Singular Perturbations in Coupled Stochastic Differential Equations

by

Seyed Naser Hashemi

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2001

0-612-65244-0

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

In this thesis we study the coupled system of stochastic integral equations

$$x^\epsilon(t) \triangleq x_0 + \epsilon \int_0^t F(x^\epsilon(s), y^\epsilon(s))ds + \epsilon^{1/2} \int_0^t G(x^\epsilon(s))dw(s), \qquad (0.0.1)$$

$$y^\epsilon(t) \triangleq y_0 + \int_0^t b(x^\epsilon(s), y^\epsilon(s))ds + \int_0^t \sigma(x^\epsilon(s), y^\epsilon(s))d\beta(s), \qquad (0.0.2)$$

in which $\epsilon > 0$ is a *small* parameter, $\{x^\epsilon(t)\}$ is an $\mathbb{R}^d$-valued *slow process*, and $\{y^\epsilon(t)\}$ is an $\mathbb{R}^D$-valued *fast process*. Our general goal is to characterize asymptotic properties of the slow process $\{x^\epsilon(t)\}$ over intervals of the form $0 \le t \le T/\epsilon$, for a fixed constant $T \in (0, \infty)$, as $\epsilon \to 0$. The motivation for studying this question is a result of Khas'minskii ("On the Averaging Principle for Itô Stochastic Differential Equations", *Kybernetika*, V. 4(3) : 260–279, 1968 (Russian), also stated as Theorem 9.1 on page 264 of the book *Random Perturbations of Dynamical Systems* by Freidlin and Wentzel, Springer-Verlag, 1984), which basically goes as follows: suppose that the auxiliary stochastic differential equation

$$d\xi(t) = b(x, \xi(t))dt + \sigma(x, \xi(t))d\beta(t) \qquad (0.0.3)$$

(which is really just (0.0.2), but with the slow variables $x^\epsilon(s)$ "frozen" at some fixed $x \in \mathbb{R}^d$) is "stable", in the sense that the Markov process arising from (0.0.3) has a *unique* invariant probability measure π_x, for each $x \in \mathbb{R}^d$. Define the "averaged" drift

$$\bar{F}(x) \triangleq \int_{\mathbb{R}^D} F(x, \xi) d\pi_x(\xi), \qquad (0.0.4)$$

and use this to write the "averaged" version of (0.0.1) namely

$$\bar{x}^\epsilon(t) = x_0 + \epsilon \int_0^t \bar{F}(\bar{x}^\epsilon(s))\, ds + \epsilon^{1/2} \int_0^t G(\bar{x}^\epsilon(s))\, dw(s). \qquad (0.0.5)$$

It is shown by Khas'minskii *op. cit.* that the solution $\{x^\epsilon(t)\}$ of (0.0.1) converges in probability to the solution $\{\bar{x}^\epsilon(t)\}$ of (0.0.5), as $\epsilon \to 0$, namely

$$\lim_{\epsilon \to 0} P\left\{ \max_{0 \leq t \leq \frac{T}{\epsilon}} |x^\epsilon(t) - \bar{x}^\epsilon(t)| \geq \delta \right\} = 0, \tag{0.0.6}$$

for each $\delta > 0$, One can regard this convergence as a type of *weak law of large numbers*. Our goal in this thesis is to establish a rate of convergence for this weak law of large numbers in the form of a result which may be regarded as a complementary *central limit theorem*. To be more precise, we are going to study the *normalized discrepancy process* $\{z^\epsilon(t)\}$ defined by

$$z^\epsilon(t) \triangleq \frac{x^\epsilon(t) - \overline{x}^\epsilon(t)}{\epsilon^{1/2}}, \quad \forall 0 \leq t \leq \frac{T}{\epsilon}. \tag{0.0.7}$$

Using the method of martingale problems we establish that $\{z^\epsilon(t), 0 \leq t \leq T/\epsilon\}$ converges weakly to a limit as $\epsilon \to 0$, and we shall characterize this limit.

Acknowledgements

I would like to deeply express my best and sincere thanks to my supervisor, Professor A. J. Heunis, for his invaluable help, support, encouragement and patience throughout this work. I express my thanks to Professor T. Konstantopouls and Professor C. Cutler who have reviewed my thesis, for their helpful comments. My special thanks goes to my parents who spiritually supported my study. I also greatly acknowledge the patience of my wife, who helped to make this thesis possible. Finally, I am appreciative of the Ministry of Culture and Higher Education of Iran which supported me in the beginning of my program.

This thesis is dedicated to Hazrate Mahdi(S)
and Fatemeh Zahra(S).

Contents

Basic Notation and Terminology

I. Matrices, Vectors, Norms, Balls: $\mathbb{R}^{m\otimes n}$ denotes the vector space of matrices with m-rows, n-columns and real entries. If $A \in \mathbb{R}^{m\otimes n}$ then $|A|$ denotes the *Frobenius norm* of matrix A, namely

$$|A| \triangleq [\sum_{i=1}^{m}\sum_{j=1}^{n}(a^{ij})^2]^{1/2}.$$

The spaces $\mathbb{R}^m$ and $\mathbb{R}^{m\otimes 1}$ are considered identical, thus $|x|$ is the *Euclidean length* of the vector $x \in \mathbb{R}^m$. For any $R \in (0,\infty)$ let S_R^m denote the closed ball of radius R centered at the origin of $\mathbb{R}^m$, thus

$$S_R^m \triangleq \{x \in \mathbb{R}^m : \ |x| \leq R\}.$$

II. Polynomially bounded functions A function $(x,y) \to \Theta(x,y) : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ is called *polynomially y-bounded of order r locally in x*, when there is a constant $r \in [0,\infty)$, and, for each $R \in [0,\infty)$, there is a constant $C(R) \in [0,\infty)$, such that

$$|\Theta(x,y)| \leq C(R)[1+|y|^r], \quad \forall (x,y) \in S_R^d \otimes \mathbb{R}^D.$$

When the value of the constant $r \in [0,\infty)$ is unimportant then $\Theta(x,y)$ is just called *polynomially y-bounded locally in x*, and $\Theta(x,y)$ is called *uniformly y-bounded locally in x* when it is polynomially y-bounded of order $r = 0$ locally in x.

III. Function Spaces and Probability Distributions: Let $C[0,1]$ denote the Banach space of all $\mathbb{R}^d$-valued continuous functions defined over the unit interval $[0,1]$ with the usual supremum norm. Put

$$\Omega^* \triangleq C[0,1] \otimes C[0,1], \qquad \mathcal{F}^* \triangleq \mathcal{B}[\Omega^*], \tag{0.0.8}$$

where Ω^* is regarded as a metric space with the usual product metric, and $\mathcal{B}[\Omega^*]$ is the Borel σ-algebra in Ω^*. We shall use (X,Z), with $X, Z \in C[0,1]$, for a generic member of Ω^*, and define the usual natural filtration $\{\mathcal{B}_\tau,\ \tau \in [0,1]\}$ in $(\Omega^*, \mathcal{F}^*)$ by

$$\mathcal{B}_\tau \triangleq \sigma\{X(s), Z(s),\ s \in [0,\tau]\}, \quad \forall \tau \in [0,1]. \tag{0.0.9}$$

Moreover, if $\{\tilde{X}(\tau),\ \tau \in [0,1\}$ and $\{\tilde{Z}(\tau),\ \tau \in [0,1\}$ are two $\mathbb{R}^d$-valued processes with continuous sample-paths on some common probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ then $\mathcal{L}(\tilde{X}, \tilde{Y})$ will denote the probability distribution measure over $(\Omega^*, \mathcal{F}^*)$ for the joint process $\{(\tilde{X}(\tau), \tilde{Z}(\tau)),\ \tau \in [0,1]\}$.

IV. Derivatives and Gradients: Suppose that the mapping $(x,y) \to F(x,y) : \mathbb{R}^n \otimes \mathbb{R}^r \to \mathbb{R}^m$ is such that $x \to F(x,y)$ is differentiable on $\mathbb{R}^n$ for each $y \in \mathbb{R}^r$. Then $(\partial_x F)(x,y)$ denotes the matrix with m-rows and n-columns whose (i,j)-entry is given by the partial derivative $(\partial_{x^j} F^i)(x,y)$ for each $i = 1,2,\ldots,m$, $j = 1,2,\ldots,n$. In particular, when $m = 1$ (i.e. F is *real*-valued) then $(\partial_x F)(x,y))$ is a *row vector* of length n, with j-th entry given by $(\partial_{x^j} F)(x,y)$. Also, if $y \to F(x,y)$ is a differentiable mapping on $\mathbb{R}^r$ for each $x \in \mathbb{R}^n$, then the notation $(\partial_y F)(x,y)$ has an obviously analogous interpretation, as does the notation $(\partial_x F)(x)$ when $F : \mathbb{R}^n \to \mathbb{R}^m$ is a differentiable mapping.

