Ministère de l'Enseignement Supérieur
et de la Recherche Scientifique
UNIVERSITE DE COCODY

République de Côte d'Ivoire
Union-Discipline-Travail

3.69 / 2002

THESE

Présentée à l'UFR de Mathématiques et Informatique
de l'Université de Cocody
pour obtenir le grade de

DOCTEUR TROISIEME CYCLE

Spécialité : Mathématiques Appliquées
Option : **PROBABILITES**
Par :
KOUAME N'GORAN Louis

AVERAGING PRINCIPLE FOR STOCHASTIC DIFFERENTIAL EQUATIONS

Soutenue publiquement le 23 Novembre 2002, devant la commission d'examen composée de :

Président	Assohoun **ADJE**	*Maître de Conférences à l'Université de Cocody, Abidjan*
Rapporteurs	Youssef **OUKNINE**	*Professeur à l'Université Cadi Ayyad, Marrakech*
	Ouagnina **HILI**	*Maître de Conférences à l'INPHB, Yamoussoukro*
Examinateur	Boubacar **DIALLO**	Maître Assistant à *l'Université de Cocody, Abidjan*
Directeur de Thèse	Modeste **N'ZI**	*Maître de Conférences à l'Université de Cocody, Abidjan*

Acknowledgement

I would like to express all my gratitude to Mr Modeste N'ZI who proposed the topic of this thesis to me. His everlasting support, many suggestions and useful advice allowed me to take the present work to an end. I learnt a lot from him and I am grateful to him.

I greatlty thank Mr Assouhoun ADJE who kindly accepted to chair the board of examiners.

I am keenly grateful to Mr Youssef OUKNINE and Mr Ouagnina HILI who took an interest in this work by accepting to make reports on it.

I am really happy to have got Mr Boubacar Diallo as a member of this thesis committee. I thank him for being at my disposal.

I thank very much all the members of the seminar of probability and statistic, especially Mr Abouo ELOUAFLIN for his inestimable help and useful discussions about this work.

I also thank Mr Adama COULIBALY warmly for his help in the working-out of this thesis.

Finally, I would like to tell my beloved wife, Liliane Denise BRIME, how grateful I am to her for her constant support all along this work; Thousand thanks to her.

Remerciements

Je voudrais exprimer toute ma gratitude envers le Professeur Modeste N'ZI, Maître de Conférence à l'Université de Cocody qui m'a proposé le sujet de cette thèse. Son soutien sans faille, ses nombreuses suggestions et conseils éclairés m'ont permis de mener ce travail à son terme. J'ai beaucoup appris à ses côtés. Qu'il trouve ici l'expression de ma profonde reconnaissance.

Je remercie très vivement le Professeur Assouhoun ADJE, Maître de Conférences à l'Université de Cocody qui a accepté gentillement de présider le jury de cette thèse.

Je tiens à exprimer ma profonde reconnaissance à Messieurs Youssef OUKNINE, Professeur titulaire à l'Université Cadi Ayyad de Marrakech (Maroc) et Ouagnina HILI, Maître de Conférences à l'INPHB de Yamoussokro qui ont bien voulu s'intéresser à ce travail en acceptant d'être les rapporteurs.

Je suis heureux d'avoir comme examinateur, M. Boubacar Diallo, Maître assistant à l'Université de Cocody. Je le remercie pour sa grande disponibilité.

Mes remerciements vont aussi à l'endroit de tous les membres du séminaire de probabilités et statistique, en particulier M. Abouo ELOUAFLIN qui m'a été d'un soutien inestimable. J'apprécie à sa juste valeur les nombreuses discussions que nous avons eu sur le principe de moyennisation en général et sur le sujet de cette thèse en particulier.

Je remercie M. Adama COULIBALY, assistant à l'Université de Cocody qui m'a beaucoup aidé dans la mise en forme de cette thèse.

Enfin, je ne saurais terminer sans remercier celle qui m'a constamment soutenu tout au long de ce travail, il s'agit de mon épouse: Liliane Denise BRIME. Milles fois Merci.

Abstract

In this thesis, we study the averaging problem for multivalued stochastic differential equations and for stochatic differential equations driven by double stochastic integrals. The averaging principle consists in showing that the solution of a differential equation whose coefficients are pertubed by a process describing a fast motion can be approxiamted by that of some unperturbed system obtained by averaging out the fluctuations arising from the fast motion. This principle play an important role in celestial mechanics, oscillation theory, radiophysic and in many others areas. The first mathematical rigorous justification of this principle go back to Bogolyubov (1945). Later on, this principle has attracted much attention of many researchers and nowdays it is a regular area of research.

In this work, Chapter 1 deals with a brief historical overview of previous works.

In Chapter 2, we extend a result of Liptser and Stoyanov (1990) to reflected stochastic differential equations whose solution is constrained to stay in the domain of a convexe function.

Chapter 3 is devoted to the averaging principle for doube Itô stochastic differential equations. Under some conditions introduced in Hashemi and Heunis (1998), we prove that the solution of the pertubed equation can be approximated in L_2 by that of the averaged one. We hope that the L_2-convergence can be improved to almost sure one.

In an Appendix , we collect some results needed in Chapter 3.

Résumé

Dans cette thèse, on étudie le problème de moyennisation pour les équations différentielles stochastiques multivoques ainsi que pour les équations différentielles stochastiques dirigées par des intégrales stochastiques doubles. Le principe de moyennisation consiste à approcher la solution d'une équation différentielle dont les coefficients sont pertubés par un processus décrivant un mouvement rapide, par celle de l'équation obtenue en moyennisant les coefficients, ce qui a pour effet de faire disparaître les oscillations dûes au processus pertubateur. Ce principe joue un rôle très important en mécanique céleste, en théorie des oscillations, en radiophysique et dans d'autres domaines. La première justification mathématique rigoureuse de ce principe est dûe à Bogolyubuv (1945). Depuis son introduction, le principe de moyennisation a attiré l'attention de nombreux chercheurs et constitue aujourd'hui un domaine de recherche très actif.

Dans ce travail, le Chapitre 1 est consacré à un bref aperçu historique des travaux antérieurs.

Dans le Chapitre 2, nous étendons un résultat de Liptser et Stoyanov (1990) aux équations différentielles stochastiques réfléchies dont la solution est contrainte à rester dans le domaine d'une fonction convexe.

Le Chapitre 3 traite du principe de moyennisation pour les équations différentielles stochastiques doubles d'Itô. Sous des conditions introduites dans Hashemi et Heunis (1998), nous établissons une convergence dans L_2 de la solution de l'équation de départ vers celle de l'équation moyennisée. Nous espérons plus tard pouvoir remplacer la convergence en probabilité par une convergence presque sûre.

Dans un Appendice, nous avons rassemblé quelques résultats dont nous avons besoin dans le Chapitre 3.

Contents

[1] *Published in Random Operators and Stochastic Equations 9(4). 399-407.*
[2] *Submitted for publication*

Chapter 1

GENERAL INTRODUCTION

Let us consider the differential equation perturbed by a function ξ_t, $t \geq 0$, assuming values in $\mathbb{R}^d$:

$$\dot{Z}_t^\varepsilon = \varepsilon a(Z_t^\varepsilon, \xi_t), \quad Z_0^\varepsilon = x_0, \quad t \geq 0 \tag{1.1}$$

where $a(x, y)$ is a function jointly continuous in its two arguments. If the function $a(x, y)$ does not increase too fast, then it is clear that $(Z_t^\varepsilon)_{t\geq 0}$ converges in any reasonable sense to the constant x_0. Thus, one may ask the following question:

How to transform the time-parameter t in order to get a nondegenerate limit function? To this end, it is convenient to put

$$X_t^\varepsilon = Z_{t/\varepsilon}^\varepsilon, \quad t \geq 0.$$

Then the equation (1.1) becomes

$$\dot{X}_t^\varepsilon = a(X_t^\varepsilon, \xi_{t/\varepsilon}), \quad X_0^\varepsilon = x_0, \quad t \geq 0. \tag{1.2}$$

Indeed, the behaviour of Z_t^ε on time intervals of order ε^{-1} is usually of main interest because over such intervals occur significant changes in system (1.1). For example, exit from the neighbourhood of an equilibrium position or of a periodic trajectory. The study of the system (1.1) on interval of the form $[0, T/\varepsilon]$ is equivalent to that of system (1.2) on finite interval $[0, T]$.

The averaging principle consists in showing that the solution X^ε of the perturbed system can be approximated by the solution of some unperturbed system as ε goes to zero.

Although the averaging principle has long been applied to problems of celestial mechanics, oscillation theory and radiophysics (see Arnold 1975), no mathematically rigorous justification of it had existed for a long time. The first general result in this area was obtained by Bogolyubov (1945), Gikhman (1952), Krasnosel'skii and Krein (1955) who studied the averaging principle for ordinary differential equation:

$$\dot{X}_t^\varepsilon = F(X_t^\varepsilon, t/\varepsilon), \; X_0^\varepsilon = x_0. \tag{1.3}$$

They proved the following: If the function $F(x,t)$ is Lipchitz continuous in x uniformly with respect to t, bounded and

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T F(x,t)dt = \bar{F}(x) \tag{1.4}$$

exists uniformly in x, then the solution X^ε of (1.3) converges to the solution $\overline{X}$ of the unpertubed system

$$\dot{\overline{X}}_t = \bar{F}(\overline{X}_t), \; \overline{X}_0 = x_0.$$

We refer the reader to Bogolyubov and Zubarev (1955), Bogolyubov and Mitropolskii (1961), Volosov (1962), Neishtadt (1975, 1976) and Sanders and Verhulst (1985), for a detail survey of these results and many extensions.

After the pioneers, several authors dealt with the averaging problem. Thus, nowdays, several extensions and connected problems can be found in the litterature. For example, stochastic versions of Bogolyubov classical averaging principle has been developed. Freidlin and Wentzell (1979) treated the averaging problem for stochastic differential equations without drift

$$\dot{X}_t^\varepsilon = a(X_t^\varepsilon, \xi_{t/\varepsilon}), \quad X_0^\varepsilon = x_0, \quad t \geq 0,$$

where $(\xi_t)_{t\geq 0}$ is a stochastic process. Under a convergence such as that in (1.4), they showed that the solution of this pertubed equation can be approximated in some sense by the non-random solution $\overline{X}$ of the averaged equation

$$\dot{\overline{X}}_t = \bar{a}(\overline{X}_t), \quad \overline{X}_0 = x_0.$$

More precisely,they proved that $\sup_{0\leq t\leq T} |X_t^\varepsilon - \overline{X}_t|$ converges in probability to zero as $\varepsilon \longrightarrow 0$ if

$$\frac{1}{T}\int_0^T a(x,\xi_t)dt \xrightarrow{\mathbb{P}} \bar{a}(x) \quad \text{as } T \to +\infty.$$

Liptser and Stoyanov (1990) generalized this result to stochastic differential equations with diffusion:

$$dX_t^\varepsilon = a(X_t^\varepsilon, \xi_{t/\varepsilon})dt + \sqrt{\varepsilon}b(X_t^\varepsilon)dW_{t/\varepsilon}, \quad X_0^\varepsilon = x_0$$

where W is a Wienner process. This kind of asymptotic behaviour of the pertubed system can be viewed as a weak law of large numbers. So it was natural to search for a strong large numbers theorem. Indeed, later on, the convergence in probability was improved to almost sure convergence by Heunis and Kouritzin (1994) for stochastic differential equations without drift and Hashemi and Heunis (1998) for those with diffusion. Let us note that a functional central limit type theorem was obtained by Khasminskii (1966) who studied the limit behaviour of $\varepsilon^{-\frac{1}{2}}(X^\varepsilon - \overline{X})$.

The first work on the averaging principle for partial differential equations is due to Khasminskii (1963). The case of stochastic partial differential equations was treated by Makhno (1980) and Bondarev (1990) among others. Recently, averaging principle for random operators received a great attention (see Campillo *et al.* 2001, Pardoux and Piatnitski 2001, Kleptsyna and Piatnitski 1997, 2002).

Lipschitz condition in x uniformly with respect to y :

$$|a(x,y) - a(x',y)| \leq L|x - x'|.$$

Let us assume that

$$\lim_{T \longrightarrow \infty} \frac{1}{T} \int_0^T a(x, \xi_s) ds = \bar{a}(x) \tag{1.6}$$

exists for every $x \in \mathbb{R}^d$. This is the case, for example, if ξ_t is periodic or is a sum of periodic functions.

We are going to give an heuristic proof of the averaging principle (see Freidlin and Wentzell 1984).

The displacement of the trajectory X_t^ε over a small time Δ can be written in the form

$$\begin{aligned} X_\Delta^\varepsilon - x &= \int_0^\Delta a(X_s^\varepsilon, \xi_{s/\varepsilon}) ds \\ &= \int_0^\Delta a(x, \xi_{s/\varepsilon}) ds + \int_0^\Delta [a(X_s^\varepsilon, \xi_{s/\varepsilon}) - a(x, \xi_{s/\varepsilon})] ds \\ &= \Delta \left(\frac{\varepsilon}{\Delta} \int_0^{\Delta/\varepsilon} a(x, \xi_s) ds \right) + \rho_\varepsilon(\Delta). \end{aligned}$$

In view of (1.6), the coefficient of Δ in the first term of the right member converges to $\bar{a}(x)$ as ε/Δ goes to zero. Since the function $a(x,y)$ is bounded, we have $|\rho_\varepsilon(\Delta)| \leq K\Delta^2$. It follows that the displacement of the trajectory X_t^ε over a small time differs from the displacement of the trajectory $\overline{X}_t$ of the differential equation

$$\dot{\overline{X}}_t = \bar{a}(\overline{X}_t), \quad \overline{X}_0 = x$$

only by an infinitely small quantity compared to Δ as $\Delta \to 0$, $\varepsilon/\Delta \to 0$.

Now, let $(\xi_t)_{t \geq 0}$ be a stochastic process. We assume that there exists a vector field $\bar{a}(x)$ such that

$$\frac{1}{T} \int_0^T a(x, \xi_s) ds \xrightarrow{\mathbb{P}} \bar{a}(x), \quad \text{as } T \to +\infty. \tag{1.7}$$

Let us note that (1.7) is satified if the process $(\xi_t)_{t\geq 0}$ is ergodic. In this case there exists a probability measure ν on $\mathbb{R}^d$ such that

$$\bar{a}(x) = \int_{\mathbb{R}^d} a(x,y)\nu(dy).$$

If in addition, $(\xi_t)_{t\geq 0}$ is a strictly stationnary process then

$$\bar{a}(x) = \mathbb{E}\left[a(x,\xi_s)\right] = \mathbb{E}\left[a(x,\xi_0)\right].$$

Under the convergence (1.7), we have

$$\sup_{t\leq T}\left|X_t^\varepsilon - \overline{X}_t\right| \xrightarrow{\mathbb{P}} 0, \qquad \text{as } \varepsilon \to 0. \tag{1.8}$$

Let us give a sketch of proof for (1.8). To this end let us put

$$\Delta_t^\varepsilon = \sup_{s\leq t}|X_s^\varepsilon - \overline{X}_s|.$$

We have

$$\begin{aligned}
\Delta_t^\varepsilon &\leq \int_0^t \left|a(X_s^\varepsilon,\xi_{s/\varepsilon}) - a(\overline{X}_s,\xi_{s/\varepsilon})\right| ds + \left|\int_0^t [a(\overline{X}_s,\xi_{s/\varepsilon}) - \overline{a}(\overline{X}_s)]ds\right| \\
&\leq K\int_0^t \Delta_s^\varepsilon ds + \sup_{s\leq t}\left|\int_0^s [a(\overline{X}_u,\xi_{u/\varepsilon}) - \overline{a}(\overline{X}_u)]du\right|.
\end{aligned}$$

By virtue of Gronwall inequality, we deduce that

$$\sup_{t\leq T}\left|X_t^\varepsilon - \overline{X}_t\right| \leq e^{LT}\sup_{t\leq T}\left|\int_0^t [a(\overline{X}_s,\xi_{s/\varepsilon}) - \overline{a}(\overline{X}_s)]ds\right|.$$

The end of the proof is based on the fact that (see Freidlin and Wentzell 1984, p. 48-51 or Liptser and Shiryaev 1989):

$$\sup_{t\leq T}\left|\int_0^t [a(\overline{X}_s,\xi_{s/\varepsilon}) - \overline{a}(\overline{X}_s)]ds\right| \xrightarrow{\mathbb{P}} 0, \qquad \text{as } \varepsilon \to 0.$$

Heunis and Kouritzin (1994) have improved the convergence in (1.8). They considered a random differerential equation

$$\dot{X}_t^{\varepsilon} = F\left(X_t^{\varepsilon}, \frac{t}{\varepsilon}\right), \qquad X_0^{\varepsilon} = x_0$$

where $\{F(x,t), t \geq 0\}$ is a strong mixing process for each x and for each ω, the function $(x,t) \longrightarrow F(x,t,\omega)$ is regular enough to ensure the existence of a unique solution of this equation over $0 \leq t \leq T$ for all $\varepsilon > 0$. Under some conditions and replacing (1.6) by

$$\lim_{T \longrightarrow \infty} \frac{1}{T} \int_0^T \mathbb{E}[F(x,t)]dt = \overline{F}(x),$$

they showed that

$$\sup_{t \leq T} \left| X_t^{\varepsilon} - \overline{X}_t \right| \stackrel{a.s.}{\rightarrow} 0, \qquad \text{as } \varepsilon \rightarrow 0$$

where $\overline{X}$ is the solution of the non-random averaged differential equation

$$\dot{\overline{X}}_t = \overline{F}(\overline{X}_t), \qquad \overline{X}_0 = x_0.$$

Now we give an application of the averaging principle in oscillation theory.

