS

G. Yin\* and K.M Ramachandran\*\*

*Abstract.* A differential delay equation with a small parameter and random noise perturbations is considered in this paper. Asymptotic properties are developed. The martingale averaging techniques are adopted to treat our problem, and the method of weak convergence is employed. The random fluctuation is assumed to be of the wideband noise type, which is quite realistic for various applications. It is shown that as  $\epsilon \rightarrow 0$ , the random process converges weakly to a diffusion process which satisfies a stochastic differential delay equation.

*Keywords.* differential delay equation, wideband noise, weak convergence, martingale problem.

*AMS (MOS) subject classifications.* 60F5, 60F17, 60J60.

## 1. Introduction

In this paper, we investigate the problem of a stochastic differential delay equation with wideband noise perturbations. Let  $x^\epsilon(\cdot), \xi^\epsilon(\cdot) \in \mathbb{R}^r$ , consider the following equations:

$$\dot{x}^\epsilon(t) = a(x^\epsilon(t), x^\epsilon(t - \delta), \xi^\epsilon(t)) + \frac{1}{\epsilon} b(x^\epsilon(t), x^\epsilon(t - \delta), \xi^\epsilon(t)), \quad t > 0 \quad (1.1)$$

$$x^\epsilon(t) = x_0(t) \text{ for } t \in [-\delta, 0]. \quad (1.2)$$

Where  $\epsilon > 0$  is a small parameter,  $x_0(t)$  is a given continuous deterministic function on  $[-\delta, 0]$ ,  $\delta > 0$  represents the delay and  $\xi^\epsilon(\cdot)$  is a wideband noise process. In case  $\delta = 0$ , (1.1) (1.2) reduce to the case of an ordinary differential equation under random perturbations.

In many applications, one assumes that the system under consideration is governed by the principle of causality. One thus considers the model of either an ordinary differential equation or a partial differential equation, or such equations with random perturbations. However, under close scrutiny, it becomes clear that such model is often only a first approximation to the real world problem. More often than not, time delay comes into play. Thus, differential delay equations and such equations with random perturbations arise naturally in many applied field. As a result, to study the behavior of such equations is not only interesting from theoretical considerations, but also necessary from an application point of view.

In the present paper, our main goal is to study the asymptotic properties of equation (1.1). Under suitable conditions, we show that  $x^\epsilon(\cdot)$  converges weakly to a Gauss-Markov diffusion process which satisfies an appropriate stochastic differential delay equation.

---

\* Department of Mathematics, Wayne State University, Detroit, MI 48202. This work was supported in part by Wayne State University under the Wayne State University Research Award.

\*\* Department of Mathematics, University of South Florida, Tampa, FL 33620.

In [10], White developed a theorem for the weak convergence of a process  $\{x^\epsilon(\cdot)\}$  satisfying the following differential delay equation

$$\dot{x}^\epsilon(t) = a(x^\epsilon(t), x^\epsilon(t - \delta), \xi^\epsilon(t)). \quad (1.3)$$

He proved that as  $\epsilon \rightarrow 0$ ,  $x^\epsilon(\cdot) \Rightarrow \bar{x}(\cdot)$ , (" $\Rightarrow$ " denotes weak convergence), such that  $\bar{x}(\cdot)$  satisfies a deterministic differential delay equation and the random fluctuation

$$\frac{1}{\sqrt{\epsilon}}(x^\epsilon(\cdot) - \bar{x}(\cdot)) \Rightarrow \text{a diffusion process.}$$

His approach is an extension of [5]. Despite the interesting results, enormous detailed estimates and calculations are needed.

(1.1) is a generalization of (1.3). When  $b(\cdot) = 0$ , (1.1) formally reduces to (1.3). Although (1.1) (1.3) have certain similarity, the limit behavior is somewhat different. In contrary to the limit of (1.3), the limiting process  $x(\cdot)$  in our formulation is no longer a solution of a deterministic differential delay equation, but rather a solution of a stochastic differential delay equation.

As we mentioned before, we assume that the noise process in (1.1) is wideband. Roughly speaking, a wideband noise is one such that it approximates the "white noise". Let  $R^\epsilon(s)$  be the correlation of  $\xi^\epsilon(\cdot)$ , i.e.,  $R^\epsilon(s) = E\xi^\epsilon(t+s)\xi^{\epsilon'}(t)$ . Let  $S^\epsilon(\cdot)$  be the power spectral density (we assume it exists),

$$S^\epsilon(\mu) = \int_{-\infty}^{\infty} e^{i\mu s} R^\epsilon(s) ds.$$

If  $\xi^\epsilon(\cdot)$  is wideband, then  $S^\epsilon(\mu)$  is effectively band limited, i.e.,  $S^\epsilon(\mu) = 0$  for  $\mu$  outside of a certain interval, and the length of this interval is wide enough. In fact, we shall assume  $\xi^\epsilon(t) = \xi(\frac{t}{\epsilon^2})$  throughout the paper. As a consequence, the spectral density is  $S^\epsilon(\mu) = S(\epsilon^2\mu)$ .  $S^\epsilon(\mu) = 0$  for all  $\mu$  satisfying  $|\mu| > \epsilon^{-2}\mu_0$ , and for some  $\mu_0 > 0$ . As  $\epsilon$  is getting smaller and smaller, the bandwidth is getting wider and wider. As a result, the bandwidth of  $S^\epsilon(\mu)$  tends to infinity as  $\epsilon \rightarrow 0$ , and the spectral density tends to that of the white or Gaussian noise as  $\epsilon \rightarrow 0$ .

In lieu of working on the  $C^r[0, \infty]$  space as in [10], we shall carry out the analysis in  $D^r[0, \infty)$ . The recent result of weak convergence and averaging techniques in [7-8] will be employed. Instead of using the traditional approach [5], the idea of martingale problem formulation will be utilized (cf. [8-9]). As a result, the proof is shorter and more illuminating than that of [10].

The remainder of this paper is organized as follows. Problem formulation and statement of the limit theorems for a fixed delay is given next. The proof is then presented in section 3. Extension of the limit theorem for random delays is indicated in section 4.

For general theory of weak convergence, the readers are referred to [1] [3] [8]. Terms such as weak convergence, Skorohod topology, Skorohod imbedding etc. will be used without any specific mention.

To proceed, a word about the notations is in order. In the following,  $z'$  stands for the transpose of  $z$  (a vector or a matrix);  $a_x$  denotes the gradient of  $a$ , similarly for the second derivatives.