Suppose that $H : \mathbb{R}^n \to \mathbb{R}^{n \otimes m}$ is differentiable at each $x \in \mathbb{R}^n$. Then, for each $x, z \in \mathbb{R}^n$, use $(\partial_x H)(x)[z]$ to denote the n by m-matrix whose (i,j)-entry is the scalar product $(\partial_x H)^{i,j}(x)z$, for each $i = 1,2,\ldots,n$, $j = 1,2,\ldots,m$.

V. Borel σ-algebra, sets of probability measures: Suppose that (S,ρ) is a complete separable metric space. Then $\mathcal{B}(S)$ denotes the Borel σ-algebra in S, and $\mathcal{P}(S)$ denotes the set of all probability measures on $\mathcal{B}(S)$.

VI. Function Spaces: $C(\mathbb{R}^n)$ denotes the collection of all real-valued continuous mappings on $\mathbb{R}^n$, and $C^r(\mathbb{R}^n)$, for some positive integer r, denotes the set of all members of $C(\mathbb{R}^n)$ which are r-times continuously differentiable on $\mathbb{R}^n$; $C^\infty(\mathbb{R}^n)$ denotes the set which is the intersection of $C^r(\mathbb{R}^n)$ over all positive integers r. Also $C_c(\mathbb{R}^n)$ [respectively $C_c^r(\mathbb{R}^n)$, $C_c^\infty(\mathbb{R}^n)$] denotes the set of all members of $C(\mathbb{R}^n)$ [respectively $C^r(\mathbb{R}^n)$, $C^\infty(\mathbb{R}^n)$] which have compact support. $C^{2,2}(\mathbb{R}^d \otimes \mathbb{R}^D)$ denotes the collection of all continuous functions $\Theta : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ such that (i) the partial derivative functions $(\partial_{x^l}\Theta)(x,y)$ and $(\partial_{x^l}\partial_{x^k}\Theta)(x,y)$, $l,k = 1,2,\ldots,d$, exist and are continuous at all $(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$, and (ii) the partial derivative functions $(\partial_{y^l}\Theta)(x,y)$ and $(\partial_{y^l}\partial_{y^k}\Theta)(x,y)$, $l,k = 1,2,\ldots,D$, exist and are continuous at all $(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Notice that there is no requirement that the mixed partial derivatives $(\partial_{x^l}\partial_{y^k}\Theta)(x,y)$, $l = 1,2,\ldots,d$, $k = 1,2,\ldots,D$, need exist for members of $C^{2,2}(\mathbb{R}^d \otimes \mathbb{R}^D)$..

Chapter 1

Historical Motivation

1.1 Introductory Remarks

Our research goal is to study the *averaging principle* for systems of stochastic differential equations. Roughly speaking, the term "averaging principle" refers to a family of results where one has a system of differential equations (either ordinary or stochastic) typically having "complex" right-hand sides, but where it is possible to approximate the solution (of the system) by solving a related but simpler system of equations in which these complex right-hand sides are replaced with simpler right-hand sides obtained by some form of "averaging" of the original complex system of equations. To make this rather vague statement more precise, and to place our own research problem in clearer perspective, we shall devote the present chapter to a short introduction on the subject of averaging principles, emphasizing in particular the historical development of ideas. Our goal in this introduction is to avoid technicalities and communicate only the basic intuition, and so we shall adopt a level of discussion which is rather heuristic, the precise statements of conditions and results being left to the detailed literature review in Chapter 2. In the course of this discussion we shall also draw attention to our own research problem, which is then taken up in Chapter 3.

1.2 Introduction to Averaging Principles - The Deterministic Case

The method of averaging in ordinary differential equations has its origins in the study of celestial mechanics (i.e. planetary motion) during the late eighteenth century. Here it proved necessary to study systems of differential equations having the form

$$\frac{dx^\epsilon(t)}{dt} = \epsilon F\left(x^\epsilon(t), t\right), \quad x^\epsilon(0) \stackrel{\triangle}{=} x_0, \tag{1.2.1}$$

where $\epsilon > 0$ is a *small* parameter appearing multiplicatively in the equation, and the problem is to determine the solution $x^\epsilon(\cdot)$ of the system over intervals of the form $0 \leq t \leq T/\epsilon$, where $T > 0$ is some finite constant. Notice that, because $\epsilon > 0$ is small in (1.2.1), it is only over such "large but finite" intervals that the solution exhibits significant variation. In (1.2.1) the mapping $F : \mathbb{R}^d \otimes [0, \infty) \rightarrow \mathbb{R}^d$ is typically fairly complex, although regular enough to ensure that (1.2.1) has a unique solution $\{x^\epsilon(t), t \in [0, \infty)\}$. In general one can say nothing about the solution of (1.2.1) without explicitly solving this system over the interval of interest, which is usually a difficult task in view of the complexity of the mapping $F(\cdot, \cdot)$. However, suppose the function $t \rightarrow F(x, t)$ enjoys the further property of having a well-defined *average* $\bar{F}(x)$ for each $x \in \mathbb{R}^d$, i.e. suppose that the limit

$$\bar{F}(x) \stackrel{\triangle}{=} \lim_{t \to \infty} \frac{1}{t} \int_0^t F(x, s) ds \tag{1.2.2}$$

exists in $\mathbb{R}^d$ for each $x \in \mathbb{R}^d$. This is the case if, for example, the mapping $t \rightarrow F(x, t)$ is periodic for each $x \in \mathbb{R}^d$, a situation which often arises in celestial mechanics where one is studying the movement of a system of planets around the sun. Now one can introduce an *averaged differential equation* namely

$$\frac{d\bar{x}^{\epsilon}(t)}{dt} = \epsilon \bar{F}\left(\bar{x}^{\epsilon}(t)\right), \quad \bar{x}(0) = x_0. \tag{1.2.3}$$

Setting aside the issue of existence and uniqueness of solutions, and supposing that (1.2.3) has a unique solution $\{\bar{x}^{\epsilon}(t),\ t \in [0, \infty)\}$, one might anticipate a relation between the mappings $t \to x^{\epsilon}(t)$ (i.e. the solution of (1.2.1)) and $t \to \bar{x}^{\epsilon}(t)$ over the interval $[1, T/\epsilon]$. Indeed, in the course of a study on planetary motion. Gauss suggested that these two mappings are virtually indistinguishable provided that the parameter $\epsilon > 0$ in (1.2.1) is small enough, i.e. the discrepancy

$$\max_{0 \le t \le T/\epsilon} |x^{\epsilon}(t) - \bar{x}^{\epsilon}(t)| \tag{1.2.4}$$

between the two solutions tends to zero as $\epsilon \to 0$. The usefulness of a result of this kind arises from the fact that the "averaged" right hand side $x \to \bar{F}(x)$ in (1.2.3) is often quite simple, enabling one to solve this equation, and the resulting mapping $t \to \bar{x}^{\epsilon}(t)$ is a very close approximation to the solution of (1.2.1) over the interval $[0, T/\epsilon]$ as long as $\epsilon > 0$ is small enough. A statement of this kind, where one uses the solution of a simpler "averaged" ordinary differential equation to approximate the solution of a given complex differential equation when an underlying multiplicative parameter is sufficiently small, is called an "averaging principle". Of course, this is not a rigorously established theorem but merely a useful method of approximation which seems to be based upon a plausible intuition. Despite the evident lack of solid mathematical justification, the averaging principle quickly became widely used in diverse areas of science including celestial mechanics, cosmology, statistical mechanics and the study of mechanical vibrations. Nevertheless, the lack of a rigorously established theorem, which could be used to justify application of the averaging principle, became a source of difficulty in certain challenging problems of physics. For example, in their book on averaging in systems of differential equations, Sanders and Verhulst [31] indicate a problem in cosmology [17] where an unjustified use of the averaging principle gave rise to erroneous conclusions (see page 19 in [31]). Problems of this kind stimulated efforts to place the averaging principle on a rigorous mathematical basis, and to establish theorems whose conditions identify when one is entitled to use the averaging principle. This task was undertaken by the Russian school of analysts during the 1930's and 1950's, especially Bogoliubov, Krylov and Mitropolskii [6]. In Section 2.1 we discuss the conditions for their main theorem and present a complete statement of it (see Theorem 2.1.1 on page 12).