Example: Van Der Pol equation :

Let us consider the Van Der Pol equation :

$$\ddot{x}_t + \omega^2 x_t = \varepsilon(1 - x_t^2)\dot{x}_t$$

where ε is a small numerical parameter.

This equation describes correctly the behavior of some electronic oscillators. When there is not a perturbation ($\varepsilon = 0$), we obtain the equation of a harmonic oscillator

$$\ddot{x}_t + \omega^2 x_t = 0.$$

The solutions of this equation are $x_t = r\cos(\omega t + \theta)$.

Let us put $y_t = \dot{x}_t$, then we have $y_t = -r\omega\sin(\omega t + \theta)$. So in the phase plane $(x, \dot{x})$, the solutions of this equation are the ellipses $x_t = r\cos(\omega t + \theta)$, $\dot{x}_t = -r\omega\sin(\omega t + \theta)$. In this case the phase point rotates with constant angular velocity ω and the amplitude r does not change with time, it is determined only by the initial conditions.

In the case with perturbation ($\varepsilon \neq 0$), let us put $f(x_t, \dot{x}_t, t) = (1 - x_t^2)\dot{x}_t$. Here, in general r and θ are not constant. Nevertheless, one may expect that the rate of change of them is small provided that ε is small.

We have

$$\begin{cases} \dot{r}_t^\varepsilon &= \varepsilon F_1(\psi_t, r_t^\varepsilon, t), \quad r_0^\varepsilon = r_0 \\ \dot{\theta}_t^\varepsilon &= \varepsilon F_2(\psi_t, r_t^\varepsilon, t), \quad \theta_0^\varepsilon = \theta_0. \end{cases} \tag{1.9}$$

where $\psi_t = \omega t + \theta_t^\varepsilon$ and

$$F_1(s, r, t) = -\frac{1}{\omega} f(r\cos s, -r\sin s, t)\sin s, \;\; F_2(s, r, t) = -\frac{1}{r\omega} f(r\cos s, -r\sin s, t)\cos s.$$

Therefore, in the van der Pol variables (r, θ) the Van Der Pol equation can be written as the form (1.5). If $f(x, y, t)$ does not depend explicitly on t, then $F_1(\omega t + \theta, r)$ and $F_2(\omega t+\theta, r)$ are periodic in t and condition (1.6) is satisfied. Then, the averaging principle is applicable to system (1.9). We refer the reader to Bogolyubov and Mitropolski (1961) for details on this example.

An averaging principle can be formulated in a more general situation. For example, Bogolyubov and Mitropolski (1961), Volosov (1962) and Neishtadt (1975, 1976) considered systems of the type

$$\begin{aligned} \dot{X}_t^\varepsilon &= b_1(X_t^\varepsilon, \xi_t^\varepsilon), \qquad X_0^\varepsilon = x \\ \dot{\xi}_t^\varepsilon &= \varepsilon^{-1} b_2(X_t^\varepsilon, \xi_t^\varepsilon), \quad \xi_0^\varepsilon = y. \end{aligned} \tag{1.10}$$

The velocity of the trajectory ξ_t has order ε^{-1} as $\varepsilon \to 0$. Therefore, the motion $(\xi_t)_{t\geq 0}$ is said to be fast. $(X_t^\varepsilon)_{t\geq 0}$ is called the slow motion.

1.1.2 Averaging Principle in Models with Diffusion.

Now, we consider stochastic differential equations with diffusion that is equations of the form

$$Z_t^\varepsilon = x_0 + \varepsilon \int_0^t a(Z_s^\varepsilon, \xi_s)ds + \sqrt{\varepsilon}\int_0^t b(Z_s^\varepsilon)dW_s$$

Let us put $X_t^\varepsilon = Z_{t/\varepsilon}^\varepsilon$, then the process $X^\varepsilon = (X_t^\varepsilon)_{t\geq 0}$ is solution of the equation

$$X_t^\varepsilon = x_0 + \int_0^t a(X_s^\varepsilon, \xi_{s/\varepsilon})ds + \int_0^t b(X_s^\varepsilon)dW_s^\varepsilon$$

where $W^\varepsilon = \sqrt{\varepsilon}W_{\cdot/\varepsilon}$.

We assume that $(\xi_t)_{t\geq 0}$ is a strictly stationary and ergodic process independent of the Wiener process W, and the functions $a(z,x)$, $b(x)$ are measurable satisfying linear growth and Lipschitz conditions in x uniformly with respect to z.

Under the assumptions given above, Liptser and Stoyanov (1990) proved that for any fixed $T > 0$

$$\sup_{t\leq T}\left|X_t^\varepsilon - \overline{X}_t^\varepsilon\right| \xrightarrow{\mathbb{P}} 0 \quad \text{as } \varepsilon \longrightarrow 0, \tag{1.11}$$

where the process $(\bar{X}_t^\varepsilon)_{t\geq 0}$ is the unique solution of the averaged equation

$$\overline{X}_t^\varepsilon = x_0 + \int_0^t \overline{a}(\overline{X}_s^\varepsilon)ds + \int_0^t b(\overline{X}_s^\varepsilon)dW_s^\varepsilon$$

with $\overline{a}(x) = \mathbb{E}(a(x,\xi_0))$.

Let us note that for each $\varepsilon > 0$ the process $\left(\overline{X}_t^\varepsilon\right)_{t\geq 0}$ coincides in the sense of distributions with the process $(X_t)_{t\geq 0}$ which is the unique solution of the equation

$$X_t = x_0 + \int_0^t \overline{a}(X_s)ds + \int_0^t b(X_s)dW_s.$$

Under additional assumptions on the functions a and b and on the process ξ, they also proved that the process $Y_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}}(X_t^\varepsilon - \overline{X}_t^\varepsilon)$ converges in the distribution sense as $\varepsilon \longrightarrow 0$ to a stochastic process Y which is solution of the following equation

$$Y_t = \int_0^t \bar{a}'(X_s)Y_s ds + \int_0^t b'(X_s)Y_s dW_s + \int_0^t \gamma(X_s)d\overline{W}_s,$$

where $\overline{W}$ is some Wiener process, $\bar{a}'(x) = E\left[a'_x(x,\xi_0)\right]$, $a'$and $b'$are respectively the derivatives of a and b, and

$$\gamma^2(x) = 2\int_0^{+\infty} \mathbb{E}[(a(x,\xi_t) - \bar{a}(x))(a(x,\xi_0) - \bar{a}(x))]dt.$$

This second result provides a second-order approximation of the process X^ε and can be regard as a functional central limit theorem.

Later, by strengthening the ergodicity condition on ξ to that of strong mixing, Hashemi and Heunis (1998) showed that the convergence in probability can be improved to almost sure convergence. When $b \equiv 0$, such a result is in Liptser and Shiryaev (1989) but with a stationary ergodic process ξ.

Let us note that it is possible to pertube also the diffusion coefficient by the process ξ. In this case, one can prove only weak convergence rather than convergence in probability (see Kashminskii 1966,1968) or the books of Freidlin and Wentzell (1984, pages 263-269) and Freidlin (1985, section 4.3).

One can also consider more general systems as in (1.10) but with diffusion as well in the slow motion as the fast one (see Veretennikov 1991). The averaging principle for Volterra equations was studied by Kleptsyna (1996).

1.2 Averaging Principle for Stochastic Partial Differential Equations

In this section we study the behavior as $\varepsilon \longrightarrow 0$ of solution of boundary value problems for elliptic or parabolic differential equations with a small parameter.

Let us consider for example, the Cauchy problem

$$\frac{\partial u}{\partial s} + \varepsilon \left[\sum_{i,j=1}^{d} a_{ij}(x,s) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x,s) \frac{\partial u}{\partial x_i} + c(x,s)u + d(x,s) \right] = 0 \qquad (1.12)$$

in the region $\mathbb{R}^d \times [0, +\infty)$. Let

$$\begin{aligned} \overline{a}_{ij}(x) &= \lim_{T \longrightarrow \infty} \frac{1}{T} \int_0^T a_{ij}(x,s)ds, \quad \overline{b}_i(x) = \lim_{T \longrightarrow \infty} \frac{1}{T} \int_0^T b_i(x,s)ds \\ \overline{c}(x) &= \lim_{T \longrightarrow \infty} \frac{1}{T} \int_0^T c(x,s)ds, \quad \overline{d}(x) = \lim_{T \longrightarrow \infty} \frac{1}{T} \int_0^T d(x,s)ds \end{aligned} \qquad (1.13)$$

and consider the average equation

$$\frac{\partial u}{\partial s} + \sum_{i,j=1}^{d} \overline{a}_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \overline{b}_i(x) \frac{\partial u}{\partial x_i} + \overline{c}(x)u + \overline{d}(x) = 0. \qquad (1.14)$$

Let the coefficients of these equations satisfy conditions:

(C1). The matrix $((a_{ij}))$ is non-negative definite for $(x,s) \in \mathbb{R}^d \times [0, +\infty)$, and all coefficients are continuous with respect to (x,s), are bounded for $s > 0$ and are sufficiently smooth so that solutions to equations (1.12) and (1.14) exist.

(C2). The limits in (1.13) are uniform in x

(C3). All coefficients are uniformly continuous in x with respect to $(x,s) \in \mathbb{R}^d \times [0, +\infty)$.

(C4). A solution of the Cauchy problem (1.14) exists.

Khasminskii (1963) proved the following

Theorem 1.1 *Let conditions (C1)-(C4) be satisfied, let $u_\varepsilon(x,s)$ be a solution of equation (1.12) in the region $\mathbb{R}^d \times (0, T/\varepsilon)$, satisfying the condition*

$$\lim_{s \to T/\varepsilon} u_\varepsilon(x,s) = f(x),$$

where $f(x)$ is a continuous bounded function in $\mathbb{R}^d$. Let $v(x,s)$ be a solution of equation (1.14) in the region $\mathbb{R}^d \times (0,T)$ satisfying the condition $v(x,T) = f(x)$. Then

$$\lim_{\varepsilon \to 0} \sup_{(x,s) \in \mathbb{R}^d \times (0,T)} \left| u_\varepsilon\left(x, \frac{s}{\varepsilon}\right) - v(x,s) \right| = 0.$$

Let us mention that other versions of the Averaging Principle for stochastic partial differential equation have been studied:

Makhno (1980) and Bondarev (1990) studied stochastic PDEs: For any $(t,x) \in (0,+\infty) \times \mathbb{R}^d$

$$\begin{cases} \dfrac{\partial X^\varepsilon}{\partial t}(t,x) &= \varepsilon[L_{t,x}X^\varepsilon(t,x)dt + A(t,x,X^\varepsilon(t,x))dt + \sum_{i=1}^d \sigma_i(t,x,X^\varepsilon(t,x))dW_i(t)] \\ X^\varepsilon(0,x) &= \varphi(x) \end{cases} \tag{1.15}$$

where ε is a small parameter, and

$$L_{t,x}u = \sum_{i,j=1}^d a_{ij}(t,x)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(t,x)\frac{\partial u}{\partial x_i} + C(t,x)u.$$

The averaging principle for equation (1.15) was justified in the first paper and an exponential estimate for the deviations of the solution to PDE (1.15) from the solution of the determinate averaged equation is obtained in the last one.

Let us note that the averaged equation is

$$\begin{cases} \dfrac{\partial U}{\partial t}(t,x) &= \overline{L}_x U + \overline{A}(x, U(t,x)) \\ U(0,x) &= \varphi(x) \end{cases}$$

where

$$\overline{L}_x u = \sum_{i,j=1}^{d} \overline{a}_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \overline{b}_i(x) \frac{\partial u}{\partial x_i} + \overline{C}(x) u$$

with the coefficients defined as in (1.13).

Recently, averaging principle for random operators received great attentions. For example, Kleptsyna and Piatnitski (2002) considered non-selfadjoint parabolic equations with random evolution. They dealt with parabolic operators involving rapidly oscillating random in time and periodic in spatial variables coefficients that is the Cauchy problem

$$\begin{aligned} \frac{\partial u^\varepsilon}{\partial t}(x,t) &= \frac{\partial a_{ij}}{\partial x_i}\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^2}\right) \frac{\partial u^\varepsilon}{\partial x_j}(x,t) + \frac{1}{\varepsilon} b_i \left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^2}\right) \frac{\partial u^\varepsilon}{\partial x_i}(x,t) \\ u^\varepsilon(x,0) &= u_0(x) \end{aligned}$$

where $(\xi_t)_{t\geq 0}$ is a stationary ergodic process taking values in $\mathbb{R}^d$.

This kind of equations was previously investigated by Campillo et al. (2001) who considered the asymptotic behaviour of the solution of the following cauchy problem: for any $(x,t) \in \mathbb{R}^d \times [0,T]$

$$\begin{cases} \dfrac{\partial u^\varepsilon}{\partial t}(t,x) &= \operatorname{div}\left[a(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}) \nabla u^\varepsilon(t,x)\right] + \dfrac{1}{\varepsilon^{1\wedge\frac{\alpha}{2}}} c(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}) u^\varepsilon(t,x) \\ u^\varepsilon(0,x) &= u_0(x) \end{cases}$$

where α is a parameter.

The case of nonlinear operator was treated by Pardoux and Piatnitski (2001). Indeed, they studied the equation

$$\frac{\partial u^\varepsilon(t,x)}{\partial t} = \frac{\partial a_{ij}}{\partial x_i}\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^2}\right) \frac{\partial u^\varepsilon}{\partial x_j}(t,x) + \frac{1}{\varepsilon} g(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^2}, u^\varepsilon(t,x)), \quad u^\varepsilon(0,x) = u_0(x).$$

1.3 Averaging Principle for Backward Stochastic Differential Equations

Backward stochastic differential equations (BSDEs) has been introduced by Pardoux and Peng (1990) in order to give a probabilistic representation of solutions of nonlinear PDEs. Thanks to this connection between the two theories, one can expect that averaging principle for PDEs could be obtained via that of BSDEs. . It seems that the first result using this idea is due to Pardoux and Veretennikov (1996). They considered the averaging problem of BSDE's where the coefficient in front of the Brownian motion does not enter the nonlinear term. This corresponds to semilinear PDE's where the non-linear term is a function of the solution, not of its gradient.