## 2. Equation with a fixed delay

Following [3] [6], we first define the notion of " $p$ -lim", and an operator  $\hat{A}^\epsilon$  as follows. Let  $M$  be the set of real valued measurable functions of  $(\omega, t)$  that are non-zero only on a bounded  $t$ -interval. Let  $\mathcal{F}_t^\epsilon$  be the  $\sigma$ -algebra generated by  $\{x^\epsilon(s); s \leq t\}$ , and

$$\bar{M}^\epsilon = \{f \in M; \sup_t E|f(t)| < \infty \text{ and } f(t) \text{ is } \mathcal{F}_t^\epsilon \text{ measurable}\}.$$

Let  $f(\cdot), f^\Delta(\cdot) \in \bar{M}^\epsilon$ , for each  $\Delta > 0$ .  $f = p\text{-}\lim_{\Delta} f^\Delta$  iff  $\sup_{t, \Delta} E|f^\Delta(t) - f(t)| = 0$ , for each  $t$ . We say that  $f(\cdot) \in \mathcal{D}(\hat{A}^\epsilon)$ , the domain of  $\hat{A}^\epsilon$ , and  $\hat{A}^\epsilon f = g$ , if  $f, g \in \bar{M}^\epsilon$  and

$$p\text{-}\lim_{\Delta \rightarrow 0} \left( \frac{E_t^\epsilon f(t + \Delta) - f(t)}{\Delta} - g(t) \right) = 0.$$

Where  $E_t^\epsilon$  is a short hand notation for the conditional expectation on the  $\sigma$ -algebra  $\mathcal{F}_t^\epsilon$ .

In what follows, we also need the notion of  $N$ -truncation. For each  $N > 0$ , let  $S_N = \{x; |x| \leq N\}$  be the  $N$ -ball and let  $x^{\epsilon, N}(0) = x^\epsilon(0)$ ,  $x^{\epsilon, N}(t) = x^\epsilon(t)$  up until the first exit from  $S_N$ , and

$$\lim_{m \rightarrow \infty} \limsup_{\epsilon} P(\sup_{t \leq T} |x^{\epsilon, N}(t)| \geq m) = 0 \text{ for each } T < \infty.$$

$x^{\epsilon, N}(t)$  is said to be the  $N$ -truncation of  $x^{\epsilon, N}(\cdot)$ .

In lieu of (1.1), we consider the following truncated version of the equation.

$$\dot{x}^{\epsilon, N}(t) = a^N(x^{\epsilon, N}(t), x^{\epsilon, N}(t - \delta), \xi^\epsilon(t)) + \frac{1}{\epsilon} b^N(x^{\epsilon, N}(t), x^{\epsilon, N}(t - \delta), \xi^\epsilon(t)) \quad (2.1)$$

where

$$\begin{aligned} & a^N(x^{\epsilon, N}(t), x^{\epsilon, N}(t - \delta), \xi^\epsilon(t)) \\ &= a(x^{\epsilon, N}(t), x^{\epsilon, N}(t - \delta), \xi^\epsilon(t)) q^{\epsilon, N}(x^{\epsilon, N}(t), x^{\epsilon, N}(t - \delta)) \\ & b^N(x^{\epsilon, N}(t), x^{\epsilon, N}(t - \delta), \xi^\epsilon(t)) \\ &= b(x^{\epsilon, N}(t), x^{\epsilon, N}(t - \delta), \xi^\epsilon(t)) q^{\epsilon, N}(x^{\epsilon, N}(t), x^{\epsilon, N}(t - \delta)) \\ & q^{\epsilon, N}(y, z) = \begin{cases} 1; & \text{when } |y| \leq N \text{ or } |z| \leq N \\ 0; & \text{otherwise.} \end{cases} \end{aligned}$$

We define  $x_\delta = x(t - \delta)$ . Thus,  $x^{\epsilon, N}(t - \delta) = x_\delta^{\epsilon, N}(t)$ . For notational simplicity, we shall omit the  $N$  in  $a^N(\cdot)$  and  $b^N(\cdot)$  in the sequel. We shall also write  $x, x_\delta$  for  $x^{\epsilon, N}(t)$  and

$x_\delta^{\epsilon, N}(t)$  respectively whenever it is possible. In the sequel, we shall use  $E_t^\epsilon$  to denote the conditioning on the  $\sigma$ -algebra  $\{\xi^\epsilon(u); u \leq t\}$ .

The following assumptions are needed in the subsequent development.

(A1)  $a(\cdot, \cdot, \cdot)$ ,  $b(\cdot, \cdot, \cdot)$  and  $b_x(\cdot, \cdot, \cdot)$  are continuous,  $a_x(\cdot, \cdot, \xi)$ ,  $b_{xx}(\cdot, \cdot, \xi)$  are continuous in  $x$  for each  $\xi$  and bounded on each bounded  $x$ -set.

(A2)  $\xi^\epsilon(t) = \xi(\frac{t}{\epsilon^2})$ .  $\xi(\cdot)$  is bounded, right continuous stationary  $\phi$ -mixing process, with mixing rate  $\phi(\cdot)$ , such that  $\int_0^\infty \phi(t)dt < \infty$ .  $E b(x, x_\delta, \xi(t)) = 0$  for each  $x, x_\delta$ .

(A3) There exist continuous functions  $S_1(\cdot)$ ,  $\bar{b}(\cdot)$ , continuously differentiable function  $\bar{a}(\cdot)$ , such that for each bounded  $x, x_\delta$ , each  $t < T$ ,

$$\int_t^T (E_t^\epsilon a(x, x_\delta, \xi(u)) - \bar{a}(x, x_\delta)) du \xrightarrow{\epsilon} 0 \text{ in probability} \quad (2.2)$$

and as  $T_1, T_2 \rightarrow \infty$ , and  $T_2 - T_1 \rightarrow \infty$ ,

$$\int_{T_1}^{T_2} E b_x(x, x_\delta, \xi(u)) b(x, x_\delta, \xi(T_1)) du \rightarrow \bar{b}(x, x_\delta) \quad (2.3)$$

$$\int_{T_1}^{T_2} E b(x, x_\delta, \xi(u)) b'(x, x_\delta, \xi(T_1)) du \rightarrow \frac{1}{2} S_1(x, x_\delta). \quad (2.4)$$

Define  $S(x, x_\delta) = \frac{1}{2}(S_1(x, x_\delta) + S_1'(x, x_\delta))$ . There exists a  $\Phi(x, x_\delta)$ , such that

$$S(x, x_\delta) = \Phi(x, x_\delta) \Phi'(x, x_\delta).$$

(A4) (2.5) has a unique solution (unique in the sense of distribution) on  $[0, \infty)$  for each deterministic function prescribed on  $[-\delta, 0]$ .

$$dx(t) = (\bar{a}(x(t), x_\delta(t)) + \bar{b}(x(t), x_\delta(t))dt + \Phi(x(t), x_\delta(t))dw(t) \quad (2.5)$$

where  $w(\cdot)$  is a standard Brownian motion process.