1.3 Averaging in Random Differential Equations

Continuing with our rather heuristic discussion of averaging in differential equations, we consider next the situation where one has a *random* ordinary differential equation, namely

$$\frac{dx^\epsilon(t)}{dt} = \epsilon F\left(x^\epsilon(t), \xi(t)\right), \quad x^\epsilon(0) \triangleq x_0. \tag{1.3.5}$$

Here $\epsilon \in (0, \infty)$ is a small multiplicative parameter, exactly as in (1.2.1), $\{\xi(t),\ t \in [0, \infty)\}$ is an $\mathbb{R}^D$-valued strictly stationary ergodic process on some probability space $(\Omega, \mathcal{F}, P)$, and the mapping $F : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^d$ is sufficiently regular to ensure that (1.3.5) has a unique solution $\{x^\epsilon(t, \omega),\ t \in [0, \infty)\}$ for each $\omega \in \Omega$. Such random differential equations are of considerable importance in many areas of physics and engineering. Motivated by the (nonrandom) averaging principle introduced earlier, one might expect that if an "averaged" right-hand side is defined by

$$\bar{F}(x) \triangleq EF\left(x, \xi(0)\right), \quad \forall x \in \mathbb{R}^d, \tag{1.3.6}$$

and the "averaged" differential equation

$$\frac{d\bar{x}^\epsilon(t)}{dt} = \epsilon \bar{F}\left(\bar{x}^\epsilon(t)\right), \quad \bar{x}(0) = x_0, \tag{1.3.7}$$

has a unique solution $\{\bar{x}^\epsilon(t),\ t \in [0,\infty)\}$, then the discrepancy

$$\max_{0 \leq t \leq T/\epsilon} |x^\epsilon(t) - \bar{x}^\epsilon(t)| \tag{1.3.8}$$

tends to zero (in some sense) as $\epsilon \to 0$. The task of making this intuition rigorous, and thus placing this "stochastic averaging principle" on a sound mathematical basis was again undertaken by the Russian school of probabilists and analysts. In Section 2.2 we present perhaps the most general theorem pertaining to the stochastic averaging principle, taken from the book of Liptser and Shiryayev [24] (and stated in full as Theorem 2.2.1 on page 13in Chapter 2 of this thesis), which shows that, under fairly general conditions, the quantity in (1.3.8) converges to zero *almost surely* as $\epsilon \to 0$, namely

$$P\left\{\lim_{\epsilon \to 0}\left[\max_{0 \leq t \leq \frac{T}{\epsilon}} |x^\epsilon(t) - \bar{x}^\epsilon(t)|\right] = 0\right\} = 1. \tag{1.3.9}$$

In the preceding paragraph we introduced an averaging principle for the random ordinary differential equation (1.3.5) in which the dynamics are perturbed by a strictly stationary process $\{\xi(t),\ t \in [0,\infty)\}$. This raises the question of possibly extending this result by considering, in place of the perturbed ordinary differential equation (1.3.5), a perturbed *stochastic differential equation*, namely

$$dx^\epsilon(t) = \epsilon F\left(x^\epsilon(t), \xi(t)\right) dt + \epsilon^{1/2} G\left(x^\epsilon(t)\right) dw(t), \quad x^\epsilon(0) \stackrel{\Delta}{=} x_0, \tag{1.3.10}$$

(observe that (1.3.10) reduces to (1.3.5) in the special case where the covariance function $G(\cdot)$ in (1.3.10) is identically zero). Here $\epsilon \in (0,\infty)$ is a small parameter, $\{\xi(t),\ t \in [0,\infty)\}$ is an $\mathbb{R}^D$-valued strictly stationary ergodic process, $\{w(t),\ t \in [0,\infty)\}$ is an $\mathbb{R}^M$-valued standard Wiener process, both defined on $(\Omega, \mathcal{F}, P)$ with $\{\xi(t),; t \in [0,\infty)\}$ and $\{w(t),; t \in [0,\infty)\}$ being independent. The mappings $F : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^d$ and $G : \mathbb{R}^D \to \mathbb{R}^{d \otimes M}$ are sufficiently regular to ensure that (1.3.10) has a pathwise unique strong solution $\{x^\epsilon(t),\ t \in [0,\infty)\}$. The stochastic averaging principle of the preceding paragraph suggests that, if one defines $\bar{F}(x)$ as in (1.3.6), and introduces an "averaged" stochastic differential equation

$$d\bar{x}^{\epsilon}(t) = \epsilon \bar{F}\left(\bar{x}^{\epsilon}(t)\right) dt + \epsilon^{1/2} G\left(\bar{x}^{\epsilon}(t)\right) dw(t), \quad \bar{x}^{\epsilon}(0) \triangleq x_0, \tag{1.3.11}$$

then it is reasonable to expect that the discrepancy between the solutions of (1.3.10) and (1.3.11), namely

$$\max_{0 \le t \le T/\epsilon} |x^{\epsilon}(t) - \bar{x}^{\epsilon}(t)| \tag{1.3.12}$$

should in some sense be small when the parameter ϵ is small. A result of this kind is an averaging principle for stochastic differential equations, and has been established rigorously by Liptser and Stoyanov [25], who show that the quantity in (1.3.12) goes to zero *in probability* as $\epsilon \to 0$, that is for each $\delta \in (0, \infty)$,

$$\lim_{\epsilon \to 0} P\left\{ \max_{0 \le t \le \frac{T}{\epsilon}} |x^{\epsilon}(t) - \bar{x}^{\epsilon}(t)| > \delta \right\} = 0. \tag{1.3.13}$$

Section 2.3 is devoted to a precise statement of this result. Liptser and Stoyanov [25] also address another crucial aspect of the averaging principle for the stochastic differential equation (1.3.10), namely the issue of a *rate of convergence* (to zero) of the discrepancy (1.3.12) as $\epsilon \to 0$. The idea here is to *normalize* the error in (1.3.12) by defining a process

$$z^{\epsilon}(t) \triangleq \frac{x^{\epsilon}(t) - \bar{x}^{\epsilon}(t)}{h(\epsilon)}, \tag{1.3.14}$$

where $h : (0, \infty) \to (0, \infty)$ is some function such that $\lim_{\epsilon \to 0} h(\epsilon) = 0$. If it can be shown that the stochastic process $\{z^{\epsilon}(t),\ 0 \le t \le T/\epsilon\}$ converges to some sensible limit (in the sense of weak convergence) as $\epsilon \to 0$ then this suggests that the solution $\{x^{\epsilon}(t),\ 0 \le t \le T/\epsilon\}$ of (1.3.10) approaches the solution $\{\bar{x}^{\epsilon}(t),\ 0 \le t \le T/\epsilon\}$ of the averaged equation (1.3.11) at a speed given by $0(h(\epsilon))$, at least in the sense of weak convergence. Liptser and Stoyanov [25] establish this result in the case where

$$h(\epsilon) \triangleq \epsilon^{1/2}, \tag{1.3.15}$$

and a detailed presentation of their theorem is given in Section 2.3.1.

1.4 Averaging in Coupled Stochastic Differential Equations

From the point of view of applications, a disadvantage of the system (1.3.10) is that the perturbation process $\{\xi(t),\ t \in [0,\infty)\}$ in the drift term evolves *unilaterally*, and is in no way conditioned by the solution $x^\epsilon(t)$ of (1.3.10). In many applications the appropriate mathematical model is a stochastic differential system similar to (1.3.10), but with the extra feature that the solution $x^\epsilon(t)$ "feeds back" and in turn influences the perturbation process $\{\xi(t),\ t \in [0,\infty)\}$ appearing in the drift. One way of modeling this dependence is to write the perturbation process as the solution of a second stochastic differential equation whose drift and covariance functions in turn depend on $x^\epsilon(t)$. Thus, instead of the single perturbed stochastic differential equation (1.3.10), we are led to consider the pair of equations

$$dx^\epsilon(t) = \epsilon F\left(x^\epsilon(t), y^\epsilon(t)\right) dt + \epsilon^{1/2} G\left(x^\epsilon(t)\right) dw(t), \quad x^\epsilon(0) \triangleq x_0, \tag{1.4.16}$$

$$dy^\epsilon(t) = b\left(x^\epsilon(t), y^\epsilon(t)\right) dt + \sigma\left(x^\epsilon(t), y^\epsilon(t)\right) d\beta(t), \quad y^\epsilon(0) \triangleq y_0. \tag{1.4.17}$$