More precisely, let us consider

$$\begin{aligned} dX_t^{1,\varepsilon} &= \varepsilon^{-1}F(X_t^{1,\varepsilon}, X_t^{2,\varepsilon})dt + G(X_t^{1,\varepsilon}, X_t^{2,\varepsilon})dt, \quad X_0^{1,\varepsilon} = x_0^1 \\ dX_t^{2,\varepsilon} &= \varepsilon^{-2}H(X_t^{2,\varepsilon})dt + \varepsilon^{-1}K(X_t^{2,\varepsilon})dW_t, \quad X_0^{2,\varepsilon} = x_0^2 \end{aligned}$$

where $X^{1,\varepsilon} \in \mathbb{R}^d, X^{2,\varepsilon} \in \mathbb{R}^l, F, G, H, K$ are measurable functions with values in $\mathbb{R}^d, \mathbb{R}^l$ and $\mathbb{R}^d \otimes \mathbb{R}^l$ correspondently, $(W_t)_{t\geq 0}$is an $l-$dimensional Wiener process.

Assume that the coefficients F, G, H, K are periodic (of period one in each direction) functions of the variable x_2, so that the the process $(X_t^{2,\varepsilon})_{t\geq 0}$ can be considered as taking values in the $l-$dimensional torus T^l.

We also make a serie of assumptions (see Pardoux and Veretennikov 1996) which imply among others things that the process $(X_t^{2,\varepsilon})_{t\geq 0}$ admits a unique invariant probability measure μ on T^l,

$$\int_{T^l} F(x_1, x_2)\mu(dx_2) = 0, \quad \forall x_1 \in \mathbb{R}^d,$$

and $X^{1,\varepsilon}$ converges in distribution to a d-dimensional diffusion processX^1 with generator

$$\overline{\pounds} = \frac{1}{2}\sum_{i,j=1}^{d} \overline{a}_{ij}(x_1)\frac{\partial^2}{\partial x_{1i}\partial x_{1j}} + \sum_{i=1}^{d} \overline{b}_i(x_1)\frac{\partial}{\partial x_{1i}}$$

where

$$\overline{a}_{ij}(x_1) = \int_{T^l} [F_i(x_1,x_2)J_j(x_1,x_2) + F_j(x_1,x_2)J_i(x_1,x_2)]\,\mu(dx_2),$$

$$\overline{b}_i(x_1) = \int_{T^l} G_i(x_1,x_2)\mu(dx_2) + \int_{T^l} \langle F(x_1,x_2), \nabla_{x_1} J_i(x_1,x_2)\rangle\,\mu(dx_2)$$

$$L_2 J_i(x_1,\cdot))(x_2) = F_i(x_1,x_2), \quad x_2 \in T^l,$$

L_2 being the inifinitesimal generator of the diffusion process $(X_t^{2,\varepsilon})_{t\geq 0}$ in case $\varepsilon = 1$.

Therefore, there exists a d-dimensional Brownian motion $\{B_t, t \geq 0\}$ such that

$$X_t^1 = x_0^1 + \int_0^t \overline{b}(X_s^1)ds + \int_0^t \overline{\sigma}(X_s^1)dB_s$$

where $\overline{\sigma}(x_1) = [\overline{a}(x_1)]^{\frac{1}{2}}$.

By using averaging principle for the semi-linear BSDE

$$Y_t^\varepsilon = g(X_T^{1,\varepsilon}) + \int_t^T f(X_s^{1,\varepsilon}, X_s^{2,\varepsilon}, Y_s^\varepsilon)ds - \int_t^T Z_s^\varepsilon dW_s \tag{1.16}$$

the authors solved an averaging problem for the semi-linear PDE

$$\begin{aligned} \frac{\partial u^\varepsilon}{\partial t}(t,x) &= \pounds_\varepsilon u^\varepsilon(t,x) + f(x,u(t,x)), \quad 0 \leq t \leq T, \quad x \in \mathbb{R}^{d+l} \\ u^\varepsilon(0,x) &= g(x_1).\, x \in \mathbb{R}^{d+l}, \end{aligned} \tag{1.17}$$

where $\pounds_\varepsilon$ denote the infinitesimal generator of the diffusion process $(X^{1,\varepsilon}, X^{2,\varepsilon})_{t\geq 0}$ i.e.

$$\pounds_\varepsilon = \varepsilon^{-2}L_2 + \sum_{i=1}^{d}\left[\varepsilon^{-1}F_i(x) + G_i(x)\right]\frac{\partial}{\partial x_{1i}}.$$

Let us note that the corrresponding averaged equations for (1.16) and (1.17)are

$$Y_t = g(X_T^1) + \int_t^T \overline{f}(X_s^1, Y_s)ds - \int_t^T Z_s dB_s$$

and

$$\begin{aligned} \frac{\partial u}{\partial t}(t,x) &= \overline{\mathcal{L}}u(t,x) + \overline{f}(x, u(t,x)),\ 0 \leq t \leq T,\ x \in \mathbb{R}^d \\ u(0,x) &= g(x), x \in \mathbb{R}^d \end{aligned}$$

where

$$\overline{f}(x_1, y) = \int_{T^l} f(x_1, x_2, y)\mu(dx_2).$$

Chapter 2

AVERAGING PRINCIPLE FOR MULTIVALUED STOCHASTIC DIFFERENTIAL EQUATIONS[1]

Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$ be the underlying stochastic basis, $W = \{W_t = (W_t^1, \cdots, W_t^d) : t \geq 0\}$ a $d-$dimensional standard Wiener process and $\xi = \{\xi_t : t \geq 0\}$ a $r-$dimensional strictly stationary ergodic process independent of W. Liptser and Stoyanov(1990) studied the limit behaviour of the family of stochastic processes $(x^\varepsilon : \varepsilon > 0)$, as $\varepsilon \longrightarrow 0$, where x^ε is the solution of an Itô's stochastic differential equation

$$x_t^\varepsilon = x_0 + \varepsilon \int_0^t a(x_s^\varepsilon, \xi_s)ds + \varepsilon^{1/2} \int_0^t b(x_s^\varepsilon)dW_s, \quad 0 \leq t \leq \varepsilon^{-1}. \tag{2.1}$$

More precisely, under some regularity conditions on the coefficients a and b, they proved that $\sup_{0\leq t\leq\varepsilon^{-1}} |x_t^\varepsilon - \overline{x}_t^\varepsilon|$ converges in probability towards zero as $\varepsilon \longrightarrow 0$, where $\overline{x}^\varepsilon$ is the solution of the Itô's stochastic differential equation obtained from the equation (2.1) by averaging out the fluctuations in the drift term arising from the stochastic process ξ:

[1] *Published in Random Operators and Stochastic Equations 9(4). 399-407.*

$$\overline{x}_t^\varepsilon = x_0 + \varepsilon \int_0^t \overline{a}(\overline{x}_s^\varepsilon)ds + \varepsilon^{1/2} \int_0^t b(\overline{x}_s^\varepsilon)dW_s,\ 0 \le t \le \varepsilon^{-1},$$

with $\overline{a}(x) = \mathbb{E}\left(a(x,\xi_0)\right)$.

The above result can be considered as stochastic version of the classical Bogolyubov averaging principle(see Bogolyubov and Mitropol'skii(1961)).

In this paper we consider multivalued stochastic differential equations

$$\begin{cases} dy_t^\varepsilon + \varepsilon A(y_t^\varepsilon)dt \ni \varepsilon a(y_t^\varepsilon, \xi_t)dt + \varepsilon^{1/2} b(y_t^\varepsilon)dW_t, \quad 0 \le t \le \varepsilon^{-1} \\ \\ y_0^\varepsilon = y_0 \end{cases} \tag{2.2}$$

$$\begin{cases} d\overline{y}_t^\varepsilon + \varepsilon A(\overline{y}_t^\varepsilon)dt \ni \varepsilon \overline{a}(\overline{y}_t^\varepsilon)dt + \varepsilon^{1/2} b(\overline{y}_t^\varepsilon)dW_t, \quad 0 \le t \le \varepsilon^{-1} \\ \\ \overline{y}_0^\varepsilon = y_0 \end{cases} \tag{2.3}$$

where A is a maximal monotone multivalued operator on $\mathbb{R}^d$.

Our goal is to study the asymptotic behaviour of $\sup_{0\le t\le\varepsilon^{-1}} |y_t^\varepsilon - \overline{y}_t^\varepsilon|$ as ε goes to zero.

The paper is organized as follows. In section 2.1 we give some notations and make assumptions used throughout. In particular, we recall some definitions and results about multivalued maximal monotone operators. Section 2.2 is devoted to the result.

2.1 Notations and Assumptions

For any matrix B we put $||B||^2 = trace(BB^t)$ where B^t stands for the transpose of B. We shall need the following assumptions .

(H1) $a : \mathbb{R}^d \times \mathbb{R}^r \longrightarrow \mathbb{R}^d$ and $b : \mathbb{R}^d \longrightarrow \mathbb{R}^d \otimes \mathbb{R}^d$ are Borel measurable mappings such that there exists a constant $L > 0$ satisfying

$$||a(x,z) - a(x',z)|| \leq L||x - x'||, \quad ||b(x) - b(x')|| \leq L||x - x'||$$

$$||a(x,z)|| \leq L(1 + ||x||), \; ||b(x)|| \leq L(1 + ||x||)$$

for all $x, x' \in \mathbb{R}^d$ and $z \in \mathbb{R}^r$

(H2) the initial condition $y_0 \in \mathbb{R}^d$ is deterministic

(H3) $\{\xi_t : t \geq 0\}$ is an $\mathbb{R}^r-$valued, $\mathcal{F}_t-$progressively measurable process

(H4) W and ξ are two independent processes

(H5) ξ is a strictly stationary process

Now, let us state some definitions and results about multivalued operators on $\mathbb{R}^d$.

Definition 2.2 *A multivalued operator A on $\mathbb{R}^d$ is an $\mathcal{P}(\mathbb{R}^d)-$valued map defined on $\mathbb{R}^d$, where $\mathcal{P}(\mathbb{R}^d)$ stands for the collection of subsets of $\mathbb{R}^d$.*

We respectively denote by $Gr(A)$ and $D(A)$ the graph and the domain of A :

$$Gr(A) = \{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d : y \in A(x)\}$$

$$D(A) = \{x \in \mathbb{R}^d : A(x) \neq \emptyset\}.$$

Definition 2.3 *A multivalued operator A on $\mathbb{R}^d$ is said monotone if*

$$\forall (x_1, y_1), (x_2, y_2) \in Gr(A) \qquad \langle x_1 - x_2, y_1 - y_2 \rangle \geq 0,$$

where $\langle \bullet, \bullet \rangle$ is the Euclidian inner product on $\mathbb{R}^d$.

Definition 2.4 *A multivalued monotone operator A on $\mathbb{R}^d$ is said maximal monotone if there is no multivalued monotone operator whose graph strictly contains the graph of*

A,that is: for every $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ *such that* $\langle x-u, y-v\rangle \geq 0$ *for all* $(u, v) \in Gr(A)$, *we have* $y \in A(x)$.

Now, let us introduce the Yosida approximation of A which is a sequence $(A_n)_{n\geq 1}$ of one valued maximal monotone maps defined on $\mathbb{R}^d$ by :

$$A_n = n(I - J_n) \text{ where } J_n = (I + \frac{1}{n}A)^{-1} \text{ and } I \text{ is the identity operator on } \mathbb{R}^d.$$

We have

- For all $n \geq 0$, A_n is a Lipschitz continuous map.
- For all $x \in D(A)$, the set $A(x)$ is closed and convex and hence there exists a unique point $A^0(x)$, such that $|A^0(x)| = \min\{|y| : y \in A(x)\}$.
- For all $x \in D(A)$, $|A_n(x)| \leq |A^0(x)|$ and $A_n(x) \longrightarrow A^0(x)$ as $n \longrightarrow +\infty$.

For more details on multivalued maximal monotone operators, the reader can see the book of Brézis(1973).

From now on , A is a maximal monotone operator on $\mathbb{R}^d$ such that the interior of $D(A)$ is non empty. Under the assumptions (H1)–(H5), the multivalued stochastic differential equations (2.2) and (2.3) admit unique solutions which we denote respectively by $(y^\varepsilon, k^\varepsilon)$ and $(\overline{y}^\varepsilon, \overline{k}^\varepsilon)$ in the sense that:

- $\{ y_t^\varepsilon : t \geq 0\}$(resp. $\{ \overline{y}_t^\varepsilon : t \geq 0\}$) is a continuous $\mathcal{F}_t$−adapted process with values almost surely in the closure of $D(A)$.
- $\{k_t^\varepsilon : t \geq 0\}$(resp. $\{ \overline{k}_t^\varepsilon : t \geq 0\}$) is a continuous $\mathbb{R}^d$−values, $\mathcal{F}_t$−adapted process, with finte variation such that $k_0^\varepsilon = 0$(resp. $\overline{k}_0^\varepsilon = 0$) almost surely.
-

$$\begin{cases} dy_t^\varepsilon = \varepsilon a(y_t^\varepsilon, \xi_t)dt + \varepsilon^{1/2} b(y_t^\varepsilon)dW_t - \varepsilon dk_t^\varepsilon, \quad t \geq 0, \quad \text{a.s.} \\ \\ y_0^\varepsilon = y_0 \end{cases}$$

$$\left(\text{resp.} \begin{cases} d\overline{y}_t^\varepsilon = \varepsilon \overline{a}(\overline{y}_t^\varepsilon)dt + \varepsilon^{1/2} b(\overline{y}_t^\varepsilon)dW_t - \varepsilon d\overline{k}_t^\varepsilon, & t \geq 0, \quad \text{a.s.} \\ \overline{y}_0^\varepsilon = y_0 \end{cases} \right)$$

- For all $\mathcal{F}_t$−adapted couple of continuous processes (α, β) such that for all $t \geq 0$ $(\alpha_t, \beta_t) \in Gr(A)$, the measure $\langle y_t^\varepsilon - \alpha_t, dk_t^\varepsilon - \beta_t dt \rangle$ (resp. $\langle \overline{y}_t^\varepsilon - \alpha_t, d\overline{k}_t^\varepsilon - \beta_t dt \rangle$) is almost surely positive on $\mathbb{R}^+$.

The proof of existence and uniqueness of solutions of equations (2.2) and (2.3) is a straightforward adaptation of methods in Cépa(1994) or Pettersson(1995).