Define an operator  $\mathcal{A}$  by

$$\mathcal{A}f(x) = \sum_i (\bar{a}_i(x, x_\delta) + \bar{b}_i(x, x_\delta)) \frac{\partial f(x)}{\partial x_i} + \frac{1}{2} \sum_{i,j} S_{ij}(x, x_\delta) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \quad (2.6)$$

where  $\bar{a}_i, \bar{b}_i, S_{ij}$  are the  $i$ -th components of  $\bar{a}, \bar{b}$  and  $ij$ -th entry of  $S$  respectively. We say that  $x(\cdot)$  solves the martingale problem for operator  $\mathcal{A}$  if

$$M_f(t) = f(x(t)) - f(x_0(t)) - \int_0^t \mathcal{A}f(x(s))ds \quad (2.7)$$

is a martingale for each  $f(\cdot) \in C_0^3$  ( $C^3$  functions with compact support).

By  $\mathcal{A}^N$ , we mean the operator  $\mathcal{A}$  with  $x, x_\delta, \bar{a}, \bar{b}, S$  replaced by  $x^N, x_\delta^N, \bar{a}^N, \bar{b}^N$  and  $S^N$  respectively.

Our conditions do not seem to be restrictive. (A1) and (2.2) imply that

$$E \sup_{x \in G} \left| \int_t^T (E_t^\epsilon a(x, x_\delta, \xi^\epsilon(u)) - \bar{a}(x, x_\delta)) du \right| \rightarrow 0$$

where  $G$  is a compact set. This equation is needed in the subsequent development.  $\phi$ -mixing (or strong mixing) processes constitute a large class of processes which have "decreasing dependence" property. Actually, we can work with an even larger class of processes that are approximated by functions of  $\phi$ -mixing type (cf. [1]). (A4) implies that the martingale problem with operator  $\mathcal{A}$  defined in (2.6) (2.7) has a unique solution for each prescribed deterministic function  $x_0(t)$  on  $[-\delta, 0]$ .

We are now in a position to state our main theorems.

**Theorem 1.** *If (A1)-(A4) hold, then for each  $N$ ,  $\{x^{\epsilon, N}(\cdot)\}$  is tight in  $D^r[0, \infty)$ .*

**Theorem 2.** *Under the conditions of Theorem 1,  $\{x^\epsilon(\cdot)\}$  is tight in  $D^r[0, \infty)$ , and the limit of any weakly convergent subsequence satisfies equation (2.5), with  $x_0(t)$  given by (1.2).*

### 3. Proofs of the theorems

The steps are as follows, we first prove the tightness of  $\{x^{\epsilon, N}(\cdot)\}$ , and then characterize its limit process.

For any  $f \in C_0^3$ ,

$$\hat{\mathcal{A}}^\epsilon f(x) = f'_x(x) \left( a(x, x_\delta, \xi^\epsilon(t)) + \frac{1}{\epsilon} b(x, x_\delta, \xi^\epsilon(t)) \right). \quad (3.1)$$

In accordance with the approach of [8] [9], we introduce the perturbed test functions as follows. Define

$$f_1^\epsilon(t) = \frac{1}{\epsilon} \int_t^T E_t^\epsilon f'_x(x) b(x, x_\delta, \xi^\epsilon(u)) du \quad (3.2)$$

and define  $f^\epsilon(t) = f(x) + f_1^\epsilon(t)$ .

The reason for introducing the perturbed test functions is that such functions allow us to eliminate the noise terms through averaging, and to obtain the desired terms in the limit. A distinct feature of this averaging procedure is that only the noise is averaged out.  $x, x_\delta$  are treated as parameters. To prove the tightness, we need the following results.

**Lemma 1.** *(cf [1], [8]) Let  $\zeta(\cdot)$  be  $\phi$ -mixing, and let  $h(\cdot)$  be a function of  $\zeta$  which is bounded and measurable on  $\mathcal{F}_0^\infty$ . Then for some  $K_i > 0, i = 1, 2, 3$ ,*

$$|E(h(t+s)|\mathcal{F}_0^t) - Eh(t+s)| \leq K_1 \phi(s). \quad (3.3.a)$$

If  $t < u < v$ , and  $Eh(s) = 0$  for all  $s$ , then

$$|E(h(u)h(v)|\mathcal{F}_0^t) - Eh(u)h(v)| \leq \begin{cases} K_2\phi(v-u) \\ K_3\phi(u-t) \end{cases} \quad (3.3.b)$$

where  $\mathcal{F}_0^t = \sigma\{\zeta(s); 0 \leq s \leq t\}$ .

**Lemma 2.** (Kushner [8]) Let  $z^\epsilon(\cdot) \in D^r[0, \infty)$  and

$$\lim_{M \rightarrow \infty} \limsup_{\epsilon} P\{\sup_{t \leq T} |z^\epsilon(t)| \geq M\} = 0$$

for each  $T < \infty$ . For each  $f(\cdot) \in C_0^3$  and  $T < \infty$ , let there be a sequence  $\{f^\epsilon(\cdot)\}$  such that  $f^\epsilon(\cdot) \in \mathcal{D}(\hat{\mathcal{A}}^\epsilon)$  and that  $\{\hat{\mathcal{A}}^\epsilon f^\epsilon(t); \epsilon > 0, t \leq T\}$  is uniformly integrable and

$$\lim_{\epsilon \rightarrow 0} P\{\sup_{t \leq T} |f^\epsilon(t) - f(z^\epsilon(t))| \geq \alpha\} = 0 \text{ for each } \alpha > 0. \quad (3.4)$$

Then  $\{z^\epsilon(\cdot)\}$  is tight in  $D^\epsilon[0, \infty)$ .