Here the first equation (1.4.16) is the analogue of (1.3.10), but now the perturbation process appearing in the drift (which we prefer to denote by $\{y^\epsilon(t),\ t \in [0,\infty)\}$ rather than by $\{\xi(t),\ t \in [0,\infty)\}$, as in (1.3.10)) is the solution of the second equation (1.4.17) whose drift $b(\cdot,\cdot)$ and covariance $\sigma(\cdot,\cdot)$ are allowed to depend upon the solution of the first equation (1.4.16). In the pair (1.4.16) and (1.4.17) the processes $\{w(t),\ t \in [0,\infty)\}$ and $\{\beta(t),\ t \in [0,\infty)\}$ are standard independent Wiener processes, of appropriate dimensions, defined on the probability space $(\Omega, \mathcal{F}, P)$. The pair of equations (1.4.16) and (1.4.17) is at the root of the so-called *Smoluchowski-Kramers approximation* in mathematical physics, where it has long been used without very much solid theoretical justification. More recently, it has also been argued in Sastry [32] that this pair of equations is a realistic model for microelectronic circuits subject to wide-band thermal noise. From the physical viewpoint, the two equations (1.4.16) and (1.4.17) become a useful mathematical model only when the second equation (1.4.17) possesses enough intrinsic "stability" to cause the perturbing process $\{y^\epsilon(t),\ t \in [0,\infty)\}$ to "average" the drift in (1.4.16). To make this idea clearer, suppose that the auxiliary stochastic differential equation

$$d\xi(t) = b\left(x, \xi(t)\right) dt + \sigma\left(x, \xi(t)\right) d\beta(t), \tag{1.4.18}$$

has a unique invariant probability measure $\pi_x(\cdot)$ for each $x \in \mathbb{R}^d$. We then define an averaged drift term by

$$\bar{F}(x) \triangleq \int_{\mathbb{R}^D} F(x,\xi) d\pi_x(\xi), \quad \forall x \in \mathbb{R}^d, \tag{1.4.19}$$

and use this to write an "averaged" version of (1.4.16), namely

$$d\bar{x}^\epsilon(t) = \epsilon \bar{F}\left(\bar{x}^\epsilon(t)\right) dt + \epsilon^{1/2} G\left(\bar{x}^\epsilon(t)\right) dw(t), \quad x^\epsilon(0) \triangleq x_0. \tag{1.4.20}$$

The intuition here is that, when ϵ is small enough, then $y^\epsilon(t)$ is varying so rapidly in comparison with $x^\epsilon(t)$ that $F(x^\epsilon(t), y^\epsilon(t))$ is virtually indistinguishable from $\bar{F}(x^\epsilon(t))$. This in turn suggests that the solution $\{x^\epsilon(t),\ t \in [0,\infty)\}$ of the rather complex equation (1.4.16) should be nicely approximated by the solution of the simpler averaged equation (1.4.20) when ϵ is small, at least over intervals of the form $[0, T/\epsilon]$. Plainly, a theorem is needed to make this rather delicate intuition precise, and such a result has been established by Khas'minskii [20], who used a rather intricate argument to show that the discrepancy

$$\max_{0 \le t \le \frac{T}{\epsilon}} |x^\epsilon(t) - \overline{x}^\epsilon(t)| \tag{1.4.21}$$

between the solutions of (1.4.16) and (1.4.20) converges to zero in probability as $\epsilon \to 0$.

1.5 Brief Summary of Research Problem

The preceding remarks provide a very condensed account of the main developments in averaging principles from their inception by Gauss up to the present time, and give enough background for us to briefly outline the main research problem of this thesis.

Goal: Our research goal concerns the pair of equations (1.4.16) and (1.4.17) which, with the exception of the early work of Khas'minskii [20], has not received very much attention, despite its importance as a mathematical model in several applications. The objective here is to establish a rate of convergence, motivated by the rate of convergence established by Liptser and Stoyanov [25] for the stochastic differential equation (1.3.10). That is, for

$$z^\epsilon(t) \triangleq \frac{x^\epsilon(t) - \overline{x}^\epsilon(t)}{\epsilon^{1/2}}, \quad \forall 0 \le t \le \frac{T}{\epsilon}, \tag{1.5.22}$$

where $\{x^\epsilon(t),\ 0 \le t \le T/\epsilon\}$ is given by (1.4.16) and (1.4.17) and $\{\overline{x}^\epsilon(t),\ 0 \le t \le T/\epsilon\}$ is given by (1.4.20), we will establish that $\{z^\epsilon(t),\ 0 \le t \le T/\epsilon\}$ converges weakly to a limit as $\epsilon \to 0$, and will charactrize this limit. Because of the extensive dependencies allowed by the model (1.4.16) and (1.4.17) this turns out to be a substantially more challenging problem than that addressed by Liptser and Stoyanov [25] and requires significantly different methods of analysis. In Chapter 3 we shall use a martingale-based method, suggested by ideas of Kurtz [10] and Papanicolaou, Stroock and Varadhan [29] to study this question.

Chapter 2

Literature Review

In this chapter we consider again the main results that were informally summarized in Chapter 1. However, we shall now pay much more attention to the careful enunciation of these results, including in particular the precise statement of the hypotheses and conclusions.

2.1 The Averaging Principle of Krylov-Bogoliubov-Mitropolskii

The *Averaging Principle* of Krylov-Bogoliubov-Mitropolskii (see e.g. [6]), is a result in the theory of non-linear ordinary differential equations which has several applications in science and engineering.

To see the idea of the averaging principle suppose that we have a differential equation in which a small parameter $\epsilon \in (0, \infty)$ appears multiplicatively as follows:

$$\frac{dx^\epsilon(t)}{dt} = \epsilon F\left(x^\epsilon(t), t\right), \quad x^\epsilon(0) \triangleq x_0. \tag{2.1.1}$$

Here $F : \mathbb{R}^d \otimes [0,\infty) \to \mathbb{R}^d, t \in [0,\infty), x_0 \in \mathbb{R}^d$, and the function $F(x,t)$ is regular enough to ensure that the differential equation (2.1.1) has a unique solution $x^\epsilon(.)$ on interval $[0,\infty)$, for each $\epsilon \in (0,\infty)$. One is usually interested in the asymptotic behaviour of $x^\epsilon(.)$ over the interval $[0, T/\epsilon]$, for some fixed $T \in (0,\infty)$, as $\epsilon \to 0$, since significant changes in the solution of (2.1.1) only occur on intervals $[0, T/\epsilon]$, when ϵ is small. Throughout this chapter, with no loss of generality, we assume $T = 1$. To investigate the limit behaviour of $x^\epsilon(t)$ over $[0, 1/\epsilon]$, it is useful to change time-scales as follows: Put

$$X^\epsilon(\tau) \triangleq x^\epsilon(\tau/\epsilon), \quad \forall \tau \in [0,1]. \tag{2.1.2}$$

Differentiating (2.1.2), and using (2.1.1), results in

$$\begin{aligned} \frac{dX^\epsilon(\tau)}{d\tau} &= \frac{d}{d\tau} x^\epsilon(\tau/\epsilon) \\ &= \frac{1}{\epsilon}\left(\frac{d}{dt}x^\epsilon(t)\right) \text{ at } t = \tau/\epsilon \\ &= \frac{1}{\epsilon}\left[\epsilon F\left(x^\epsilon(\tau/\epsilon), \tau/\epsilon\right)\right] \end{aligned} \tag{2.1.3}$$

i.e. we have

$$\frac{dX^\epsilon(\tau)}{d\tau} = F\left(X^\epsilon(\tau), \tau/\epsilon\right), \quad X^\epsilon(0) = x_0, \ \forall \tau \in [0,1]. \tag{2.1.4}$$

We observe that $\{X^\epsilon(\tau),\ 0 \le \tau \le 1\}$, the solution of (2.1.4), is equivalent to $\{x^\epsilon(t),\ 0 \le t \le 1/\epsilon\}$, the solution of (2.1.1), via the relation (2.1.2).

Suppose that the time average $\bar{F}(x)$ of the function $t \longmapsto F(x,t)$ exists for each $x \in \mathbb{R}^d$, namely

$$\bar{F}(x) \triangleq \lim_{t\to\infty} \frac{1}{t} \int_0^t F(x,s)ds. \tag{2.1.5}$$

The idea of the averaging principle of Krylov, Bogoliubov and Mitropolskii [6] is that the solution $X^\epsilon(\tau)$ of (2.1.4) can be approximated by the solution $\bar{X}(\tau)$ of the following "averaged" differential equation

$$\frac{d\bar{X}(\tau)}{d\tau} = \bar{F}\left(\bar{X}(\tau)\right), \quad \bar{X}(0) \triangleq x_0, \quad \forall \tau \in [0,1], \tag{2.1.6}$$

in a sense which is made precise by the next result:

Theorem 2.1.1. *suppose that the function* $(x, t) \mapsto F(x, t) : \mathbb{R}^d \otimes [0, \infty) \to \mathbb{R}^d$ *satisfies the following conditions:*

1. $(x, t) \to F(x, t)$ *is Borel measurable on* $\mathbb{R}^d \otimes [0, \infty)$.