2.2 Result

Let us put

$$Y_t^\varepsilon = y_{t/\varepsilon}^\varepsilon, \quad \overline{Y}_t^\varepsilon = \overline{y}_{t/\varepsilon}^\varepsilon, \quad K_t^\varepsilon = \varepsilon k_{t/\varepsilon}^\varepsilon, \quad \overline{K}_t^\varepsilon = \varepsilon \overline{k}_{t/\varepsilon}^\varepsilon \text{ and } W_t^\varepsilon = \varepsilon^{1/2} W_{t/\varepsilon}.$$

One can prove that $(Y^\varepsilon, K^\varepsilon)$ (resp. $(\overline{Y}^\varepsilon, \overline{K}^\varepsilon)$) is the unique solution of the multivalued stochastic differential equation

$$\begin{cases} dY_t^\varepsilon + A(Y_t^\varepsilon)dt \ni a(Y_t^\varepsilon, \xi_{t/\varepsilon})dt + b(Y_t^\varepsilon)dW_t^\varepsilon, & 0 \leq t \leq 1, \quad \text{a.s.} \\ Y_0^\varepsilon = y_0 \end{cases} \tag{2.4}$$

$$\left(\text{resp.} \begin{cases} d\overline{Y}_t^\varepsilon + A(\overline{Y}_t^\varepsilon)dt \ni \overline{a}(\overline{Y}_t^\varepsilon)dt + b(\overline{Y}_t^\varepsilon)dW_t^\varepsilon, \ 0 \leq t \leq 1 \ , & \text{a.s.} \\ \overline{Y}_0^\varepsilon = y_0 \end{cases} \right) \tag{2.5}$$

For all $n \geq 1$ and $\varepsilon > 0$, let $Y^{\varepsilon,n}$ (resp. $\overline{Y}^{\varepsilon,n}$) be the unique solution of the stochastic differential equation

$$\begin{cases} dY_t^{\varepsilon,n} = a(Y_t^{\varepsilon,n}, \xi_{t/\varepsilon})dt + b(Y_t^{\varepsilon,n})dW_t^\varepsilon - A_n(Y_t^{\varepsilon,n})dt, \ 0 \le t \le 1 \\ \\ Y_0^{\varepsilon,n} = y_0 \end{cases}$$

$$\left(\text{resp.} \quad \begin{cases} d\overline{Y}_t^{\varepsilon,n} = \overline{a}(\overline{Y}_t^{\varepsilon,n})dt + b(\overline{Y}_t^{\varepsilon,n})dW_t^\varepsilon - A_n(\overline{Y}_t^{\varepsilon,n})dt, \ 0 \le t \le 1 \\ \\ \overline{Y}_0^{\varepsilon,n} = y_0 \end{cases} \right)$$

We begin by some preliminary results.

Lemma 2.5 *Under assumptions (H1)-(H5), we have*

$$\sup_{\varepsilon>0} \sup_{n\ge0} \mathbb{E}\left(\sup_{0\le t\le1} |Y_t^{\varepsilon,n}|^2 \right) < +\infty \tag{2.6}$$

and

$$\sup_{\varepsilon>0} \sup_{n\ge0} \mathbb{E}\left(\sup_{0\le t\le1} |\overline{Y}_t^{\varepsilon,n}|^2 \right) < +\infty. \tag{2.7}$$

Proof. For all $N \ge 0$, $\varepsilon > 0$ and $n \ge 1$, let us put

$$\tau_N^\varepsilon = \inf\{0 \le t \le 1 : |Y_t^{\varepsilon,n}| > N\}$$

with the convention $\inf \emptyset = +\infty$.

By virtue of Lemma 5.4 in Cépa(1994), there exists $\alpha \in \mathbb{R}^d$ and two positive constants β and μ such that for all $x \in \mathbb{R}^d, n \ge 1$

$$\langle A_n(x), x - \alpha \rangle \ge \beta |A_n(x)| - \mu |x - \alpha| - \beta\mu \tag{2.8}$$

Now, by applying Itô's formula to the process $|Y_{t\wedge\tau_N^\varepsilon}^{\varepsilon,n} - \alpha|^2$, we obtain

$$|Y_{t\wedge\tau_N^\varepsilon}^{\varepsilon,n} - \alpha|^2 = |Y_0 - \alpha|^2 + 2\int_0^{t\wedge\tau_N^\varepsilon} \langle a(Y_s^{\varepsilon,n}, \xi_{s/\varepsilon}), Y_s^{\varepsilon,n} - \alpha \rangle ds$$

$$+2\int_0^{t\wedge\tau_N^\varepsilon}\langle Y_s^{\varepsilon,n}-\alpha, b(Y_s^{\varepsilon,n})dW_s^\varepsilon\rangle - 2\int_0^{t\wedge\tau_N^\varepsilon}\langle A_n(Y_s^{\varepsilon,n}), Y_s^{\varepsilon,n}-\alpha\rangle ds$$

$$+\int_0^{t\wedge\tau_N^\varepsilon}||b(Y_s^{\varepsilon,n})||^2 ds.$$

By using the inequality (2.8) and assumption (H1), we derive that for all $0\le t\le 1$

$$|Y_{t\wedge\tau_N^\varepsilon}^{\varepsilon,n}-\alpha|^2 \le |Y_0-\alpha|^2 + 2L\int_0^{t\wedge\tau_N^\varepsilon}|Y_s^{\varepsilon,n}-\alpha|\,(1+|Y_s^{\varepsilon,n}|)ds$$

$$+2\left|\int_0^{t\wedge\tau_N^\varepsilon}\langle Y_s^{\varepsilon,n}-\alpha, b(Y_s^{\varepsilon,n})dW_s^\varepsilon\rangle\right| - 2\beta\int_0^{t\wedge\tau_N^\varepsilon}|A_n(Y_s^{\varepsilon,n})|ds$$

$$+2\mu\int_0^{t\wedge\tau_N^\varepsilon}|Y_s^{\varepsilon,n}-\alpha|\,ds + 2\beta\mu + L^2\int_0^{t\wedge\tau_N^\varepsilon}(1+|Y_s^{\varepsilon,n}|)^2 ds.$$

Now, the elementary inequality $xy\le \dfrac{x^2}{2}+\dfrac{y^2}{2}$ implies that there exists $C=C(\alpha,\mu,\beta,L)$ such that

$$|Y_{t\wedge\tau_N^\varepsilon}^{\varepsilon,n}-\alpha|^2 \le C\left(1+\int_0^{t\wedge\tau_N^\varepsilon}|Y_s^{\varepsilon,n}-\alpha|^2\,ds\right)+\left|\int_0^{t\wedge\tau_N^\varepsilon}\langle Y_s^{\varepsilon,n}-\alpha, b(Y_s^{\varepsilon,n})dW_s^\varepsilon\rangle\right|.$$

It follows that

$$\mathbb{E}\left(\sup_{0\le s\le t}|Y_{s\wedge\tau_N^\varepsilon}^{\varepsilon,n}-\alpha|^2\right)\le C\left(1+\int_0^t\mathbb{E}\left(\sup_{0\le u\le s}|Y_{u\wedge\tau_N^\varepsilon}^{\varepsilon,n}-\alpha|^2\right)ds\right)$$

$$+\mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^{s\wedge\tau_N^\varepsilon}\langle Y_u^{\varepsilon,n}-\alpha, b(Y_u^{\varepsilon,n})dW_u^\varepsilon\rangle\right|\right). \tag{2.9}$$

In view of Burkhölder-Gundy inequality for stochastic integrals in Barlow and Protter(1989), we have

$$\mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^{s\wedge\tau_N^\varepsilon}\langle Y_u^{\varepsilon,n}-\alpha, b(Y_u^{\varepsilon,n})dW_u^\varepsilon\rangle\right|\right)\le C\left(\mathbb{E}\left(\sup_{0\le s\le t}|Y_{s\wedge\tau_N^\varepsilon}^{\varepsilon,n}-\alpha|^2\right)\right)^{1/2}$$

$$\times \left(\mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^{s\wedge\tau_N^\varepsilon} b(Y_u^{\varepsilon,n})dW_u^\varepsilon\right|^2\right)\right)^{1/2}.$$

By virtue of Doob's inequality, we deduce that

$$\mathbb{E}\left(\sup_{0\le s\le t}\left(\left|\int_0^{s\wedge\tau_N^\varepsilon}\langle Y_u^{\varepsilon,n}-\alpha, b(Y_u^{\varepsilon,n})dW_u^\varepsilon\rangle\right|\right)\right) \le C\left(\mathbb{E}\left(\sup_{0\le s\le t}|Y_{s\wedge\tau_N^\varepsilon}^{\varepsilon,n}-\alpha|^2\right)\right)^{1/2}$$
$$\times\left(\mathbb{E}\left(\int_0^t ||b(Y_{u\wedge\tau_N^\varepsilon}^{\varepsilon,n})||^2 du\right)\right)^{1/2}.$$

So, it follows from (2.9) and assumption (H1) that

$$\mathbb{E}\left(\sup_{0\le s\le t}|Y_{s\wedge\tau_N^\varepsilon}^{\varepsilon,n}-\alpha|^2\right) \le C\left(1+\int_0^t \mathbb{E}\left(\sup_{0\le u\le s}|Y_{u\wedge\tau_N^\varepsilon}^{\varepsilon,n}-\alpha|^2\right)ds\right).$$

Now, Gronwall lemma implies that there exists a constant $C = C(\alpha,\mu,\beta,L)$ such that

$$\sup_{\varepsilon>0}\sup_{n\ge1}\left[\mathbb{E}\left(\sup_{0\le t\le 1}|Y_{t\wedge\tau_N^\varepsilon}^{\varepsilon,n}-\alpha|^2\right)\right] \le C < +\infty.$$

By letting $N \longrightarrow +\infty$, we conlude that

$$\sup_{\varepsilon>0}\sup_{n\ge1}\left[\mathbb{E}\left(\sup_{0\le t\le 1}|Y_t^{\varepsilon,n}-\alpha|^2\right)\right] < +\infty,$$

which yields (2.6). The inequality (2.7) can be derive by the same techniques. ■

We shall need the following integrability condition

$$\sup_{\varepsilon>0}\sup_{n\ge1}\mathbb{E}\left(\int_0^t |A_n(Y_s^{\varepsilon,n})|^2 ds\right) < +\infty. \tag{2.10}$$

Let us give some situations where condition (2.10) is satisfied.

a) If $D(A)=\mathbb{R}^d$ and $|A^0(x)| \le L(1+|x|)$ for all $x\in\mathbb{R}^d$, one can prove by using Lemma 4 that condition (2.10) is satisfied.

b) If A is the subdifferential operator of a proper lower semicontinuous convex function $\Phi : \mathbb{R}^d \longrightarrow [-\infty, +\infty)$ then by applying Itô's formula to the semimartingale $\{\Phi(Y_t^{\varepsilon,n}) : 0 \leq t \leq 1\}$ and using Lemma 4, one can prove that condition (2.10) is also satisfied.

Proposition 2.6 *Under assumptions (H1)-(H5), if we assume that condition (2.10) is satisfied, then there exists a positive constant C such that for all $n, m \geq 1$*

$$\sup_{\varepsilon>0} \mathbb{E}\left(\sup_{0\leq t\leq 1} |Y_t^{\varepsilon,n} - Y_t^{\varepsilon,m}|^2 \right) \leq C\left(\frac{1}{n} + \frac{1}{m}\right) \tag{2.11}$$

$$\sup_{\varepsilon>0} \mathbb{E}\left(\sup_{0\leq t\leq 1} |\overline{Y}_t^{\varepsilon,n} - \overline{Y}_t^{\varepsilon,m}|^2 \right) \leq C\left(\frac{1}{n} + \frac{1}{m}\right). \tag{2.12}$$

Proof. By virtue of Itô's formula, we have

$$\begin{aligned} |Y_t^{\varepsilon,n} - Y_t^{\varepsilon,m}|^2 = 2\int_0^t &\langle a(Y_s^{\varepsilon,n}, \xi_{s/\varepsilon}) - a(Y_s^{\varepsilon,m}, \xi_{s/\varepsilon}), Y_s^{\varepsilon,n} - Y_s^{\varepsilon,m}\rangle ds \\ &-2\int_0^t \langle A_n(Y_s^{\varepsilon,n}) - A_m(Y_s^{\varepsilon,m}), Y_s^{\varepsilon,n} - Y_s^{\varepsilon,m}\rangle ds \\ &+2\int_0^t \langle Y_s^{\varepsilon,n} - Y_s^{\varepsilon,m}, (b(Y_s^{\varepsilon,n}) - b(Y_s^{\varepsilon,m}))\, dW_s^{\varepsilon}\rangle \\ &+\int_0^t \|b(Y_s^{\varepsilon,n}) - b(Y_s^{\varepsilon,m})\|^2\, ds \end{aligned}$$

By using the fact that $I = J_n + \frac{1}{n}A_n = J_m + \frac{1}{m}A_m$, $A_n(Y_s^{\varepsilon,n}) \in A(J_n(Y_s^{\varepsilon,n}))$ and $A_m(Y_s^{\varepsilon,m}) \in A(J_m(Y_s^{\varepsilon,m}))$, we obtain

$$\begin{aligned} -\langle A_n(Y_s^{\varepsilon,n}) - A_m(Y_s^{\varepsilon,m}), Y_s^{\varepsilon,n} - Y_s^{\varepsilon,m}\rangle = -\langle &A_n(Y_s^{\varepsilon,n}) - A_m(Y_s^{\varepsilon,m}), J_n(Y_s^{\varepsilon,n}) \\ &+\frac{1}{n}A_n(Y_s^{\varepsilon,n}) - J_m(Y_s^{\varepsilon,m}) - \frac{1}{m}A_m(Y_s^{\varepsilon,m})\rangle \end{aligned}$$

$$= -\langle A_n(Y_s^{\varepsilon,n}) - A_m(Y_s^{\varepsilon,m}), J_n(Y_s^{\varepsilon,n}) - J_m(Y_s^{\varepsilon,m})\rangle$$

$$-\langle A_n(Y_s^{\varepsilon,n}) - A_m(Y_s^{\varepsilon,m}), \frac{1}{n}A_n(Y_s^{\varepsilon,n}) - \frac{1}{m}A_m(Y_s^{\varepsilon,m})\rangle$$

$$\le -\langle A_n(Y_s^{\varepsilon,n}) - A_m(Y_s^{\varepsilon,m}), \frac{1}{n}A_n(Y_s^{\varepsilon,n}) - \frac{1}{m}A_m(Y_s^{\varepsilon,m})\rangle$$

$$\le -\frac{1}{n}\left|A_n(Y_s^{\varepsilon,n})\right|^2 - \frac{1}{m}\left|A_m(Y_s^{\varepsilon,m})\right|^2$$

$$+\left(\frac{1}{n}+\frac{1}{m}\right)\langle A_n(Y_s^{\varepsilon,n}), A_m(Y_s^{\varepsilon,m})\rangle$$

$$\le \frac{1}{4n}\left|A_m(Y_s^{\varepsilon,m})\right|^2 + \frac{1}{4m}\left|A_n(Y_s^{\varepsilon,n})\right|^2.$$

Therefore, Burkhölder-Gundy inequality for stochastic integrals in Barlow and Protter(1989) and assumptions (H1)-(H5) imply that there exists a positive constant C such that for all $\varepsilon > 0, n, m \ge 1$

$$\mathbb{E}\left(\sup_{0\le s\le t}|Y_s^{\varepsilon,n} - Y_s^{\varepsilon,m}|^2\right) \le C\int_0^t \mathbb{E}\left(|Y_s^{\varepsilon,n} - Y_s^{\varepsilon,m}|^2\right)ds + \frac{1}{2}\mathbb{E}\left(\sup_{0\le s\le t}|Y_s^{\varepsilon,n} - Y_s^{\varepsilon,m}|^2\right)$$

$$+C\left(\frac{1}{n}+\frac{1}{m}\right).$$

It follows that for all $n, m \ge 1$

$$\sup_{\varepsilon>0}\mathbb{E}\left(\sup_{0\le s\le t}|Y_s^{\varepsilon,n} - Y_s^{\varepsilon,m}|^2\right) \le C\left(\int_0^t \sup_{\varepsilon>0}\mathbb{E}\left(\sup_{0\le u\le s}|Y_s^{\varepsilon,n} - Y_s^{\varepsilon,m}|^2\right)ds\right) + C\left(\frac{1}{n}+\frac{1}{m}\right).$$

Now, Gronwall lemma leads to (2.11). The proof of the inequality (2.12) can be done by a similar argument. ■

Proposition 2.7 *Under assumptions (H1)-(H5), if we assume that condition (2.10) is* *satisfied, then*

$$\lim_{n\longrightarrow\infty}\left[\sup_{\varepsilon>0}\mathbb{E}\left(\sup_{0\le t\le 1}|Y_t^{\varepsilon,n}-Y_t^{\varepsilon}|^2\right)\right]=0 \quad \textit{and} \quad \lim_{n\longrightarrow\infty}\left[\sup_{\varepsilon>0}\mathbb{E}\left(\sup_{0\le t\le 1}|\overline{Y}_t^{\varepsilon,n}-\overline{Y}_t^{\varepsilon}|^2\right)\right]=0.$$