Proof of Theorem 1: To prove the tightness of  $\{x^\epsilon(\cdot)\}$ , we need only verify that all the conditions of Lemma 2 are satisfied. In fact, we need only show  $\{\hat{\mathcal{A}}^\epsilon f^\epsilon(\cdot)\}$  is uniformly integrable,  $f^\epsilon(\cdot) \in \mathcal{D}(\hat{\mathcal{A}}^\epsilon)$  and (3.4) holds.

Making change of variable  $\frac{u}{\epsilon^2} \rightarrow u$  in (3.2), we have

$$f_1^\epsilon(t) = \epsilon \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E_t^\epsilon f'_x(x) b(x, x_\delta, \xi(u)) du.$$

In view of (3.3) and (A2), for some  $K > 0$ ,

$$\begin{aligned} \sup_{t \leq T} |f_1^\epsilon(t)| &= \epsilon \sup_{t \leq T} \left| \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} f'_x(x) \left( E_t^\epsilon b(x, x_\delta, \xi(u)) - Eb(x, x_\delta, \xi(u)) \right) du \right| \\ &\leq K \epsilon \sup_{t \leq T} \left( \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} \phi(u - \frac{t}{\epsilon^2}) du \right) = O(\epsilon). \end{aligned} \quad (3.5)$$

Hence,  $\lim_{\epsilon} E \sup_{t \leq T} |f_1^\epsilon(t)| = 0$ . As a consequence, (3.4) is verified.

To prove the uniform integrability of  $\{\hat{\mathcal{A}}^\epsilon f^\epsilon(\cdot)\}$  and  $f^\epsilon(\cdot) \in \mathcal{D}(\hat{\mathcal{A}}^\epsilon)$ , we need to compute  $\hat{\mathcal{A}}^\epsilon f^\epsilon(t)$ . The delay term causes much of the troubles. When we calculate  $\hat{\mathcal{A}}^\epsilon f_1^\epsilon(\cdot)$  directly, we will introduce an additional time lag. In order to overcome the difficulties, we note the following. The continuity of  $x_\delta(\cdot)$  implies that  $x_\delta(t)$  can be approximated by  $x_\delta^\eta(t)$  which takes only finitely many values (say  $x_{\delta,1}, \dots, x_{\delta,p}$ , for some  $p > 0$ ). To be more precise, for each  $\epsilon > 0$ , choose  $\eta = \eta(\epsilon) > 0$ , such that  $\eta \xrightarrow{\epsilon} 0$ ,  $\frac{\eta}{\epsilon^2} \xrightarrow{\epsilon} 0$ , and

$$|x_\delta(t) - x_\delta^\eta(t)| < \eta. \quad (3.6)$$

Owing to (3.2),

$$\begin{aligned} f_1^\epsilon(t) &= \frac{1}{\epsilon} \int_t^T E_t^\epsilon f'_x(x) b(x, x_\delta^\eta, \xi^\epsilon(u)) du \\ &\quad + \frac{1}{\epsilon} \int_t^T E_t^\epsilon f'_x(x) \left( b(x, x_\delta, \xi^\epsilon(u)) - b(x, x_\delta^\eta, \xi^\epsilon(u)) \right) du. \end{aligned} \quad (3.7)$$

The  $N$ -truncation and (A1) imply that  $b(x, \cdot, \xi)$  is Lipschitz continuous. Consequently,

$$\frac{1}{\epsilon} \left| \int_t^T E_t^\epsilon f'_x(x) \left( b(x, x_\delta, \xi^\epsilon(u)) - b(x, x_\delta^\eta, \xi^\epsilon(u)) \right) du \right| \leq O\left(\frac{\eta(\epsilon)}{\epsilon}\right) \quad (3.8)$$

and hence  $f_1^\epsilon(\cdot)$  can be written as

$$f_1^\epsilon(t) = \hat{f}_1^\epsilon(t) + o(\epsilon)$$

where  $\hat{f}_1^\epsilon(t)$  is equal to the first term on the right hand side of (3.7).

Using the idea of  $p$ -limit to compute  $\hat{A}^\epsilon f_1^\epsilon(t)$ , we have

$$\begin{aligned} \hat{A}^\epsilon f_1^\epsilon(t) &= p - \lim_{\delta \rightarrow 0} \left( \frac{E_t^\epsilon f_1^\epsilon(t + \delta) - f_1^\epsilon(t)}{\delta} \right) \\ &= O(\eta(\epsilon)/\epsilon) - \frac{1}{\epsilon} f'_x(x) b(x, x_\delta^\eta, \xi^\epsilon(t)) \\ &\quad + \epsilon \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E_t^\epsilon (f_{xx}(x) b(x, x_\delta^\eta, \xi(u)))' du a(x, x_\delta, \xi^\epsilon(t)) \\ &\quad + \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E_t^\epsilon (f_{xx}(x) b(x, x_\delta^\eta, \xi(u)))' b(x, x_\delta, \xi^\epsilon(t)) du \\ &\quad + \epsilon \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E_t^\epsilon f'_x(x) b_x(x, x_\delta^\eta, \xi(u)) du a(x, x_\delta, \xi^\epsilon(t)) \\ &\quad + \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E_t^\epsilon f'_x(x) b_x(x, x_\delta^\eta, \xi(u)) b(x, x_\delta, \xi^\epsilon(t)) du. \end{aligned} \quad (3.9)$$

The third and fifth term on the right hand side of (3.9) tend to 0 as  $\epsilon \rightarrow 0$ . Hence, by virtue of (3.6) and the definition of  $f^\epsilon(t)$ ,

$$\begin{aligned} \hat{A}^\epsilon f^\epsilon(t) &= o(1) + f'_x(x) a(x, x_\delta, \xi^\epsilon(t)) \\ &\quad + \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E_t^\epsilon (f_{xx}(x) b(x, x_\delta^\eta, \xi(u)))' du b(x, x_\delta^\eta, \xi^\epsilon(t)) \\ &\quad + \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E_t^\epsilon f'_x(x) b_x(x, x_\delta^\eta, \xi(u)) b(x, x_\delta^\eta, \xi^\epsilon(t)) du \end{aligned} \quad (3.10)$$



where  $p - \lim_{\epsilon} o(1) = 0$  uniformly in  $t$ .

The uniform integrability then follows from (3.10), the  $N$ -truncation, and the assumptions. Moreover,  $f^\epsilon(\cdot) \in \mathcal{D}(\hat{\mathcal{A}}^\epsilon)$ . As a consequence,  $\{x^{\epsilon, N}(\cdot)\}$  is tight.  $\square$

Next, we show the weak convergence and characterize the weak limit. To do so, we shall apply the following lemma.