2. *there is a constant* $M \in (0, \infty)$ *such that*

$$|F(x, t)| \leq M, \quad \forall x \in \mathbb{R}^d, \ \forall t \in [0, \infty), \tag{2.1.7}$$

3. *the function* $x \longmapsto F(x, t)$ *satisfies a Lipschitz condition, i.e.*

$$|F(x_1, t) - F(x_2, t)| \leq \lambda |x_1 - x_2| \tag{2.1.8}$$

for some constant $\lambda \in (0, \infty), \forall t \in [0, \infty)$ *and* $\forall x_1, x_2 \in \mathbb{R}^d$,

4. *the time average* $\bar{F}(x)$ *of (2.1.5) exists uniformly for each* $x \in \mathbb{R}^d$, *i.e.*

$$\lim_{t \to \infty} \left\{ \sup_x \left| \bar{F}(x) - 1/t \int_0^t F(x, s) ds \right| \right\} = 0. \tag{2.1.9}$$

Then (2.1.6) has a unique solution $\{\bar{X}(\tau), \ 0 \leq \tau \leq 1\}$ *and*

$$\lim_{\epsilon \to 0} \max_{\tau \in [0,1]} \left| X^\epsilon(\tau) - \bar{X}(\tau) \right| = 0. \tag{2.1.10}$$

where $\{X^\epsilon(\tau), \ 0 \leq \tau \leq 1\}$ *is the unique solution of (2.1.4) for each* $\epsilon \in (0, 1)$.

A detailed proof of this result can be found in Krylov-Bogoliubov-Mitropolskii [6], Gihman [13], and Besjes [2]. We have taken the preceding statement from Besjes ([2], Theorem 1*, page 358).

2.2 Stochastic Version of the Averaging Principle

Here we introduce a stochastic analogue of the averaging principle of Section 2.1. Let $F : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^d$ be $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^D)/\mathcal{B}(\mathbb{R}^d)$-measurable and suppose that

$\{\xi(t), 0 \leq t < \infty\}$ is an $\mathbb{R}^D$-valued jointly measurable strictly stationary random process defined on a probability space $(\Omega, \mathcal{F}, P)$, such that $(x, t) \longmapsto F(x, \xi(t))$ is regular enough to ensure that the "random" differential equation

$$\frac{dx^\epsilon(t)}{dt} = \epsilon F\left(x^\epsilon(t), \xi(t)\right), \quad x^\epsilon(0) \triangleq x_0, \tag{2.2.11}$$

has a unique solution $\{x^\epsilon(t),\ t \in [0, \infty)\}$ for each $\epsilon \in (0, \infty)$ and $\omega \in \Omega$. Using (2.1.2), as before, we first change time-scales in the equation (2.2.11) as follows: Put

$$X^\epsilon(\tau, \omega) \triangleq x^\epsilon(\tau/\epsilon, \omega), \quad \forall \tau \in [0, 1], \ \forall \omega \in \Omega. \tag{2.2.12}$$

Then, in the same way that (2.1.3) followed, we see that

$$\frac{dX^\epsilon(\tau)}{d\tau} = F\left(X^\epsilon(\tau), \xi(\tau/\epsilon)\right), \quad X^\epsilon(0) \triangleq x_0, \ \forall \tau \in [0, 1]. \tag{2.2.13}$$

Theorem 2.1.1 on page 12 seems to suggest that the solution $X^\epsilon(\tau)$ of the equation (2.2.13) is approximated, for small values of $\epsilon \in (0, \infty)$, by the solution $\bar{X}(\tau)$ of the non-random differential equation

$$\frac{d\bar{X}(\tau)}{d\tau} = \bar{F}\left(\bar{X}(\tau)\right), \quad X(0) \triangleq x_0, \ \forall \tau \in [0, 1], \tag{2.2.14}$$

where $\bar{F}(x)$ is the "averaged" value of the right-hand side of (2.2.13), defined by

$$\bar{F}(x) \triangleq EF\left(x, \xi(t)\right), \quad \forall x \in \mathbb{R}^d. \tag{2.2.15}$$

Under certain conditions this intuition has been made precise by Liptser and Shiryayev (see [24], page 722, Theorem 1) who showed:

Theorem 2.2.1. *Let $\{\xi(t), t \in [0, \infty)\}$ be a strictly stationary jointly measurable ergodic process defined on a probability space $(\Omega, \mathcal{F}, P)$, taking values in $\mathbb{R}^D$, and let the function $(x, y) \mapsto F(x, y) : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^d$ satisfy the following conditions:*

1. *the function $(x, y) \mapsto F(x, y)$ is $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^D)/\mathcal{B}(\mathbb{R}^d)$- measurable.*

2. There is a constant $L \in [0, \infty)$ *such that*

$$|F(x_1, y) - F(x_2, y)| \le L\,|x_1 - x_2|, \tag{2.2.16}$$

$\forall x_1, x_2 \in \mathbb{R}^d$ *and* $y \in \mathbb{R}^D$.

3.

$$E\,|F\,(x, \xi(0))| < \infty, \quad \forall x \in \mathbb{R}^d. \tag{2.2.17}$$

Then we have

$$\lim_{\epsilon \to 0} \left\{ \max_{\tau \in [0,1]} |X^\epsilon(\tau) - \bar{X}(\tau)| \right\} = 0, \quad a.s. \tag{2.2.18}$$

2.3 The Averaging Principle for Stochastic Differential Equations

Liptser and Stoyanov [25] extend the stochastic version of the averaging principle in Section 2.2 to Ito stochastic differential equations of the form

$$x^\epsilon(t) = x_0 + \epsilon \int_0^t F\,(x^\epsilon(s), \xi(s))\,ds + \epsilon^{1/2} \int_0^t G\,(x^\epsilon(s))\,dw(s), \tag{2.3.19}$$

where $\epsilon \in (0, \infty)$ is a small parameter, the functions $F\,(x, \xi(t))$ and $G(x)$ are regular enough to ensure that (2.3.19) has a pathwise unique strong solution for each $\epsilon \in (0, \infty)$, and $\{\xi(t), t \in [0, \infty)\}$ is a strictly stationary ergodic process and independent of the Wiener process $\{w(t), t \in [0, \infty)\}$. Among their results(see Theorem 2.3.1 on page 15 which follows) is one which shows that the solution $\{x^\epsilon(t), t \in [0, 1/\epsilon]\}$ of (2.3.19) is approximated by the solution $\{\bar{x}^\epsilon(t), t \in [0, 1/\epsilon]\}$ defined by the "averaged" stochastic differential equation of the form:

$$\bar{x}^\epsilon(t) = x_0 + \epsilon \int_0^t \bar{F}\,(\bar{x}^\epsilon(s))\,ds + \epsilon^{1/2} \int_0^t G\,(\bar{x}^\epsilon(s))\,dw(s), \tag{2.3.20}$$

which differs from (2.3.19) in that the function $\bar{F}(.)$ is now non-random and defined by "averaging" in the sense of

$$\bar{F}(x) = EF\left(\xi(t), x\right), \quad \forall x \in \mathbb{R}^d. \tag{2.3.21}$$

Before formulating this result in precise terms, we normalize from a time scale of $[0, 1/\epsilon]$ to the finite interval $[0,1]$ by defining

$$X^\epsilon(\tau,\omega) \triangleq x^\epsilon(\tau/\epsilon,\omega) \quad \text{and} \quad \bar{X}^\epsilon(\tau,\omega) \triangleq \bar{x}^\epsilon(\tau/\epsilon,\omega), \ \forall \tau \in [0,1], \ \forall \omega \in \Omega, \tag{2.3.22}$$

and

$$W^\epsilon(\tau,\omega) \triangleq \epsilon^{1/2} w(\tau/\epsilon,\omega), \quad \forall \tau \in [0,1], \ \forall \omega \in \Omega. \tag{2.3.23}$$

Then we can re-write (2.3.19) and (2.3.20) as

$$X^\epsilon(\tau) = x_0 + \int_0^\tau F\left(X^\epsilon(s), \xi(s/\epsilon)\right) ds + \int_0^\tau G\left(X^\epsilon(s)\right) dW^\epsilon(s), \quad \forall \tau \in [0,1], \tag{2.3.24}$$

and

$$\bar{X}^\epsilon(\tau) = x_0 + \int_0^\tau \bar{F}\left(\bar{X}^\epsilon(s)\right) ds + \int_0^\tau G\left(\bar{X}^\epsilon(s)\right) dW^\epsilon(s), \quad \forall \tau \in [0,1], \tag{2.3.25}$$

respectively. It is well-known from the re-scaling property that $\{W^\epsilon(\tau),\ \tau \in [0,1]\}$ in (2.3.23) is a standard Wiener process on $(\Omega, \mathcal{F}, P)$, for each $\epsilon \in (0,\infty)$.