Proof. It follows by proposition 2.6 and the uniqueness of solutions of multivalued stochastic differential equations (2.4) and (2.5). ■

Theorem 2.8 *Under assumptions (H1)-(H5), if we assume that condition (2.10) is satisfied, then for all* $\delta\ge 0$

$$\lim_{\varepsilon\longrightarrow 0}\mathbb{P}\left(\sup_{0\le t\le\varepsilon^{-1}}|y_t^{\varepsilon}-\overline{y}_t^{\varepsilon}|\ge\delta\right)=0.$$

Proof. us note that

$$\sup_{0\le t\le\varepsilon^{-1}}|y_t^{\varepsilon}-\overline{y}_t^{\varepsilon}|=\sup_{0\le t\le 1}|Y_t^{\varepsilon}-\overline{Y}_t^{\varepsilon}|.$$

By virtue of Proposition 2.7, for all $\delta>0$, we have

$$\lim_{n\longrightarrow\infty}\left(\sup_{\varepsilon>0}\mathbb{P}\left(\sup_{0\le t\le 1}|Y_t^{\varepsilon,n}-Y_t^{\varepsilon}|\right)\ge\delta\right)=0 \text{ and } \lim_{n\longrightarrow\infty}\left(\sup_{\varepsilon>0}\mathbb{P}\left(\sup_{0\le t\le 1}|\overline{Y}_t^{\varepsilon,n}-\overline{Y}_t^{\varepsilon}|\right)\ge\delta\right)=0.$$

Therefore for all $\alpha,\delta>0$, there exists $n_0\ge 0$ such that for all $\varepsilon>0$

$$\mathbb{P}\left(\sup_{0\le t\le 1}|Y_t^{\varepsilon,n_0}-Y_t^{\varepsilon}|\ge\frac{\delta}{3}\right)\le\frac{\alpha}{2} \text{ and } \mathbb{P}\left(\sup_{0\le t\le 1}|\overline{Y}_t^{\varepsilon,n_0}-\overline{Y}_t^{\varepsilon}|\ge\frac{\delta}{3}\right)\le\frac{\alpha}{2}.$$

Since A_{n_0} is Lipschitz, by virtue of Theorem 1 in Liptser and Stoyanov(1990),

$$\lim_{\varepsilon\longrightarrow 0}\mathbb{P}\left(\sup_{0\le t\le 1}|Y_t^{\varepsilon,n_0}-\overline{Y}_t^{\varepsilon,n_0}|\ge\frac{\delta}{3}\right)=0\ .$$

Now, by using the inequality

$$|Y_t^{\varepsilon}-\overline{Y}_t^{\varepsilon}|\le|Y_t^{\varepsilon}-Y_t^{\varepsilon,n_0}|+|Y_t^{\varepsilon,n_0}-\overline{Y}_t^{\varepsilon,n_0}|+|\overline{Y}_t^{\varepsilon,n_0}-\overline{Y}_t^{\varepsilon}|,$$

we deduce

$$\limsup_{\varepsilon \longrightarrow 0} \mathbb{P}\left(\sup_{0\leq t\leq 1} |Y_t^\varepsilon - \overline{Y}_t^\varepsilon| \geq \delta\right) \leq \alpha.$$

Since α is arbitrary, we conclude that

$$\lim_{\varepsilon \longrightarrow 0} \mathbb{P}\left(\sup_{0\leq t\leq 1} |Y_t^\varepsilon - \overline{Y}_t^\varepsilon| \geq \delta\right) = 0.$$

■.

Chapter 3

AVERAGING PRINCIPLE FOR DOUBLE ITÔ STOCHASTIC PROCESSES [1]

The averaging principle for dynamic sytems plays an important role in problems of celestial mechanics, oscillation theory, control theory, radiophysics, and many others areas. The first rigorous result on this subject was given by Bogolyubov (1945), who considered the system of ordinary differential equations

$$\dot{X}_t = \varepsilon F(t, X_t), \qquad X_0 = x_0.$$

He formulated a general principle according to which, for $\varepsilon \to 0$, a solution of this system on a time interval of length $O(1/\varepsilon)$ can be approximated arbitrarily closely by the solution of the averaged equation

$$\dot{\overline{X}}_t = \varepsilon \overline{F}(X_t), \qquad \overline{X}_0 = x_0,$$

[1] *Submitted for publication*

if the limit

$$\overline{F}(x) = \lim_{T \to +\infty} \frac{1}{T} \int_0^T F(t,x)dt$$

exists, and the function $F(t,x)$ is bounded and statisfies a Lipschitz condition with respect to the space variable.

Since its introduction, the averaging principle has attracted much attention of many authors. For example, Liptser and Stoyanov (1990) studied the asymptotic behaviour of an Itô's process $(x_t^\varepsilon)_{t\geq 0}$ whose drift is pertubed by an ergodic stionnary process $(\xi_t)_{t\geq 0}$:

$$x_t^\varepsilon = x_0 + \varepsilon \int_0^t a\left(x_s^\varepsilon, \xi_s\right) ds + \sqrt{\varepsilon} \int_0^t b(x_s^\varepsilon) dW_s, 0 \leq t \leq \varepsilon^{-1}. \tag{3.1}$$

More precisely, under some regularity conditions on the coefficients a and b, they proved that $\sup_{0\leq t\leq \varepsilon^{-1}} |x_t^\varepsilon - \overline{x}_t^\varepsilon|$ converges in probability towards zero as $\varepsilon \longrightarrow 0$, where $\overline{x}_t^\varepsilon$ is the solution of the Itô's stochastic differential equation, obtained from the equation (3.1) by averaging out the fluctuations in the drift term arising from the stochastic process ξ:

$$\overline{x}_t^\varepsilon = x_0 + \varepsilon \int_0^t \overline{a}\ \left(\overline{x}_t^\varepsilon\right) ds + \sqrt{\varepsilon} \int_0^t b(\overline{x}_t^\varepsilon) dW_s.$$

This result was generalized by Hashemi and Heunis (1998), who improved the convergence in probability to an almost surely convergence when the ergodicity hypothesis for the perturbing process $(\xi_t)_{t\geq 0}$ is strengthened to that of strong mixing.

On the other hand, the multiple stochastic integrals with respect to a particular class of martingales and for random integrands have been introduced by Meyer (1976). These integrals were extended by Ruiz de Chavez (1985) for a class of semimartingales. Recently, C.Tudor and M.Tudor (2002) considered Double Itô processes that is solutions to stochastic differential equations driven by multiple stochastic integrals (see also C. Tudor and M. Tudor 1997). This kind of equations include the classical Itô equations, integro-differential and some classes of Volterra equations. The asymptotic behaviour of double Itô processes was first considered in Pérez-Abreu and Tudor (2001) who proved a large deviations principle for these processes.

In this paper, we consider the following double Itô's stochastic differential equation

$$\begin{aligned} x_t^\varepsilon &= x_0 + \varepsilon \int_0^t F\left(\xi_s, x_s^\varepsilon\right) ds + \sqrt{\varepsilon} \int_0^t G(x_s^\varepsilon) dW_s + \varepsilon^2 \int_{C_2(t)} H\left(\varepsilon t_1, \varepsilon t_2, \xi_{t_1}, x_{t_1}^\varepsilon\right) dt_1 dt_2 \\ &+ \varepsilon \int_{C_2(t)} K(\varepsilon t_1, \varepsilon t_2, x_{t_1}^\varepsilon) d\widetilde{W}_{t_1} d\widetilde{W}_{t_2}, \end{aligned} \tag{3.2}$$

where $\xi = (\xi_t)_{t\geq 0}$ is a strictly stationary process and $C_2(t) = \{(t_1, t_2) \in \mathbb{R}_+^2 \mid 0 \leq t_1 \leq t_2 \leq t\}$.

The averaged equation corresponding to (3.2) is the following equation

$$\begin{aligned} \overline{x}_t^\varepsilon &= x_0 + \varepsilon \int_0^t \overline{F}\left(\overline{x}_s^\varepsilon\right) ds + \sqrt{\varepsilon} \int_0^t G(\overline{x}_s^\varepsilon) dW_s + \varepsilon^2 \int_{C_2(t)} \overline{H}\left(\varepsilon t_1, \varepsilon t_2, \overline{x}_{t_1}^\varepsilon\right) dt_1 dt_2 \\ &+ \varepsilon \int_{C_2(t)} K(\varepsilon t_1, \varepsilon t_2, \overline{x}_{t_1}^\varepsilon) d\widetilde{W}_{t_1} d\widetilde{W}_{t_2}, \end{aligned} \tag{3.3}$$

where

$$\overline{F}\left(x\right) = \mathbb{E}\left[F\left(\xi_t, x\right)\right] = \mathbb{E}\left[F\left(\xi_0, x\right)\right]$$

and

$$\overline{H}\left(t_1, t_2, x\right) = \mathbb{E}\left[H\left(t_1, t_2, , \xi_{t_1}, x\right)\right] = \mathbb{E}\left[H\left(t_1, t_2, , \xi_0, x\right)\right].$$

Our goal is to study the asymptotic behaviour of $\sup_{0\leq t\leq \varepsilon^{-1}} \left|x_t^\varepsilon - \overline{x}_t^\varepsilon\right|$ as $\varepsilon \longrightarrow 0$.

The paper is organized as follows. In section 3.1, we give some notations and make assumptions used throughout. Section 3.2 contains a preliminary result on the estimate of the moments of solutions to double Itô stochatic differential equations. Section 3.3 is devoted to the the main result. The last section is an appendix wherein, we have collected some results needed for the proof of the main result.

3.1 Notations and Assumptions

Let $\left(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0}\right)$ be the underlying stochastic basis, $(W_t)_{t\geq 0}$ and $(\widetilde{W}_t)_{t\geq 0}$ two linear $\mathcal{F}_t$−Wiener processes and $\xi = \{\xi_t : t \geq 0\}$ a one dimensional strictly stationary process.

For all $t \geq 0$ we put

$$C_2(t) = \{(t_1, t_2) \in \mathbb{R}_+^2 : 0 \leq t_1 \leq t_2 \leq t\}$$

and

$$C_2 = \{(t_1, t_2) \in \mathbb{R}_+^2 : 0 \leq t_1 \leq t_2\}.$$

Let LS denote the space of functions $\alpha : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ such that α is strictly increasing , continuous , concave and $\int_{0+}^{1} \frac{du}{\alpha(u)} = +\infty$.

It is clear that $\alpha_1(u) = Lu$, $L > 0$, $\alpha_2(u) = u|\log u|^{1-\varepsilon}$, $\alpha_3(u) = u|\log u||\log|\log u||^{1-\varepsilon}$, $0 < \varepsilon < 1$, belong to LS and α_1, α_2, α_3 are not Lipschitz. Also, if α_1, $\alpha_2 \in LS$, c_1, $c_2 \geq 0$, $c_1 + c_2 > 0$, then $\alpha_1 c_1 + \alpha_2 c_2 \in LS$.

Let $F(t,x), G(t,x) : \mathbb{R}_+ \times \mathbb{R} \longrightarrow \mathbb{R}$, $H(t_1,t_2,y,x) : C_2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$, $K(t_1,t_2,x) : C_2 \times \mathbb{R} \longrightarrow \mathbb{R}$ be mesurable functions such that

(H1) (Growth condition). There exists a constant $L > 0$ satisfying

$$\begin{aligned} \forall x &\in \mathbb{R}, \forall y \in \mathbb{R},\ |F(y,x)|^2 + |G(x)|^2 \leq L(1 + |x|^2) \\ \forall (t_1, t_2) &\in C_2,\ \forall x \in \mathbb{R},\ \forall y \in \mathbb{R},\ |H(t_1,t_2,y,x)|^2 + |K(t_1,t_2,x)|^2 \leq L(1 + |x|^2). \end{aligned}$$

(H2) (Hölder-type condition on F, G). There exists $\alpha \in LS$ such that: $\forall t \in \mathbb{R}_+$, $\forall x, y \in \mathbb{R}$,

$$|F(t,x) - F(t,y)|^2 + |G(x) - G(y)|^2 \leq \alpha(|x-y|^2).$$

(H3) (Hölder-type condition on H, K). There exists $\widetilde{\alpha} \in LS$ such that: $\forall (t_1, t_2) \in C_2$, $\forall z \in \mathbb{R}$, $\forall x, y \in \mathbb{R}$,

$$|H(t_1, t_2, z, x) - H(t_1, t_2, z, y)|^2 + |K(t_1, t_2, x) - K(t_1, t_2, y)|^2 \leq \widetilde{\alpha}(|x-y|^2).$$

(H4) The initial condition $x_0 \in \mathbb{R}$ is deterministic.

(H5) $(\xi)_{t \geq 0}$ is independent of the Brownian motions $(W_t)_{t\geq 0}$, $\left(\widetilde{W}_t\right)_{t \geq 0}$, strictly stationary and satisfies the following strongly mixing condition:
if

$$\mathcal{G}_s^t = \sigma\{\xi_u, u \in [s,t]\}, \; \mathcal{G}_s^\infty = \sigma\{\xi_u, u \in [s, \infty)\}, \; \forall 0 \leq s \leq t < \infty,$$

and $\gamma(u)$ is a function defined for all $u \in [0, \infty)$ by

$$\gamma(u) = \sup_{t \geq 0} \sup_{\substack{A \in \mathcal{G}_0^t \\ B \in \mathcal{G}_{t+u}^\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$

then $\gamma(u) \longrightarrow 0$ as $u \longrightarrow \infty$.

(H6) We suppose that the Rosenblatt mixing coefficient γ defined in **(H5)** satisfies the following condition: there are constants $\theta \in (1, \infty)$ and $\eta \in (0, \infty)$ such that

$$\forall u \in [1, \infty) \;\; \gamma(u) \leq \eta u^{-\theta}.$$

It is clear that the functions $\overline{F}$ and $\overline{H}$ satisfy growth and Hölder type conditions: $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}$, $\forall (t_1, t_2) \in C_2$,

$$\begin{aligned} |\overline{F}(x)|^2 &\leq L(1 - |x|^2), \quad \left|\overline{H}(t_1, t_2, x)\right|^2 \leq L(1 - |x|^2) \\ \left|\overline{F}(x) - \overline{F}(y)\right|^2 &\leq \alpha(|x-y|^2), \left|\overline{H}(t_1, t_2, x) - \overline{H}(t_1, t_2, y)\right|^2 \leq \widetilde{\alpha}(|x-y|^2). \end{aligned}$$

Let us put

$$X_t^\varepsilon = x_{t/\varepsilon}^\varepsilon,\ W_t^\varepsilon = \sqrt{\varepsilon} W_{t/\varepsilon} \quad \text{and} \quad \widetilde{W}_t^\varepsilon = \sqrt{\varepsilon}\, \widetilde{W}_{t/\varepsilon}.$$

Then, equations (3.2) and (3.3) can be written respectively

$$\begin{aligned} X_t^\varepsilon &= x_0 + \int_0^t F(\xi_{s/\varepsilon}, X_s^\varepsilon) ds + \int_0^t G(X_s^\varepsilon) dW_s^\varepsilon + \int_{C_{2(t)}} H(t_1, t_2, \xi_{t_1/\varepsilon}, X_{t_1}^\varepsilon) dt_1 dt_2 \\ &\quad + \int_{C_{2(t)}} K(t_1, t_2, X_{t_1}^\varepsilon) d\widetilde{W}_{t_1}^\varepsilon \widetilde{W}_{t_2}^\varepsilon \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} \overline{X}_t^\varepsilon &= x_0 + \int_0^t \overline{F}(\overline{X}_s^\varepsilon) ds + \int_0^t G(\overline{X}_s^\varepsilon) dW_s^\varepsilon + \int_{C_{2(t)}} \overline{H}(t_1, t_2, \overline{X}_{t_1}^\varepsilon) dt_1 dt_2 \\ &\quad + \int_{C_{2(t)}} K(t_1, t_2, \overline{X}_{t_1}^\varepsilon) d\widetilde{W}_{t_1}^\varepsilon \widetilde{W}_{t_2}^\varepsilon. \end{aligned} \tag{3.5}$$

Let $\mathcal{C}([0,1])$ denote the vector space of all $\mathbb{R}$-valued continuous functions defined over the unit interval endowed with the uniform norm

$$\|\Phi\|_\infty = \sup_{0 \le t \le 1} |\Phi_t|, \quad \forall\, \Phi \in \mathcal{C}([0,1]).$$

We put

$$\Delta_t^\varepsilon = X_t^\varepsilon - \overline{X}_t^\varepsilon, \forall t \in [0,1].$$

Let us note that

$$\|\Delta^\varepsilon\|_\infty = \sup_{0 \le t \le 1} \left| X_t^\varepsilon - \overline{X}_t^\varepsilon \right| = \sup_{0 \le t \le \varepsilon^{-1}} \left| x_t^\varepsilon - \overline{x}_t^\varepsilon \right|.$$

Under conditions **(H1)**-**(H5)**, C. Tudor and M. Tudor (2002) have proved the existence and uniqueness of solutions to equations (3.4) and (3.5).