**Lemma 3.** *Let  $\{z^\epsilon(\cdot)\}$  be tight in  $D^r[0, \infty)$ , and suppose that (A4) holds. If for each  $f(\cdot) \in C_0^3$ , and each  $T < \infty$ , then there exists  $f^\epsilon(\cdot) \in \mathcal{D}(\hat{\mathcal{A}}^\epsilon)$ , such that*

$$p - \lim_{\epsilon} [f^\epsilon(\cdot) - f(z^\epsilon(\cdot))] = 0 \quad (3.11)$$

and

$$p - \lim_{\epsilon} [\hat{\mathcal{A}}^\epsilon f^\epsilon(\cdot) - \mathcal{A}f(x^\epsilon(\cdot))] = 0. \quad (3.12)$$

Then  $z^\epsilon(\cdot) \Rightarrow z(\cdot)$ .

The proof is due to Kushner [7-8].

Proof of Theorem 2: In view of Lemma 3, it is enough to show that (3.11) and (3.12) hold.

Define

$$\begin{aligned} f_2^\epsilon(t) = & \int_t^T \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} \left\{ E_t^\epsilon (f_{xx}(x)b(x, x_\delta^\eta, \xi(u)))' b(x, x_\delta^\eta, \xi^\epsilon(v)) \right. \\ & \left. - E(f_{xx}(x)b(x, x_\delta^\eta, \xi(u)))' b(x, x_\delta^\eta, \xi^\epsilon(v)) \right\} dudv \end{aligned} \quad (3.13)$$

$$\begin{aligned} f_3^\epsilon(t) = & \int_t^T \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} \left\{ E_t^\epsilon f'_x(x)b_x(x, x_\delta^\eta, \xi(u))b(x, x_\delta^\eta, \xi^\epsilon(v)) \right. \\ & \left. - E f'_x(x)b_x(x, x_\delta^\eta, \xi(u))b(x, x_\delta^\eta, \xi^\epsilon(v)) \right\} dudv \end{aligned} \quad (3.14)$$

$$f_4^\epsilon(t) = \int_t^T E_t^\epsilon f'_x(x) \left( a(x, x_\delta^\eta, \xi^\epsilon(u)) - \bar{a}(x, x_\delta^\eta) \right) du. \quad (3.15)$$

Note that

$$\begin{aligned} f_2^\epsilon(t) = & \epsilon^2 \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} dv \int_v^{\frac{T}{\epsilon^2}} du \left\{ E_t^\epsilon (f_{xx}(x)b(x, x_\delta^\eta, \xi(u)))' b(x, x_\delta^\eta, \xi(v)) \right. \\ & \left. - E(f_{xx}(x)b(x, x_\delta^\eta, \xi(u)))' b(x, x_\delta^\eta, \xi(v)) \right\}. \end{aligned} \quad (3.16.a)$$

Similarly,

$$f_3^\epsilon(t) = \epsilon^2 \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} \int_v^{\frac{T}{\epsilon^2}} \left\{ E_t^\epsilon f'_x(x) b_x(x, x_\delta^\eta, \xi(u)) b(x, x_\delta^\eta, \xi(v)) \right. \\ \left. - E f'_x(x) b_x(x, x_\delta^\eta, \xi(u)) b(x, x_\delta^\eta, \xi(v)) \right\} dudv \quad (3.16.b)$$

$$f_4^\epsilon(t) = \epsilon^2 \int_t^T E_t^\epsilon f'_x(x) \left( a(x, x_\delta^\eta, \xi^\epsilon(u)) - \bar{a}(x, x_\delta^\eta) \right) du. \quad (3.16.c)$$

In view of Lemma 1, we can show that for  $i = 2, 3, 4$ ,

$$\lim_\epsilon E \sup_{t \leq T} |f_i^\epsilon(t)| = 0. \quad (3.17)$$

Define

$$\tilde{f}^\epsilon(t) = f(x) + \sum_{i=1}^4 f_i^\epsilon(t). \quad (3.18)$$

By virtue of (3.5), (3.16)-(3.18),

$$p - \lim_\epsilon [\tilde{f}^\epsilon(t) - f(x)] = 0. \quad (3.19)$$

To proceed, we compute  $\hat{A}^\epsilon f_i^\epsilon(\cdot)$ , for  $i = 2, 3, 4$ ,

$$\hat{A}^\epsilon f_2^\epsilon(t) = \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E(f_{xx}(x) b(x, x_\delta^\eta, \xi(u)))' b(x, x_\delta^\eta, \xi^\epsilon(t)) du \\ - \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E_t^\epsilon(f_{xx}(x) b(x, x_\delta^\eta, \xi(u)))' b(x, x_\delta^\eta, \xi^\epsilon(t)) du \\ + \epsilon^2 \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} dv \int_v^{\frac{T}{\epsilon^2}} du \left\{ E_t^\epsilon(f_{xx} b(x, x_\delta^\eta, \xi(u)))' b(x, x_\delta^\eta, \xi(v)) \right. \\ \left. - E(f_{xx}(x) b(x, x_\delta^\eta, \xi(u)))' b(x, x_\delta^\eta, \xi(v)) \right\}_x \\ \times \left( a(x, x_\delta, \xi^\epsilon(t)) + \frac{1}{\epsilon} b(x, x_\delta, \xi^\epsilon(t)) \right) \quad (3.20)$$

$$\hat{A}^\epsilon f_3^\epsilon(t) = \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E f'_x(x) b_x(x, x_\delta^\eta, \xi(u)) b(x, x_\delta^\eta, \xi^\epsilon(t)) du \\ - \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E_t^\epsilon f'_x(x) b_x(x, x_\delta^\eta, \xi(u)) b(x, x_\delta^\eta, \xi^\epsilon(t)) du \\ + \epsilon^2 \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} dv \int_v^{\frac{T}{\epsilon^2}} du \left\{ E_t^\epsilon f'_x(x) b_x(x, x_\delta^\eta, \xi(u)) b(x, x_\delta^\eta, \xi(v)) \right. \\ \left. - E f'_x(x) b_x(x, x_\delta^\eta, \xi(u)) b(x, x_\delta^\eta, \xi(v)) \right\}_x \\ \times \left( a(x, x_\delta, \xi^\epsilon(t)) + \frac{1}{\epsilon} b(x, x_\delta, \xi^\epsilon(t)) \right). \quad (3.21)$$

The  $p - \lim_{\epsilon}$  of the last term on the right hand side of (3.20) and the last term of (3.21) are both zero.