The following Theorem, established by Liptser and Stoyanov ([25], Theorem 1), formulates an averaging principle for stochastic differential equations in the preceding context:

Theorem 2.3.1. *Suppose that $\{\mathcal{F}_t,\ t \in [0,\infty)\}$ is a filtration in $(\Omega, \mathcal{F}, P)$ such that $\mathcal{F}_0$ includes all P-null events in $\mathcal{F}$, $\{(\xi(t), \mathcal{F}_t),\ t \in [0,\infty)\}$ is a strictly stationary $\mathbb{R}^D$-valued progressively measurable ergodic process, $\{(w(t), \mathcal{F}_t); t \in [0,\infty)\}$ is an $\mathbb{R}^M$-valued standard Wiener process, and the following conditions hold:*

1. *functions* $(x, y) \mapsto F(x, y) : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^d$ *and* $x \mapsto G(x) : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^M$ *are* $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^D)/\mathcal{B}(\mathbb{R}^d)$ *and* $\mathcal{B}(\mathbb{R}^d)/\mathcal{B}(\mathbb{R}^d \otimes \mathbb{R}^D)$ *measurable, respectively.*

2. $F(x, y)$ *and* $G(x)$ *are globally Lipschitz continuous and satisfy linear growth condition with respect to* x, *uniformly in* y: *namely, for some constant* $L \in (0, \infty)$, *we have*

$$\begin{aligned} |F(x_1, y) - F(x_2, y)| &\le L|x_1 - x_2|, \quad \forall x_1, x_2 \in \mathbb{R}^d, \ \forall y \in \mathbb{R}^D, \\ |F(x, y)| &\le L(1 + |x|), \quad \forall x \in \mathbb{R}^d, \ \forall y \in \mathbb{R}^D, \\ |G(x_1) - G(x_2)| &\le L|x_1 - x_2|, \quad \forall x_1, x_2 \in \mathbb{R}^d, \\ |G(x)| &\le L(1 + |x|), \quad \forall x \in \mathbb{R}^d. \end{aligned} \tag{2.3.26}$$

3.

$$E|F(x, \xi(0))| < \infty, \quad \forall x \in \mathbb{R}^d \tag{2.3.27}$$

4. *The processes* $\{\xi(t),\ t \in [0, \infty)\}$ *and* $\{w(t),\ t \in [0, \infty)\}$ *are independent*

Then for each $\delta > 0$, *one has*

$$\lim_{\epsilon \to 0} P\left\{ \max_{\tau \in [0,1]} |X^\epsilon(\tau) - \bar{X}^\epsilon(\tau)| \ge \delta \right\} = 0, \tag{2.3.28}$$

where $X^\epsilon(.)$ *and* $\bar{X}^\epsilon(.)$, *for each* $\epsilon \in (0, \infty)$, *are the unique strong solutions of (2.3.24) and (2.3.25), respectively.*

Remark 2.3.2. Notice that the conditions postulated in Theorem 2.3.1 on page 15 are sufficient to ensure that (2.3.24) has a pathwise unique strong solution $\{X^\epsilon(\tau),\ \tau \in [0, 1]\}$ and (2.3.25) has a pathwise unique strong solution $\{\bar{X}^\epsilon(\tau),\ \tau \in [0, 1]\}$ for each $\epsilon \in (0, \infty)$. This is ensured, for example, by Theorem 5.1.1 of Kallianpur [18].

2.3.1 Rate of Convergence in Theorem 2.3.1 on page 15

Liptser and Stoyanov (see [25], Theorem 2.1), also establish *a rate of convergence* in (2.3.28) in the following form:
Define the family of processes $\{Z^\epsilon(\tau), \quad 0 \leq \tau \leq 1\}$ by

$$Z^\epsilon(\tau) \triangleq \epsilon^{-1/2}\left[X^\epsilon(\tau) - \bar{X}^\epsilon(\tau)\right], \quad \forall\epsilon \in (0,\infty), \ \forall\tau \in [0,1], \tag{2.3.29}$$

where $X^\epsilon(.)$ and $\bar{X}^\epsilon(.)$ are the unique strong solutions of equations (2.3.24) and (2.3.25), respectively. Thus $Z^\epsilon(.)$ normalizes the discrepancy between $X^\epsilon(.)$ and $\bar{X}^\epsilon(.)$ by a factor of $\epsilon^{1/2}$. Liptser and Stoyanov (see [25], Theorem 2.1), established that the process $\{Z^\epsilon(\tau), \ 0 \leq \tau \leq 1\}$ converges weakly to some well-defined limit as $\epsilon \to 0$ (under additional assumptions for the functions $F(.,.)$ and $G(.)$ which are formulated precisely in the next paragraph). This suggests that the process $\{X^\epsilon(\tau), \tau \in [0,1]\}$ converges weakly to the simpler process $\{\bar{X}^\epsilon(\tau), \tau \in [0,1]\}$ at a "rate" which is $O(\epsilon^{1/2})$ as $\epsilon \to 0$.

We next formulate the regularity conditions which are needed for this result to hold, and then present a precise statement of the result. Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space, $\{\mathcal{F}_t, t \in (-\infty,\infty)\}$ is a filtration in this probability space such that $\mathcal{F}_0$ includes all P-null events in $\mathcal{F}$, and the following Conditions $(B0) - (B6)$ hold for (2.3.19).

(B0) Mappings $(x,y) \to F(x,y) : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^d$ and $x \to G(x) : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^M$ are Borel measurable, and there is a constant $L \in (0,\infty)$, such that

$$\begin{aligned} |F(x_1,y) - F(x_2,y)| &\leq L|x_1 - x_2|, \quad \forall x_1,x_2 \in \mathbb{R}^d, \ \forall y \in \mathbb{R}^D, && (2.3.30) \\ |G(x_1) - G(x_2)| &\leq L|x_1 - x_2|, \quad \forall x_1,x_2 \in \mathbb{R}^d && (2.3.31) \\ |F(x,y)| &\leq L(1+|x|), \quad \forall x \in \mathbb{R}^d, \ \forall y \in \mathbb{R}^D, && (2.3.32) \\ |G(x)| &\leq L(1+|x|), \quad \forall x \in \mathbb{R}^d && (2.3.33) \end{aligned}$$

(B1) The functions F and G are continuously differentiable in x such that their derivatives are bounded, and satisfy a Lipschitz condition. Namely, for some constant $L \in [0, \infty)$, we have

$$|\frac{\partial F^i(x,y)}{\partial x^k}| \leq L, \quad \forall x \in I\!R^d, \forall y \in I\!R^D, \ 1 \leq i,k \leq d \tag{2.3.34}$$

$$|\frac{\partial G^{ij}(x)}{\partial x^k}| \leq L, \quad \forall x \in I\!R^d, \ 1 \leq i,k \leq d, \ 1 \leq j \leq M, \tag{2.3.35}$$

$$|\frac{\partial F^i(x_1,y)}{\partial x^k} - \frac{\partial F^i(x_2,y)}{\partial x^k}| \leq L|x_1 - x_2|, \quad \forall x_1, x_2 \in I\!R^d, \ \forall y \in I\!R^D, \ 1 \leq i,k \leq d \tag{2.3.36}$$

$$|\frac{\partial G^{ij}(x_1)}{\partial x^k} - \frac{\partial G^{ij}(x_2)}{\partial x^k}| \leq L|x_1 - x_2|, \quad \forall x_1, x_2, x \in I\!R^d, \ 1 \leq i,k \leq d, \ 1 \leq j \leq M. \tag{2.3.37}$$

(B2) The initial condition $x_0 \in I\!R^d$ in (2.3.19) and (2.3.20) is a non-random constant.

(B3) $\{(w(t), \mathcal{F}_t); t \in [0, \infty)\}$ is an $I\!R^M$- valued standard Wiener process on $(\Omega, \mathcal{F}, P)$.

(B4) $\{(\xi(t), \mathcal{F}_t); \ t \in (-\infty, \infty)\}$ is an $I\!R^D$- valued progressively measurable process on $(\Omega, \mathcal{F}, P)$, and $\{\xi(t); t \in (-\infty, \infty)\}$ is strictly stationary and ergodic. We define

$$\bar{F}(x) \triangleq E\left[F(x, \xi(t))\right], \quad \forall x \in I\!R^d. \tag{2.3.38}$$

(B5) The processes $\{(w(t); t \in [0, \infty)\}$ and $\{\xi(t); t \in (-\infty, \infty)\}$ are independent.