3.2 Preliminary Result

Lemma 3.9 *Assume conditions **(H1)-(H5)**. Then, for every $p \in [2, +\infty)$, there exists positive constants $C(p)$ and $\overline{C}(p)$ such that*

$$\mathbb{E}\left[(\|X^{\varepsilon}\|_{\infty})^{p}\right] \leq C(p) \quad \textit{and} \quad \mathbb{E}[(\|\overline{X}^{\varepsilon}\|_{\infty})^{p}] \leq \overline{C}(p),$$

where X^{ϵ} and $\overline{X}^{\varepsilon}$ are the solutions of equations (3.4) and (3.5) respectively.

Proof. For every $p \in [1, +\infty)$ and $t \in [0, 1]$, we have

$$\begin{aligned} |X_t^{\varepsilon}|^{2p} \leq\ & 5^{2p}\left(|x_0|^{2p} + \left|\int_0^t F(\xi_{s/\varepsilon}, X_s^{\varepsilon})ds\right|^{2p} + \left|\int_0^t G(X_s^{\varepsilon})dW_s^{\varepsilon}\right|^{2p}\right. \\ & \left. + \left|\int_{C_{2(t)}} H(t_1, t_2, \xi_{t_1/\varepsilon}, X_{t_1}^{\varepsilon})dt_1 dt_2\right|^{2p} + \left|\int_{C_{2(t)}} K(t_1, t_2, X_{t_1}^{\varepsilon})d\widetilde{W}_{t_1}^{\varepsilon} d\widetilde{W}_{t_2}^{\varepsilon}\right|^{2p}\right). \end{aligned}$$

It follows that

$$\begin{aligned} \mathbb{E}\left(\sup_{0\leq t\leq 1} |X_t^{\varepsilon}|^{2p}\right) \leq\ & 5^{2p}\left[|x_0|^{2p} + \mathbb{E}\left(\sup_{0\leq t\leq 1}\left|\int_0^t F(\xi_{s/\varepsilon}, X_s^{\varepsilon})ds\right|^{2p}\right)\right. \\ & + \mathbb{E}\left(\sup_{0\leq t\leq 1}\left|\int_0^t G(X_s^{\varepsilon})dW_s^{\varepsilon}\right|^{2p}\right) \\ & + \mathbb{E}\left(\sup_{0\leq t\leq 1}\left|\int_{C_{2(t)}} H(t_1, t_2, \xi_{t_1/\varepsilon}, X_{t_1}^{\varepsilon})dt_1 dt_2\right|^{2p}\right) \\ & \left. + \mathbb{E}\left(\sup_{0\leq t\leq 1}\left|\int_{C_{2(t)}} K(t_1, t_2, X_{t_1}^{\varepsilon})d\widetilde{W}_{t_1}^{\varepsilon}\widetilde{W}_{t_2}^{\varepsilon}\right|^{2p}\right)\right]. \end{aligned} \quad (3.6)$$

In view of Cauchy-Schwarz inequality and condition **(H1)**,we have

$$\begin{aligned} \sup_{0\leq t\leq 1}\left|\int_0^t F(\xi_{s/\varepsilon}, X_s^{\varepsilon})ds\right|^{2p} \leq\ & \int_0^1 |F(\xi_{s/\varepsilon}, X_s^{\varepsilon})|^{2p} ds \\ \leq\ & L^p \int_0^1 (1 + |X_s^{\varepsilon}|^2)^p ds \end{aligned}$$

$$\leq 2^p L^p \int_0^1 (1+|X_s^\varepsilon|^{2p})ds.$$

Therefore

$$\mathbb{E}\left(\sup_{0\leq t\leq 1}\left|\int_0^t F(\xi_{s/\varepsilon}, X_s^\varepsilon)ds\right|^{2p}\right) \leq 2^p L^p \int_0^1 \left(1+\mathbb{E}\left(\sup_{0\leq u\leq s}|X_u^\varepsilon|^{2p}\right)\right)ds.$$

By virtue of Burkhölder-Davis-Gundy inequality, there exists a positive constant C(depending only on p) such that

$$\mathbb{E}\left(\sup_{0\leq t\leq 1}\left|\int_0^t G(X_s^\varepsilon)dW_s^\varepsilon\right|^{2p}\right) \leq C\mathbb{E}\left(\int_0^1 |G(X_s^\varepsilon)|^2\, ds\right)^p.$$

Therefore, by using Cauchy-Schwarz inequality and condition (H1), we obtain

$$\mathbb{E}\left(\sup_{0\leq t\leq 1}\left|\int_0^t G(X_s^\varepsilon)dW_s^\varepsilon\right|^{2p}\right) \leq 2^{2p}CL^p \int_0^1 \left(1+E\left(\sup_{0\leq u\leq s}|X_u^\varepsilon|^{2p}\right)\right)ds. \quad (3.7)$$

Now,

$$\int_{C_{2(t)}} H(t_1,t_2,\xi_{t_1/\varepsilon},X_{t_1}^\varepsilon)dt_1dt_2 = \int_0^t \left(\int_0^{t_2} H(t_1,t_2,\xi_{t_1/\varepsilon},X_{t_1}^\varepsilon)dt_1\right)dt_2.$$

So, Cauchy-Schwarz inequality and condition **(H1)** lead to

$$\begin{aligned}\sup_{0\leq t\leq 1}\left|\int_0^t \left(\int_0^{t_2} H(t_1,t_2.\xi_{t_1/\varepsilon},X_{t_1}^\varepsilon)dt_1\right)dt_2\right|^{2p} &\leq \int_0^1\int_0^{t_2} \left|H(t_1.t_2,\xi_{t_1/\varepsilon},X_{t_1}^\varepsilon)\right|^{2p}dt_1dt_2 \\ &\leq 2^pL^p\int_0^1\left(1+\sup_{0\leq u\leq s}|X_u^\varepsilon|^{2p}\right)ds.\end{aligned}$$

Therefore,

$$\mathbb{E}\left(\sup_{0\leq t\leq 1}\left|\int_{C_{2(t)}} H(t_1.t_2.\xi_{t_1/\varepsilon}.X_{t_1}^\varepsilon)dt_1dt_2\right|^{2p}\right) \leq 2^pL^p\int_0^1\left(1+E\left(\sup_{0\leq u\leq s}|X_u^\varepsilon|^{2p}\right)\right)ds. \quad (3.8)$$

Let us note that

$$\int_{C_{2(t)}} K(t_1,t_2,X^{\varepsilon}_{t_1})d\widetilde{W}^{\varepsilon}_{t_1}d\widetilde{W}^{\varepsilon}_{t_2} = \int_0^t \left(\int_0^{t_2} K(t_1,t_2,X^{\varepsilon}_{t_1})d\widetilde{W}^{\varepsilon}_{t_1}\right) d\widetilde{W}^{\varepsilon}_{t_2}$$

Since the process $t \longmapsto \int_0^t \left(\int_0^{t_2} K(t_2,X^{\varepsilon}_{t_1})d\widetilde{W}^{\varepsilon}_{t_1}\right) d\widetilde{W}^{\varepsilon}_{t_2}$ is a martingale, by applying the Burkhölder-Davis-Gundy inequality, there exists a positive constant C(depending only on p) such that

$$\mathbb{E}\left(\sup_{0\le t\le 1}\left|\int_{C_{2(t)}} K(t_1,t_2,X^{\varepsilon}_{t_1})d\widetilde{W}^{\varepsilon}_{t_1}\widetilde{W}^{\varepsilon}_{t_2}\right|^{2p}\right) \le C\mathbb{E}\left(\int_0^1\left|\int_0^{t_2} K(t_1,t_2,X^{\varepsilon}_{t_1})d\widetilde{W}^{\varepsilon}_{t_1}\right|^2 dt_2\right)^p.$$

Now, for any v the process $u \longmapsto \int_0^u K(v,X^{\varepsilon}_{t_1})d\widetilde{W}^{\varepsilon}_{t_1}$ is a martingale. Consequently, by applying Burkhölder-Davis-Gundy inequality and condition **(H1)**, there exists a constant C (depending only on p) such that

$$\mathbb{E}\left(\sup_{0\le u\le t_2}\left|\int_0^u K(t_1,t_2,X^{\varepsilon}_{t_1})d\widetilde{W}^{\varepsilon}_{t_1}\right|^{2p}\right) \le 2^pCL^p\int_0^{t_2}\left(1+\mathbb{E}\left(\sup_{u\le s}|X^{\varepsilon}_u|^{2p}\right)\right)ds.$$

Therefore,

$$\mathbb{E}\left(\sup_{0\le t\le 1}\left|\int_{C_{2(t)}} K(t_1,t_2,X^{\varepsilon}_{t_1})d\widetilde{W}^{\varepsilon}_{t_1}\widetilde{W}^{\varepsilon}_{t_2}\right|^{2p}\right) \le 2^pCL^p\int_0^1\left(1+\mathbb{E}\left(\sup_{u\le s}|X^{\varepsilon}_u|^{2p}\right)\right)ds. \tag{3.9}$$

By combining the inequalities (3.6) to (3.9), one can show that there exists a positive constant C (depending only on L,p) such that

$$\mathbb{E}\left(\sup_{0\le t\le 1}|X^{\varepsilon}_t|^{2p}\right) \le C\left(1+\int_0^1\mathbb{E}\left(\sup_{0\le u\le s}|X^{\varepsilon}_u|^{2p}\right)\right)ds.$$

Hence, by virtue of Gronwall inequality, we deduce that

$$\mathbb{E}\left(\sup_{0\le t\le 1}|X^{\varepsilon}_t|^{2p}\right) \le C(p).$$

The same calculations show that

$$\mathbb{E}\left(\sup_{0\le t\le 1}|\overline{X}_t^{\varepsilon}|^{2p}\right)\le \overline{C}(p).$$

■

3.3 Main Result

Theorem 3.10 *Assume conditions $(H1)$ to $(H6)$. Then we have*

$$\lim_{\varepsilon\to 0} E\left(\sup_{0\le t\le \varepsilon^{-1}}|x_t^{\varepsilon}-\overline{x}_t^{\varepsilon}|^2\right)=0$$

where $\{x_t^{\varepsilon}\mid 0\le t\le \varepsilon^{-1}\}$ and $\{\overline{x}_t^{\varepsilon}\mid 0\le t\le \varepsilon^{-1}\}$ are the unique strong solutions of (3.2) and (3.3) respectively.

Proof. We have

$$\begin{aligned}\Delta_t^{\varepsilon} &= X_t^{\varepsilon}-\overline{X}_t^{\varepsilon}=\int_0^t\left[F(\xi_{s/\varepsilon},X_s^{\varepsilon})-\overline{F}(\overline{X}_s)\right]ds+\int_0^t\left[G(X_s^{\varepsilon})-G(\overline{X}_s^{\varepsilon})\right]dW_s^{\varepsilon}\\ &+\int_{C_{2(t)}}\left[H(t_1,t_2,\xi_{t_1/\varepsilon},X_{t_1}^{\varepsilon})-\overline{H}(t_1,t_2,\overline{X}_{t_1}^{\varepsilon})\right]dt_1dt_2\\ &+\int_{C_{2(t)}}\left[K(t_1,t_2,X_{t_1}^{\varepsilon})-K(t_1,t_2,\overline{X}_{t_1}^{\varepsilon})\right]d\widetilde{W}_{t_1}^{\varepsilon}d\widetilde{W}_{t_2}^{\varepsilon}.\end{aligned}$$

Therefore,

$$\begin{aligned}\Delta_t^{\varepsilon} &= \int_0^t\left[F(\xi_{s/\varepsilon},X_s^{\varepsilon})-F(\xi_{s/\varepsilon},\overline{X}_s^{\varepsilon})\right]ds\\ &+\int_0^t\left[F(\xi_{s/\varepsilon},\overline{X}_s^{\varepsilon})-\overline{F}(\overline{X}_s)\right]ds+\int_0^t\left[G(X_s^{\varepsilon})-G(\overline{X}_s^{\varepsilon})\right]dW_s^{\varepsilon}\\ &+\int_{C_{2(t)}}\left[H(t_1,t_2,\xi_{t_1/\varepsilon},X_{t_1}^{\varepsilon})-H(t_1,t_2,\xi_{t_1/\varepsilon},\overline{X}_{t_1}^{\varepsilon})\right]dt_1dt_2\\ &+\int_{C_{2(t)}}\left[H(t_1,t_2,\xi_{t_1/\varepsilon},\overline{X}_{t_1}^{\varepsilon})-\overline{H}(t_1,t_2,\overline{X}_{t_1}^{\varepsilon})\right]dt_1dt_2\end{aligned}$$

$$+\int_{C_{2(t)}} \left[K(t_1,t_2,X^\varepsilon_{t_1}) - K(t_1,t_2,\overline{X}^\varepsilon_{t_1})\right] d\widetilde{W}^\varepsilon_{t_1} d\widetilde{W}^\varepsilon_{t_2}. \tag{3.10}$$

Let us put

$$f^\varepsilon(s) = F(\xi_{s/\varepsilon}, \overline{X}^\varepsilon_s) - \overline{F}(\overline{X}^\varepsilon_s) \text{ and } h^\varepsilon(t_1,t_2) = H(t_1,t_2,\xi_{t_1/\varepsilon}, \overline{X}^\varepsilon_{t_1}) - \overline{H}(t_1,t_2,\overline{X}^\varepsilon_{t_1}).$$

It follows that for every $t \in [0,1]$

$$\begin{aligned} |\Delta^\varepsilon_t|^2 \ \le\ & 6\int_0^t \left|F(\xi_{s/\varepsilon}, X^\varepsilon_s) - F(\xi_{s/\varepsilon}, \overline{X}^\varepsilon_s)\right|^2 ds + 6\left|\int_0^t f^\varepsilon(s) ds\right|^2 \\ & +6\left|\int_0^t \left[G(X^\varepsilon_s) - G(\overline{X}^\varepsilon_s)\right] dW^\varepsilon_s\right|^2 + 6\left|\int_{C_{2(t)}} h^\varepsilon(t_1,t_2) dt_1 dt_2\right|^2 \\ & +6\int_{C_{2(t)}} \left|H\left(t_1,t_2,\xi_{t_1/\varepsilon}, X^\varepsilon_{t_1}\right) - H\left(t_1,t_2,\xi_{t_1/\varepsilon}, \overline{X}^\varepsilon_{t_1}\right)\right|^2 dt_1 dt_2 \\ & +\ 6\ \left|\int_{C_{2(t)}} \left[K\left(t_1,t_2,X^\varepsilon_{t_1}\right) - K\left(t_1,t_2,\overline{X}^\varepsilon_{t_1}\right)\right] d\widetilde{W}^\varepsilon_{t_1} d\widetilde{W}_{t_2}\right|^2. \end{aligned}$$