As for  $f_4^\epsilon(\cdot)$ ,

$$\begin{aligned} \hat{A}^\epsilon f_4^\epsilon(t) &= -f'_x(x)a(x, x_\delta^\eta, \xi^\epsilon(t)) + f'_x(x)\bar{a}(x, x_\delta^\eta) \\ &\quad + \epsilon^2 \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} (E_t^\epsilon f'_x(x)a(x, x_\delta^\eta, \xi(u)) - \bar{a}(x, x_\delta^\eta))'_x a(x, x_\delta, \xi^\epsilon(t)) du \\ &\quad + \epsilon \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} (E_t^\epsilon f'_x(x)a(x, x_\delta^\eta, \xi(u)) - \bar{a}(x, x_\delta^\eta))'_x b(x, x_\delta, \xi^\epsilon(t)) du. \end{aligned} \quad (3.22)$$

The last two terms of (3.22) have  $p - \lim_{\epsilon} 0$  uniformly in  $t$ .

(3.1), (3.9), (3.20)-(3.22) yield that

$$\begin{aligned} \hat{A}^\epsilon \tilde{f}^\epsilon(t) &= o(1) + f'_x(x)\bar{a}(x, x_\delta^\eta) \\ &\quad + f'_x(x)(a(x, x_\delta, \xi^\epsilon(t)) - a(x, x_\delta^\eta, \xi^\epsilon(t))) \\ &\quad + \frac{1}{\epsilon} f'_x(x)(b(x, x_\delta, \xi^\epsilon(t)) - b(x, x_\delta^\eta, \xi^\epsilon(t))) \\ &\quad + \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E(f_{xx}(x)b(x, x_\delta^\eta, \xi(u)))'_x b(x, x_\delta^\eta, \xi^\epsilon(t)) du \\ &\quad + \int_{\frac{t}{\epsilon^2}}^{\frac{T}{\epsilon^2}} E f'_x(x)b_x(x, x_\delta^\eta, \xi(u))b(x, x_\delta^\eta, \xi^\epsilon(t)) du \end{aligned} \quad (3.23)$$

where  $p - \lim_{\epsilon \rightarrow 0} o(1) = 0$  uniformly in  $t$ .

(3.1) (3.6) (A3) together with (3.23) yield (3.12). By virtue of Lemma 3,  $x^{\epsilon, N}(\cdot) \Rightarrow x^N(\cdot)$ . And  $x^N(\cdot)$  solves the martingale problem with operator  $A^N$ , or  $x^N(\cdot)$  satisfies (2.5) with  $\bar{a}$ ,  $\bar{b}$ ,  $\Phi$  replaced by  $\bar{a}^N$ ,  $\bar{b}^N$ , and  $\Phi^N$  respectively.

Similar to the proof of Corollary 3.2.2 in [8], for any prescribed deterministic function  $x_0(t)$  on  $[-\delta, 0]$ , let  $P(\cdot)$  and  $P^N(\cdot)$  denote the measures induced by  $x(\cdot)$  and  $x^N(\cdot)$  respectively, on the Borel sets of  $D^r[0, \infty)$ . By (A4), the martingale problem has a unique solution for each prescribed deterministic function, thus  $P(\cdot)$  is unique. For each  $T < \infty$ , the uniqueness of  $P(\cdot)$  implies that  $P(\cdot)$  agree with  $P^N(\cdot)$  on all Borel sets of the set of paths in  $D^r[0, \infty)$  whose value are in  $S_N$  for each  $t \leq T$ . However,  $P\{\sup_{t \leq T} |x(t)| \leq N\} \xrightarrow{N} 1$ . This together with the weak convergence of  $x^{\epsilon, N}(\cdot)$  imply that  $x^\epsilon \Rightarrow x(\cdot)$ . Moreover, the uniqueness implies that the limit does not depend on the chosen subsequences. The proof of the Theorem 2 is thus completed.  $\square$

Remark: The proof presented in this paper can be adopted to treat the differential delay equation with random perturbations (1.3). In this case, the wideband noise has the scaling  $\xi^\epsilon(t) = \xi(t/\epsilon)$ . Under appropriate conditions, we can get exactly the same results

as in [10]. However, the proof is much simpler than that of [10]. The martingale averaging technique seems to have clear advantage over the approach in [5] and [10].

We used perturbed test function methods throughout the proofs. An alternative procedure is that after proving the tightness of  $\{x^{\epsilon, N}(\cdot)\}$ , use “direct averaging” (cf. [8]) to average out the rest of the terms and obtain the desired results.

We point out that all the previous development can be extended to the case

$$\begin{aligned} \dot{x}^{\epsilon}(t) &= a(x^{\epsilon}(t), x^{\epsilon}(t - \delta_1), \dots, x^{\epsilon}(t - \delta_p), \xi^{\epsilon}(t)) \\ &\quad + \frac{1}{\epsilon} b(x^{\epsilon}(t), x^{\epsilon}(t - \delta_1), \dots, x^{\epsilon}(t - \delta_p), \xi^{\epsilon}(t)) \end{aligned} \quad (3.24)$$

for some  $p > 0$ , and  $0 < \delta_1 < \delta_2 < \dots < \delta_p$ .

#### 4. Equation with random delays

In this section, we make remarks on what can be done if the time delay is a random process. We shall assume that the delays are independent of the state  $x$  and the noise  $\xi$  throughout this section. The random delay process is rapidly varying, and it has the same fast time scale as the noise  $\xi$ , i.e.,  $\delta^{\epsilon}(t) = \delta(t/\epsilon^2)$ . In section 4.1, we consider the case that  $\delta$  is random, but takes only finitely many possible values. In 4.2, we approximate the weakly convergent random process  $\delta^{\epsilon}(\cdot)$  by a process which takes values in a finite set. Since the basic approach is essentially the same as before, we make no attempt to spell out all the details. In fact, we shall keep the discussions in a rather informal manner, so as to make the main idea clear.