For each $x \in I\!R^d$, let $\{\eta_t(x), \ t \in [0, \infty)\}$ be an $I\!R^d$-valued process defined by

$$\eta_t(x) \triangleq F(x, \xi(t)) - \bar{F}(x), \quad \forall x \in I\!R^d. \tag{2.3.39}$$

In view of Condition (B4), the process $\eta(x) \triangleq \{\eta_t(x),\ t \in [0, \infty)\}$, for each $x \in \mathbb{R}^d$, is strictly stationary and ergodic with $E\eta_t(x) = 0$. Here is an additional assumption on the process $\xi(.)$ given in terms of the process $\eta(.)$:

(B6) For some $p > 2$, we have

$$\int_0^\infty \operatorname{ess\,sup}_x \| E\left(\eta_t(x) | \mathcal{F}_0^\xi\right) \|_p dt < \infty, \tag{2.3.40}$$

where $\|Q\|_p \triangleq (E(|Q|^p))^{1/p}$ and $\mathcal{F}_0^\xi$ denotes the σ-algebra generated by the process ξ up to time 0, i.e. $\mathcal{F}_0^\xi \triangleq \sigma\{\xi(t),\ -\infty < t \le 0\}$.

Remark 2.3.3. If there existed perfect independence of $\eta_t(x)$ and $\mathcal{F}_0^\xi$, then we would have

$$E\left(\eta_t(x)|\mathcal{F}_0^\xi\right) = E(\eta_t(x)) = 0. \tag{2.3.41}$$

An ergodic process $\{\xi(t)\}$ generally fails to have a strong independence property of this kind, and Condition $(B6)$ expresses a decaying dependence of $\eta_t(x)$ on $\mathcal{F}_0^\xi$ as $t \to \infty$, i.e. $E\left(\eta_t(x)|\mathcal{F}_0^\xi\right)$ is "getting small" as $t \to \infty$.

Remark 2.3.4. Condition $(B6)$ implies that

$$\nu(x) \triangleq \int_0^\infty E\left(\eta_t(x)\eta_0^\top(x)\right) dt + \int_0^\infty E\left(\eta_0(x)\eta_t^\top(x)\right) dt \tag{2.3.42}$$

is defined and is a symmetric $d \otimes d$ positive semi-definite matrix for each $x \in \mathbb{R}^d$. It then follows from linear algebra (see Exercise 12.46 of Noble [27]) that, for each $x \in \mathbb{R}^d$, there is a *unique* symmetric positive semi-definite $d \otimes d$ matrix $\nu^{1/2}(x)$ such that

$$\nu(x) = (\nu^{1/2}(x)).(\nu^{1/2}(x))^\top, \quad \forall x \in \mathbb{R}^d, \tag{2.3.43}$$

and the mapping $x \to \nu^{1/2}(x) : \mathbb{R}^d \to \mathbb{R}^{d\otimes d}$ is $\mathcal{B}(\mathbb{R}^d)/\mathcal{B}(\mathbb{R}^{d\otimes d})$ - measurable (see comments on page 171 of Karatzas and Shreve [19]). The square-root $\nu^{1/2}(\cdot)$ will shortly be needed when we formulate Theorem 2.3.6 on page 20 to follow on rates of convergence.

Remark 2.3.5. Let $B(x, z)$ be the $d \otimes M$ matrix defined for each $x, z \in \mathbb{R}^d$ by

$$B(x, z) \triangleq (\partial_x G)(x)[z],$$

(see Item IV in "Basic Notation and Terminology"); that is, the (i, j)th entry of $B(x, z)$ is given by the inner product

$$B^{ij}(x, z) \triangleq (\partial_x G^{ij})(x)z, \qquad (2.3.44)$$

for $i = 1, 2, ..., d,\ \ j = 1, 2, ..., M$, where we recall from Item IV of "Basic Notation and Terminology" that $(\partial_x G^{ij})(x)$ denotes the *row vector* of length d whose k-th entry is $(\partial_{x^k} G^{ij})(x)$.

Now we can formulate the following rate-of-convergence result which was established by Liptser and Stoyanov [25] for Theorem 2.3.1 on page 15:

Theorem 2.3.6. *Suppose Conditions $(B0)$ to $(B6)$ hold and recall (2.3.29). Then*

$$Z^\epsilon \xrightarrow{d} Z, \quad \text{as } \epsilon \to 0, \qquad (2.3.45)$$

where $\{Z(\tau),\ \tau \in [0, 1]\}$ is the solution of the following stochastic differential equation

$$Z(\tau) = \int_0^\tau (\partial_x \bar{F})(X(s))Z(s)ds + \int_0^\tau B(X(s), Z(s))dW(s) + \int_0^\tau \nu^{1/2}(X(s))d\beta(s), \qquad (2.3.46)$$

in which $\{\beta(\tau),\ \tau \in [0, 1]\}$ and $\{W(\tau),\ \tau \in [0, 1]\}$ are standard independent Wiener processes on a common probability space, with $\{\beta(\tau)\}$ being $\mathbb{R}^d$-valued and $\{W(\tau)\}$ being $\mathbb{R}^M$-valued.

Remark 2.3.7. The relevance of Theorem 2.3.6 on page 20 can be explained by some ideas from simple probability theory. Suppose that a sequence $\{\xi_n,\ n = 1, 2, 3, \dots\}$ of random variables on a probability space $(\Omega, \mathcal{F}, P)$ is subject to a *weak law of large numbers* of the following form: There is some number $\bar{x} \in \mathbb{R}$ such that

$$x_n \triangleq \frac{1}{n}\sum_{i=1}^{n}\xi_i \tag{2.3.47}$$

converges *in probability* to the limit $\bar{x}$, namely

$$\lim_{n\to\infty} P\{|x_n - \bar{x}| > \delta\} = 0, \tag{2.3.48}$$

for each $\delta > 0$. Then a natural question is how to characterize a possible rate of convergence of x_n to the limit $\bar{x}$ (in this weak law of large numbers). One typically does this by normalizing the discrepancy $(x_n - \bar{x})$ by a non-random factor $h(n)$, where $h(n) \to 0$, as $n \to \infty$, and asking if $h(.)$ can be chosen to ensure that

$$Z_n \triangleq \frac{x_n - \bar{x}}{h(n)} \tag{2.3.49}$$

converges *weakly* to some limiting random variable Z as $n \to 0$. Of course the possibility of doing this requires that $\{\xi_n\}$ exhibit some form of "partial independence". If this partial dependence is strong enough then one can typically show that

$$Z_n \xrightarrow{d} \text{a Gaussian limit Z} \tag{2.3.50}$$

when

$$h(n) \triangleq \frac{1}{\sqrt{n}}, \tag{2.3.51}$$

(this is just the Central Limit Theorem). Going back to the context of Theorem 2.3.1 on page 15, we have a somewhat analogous situation, in which the "function - space valued" random variables $X^\epsilon(.)$ and $\overline{X}^\epsilon(.)$ play roles similar to x_n and $\bar{x}$ in the preceding discussion, and the convergence in (2.3.28) is just a weak law of large numbers in "functional form" which is analogous to (2.3.48). The question then arises as to how to establish an analogue of (2.3.50) in the context of Theorem 2.3.1 on page 15. It is natural to proceed by defining $Z^\epsilon(.)$ as in (2.3.29), to be the "function - space" analogue of Z_n in (2.3.49) with $h(n)$ given by (2.3.51), and then to seek a weak limit of $Z^\epsilon(.)$ as $\epsilon \to 0$. This is provided by Theorem 2.3.6 on page 20, in which (2.3.45) is the analogue of (2.3.50) except that in this case the limit in (2.3.45) no longer turns out to be Gaussian as it is in (2.3.50). In summary, we can regard Theorem 2.3.6 on page 20 as a type of "central limit theorem" for the "weak law of large numbers" given by Theorem 2.3.1 on page 15.