By using conditions **(H2)** and **(H3)**, we derive that for all $t \in [0,1]$ we have

$$\begin{aligned} |\Delta^\varepsilon_t|^2 \ \le\ & 6\int_0^t \alpha\left(\left|X^\varepsilon_s - \overline{X}^\varepsilon_s\right|^2\right) ds + 6\left|\int_0^t f^\varepsilon(s) ds\right|^2 + 6\left|\int_0^t \left[G(X^\varepsilon_s) - G(\overline{X}^\varepsilon_s)\right] dW^\varepsilon_s\right|^2 \\ & +6\int_{C_{2(t)}} \widetilde{\alpha}\left(\left|X^\varepsilon_{t_1} - \overline{X}^\varepsilon_{t_1}\right|^2\right) dt_1 dt_2 + \ 6\left|\int_{C_{2(t)}} h^\varepsilon(t_1,t_2) dt_1 dt_2\right|^2 \\ & +6\left|\int_{C_{2(t)}} \left[K\left(t_1,t_2,X^\varepsilon_{t_1}\right) - K\left(t_1,t_2,\overline{X}^\varepsilon_{t_1}\right)\right] d\widetilde{W}^\varepsilon_{t_1} d\widetilde{W}^\varepsilon_{t_2}\right|^2. \end{aligned} \tag{3.11}$$

In view of Burkhölder-Davis-Gundy inequality, there exists a positive constant C such that

$$\mathbb{E}\left[\sup_{0\le s\le t}\left|\int_0^s \left[G(X^\varepsilon_u) - G(\overline{X}^\varepsilon_u)\right] dW^\varepsilon_u\right|^2\right] \le C\mathbb{E}\left[\int_0^t \left|G(X^\varepsilon_s) - G(\overline{X}^\varepsilon_s)\right|^2 ds\right].$$

By virtue of condition **(H2)**, we have

$$\int_0^t \left|G(X_s^\varepsilon) - G(\overline{X}_s^\varepsilon)\right|^2 ds \leq \int_0^t \alpha\left(\left|X_s^\varepsilon - \overline{X}_s^\varepsilon\right|^2\right) ds.$$

Now, since α is concave, we have

$$\mathbb{E}\left(\alpha\left(\sup_{0\leq u\leq s} |\Delta_u^\varepsilon|^2\right)\right) \leq \alpha\left(\mathbb{E}\left(\sup_{0\leq u\leq s} |\Delta_u^\varepsilon|^2\right)\right).$$

Therefore,

$$\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\int_0^s \left[G(X_u^\varepsilon) - G(\overline{X_u}^\varepsilon)\right] dW_u^\varepsilon\right|^2\right] \leq C\int_0^t \alpha\left(\mathbb{E}\left(\sup_{0\leq u\leq s} |\Delta_u^\varepsilon|^2\right)\right) ds. \tag{3.12}$$

In view of the concavity of $\widetilde{\alpha}$, for every $t\in[0,1]$, we have

$$\begin{aligned}
\mathbb{E}\left(\sup_{0\leq s\leq t}\int_{C_2(s)} \widetilde{\alpha}\left(\left|X_{t_1}^\varepsilon - X_{t_2}^\varepsilon\right|^2\right) dt_1 dt_2\right) &\leq \mathbb{E}\left(\int_{C_2(t)} \widetilde{\alpha}\left(\left|X_{t_1}^\varepsilon - \overline{X}_{t_1}^\varepsilon\right|^2\right) dt_1 dt_2\right) \\
&\leq \mathbb{E}\left(\int_0^t \widetilde{\alpha}\left(\sup_{0\leq u\leq s} |\Delta_u^\varepsilon|^2\right) ds\right) \\
&\leq \int_0^t \widetilde{\alpha}\left(\mathbb{E}\left(\sup_{0\leq u\leq s} |\Delta_u^\varepsilon|^2\right)\right) ds. \qquad (3.13)
\end{aligned}$$

Let us note that

$$\begin{aligned}
&\int_{C_2(s)} \left[K\left(t_1,t_2,X_{t_1}^\varepsilon\right) - K\left(t_1,t_2,\overline{X}_{t_1}^\varepsilon\right)\right] d\widetilde{W}_{t_1}^\varepsilon d\widetilde{W}_{t_2}^\varepsilon \\
= &\int_0^s \left(\int_0^{t_2} \left[K\left(t_1,t_2,X_{t_1}^\varepsilon\right) - K\left(t_1,t_2,\overline{X}_{t_1}^\varepsilon\right)\right] d\widetilde{W}_{t_1}^\varepsilon\right) d\widetilde{W}_{t_2}^\varepsilon.
\end{aligned}$$

Since the process $\left\{\int_0^s \left(\int_0^{t_2} \left[K\left(t_1,t_2,X_{t_1}^\varepsilon\right) - K\left(t_1,t_2,\overline{X}_{t_1}^\varepsilon\right)\right] d\widetilde{W}_{t_1}^\varepsilon\right) d\widetilde{W}_{t_2}^\varepsilon : s\geq 0\right\}$ is a martingale. Burkhölder-Davis-Gundy inequality implies that there exists a positive constant

C such that

$$\begin{aligned}
& \mathbb{E}\left(\sup_{0\le s\le t}\left|\int_{C_2(s)}\left[K\left(t_1,t_2,X^\varepsilon_{t_1}\right)-K\left(t_1,t_2,\overline{X}^\varepsilon_{t_1}\right)\right]d\widetilde{W}^\varepsilon_{t_1}d\widetilde{W}^\varepsilon_{t_2}\right|^2\right)\\
\le\ & C\mathbb{E}\left(\int_0^t\left(\int_0^{t_2}\left[K\left(t_1,t_2,X^\varepsilon_{t_1}\right)-K\left(t_1,t_2,\overline{X}^\varepsilon_{t_1}\right)\right]d\widetilde{W}^\varepsilon_{t_1}\right)^2dt_2\right)\\
\le\ & C\int_0^t\mathbb{E}\left(\int_0^{t_2}\left[K\left(t_1,t_2,X^\varepsilon_{t_1}\right)-K\left(t_1,t_2,\overline{X}^\varepsilon_{t_1}\right)\right]d\widetilde{W}^\varepsilon_{t_1}\right)^2dt_2\\
\le\ & C\int_0^t\mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^s\left[K\left(t_1,t_2,X^\varepsilon_{t_1}\right)-K\left(t_1,t_2,\overline{X}^\varepsilon_{t_1}\right)\right]d\widetilde{W}^\varepsilon_{t_1}\right|^2\right)dt_2.
\end{aligned}$$

Now, for every $v\in[0,1]$, the process $\left\{\int_0^s\left[K\left(t_1,v,X^\varepsilon_{t_1}\right)-K\left(t_1,v,\overline{X}^\varepsilon_{t_1}\right)\right]d\widetilde{W}^\varepsilon_{t_1}:s\ge 0\right\}$ is a martingale. Therefore, Burkhölder-Davis-Gundy inequality implies that there exists a positive constant C such that

$$\begin{aligned}
& \mathbb{E}\left(\sup_{0\le s\le t}\int_0^s\left|\left[K\left(t_1,t_2,X^\varepsilon_{t_1}\right)-K\left(t_1,t_2,\overline{X}^\varepsilon_{t_1}\right)\right]d\widetilde{W}^\varepsilon_{t_1}\right|^2\right)\\
\le\ & \mathbb{E}\left(\int_0^t|K\left(t_1,t_2,X^\varepsilon_{t_1}\right)-K\left(t_1,t_2,\overline{X}^\varepsilon_{t_1}\right)|^2dt_1\right).
\end{aligned}$$

Thanks to condition **(H3)**, it follows that

$$\begin{aligned}
& \mathbb{E}\left(\sup_{0\le s\le t}\left|\int_{C_2(s)}\left[K\left(t_1,t_2,X^\varepsilon_{t_1}\right)-K\left(t_1,t_2,\overline{X}^\varepsilon_{t_1}\right)\right]d\widetilde{W}^\varepsilon_{t_1}d\widetilde{W}^\varepsilon_{t_2}\right|^2\right)\\
& C\int_0^t\widetilde{\alpha}\left(\mathbb{E}\left(\sup_{0\le u\le s}|\Delta^\varepsilon_u|^2\right)\right)ds. \qquad (3.14)
\end{aligned}$$

In view of the concavity of α, we have

$$\mathbb{E}\left[\int_0^t\alpha\left(\left|X^\varepsilon_s-\overline{X}^\varepsilon_s\right|^2\right)ds\right]\le\int_0^t\alpha\left(\mathbb{E}\left(\sup_{0\le u\le s}|\Delta^\varepsilon_u|^2\right)\right)ds. \qquad (3.15)$$

Let

$$g^\varepsilon(t)=\mathbb{E}\left(\sup_{0\le s\le t}|\Delta^\varepsilon_s|^2\right).$$

By combining (3.11) to (3.15), we deduce that

$$g^{\varepsilon}(t) \leq g_0(\varepsilon) + C\int_0^t (\alpha + \widetilde{\alpha})(g^{\varepsilon}(s))ds.$$

where

$$g_0(\varepsilon) = \mathbb{E}\left(\sup_{0\leq s\leq 1}\left|\int_0^s f^{\varepsilon}(s)ds\right|^2\right) + \mathbb{E}\left(\sup_{0\leq s\leq 1}\left|\int_{C_2(s)} h^{\varepsilon}(s_1,s_2)ds_1ds_2\right|^2\right).$$

In view of Lemma A.13, for every $t \in [0,1]$

$$\begin{aligned} g^{\varepsilon}(t) &\leq G^{-1}\left[G\left(g_0(\varepsilon)\right) + C\int_0^t ds\right] \\ &\leq G^{-1}\left[G\left(g_0(\varepsilon)\right) + C\right]. \end{aligned}$$

where $G(r) = \int_1^r \dfrac{ds}{(\alpha+\widetilde{\alpha})(s)}$.

We have $G(r) \longrightarrow -\infty$ as $r \longrightarrow 0$ and $G^{-1}(r) \longrightarrow 0$ as $r \longrightarrow -\infty$.

Therefore, it remains to prove that $\lim_{\varepsilon\longrightarrow 0} g_0(\varepsilon) = 0$.

To this end let us note that

$$\begin{aligned} \int_{C_2(s)} h^{\varepsilon}(t_1,t_2)dt_1dt_2 &= \int_0^s \left(\int_0^{t_2} h^{\varepsilon}(t_1,t_2)dt_1\right)dt_2 \\ &= \int_0^s \left(\int_0^{t_2} \left[H\left(t_1,t_2,\xi_{t_1/\varepsilon},\overline{X}^{\varepsilon}_{t_1}\right) - \overline{H}\left(t_1,t_2,\overline{X}^{\varepsilon}_{t_1}\right)\right]dt_1\right)dt_2 \\ &= \varepsilon^2\int_0^{s/\varepsilon} \left(\int_0^{t_2} \left[H\left(\varepsilon t_1,\varepsilon t_2,\xi_{t_1},\overline{X}^{\varepsilon}_{\varepsilon t_1}\right) - \overline{H}\left(\varepsilon t_1,\varepsilon t_2,\overline{X}^{\varepsilon}_{\varepsilon t_1}\right)\right]dt_1\right)dt_2. \end{aligned}$$

For each $\Phi \in C\,[0,1]$ and $\varepsilon > 0$, we put

$$a_{\varepsilon}(t,\Phi) = \begin{cases} \varepsilon^2\int_0^t \left[H(\varepsilon t_1,\varepsilon t,\xi_{t_1},\Phi_{\varepsilon t_1}) - \overline{H}(\varepsilon t_1,\varepsilon t,\Phi_{\varepsilon t_1})\right]dt_1 & \text{if} \quad 0\leq t\leq \varepsilon^{-1} \\ 0 & \text{if} \quad \varepsilon^{-1} < t < \infty. \end{cases}$$

In view of Cauchy-Schwarz inequality and Condition **(H1)**, for every $t \in [0, \varepsilon^{-1}]$, we have

$$
\begin{aligned}
|a_\varepsilon(t,\Phi)|^2 &\le \varepsilon^4 t \int_0^t \left|H(t_1,\varepsilon t,\xi_{t_1},\Phi_{\varepsilon t_1}) - \overline{H}(t_1,\varepsilon t,\Phi_{\varepsilon t_1})\right|^2 dt_1 \\
&\le 2\varepsilon^4 t^2 \left[\sup_{0\le t_1\le t} \left|H(t_1,\varepsilon t,\xi_{t_1},\Phi_{\varepsilon t_1})\right|^2 + \sup_{0\le t_1\le t} \left|\overline{H}(t_1,\varepsilon t,\Phi_{\varepsilon t_1})\right|^2\right] \\
&\le 4L\left(1+\|\Phi\|_\infty^2\right)\varepsilon^2.
\end{aligned}
$$

So, for all $t \in [0,+\infty)$

$$(4L)^{-1/2}\left(1+\|\Phi\|_\infty^2\right)^{-1/2}\varepsilon^{-1}|a_\varepsilon(t,\Phi)| \le 1.$$

Let us put

$$\Theta(s) = (4L)^{-1/2}\left(1+\|\Phi\|_\infty^2\right)^{-1/2}\varepsilon^{-1}a_\varepsilon(t,\Phi).$$

Since $\theta > 1$ in Condition (H3), there exists $\delta \in (0,+\infty)$ such that $\theta > 1 + \dfrac{2}{\delta}$.

We have

$$\mathbb{E}(\Theta(s)) = 0 \quad \text{and} \quad \mathbb{E}(|\Theta(s)|^{2+\delta}) \le 1, \quad \forall s \ge 0.$$

By applying Lemma A.12 with

$$\Theta(s) = (4L)^{-1/2}\left(1+\|\Phi\|_\infty^2\right)^{-1/2}\varepsilon^{-1}a_\varepsilon(t,\Phi), \quad M = 1, \text{ and } \eta \text{ as in Condition } \textbf{(H3)}$$

there exists $\Gamma \in (0,+\infty)$, $\beta \in (0.+\infty)$ such that

$$\mathbb{E}\left(\left|\int_t^u \Theta(s)ds\right|^{2+\beta}\right) \le \Gamma(u-t)^{1+\frac{\beta}{2}}.$$

It follows that

$$\mathbb{E}\left(\left|\int_t^u a_\varepsilon(s,\Phi)ds\right|^{2+\beta}\right) \le \Gamma\varepsilon^{2+\beta}(4L(1+\|\Phi\|_\infty^2)(u-t))^{1+\frac{\beta}{2}}.$$

Now, let

$$Q(t) = \int_0^t a_\varepsilon(s,\Phi)ds, \quad 0 \le t \le \varepsilon^{-1}.$$

For all $0 \le t \le u < \infty$

$$\mathbb{E}[|Q(u) - Q(t)|^{2+\beta}] \le h(t,u)^{1+\frac{\beta}{2}},$$

where $h(t,u) = 4\Gamma^{\frac{2}{2+\beta}} L(1 + \|\Phi\|_\infty^2)(u-t)\varepsilon^2$.