##### 4.1 Random delays with finitely many possible values

Suppose that the delays are random, and suppose  $\delta(\cdot)$  takes only finitely many values (say  $0 < \delta_1 < \dots < \delta_p$  for some  $p > 0$ ) w.p.1. We can then rewrite (1.1) as

$$\dot{x}^{\epsilon}(t) = \sum_{i=1}^p (a(x^{\epsilon}(t), x^{\epsilon}(t - \delta_i), \xi^{\epsilon}(t)) + \frac{1}{\epsilon} b(x^{\epsilon}(t), x^{\epsilon}(t - \delta_i), \xi^{\epsilon}(t))) I_{\{\delta(t) = \delta_i\}}. \quad (4.1)$$

(4.1) in turn can be recast to the form

$$\begin{aligned} \dot{x}^{\epsilon}(t) &= A(x^{\epsilon}(t), x^{\epsilon}(t - \delta_1), \dots, x^{\epsilon}(t - \delta_p), \xi^{\epsilon}(t)) \\ &\quad + \frac{1}{\epsilon} B(x^{\epsilon}(t), x^{\epsilon}(t - \delta_1), \dots, x^{\epsilon}(t - \delta_p), \xi^{\epsilon}(t)) \end{aligned} \quad (4.2)$$

$$x^{\epsilon}(t) = x_0(t) \text{ for } t \in [-\delta_p, 0].$$

Now, the martingale averaging technique can be employed, and the desired limit theorem can be obtained similarly as in section 3. We decide not to dwell on it here.

## 4.2 Random delays taking infinitely many possible values

Let the process  $\{x^\epsilon(\cdot)\}$  satisfy equation (1.1) with  $\delta$  replaced by  $\delta^\epsilon(\cdot)$ , such that for some  $M > 0$ ,  $\delta^\epsilon(\cdot) \in [0, M]$  with probability 1. Let (1.2) be replaced by

$$x^\epsilon(t) = x_0(t) \text{ for } t \in [-M, 0] \quad (4.3)$$

where  $x_0(t)$  is a given continuous deterministic function on  $[-M, 0]$ . We assume that  $\delta^\epsilon(\cdot) \Rightarrow \delta(\cdot)$ , and  $\delta(\cdot)$  is a stationary stochastic process with a stationary measure  $P(\cdot)$ . This does not seem to be restrictive. In fact, the boundedness of  $\{\delta^\epsilon(\cdot)\}$  implies the tightness of  $\{\delta^\epsilon(\cdot)\}$ . We can thus extract convergent subsequences.

For the process  $\delta^\epsilon(t)$ , the possibility of taking infinitely many values is allowed. We shall demonstrate that this case can be reduced to the previous situation treated in 4.1.

To begin with, we note  $\forall \Delta > 0$ , any  $T < \infty$  and  $t \in [0, T]$ , there exists a  $N(\Delta)$ , such that there is a sequence of simple processes  $\{\delta^{\epsilon, n}(\cdot)\}$  (each  $\delta^{\epsilon, n}(t)$  takes only finitely many values), such that  $\forall n \geq N(\Delta)$ ,

$$\int_0^t E|\delta^\epsilon(u) - \delta^{\epsilon, n}(u)|^2 du < \Delta. \quad (4.4)$$

In the sequel, we choose  $\Delta = o(\epsilon^2)$ , select one function out of the approximating sequence, and denote it by  $\delta^{\epsilon, \Delta}(\cdot)$ , such that (4.4) holds with  $\delta^{\epsilon, n}(u)$  replaced by  $\delta^{\epsilon, \Delta}(u)$ .

Now, define another process  $\{y^\epsilon(\cdot)\}$  by

$$\dot{y}^\epsilon(t) = a(y^\epsilon(t), y^\epsilon(t - \delta^{\epsilon, \Delta}(t)), \xi^\epsilon(t)) + \frac{1}{\epsilon} b(y^\epsilon(t), y^\epsilon(t - \delta^{\epsilon, \Delta}(t)), \xi^\epsilon(t)) \quad (4.5)$$

$$y^\epsilon(t) = x_0(t) \text{ on } [-M, 0]. \quad (4.6)$$

We shall show that  $x^\epsilon(\cdot) = y^\epsilon(\cdot) + \hat{y}^\epsilon(\cdot)$ , such that  $\hat{y}^\epsilon(\cdot) \Rightarrow 0$ . Without loss of generality, we may assume that  $x^\epsilon(\cdot)$ ,  $y^\epsilon(\cdot)$ ,  $a(\cdot)$ ,  $b(\cdot)$  are all bounded, otherwise we can always use the truncation device as in the last section.

**Lemma 4.** *Suppose (A1) is satisfied. For any  $T < \infty$  and  $t \in [0, T]$ ,*

$$x^\epsilon(t) = y^\epsilon(t) + O\left(\frac{\Delta}{\epsilon^2}\right).$$

*Proof:* For each  $T < \infty$ , partition the interval  $[0, T]$  into subintervals of length  $\Delta$ . We shall show that Lemma 4 holds for each integer  $k$ , and any  $t \in [k\Delta, (k+1)\Delta]$  by induction. In the following,  $K_i$ ,  $i = 1, \dots, 5$  stand for various positive constants.

When  $t \in [0, \Delta]$ , in view of the representations of  $x^\epsilon(\cdot)$  and  $y^\epsilon(\cdot)$ ,

$$\begin{aligned}
& E|x^\epsilon(t) - y^\epsilon(t)|^2 \\
&= E \left| \int_0^t [a(x^\epsilon(u), x^\epsilon(u - \delta^\epsilon(u)), \xi^\epsilon(u)) - a(y^\epsilon(u), y^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du \right. \\
&\quad \left. + \frac{1}{\epsilon} \int_0^t [b(x^\epsilon(u), x^\epsilon(u - \delta^\epsilon(u)), \xi^\epsilon(u)) - [b(y^\epsilon(u), y^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du \right|^2.
\end{aligned} \tag{4.7}$$

Moreover,

$$\begin{aligned}
& \int_0^t [a(x^\epsilon(u), x^\epsilon(u - \delta^\epsilon(u)), \xi^\epsilon(u)) - a(y^\epsilon(u), y^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du \\
&= \int_0^t [a(x^\epsilon(u), x^\epsilon(u - \delta^\epsilon(u)), \xi^\epsilon(u)) - a(x^\epsilon(u), x^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du \\
&\quad + \int_0^t [a(x^\epsilon(u), x^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u)) - a(y^\epsilon(u), x^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du \\
&\quad + \int_0^t [a(y^\epsilon(u), x^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u)) - a(y^\epsilon(u), y^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du.
\end{aligned}$$