By applying Lemma A.11 with $Y = \mathbb{R}, T = 0$, $U = \varepsilon^{-1}$, $\mu = 2+\beta$, $\gamma = 1 + \frac{\beta}{2}$, we obtain the existence of a constant A such that

$$\mathbb{E}\left(\sup_{0\le t\le\varepsilon^{-1}} \left|\int_0^t a_\varepsilon(s,\Phi)ds\right|^{2+\beta}\right) \le A(h(0,\varepsilon^{-1}))^{1+\frac{\beta}{2}}.$$

Therefore,

$$\begin{aligned}\mathbb{E}\left(\sup_{0\le t\le\varepsilon^{-1}} \left|\int_0^t a_\varepsilon(s,\Phi)ds\right|^{2+\beta}\right) &\le A\Gamma\varepsilon^{2+\beta}(4L(1+\|\Phi\|_\infty^2)\varepsilon^{-1})^{1+\frac{\beta}{2}} \\ &\le A\Gamma(4L(1+\|\Phi\|_\infty^2)\varepsilon)^{1+\frac{\beta}{2}}.\end{aligned}$$

It follows that

$$\begin{aligned}\mathbb{E}\left(\sup_{0\le t\le\varepsilon^{-1}} \left|\int_0^t a_\varepsilon(s,\Phi)ds\right|^{2}\right) &\le \left[\mathbb{E}\left(\sup_{0\le t\le\varepsilon^{-1}} \left|\int_0^t a_\varepsilon(s,\Phi)ds\right|^{2+\beta}\right)\right]^{\frac{2}{2+\beta}} \\ &\le 4(A\Gamma)^{\frac{2}{2+\beta}} L\left(1+\|\Phi\|_\infty^2\right)\varepsilon. \qquad (3.16)\end{aligned}$$

In view of the fact that ξ is independent of the Brownian motions W and $\widetilde{W}$ by using (3.16), we have

$$\begin{aligned}\mathbb{E}\left(\sup_{0\le t\le 1} \left|\int_{C_2(t)} h^\varepsilon(t_1,t_2)dt_1dt_2\right|^{2}\right) &= \int_{C([0,1])} \mathbb{E}\left(\sup_{0\le t\le\varepsilon^{-1}} \left|\int_0^t a_\varepsilon(s,\Phi)ds\right|^{2}\right) d\mu_\varepsilon(\Phi) \\ &\le 4(A\Gamma)^{\frac{2}{2+\beta}} L\left(\int_{C([0,1])} (1+\|\Phi\|_\infty^2)d\mu_\varepsilon(\Phi)\right)\varepsilon.\end{aligned}$$

where μ_ε is the law of the process $\overline{X}^\varepsilon$ which does not depend on ε.

It follows that

$$\mathbb{E}\left(\sup_{0\le t\le 1}\left|\int_{C_2(t)} h^\varepsilon(t_1,t_2)dt_1dt_2\right|^2\right)\le C\varepsilon,$$

which leads to

$$\lim_{\varepsilon\to 0}\mathbb{E}\left(\sup_{0\le t\le 1}\left|\int_{C_2(t)} h^\varepsilon(t_1,t_2)dt_1dt_2\right|^2\right)=0.$$

For the proof of

$$\lim_{\varepsilon\to 0}\mathbb{E}\left(\sup_{0\le t\le 1}\left|\int_0^t f^\varepsilon(s)ds\right|^2\right)=0,$$

it suffices to put

$$a_\varepsilon(t,\Phi)=\begin{cases}\int_0^t\left[F(\xi_s,\Phi_{\varepsilon s})-\overline{F}(\Phi_{\varepsilon s})\right]ds & \text{if}\quad 0\le t\le \varepsilon^{-1}\\ 0 & \text{if}\quad \varepsilon^{-1}<t<\infty.\end{cases}$$

and use analogous calculations as above.■

Appendix A

Appendix

In this appendix we collect some results which are needed for the proof of the main result.

Lemma A.11 *Let* $0 < T < U < \infty$ *and* $\{Q(t), T \leq t \leq U\}$ *be a process on some probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *taking values in a separable normed vector space* Y *with norm* $\|.\|$ *such that*

a) $t \longmapsto Q(t, w)$ *is continuous on* $[T, U]$ *for a.a.*w

b) there are constants $\gamma \in (1, \infty)$ *and* $\mu \in (0, \infty)$ *such that:* $\forall T \leq t \leq u \leq U$,

$$E\,\|Q(u) - Q(t)\|^{\mu} \leq [h(t, u)]^{\gamma}$$

where $h(t, u)$ *is a non-negative continuous function defined for* $T \leq t \leq u \leq U$ *satisfying*

$$\forall T \leq t \leq u \leq U, \qquad h(t, u) + h(u, v) \leq h(t, v).$$

Then there exists some constant $A \in (0, \infty)$ *depending only on* μ *and* γ *such that :*

$$E\left[\sup_{T \leq \cdot \leq U} \|Q(t) - Q(T)\|^{\mu}\right] \leq A\,[h(T, U)]^{\gamma}.$$

Proof. It is just a Corollary of Theorem 1 in Longnecker and Serfling (1977). ■

Lemma A.12 *Suppose that* $\{\Phi(s), s \in [0,\infty)\}$ *is a zero-mean* $\mathbb{R}^d$*-valued jointly measurable process on* $(\Omega, \mathcal{F}, \mathbb{P})$ *such that*

(a) $E\left[\|\Phi(s)\|^{2+\delta}\right] \leq M$, $\forall s \in [0,\infty)$, *for some constants* $M, \delta \in (0,\infty)$;

(b) there are σ*-algebras* $\{\mathcal{G}_s^t, 0 \leq s \leq t \leq \infty\}$ *over* Ω *such that* $\mathcal{G}_s^t \subset \mathcal{G}_u^v \subset \mathcal{F}$ *for all* $0 \leq u < s < t < v \leq \infty$, *and* $s \longrightarrow \Phi(s)$ *is* $\mathcal{G}_s^s$*-measurable for each* $s \in [0,\infty)$;

(c) there are constants $\theta \in (1+2\delta^{-1},\infty)$ *and* $\eta \in (0,\infty)$ *such that* $\alpha(u) \leq \eta u^{-\theta}, \forall u \in [1,\infty)$ *where*

$$\alpha(u) = \sup_{t\geq 0} \sup_{\substack{A\in\mathcal{G}_0^t \\ B\in\mathcal{G}_{t+u}^\infty}} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\,\mathbb{P}(B)|$$

Then there are constants $\Gamma \in [0,\infty)$ *and* $\beta \in (0,\infty)$ *such that*

$$E\left\|\int_t^u \Phi(s)ds\right\|^{2+\beta} \leq \Gamma(u-t)^{1+\frac{\beta}{2}}, \forall 0 \leq t \leq u \leq \infty$$

where Γ *and* β *depend only on the constants* M, δ, η *and* θ *in (a) and (c).*

Proof. It is a straightforward adaptation of Theorem 2.1 in Sotres and Ghosh (1977).■

Lemma A.13 ***(Bihari's inequality)****. Let* u *and* v *be two continuous functions on* $[0,1]$ *and* H *another continuous function from* $\mathbb{R}^+$ *into itself which is moreover nondecreasing and such that* $H(r) > 0$ *for* $r > 0$ *. If there exists* $u_0 \in \mathbb{R}^+$ *such that :*

$$\forall t \leq 1 \;\; u(t) \leq u_0 + \int_0^t v(s)H(u(s))ds$$

then

$$u(t) \leq G^{-1}\left[G(u_0) + \int_0^t v(s)ds\right]$$

for all $t \in [0,1]$ *such that*

$$G(u_0) + \int_0^t v(s)ds \in Dom(G^{-1})$$

Bibliography

[1] V. I. Arnold (1975) Matematicheskie Metody Klassicheskoi Mekhaniki. Nauka: Moscow, 1975. [English translation: Mathematical Method of classical Mechanics. *Graduate Texts in Mathematics*, **60**. Springer-Verlag: New-York, Heidelberg, Berlin, 1978.]

[2] M.Barlow and P Protter (1989) On convergence of semimartingales. *Séminaire de Probabilités* **XXIV**. *Lecture Notes Math.*, 188-193. Springer-Verlag, New York.

[3] N. N. Bogolyubov (1945) On certain Statistical Methods in Mathematical Physics, *Izd-vo Akad. Nauk UkrSSR.*(In Russian)

[4] N. Bogolyubov and Yu. Mitropol'skii (1961) Asymptotic Methods in the Theory of Non-linear Oscillations. Hindustan Publ.

[5] N. N. Bogolyubov and D. N. Zubarev (1955) Metod asimptoticheskogo priblizheniya dlya sistem s vrashchayuschcheisya fazoi i ego primenenie k dvizheniyu zaryazhennykh chastits v magnitnom pole. *Ukrain. Mat. Zh.*,**7**, No 7.

[6] B. V. Bondarev (1990) Averaging in stochastic quasilinear parabolic equations. *Theory of Probability and Mathematical Statistics* **41**,17-24.

[7] B. Boufoussi, M. Eddahbi and M. N'zi (2002) Averaging of BSDE's with reflection, local time and applications to PDE's. (submitted).

[17] S. N. Hashemi and A. J. Heunis (1998) Averaging principle for diffusion processes. *Stochastics and Stochastics Reports.* **62**, 201-216.

[18] A. J. Heunis and M. A. Kouritzin (1994) Strong convergence in the stochastic averaging principle. *Journal of Mathematical Analysis and Applications* **187**, 134-155.

[19] R. Z. Khasminskii (1963) Principle of averaging for parabolic and elliptic differential equations and for Markov processes with small diffusion. *Theory of Probability and its Applications,* **8(1)**, 1-19.

[20] R. Z. Khasminskii (1966) On stochastic processes defined by differential equations with a small parameter. *Theory of Probability and Applications* **2**, 211-228.

[21] R. Z. Khasminskii (1968) O printsipe usredneniya dlya stokhasticheskikh differentsial'nykh uravnenii Ito. *Kybernetika,* **4**, No. 3, 260- 279.

[22] R. Khasminskii, N. Krylov (2001) On averaging principle for diffusion processes with null-recurrent fast component. *Stochastic Processes and their Applications,* **93**, 229-240.

[23] M. Kleptsyna (1996) The averaging principle and diffusion approximation for Volterra equations. *Theory of probability and its Applications* **41(2)**, 359-367.

[24] M.L. Kleptsina, R. Liptser and A.P. Serebrovski (1997) Nonlinear filtering problem with contamination (averaging principle). *Annals of Applied Probability,* **7**(4), 917-934.

[25] M. Kleptsyna and A. Piatnitski (1997) Homogenization of random parabolic operators. homogenization and application to material sciences. *GAKKUTO Internat. Ser. Math. Sci-Appl.*. **9**. 241-255. Gakkotosko, Tokyo.

[26] M. Kleptsyna and A. Piatnitski (2002) Averaging of Non-Selfadjoint Parabolic Equations with random evolution (dynamics). *Rapport de recherche, INRIA,* n° **3951**.

[27] M. A. Krasnosel'skii, S. G. Krein (1955) On the averaging principle in nonlinear mechanics. *Uspekhi Mat. Nauk*, **10(3)**,147-152. (In Russian)

[28] R. Liptser, W. J. Runggaldier and M. Taksar (1996) Deterministic Approximation for Stochastic Control Problems. *SIAM J. On Optimization and Control* **34(1)**, 161-178.

[29] R. Liptser, W. J. Runggaldier and M. Taksar (1999) Asymptotic optimality for stochastic control problems via diffusion approximation. *Theory of Probability and Application* **44(4)**, 705-737.

[30] R. Liptser and A. Shiryaev (1989) Theory of Martingales. Kluwer Acad. Publishers, Dordrecht.

[31] R. Liptser and V. Spokoiny (2000) On estimating a dynamic function of stochastic system with averaging. *Statistical Inference for Stochastic Processes* **3**, 225-349.

[32] R. Liptser, J. Stoyanov (1990) Stochastic version of the averaging principle for diffusion type processes. *Stochastics and Stochastics Reports* **32**, 145-163.

[33] M. Longnecker and R. Serfling (1977) General moment and probability inequalities for the maximum partial sum. *Acta Mathematica Academiae Scientiarum Hungarica* **30**, 129-133.

[34] S. Ya. Makhno (1980) The averaging principle for stochastic partial differential equations, *Theory of Random Processes* **8**, 113-117. "Naukova Dumka". Kiev. (In Russian).

[35] P.A.Meyer (1976) Notions sur les intégrales multiples . *Séminaire de Probabilités .X. Lecture Notes in Math* **511**, 321-331. Springer - Verlag.

[36] A. I. Neishtadt (1975 Averaging in multifrequency system. I. *Soviet Physics. Doklady*, **20(7)**, 492-494.

[37] A. I. Neishtadt (1976) Averaging in multifrequency syste[illegible]viet *Physics. Doklady*, **21(2)**, 80-82.

[38] E. Pardoux and A. Yu. Veretennikov (1997) Averaging of bac[illegible]d stochastic differential equations, with application to semi- linear PDE's. *Stoc[illegible]ics and Stochastics Reports* **60**, 255-270.

[39] E. Pardoux and S. Peng (1990) Adapted solution of a backward stochastic differential equations. *Systems & Control Letters* **14**, 55-61

[40] E. Pardoux and A. L. Piatnitski (2001) Averaging of Nonlinear Random Parabolic Operators. *Optimal Control and Partial Differential Equations.* J. L. Menaldi *et al.* (Eds) IOS Press, 268-276.

[41] V. Pérez-Abreu, C. Tudor (2001) Large deviations for double Itô equations. *Stochastic in Finite and Infinite Dimensions*, 379-399. Trends Math. Birkhäuser Boston.

[42] R. Pettersson (1995) Yosida approximations for multivalued stochastic differential equations . *Stochastics and Stochastics Reports* **52**, 107-120.

[43] Yu. V. Prokhorov, A. N. Shiryaev (1998) Probability Theory III. *Encyclopaedia of Mathematical Sciences* **45**.

[44] J.Ruiz de Chavez (1985) Sur les intégrales stochastiques multiples . In : Sém . Probab . XIX . *Lectures Notes in Math* . **1123**, 248-257. *Springer Verlag.*

[45] J. A. Sanders and F. Verhulst (1985) Averaging methods in nonlinear dynamical systems. *Applied Mathematical Sciences Series*, **59**, Springer-Verlag, New York.

[46] D. Sotres and M. Ghosh (1977) Strong convergence of linear rank statistics for mixing processes. *Sankhyā*. Series B. *The Indian Journal of Statistics* **39**. 1-11.

[47] C.Tudor and M.Tudor (1997) Approximate solutions for multiple stochastic equations with respect to semimartingales. *Zeitschrift für Analysis und ihre Anwendungen* **16(3)**, 761-768.

[48] C.Tudor and M.Tudor (2002) Some properties of solutions of double stochastic differential equations. *Journal of Theorical Probability* **15(1)**, 129-151.

[49] A. Yu. Veretennikov (1990) On large deviations in the averaging principle for stochastic differential equations with periodic coefficients. In: Grigelionis B. *et al* (Eds). Probability Theory and Mathematical Statistics, Proceedings of the Fifth Vilnius Conference, Mokslas, Vilnius, Lithuania, and VSP, Utrcht, The Netherlands, Vol **2**, 1989, pp 542-551.

[50] A. Yu. Veretennikov (1991) On the averaging principle for systems of stochastic differential equation. *Math. USSR Sbornik,* **69(1)**, .

[51] A. Yu. Veretennikov (1992) On large deviation in the averaging principle for stoshastic differential equation with periodic coefficients II. *Math USSR, Izv.* **39(1)**, 677-693.

[52] A. Yu. Veretennikov (1994) Large deviations in averaging principle for stochastic differential equations systems (Non compact case). *Stochastics and Stochastics Reports,* **48**, 83-96.

[53] A. Yu. Veretennikov (2000) On large deviations for SDEs with small diffusion and averaging. *Stochastic Processes and their Applications* **89**, 69-79.

[54] V. M.Volosov (1962) Averaging in systems of ordinary differential equations. *Russian Math. Surveys,* **17(6)**, 1-126.