Similarly,

$$\begin{aligned}
& \frac{1}{\epsilon} \int_0^t [b(x^\epsilon(u), x^\epsilon(u - \delta^\epsilon(u)), \xi^\epsilon(u)) - b(y^\epsilon(u), y^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du \\
&= \frac{1}{\epsilon} \int_0^t [b(x^\epsilon(u), x^\epsilon(u - \delta^\epsilon(u)), \xi^\epsilon(u)) - b(x^\epsilon(u), x^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du \\
&\quad + \frac{1}{\epsilon} \int_0^t [b(x^\epsilon(u), x^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u)) - b(y^\epsilon(u), x^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du \\
&\quad + \frac{1}{\epsilon} \int_0^t [b(y^\epsilon(u), x^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u)) - b(y^\epsilon(u), y^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du.
\end{aligned}$$

By virtue of the Lipschitz continuity, (4.7) and the above expressions,

$$\begin{aligned}
& E|x^\epsilon(t) - y^\epsilon(t)|^2 \\
&\leq K_1 \left(1 + \frac{1}{\epsilon^2}\right) \int_0^t E|x^\epsilon(u) - y^\epsilon(u)|^2 du \\
&\quad + K_2 \left(1 + \frac{1}{\epsilon^2}\right) \int_0^t E|x^\epsilon(u - \delta^{\epsilon, \Delta}(u)) - y^\epsilon(u - \delta^{\epsilon, \Delta}(u))|^2 du \\
&\quad + \frac{K_3}{\epsilon^2} \int_0^t E|\delta^\epsilon(u) - \delta^{\epsilon, \Delta}(u)|^2 du.
\end{aligned} \tag{4.8}$$

Since  $x^\epsilon(t) = y^\epsilon(t)$  on  $[-M, 0]$ , the second term on the right side of (4.8) is bounded above by  $K_4 \left(1 + \frac{1}{\epsilon^2}\right) \int_0^t E|x^\epsilon(u) - y^\epsilon(u)|^2 du$ . As a result, the sum of the first two terms

on the right hand side of (4.8) is bounded above by  $K_5 \left(1 + \frac{1}{\epsilon^2}\right) \int_0^t E|x^\epsilon(u) - y^\epsilon(u)|^2 du$ . Consequently, Gronwall's inequality and the choice of  $\Delta$  yield that

$$E|x^\epsilon(t) - y^\epsilon(t)|^2 \leq \frac{K_3\Delta}{\epsilon^2} \exp\left(K_5 \left(1 + \frac{1}{\epsilon^2}\right) \Delta\right).$$

Hence, Lemma 4 holds for  $t \in [0, \Delta]$ .

Suppose that Lemma 4 is true for  $t \in [k\Delta, (k+1)\Delta]$ , for some  $k \geq 1$ . We shall show that the assertion holds for  $t \in [(k+1)\Delta, (k+2)\Delta]$ .

Since for any  $t \in [(k+1)\Delta, (k+2)\Delta]$ ,

$$\begin{aligned} x^\epsilon(t) &= x^\epsilon((k+1)\Delta) + \int_{(k+1)\Delta}^t [a(x^\epsilon(u), x^\epsilon(u - \delta^\epsilon(u)), \xi^\epsilon(u)) \\ &\quad + \frac{1}{\epsilon} \int_{(k+1)\Delta}^t b(x^\epsilon(u), x^\epsilon(u - \delta^\epsilon(u)), \xi^\epsilon(u))] du \\ y^\epsilon(t) &= y^\epsilon((k+1)\Delta) + \int_{(k+1)\Delta}^t [a(y^\epsilon(u), y^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u)) \\ &\quad + \frac{1}{\epsilon} \int_{(k+1)\Delta}^t b(y^\epsilon(u), y^\epsilon(u - \delta^{\epsilon, \Delta}(u)), \xi^\epsilon(u))] du. \end{aligned}$$

By induction hypothesis,

$$E|x^\epsilon((k+1)\Delta) - y^\epsilon((k+1)\Delta)|^2 = O\left(\frac{\Delta}{\epsilon^2}\right). \quad (4.9)$$

Performing the same kind of calculation as in (4.8), and taking (4.9) into account,

$$E|x^\epsilon(t) - y^\epsilon(t)|^2 = O\left(\frac{\Delta}{\epsilon^2}\right) \text{ for } t \in [(k+1)\Delta, (k+2)\Delta]. \quad (4.10)$$

Thus, by induction, for any  $t \in [0, T]$ , Lemma 4 is true.  $\square$

By virtue of the above lemma, to consider the weak convergence of  $\{x^\epsilon(\cdot)\}$ , it suffices to study the weak convergence of  $\{y^\epsilon(\cdot)\}$ . As a result, we come back to the situation discussed in section 4.1. The perturbed test function and martingale averaging techniques now can be applied.

*Acknowledgement.* This work was completed when both authors were visiting the Institute for Mathematics and Its Applications, at the University of Minnesota. We wish to thank Prof. Avner Friedman, Prof. Willard Miller and all the IMA staff members for hospitality, various supports, and for providing us with such a nice and stimulating environment.

## References

- [1] P. Billingsley, *Convergence of Probability Measure*, J. Wiley, 1968.
- [2] G.B. Blankenship and G.C. Papanicolaou, Stability and control of stochastic systems with wideband noise disturbances, *SIAM J. Appl. Math.* 42(1978), 437-476.
- [3] E. Either and T. Kurtz, *Markov Process*, J. Wiley, 1986.
- [4] J. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, 1977.
- [5] R.Z. Khasminskii, Stochastic processes defined by differential equations, *Theory Prob. Appl.* XI(1966), 211-228.
- [6] T. Kurtz, Semigroup of conditioned shifts and approximation of Markov processes, *Ann. Prob.* 3(1975), 618-642.
- [7] H.J. Kushner, A martingale method for the convergence of a sequence of process to a jump-diffusion process, *Z. Wahr. Verw. Gebiete.* 53(1980), 207-219.
- [8] H.J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic systems Theory*, MIT Press, 1984.
- [9] G.C. Papanicolaou, D. Strook and S.R.S. Varadhan, martingale approach to some limit theorems, *Proc. of the 1976 Duke Univ. Conference on Turbulence*, 1977.
- [10] B.S. White, Some limit theorems for stochastic delay-differential equation, *Comm. Pure and Appl. Math.* XXIX(1976), 113-141.