2.4 The Averaging Principle for Coupled Itô equations

In the preceding section, we have seen an averaging principle when the perturbing process $\{\xi(t),\ t \in [0,\infty)\}$ in (2.3.19) is independent of $\{w(t),\ t \in [0,\infty)\}$. In this section, we shall introduce a case in which the perturbing process is constructed by a stochastic differential equation conditioned by the slowly varing process $\{x^\epsilon(t),\ t \in [0,\infty)\}$ and is hence conditioned by the driving Wiener process $\{w(t),\ t \in [0,\infty)\}$ in the stochastic differential equation for $\{x^\epsilon(t),\ t \in [0,\infty)\}$. To be precise we consider the *coupled* stochastic differential equations

$$x^\epsilon(t) = x_0 + \epsilon \int_0^t F(x^\epsilon(s), y^\epsilon(s))ds + \epsilon^{1/2} \int_0^t G(x^\epsilon(s))dw(s), \quad (2.4.52)$$

$$y^\epsilon(t) = y_0 + \int_0^t b(x^\epsilon(s), y^\epsilon(s))ds + \int_0^t \sigma(x^\epsilon(s), y^\epsilon(s))d\beta(s), \quad (2.4.53)$$

$\forall t \in [0,\infty)$, where $\{w(t),\ t \in [0,\infty)\}$ and $\{\beta(t),\ t \in [0,\infty)\}$ are independent $\mathbb{R}^M$ and $\mathbb{R}^N$-valued Wiener processes respectively, defined on a probability space $(\Omega, \mathcal{F}, P)$. The idea here is to postulate that the stochastic differential equation

$$d\xi(t) = b(x, \xi(t))dt + \sigma(x, \xi(t))d\beta(t), \quad (2.4.54)$$

is "stable", in the sense that the Markov process arising from (2.4.54) has a *unique* invariant probability measure π_x on $\mathbb{R}^d$, for each $x \in \mathbb{R}^d$. Now define an "averaged" drift

$$\bar{F}(x) \triangleq \int_{\mathbb{R}^D} F(x,\xi)d\pi_x(\xi), \quad \forall x \in \mathbb{R}^d, \tag{2.4.55}$$

and use this to write the "averaged" version of (2.4.52) namely

$$\overline{x}^\epsilon(t) = x_0 + \epsilon\int_0^t \bar{F}\left(\overline{x}^\epsilon(s)\right) ds + \epsilon^{1/2}\int_0^t G\left(\overline{x}^\epsilon(s)\right) dw(s), \quad \forall t \in [0,\infty). \tag{2.4.56}$$

From (2.4.52) and (2.4.53) we see that $y^\epsilon(t)$ varies much more rapidly than $x^\epsilon(t)$, which suggests that $F(x^\epsilon(s), y^\epsilon(s))$ is "close" to $\bar{F}(x^\epsilon(s))$, and hence that $\{x^\epsilon(t)\}$ arising from (2.4.52) may be well-approximated by $\{\overline{x}^\epsilon(t)\}$ arising from (2.4.56). It turns out that this intuition is correct provided that we limit attention to time intervals $t \in [0, 1/\epsilon]$. This intuition is justified by the next theorem, essentially due to Khas'minskii [20] (also given as Theorem 9.1 on page 264 of the book [11] by Freidlin and Wentzel), which says that

$$\max_{0\le t\le 1/\epsilon} |x^\epsilon(t) - \overline{x}^\epsilon(t)| \tag{2.4.57}$$

goes to zero in probability as $\epsilon \to 0$. Before presenting this result, it is convenient to normalize the time scale of $[0, 1/\epsilon]$ to the finite interval $[0, 1]$ by defining

$$\begin{aligned} X^\epsilon(\tau) &\triangleq x^\epsilon(\tau/\epsilon), \qquad \overline{X}^\epsilon(\tau) \triangleq \overline{x}^\epsilon(\tau/\epsilon), \\ Y^\epsilon(\tau) &\triangleq y^\epsilon(\tau/\epsilon), \qquad W^\epsilon(\tau) \triangleq \epsilon^{1/2}w(\tau/\epsilon), \\ &B^\epsilon(\tau) \triangleq \epsilon^{1/2}\beta(\tau/\epsilon), \end{aligned} \tag{2.4.58}$$

for all $\epsilon \in (0, 1]$ and $\tau \in [0, 1]$. Then, from (2.4.52), (2.4.53) and (2.4.56), we see that

$$X^\epsilon(\tau) = x_0 + \int_0^\tau F\left(X^\epsilon(s), Y^\epsilon(s)\right) ds + \int_0^\tau G(X^\epsilon(s))dW^\epsilon(s), \tag{2.4.59}$$

$$Y^\epsilon(\tau) = y_0 + \epsilon^{-1}\int_0^\tau b\left(X^\epsilon(s), Y^\epsilon(s)\right) ds + \epsilon^{-1/2}\int_0^\tau \sigma\left(X^\epsilon(s), Y^\epsilon(s)\right) dB^\epsilon(s), \tag{2.4.60}$$

and

$$\overline{X}^\epsilon(\tau) = x_0 + \int_0^\tau \bar{F}\left(\overline{X}^\epsilon(s)\right) ds + \int_0^\tau G(\overline{X}^\epsilon(s))dW^\epsilon(s). \tag{2.4.61}$$

For the system of equations (2.4.52), (2.4.53), we assume that the following conditions hold:

(C0) The mappings $F : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^d$, $G : \mathbb{R}^d \to \mathbb{R}^{d\otimes M}$, $b : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^{d\otimes N}$ are uniformly bounded and satisfy a Lipschitz condition, i.e. there is a constant $L \in [0, \infty)$ such that

$$\begin{aligned}
|F(x_1, y_1) - F(x_2, y_2)| &\leq L\left[|x_1 - x_2| + |y_1 - y_2|\right], \quad \forall x_1, x_2 \in \mathbb{R}^d,\ \forall y_1, y_2 \in \mathbb{R}^D \\
|G(x_1) - G(x_2)| &\leq L\left[|x_1 - x_2|\right], \quad \forall x_1, x_2 \in \mathbb{R}^d, \\
|b(x_1, y_1) - b(x_2, y_2)| &\leq L\left[|x_1 - x_2| + |y_1 - y_2|\right], \quad \forall x_1, x_2 \in \mathbb{R}^d,\ \forall y_1, y_2 \in \mathbb{R}^D \\
|\sigma(x_1, y_1) - \sigma(x_2, y_2)| &\leq L\left[|x_1 - x_2| + |y_1 - y_2|\right], \quad \forall x_1, x_2 \in \mathbb{R}^d,\ \forall y_1, y_2 \in \mathbb{R}^D
\end{aligned} \tag{2.4.62}$$

and

$$\begin{aligned}
\sup_{(x,y)\in\mathbb{R}^d\otimes\mathbb{R}^D} |F(x, y)| &< \infty, \\
\sup_{x\in\mathbb{R}^d} |G(x)| &< \infty, \\
\sup_{(x,y)\in\mathbb{R}^d\otimes\mathbb{R}^D} |b(x, y)| &< \infty, \\
\sup_{(x,y)\in\mathbb{R}^d\otimes\mathbb{R}^D} |\sigma(x, y)| &< \infty.
\end{aligned} \tag{2.4.63}$$

(C1) For each $x \in \mathbb{R}^d$, there is a unique invariant probability measure π_x on $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ for the Markov process $\{\xi(t, x),\ t \in [0, \infty)\}$ defined by the stochastic differential equation

$$d\xi(t, x) = b(x, \xi(t, x))dt + \sigma(x, \xi(t, x))d\beta_t. \tag{2.4.64}$$

(C2) The mapping $\bar{F} : \mathbb{R}^d \to \mathbb{R}^d$ defined by

$$\bar{F}(x) \triangleq \int_{\mathbb{R}^D} F(x, y)\pi_x(dy), \quad x \in \mathbb{R}^d. \tag{2.4.65}$$

is globally Lipschitz continuous, namely there is a constant $C \in [0, \infty)$ such that

$$\left|\bar{F}(x_1) - \bar{F}(x_2)\right| \leq C|x_1 - x_2|, \quad \forall x_1, x_2 \in \mathbb{R}^d. \tag{2.4.66}$$

Moreover

$$\lim_{T\to\infty} \left(\sup_{\substack{t\in[0,\infty) \\ (x,y)\in\mathbb{R}^d\otimes\mathbb{R}^D}} E\left|\frac{1}{T}\int_t^{t+T} F\left(x, \xi(\tau, x, y)\right)\, d\tau - \bar{F}(x)\right| \right) = 0, \tag{2.4.67}$$

where $\{\xi(t, x, y),\ t \in [0, \infty)\}$ is the solution of (2.4.64) with initial condition $\xi(0, x) = y$.

(C3) The processes $\{w(t),\ t \in [0, \infty)\}$ and $\{\beta(t),\ t \in [0, \infty)\}$ in (2.4.52) and (2.4.53) are independent Wiener processes on a common probability space $(\Omega, \mathcal{F}, P)$, $w(t)$ is $\mathbb{R}^M$-valued and $\beta(t)$ is $\mathbb{R}^N$-valued.

Theorem 2.4.1. *Suppose Conditions* $(C0)$, $(C1)$, $(C2)$ *and* $(C3)$ *hold for the system of equations (2.4.52), (2.4.53). Then for any* $\delta > 0$, *we have*

$$\lim_{\epsilon\to 0} P\left\{ \sup_{0\leq\tau\leq 1} \left|X^\epsilon(\tau) - \bar{X}^\epsilon(\tau)\right| > \delta \right\} = 0. \tag{2.4.68}$$

2.5 Goals and Organization of Thesis

Theorem 2.4.1 on page 25, which pertains to the coupled system (2.4.52) and (2.4.53), is an obvious analogue to Theorem 2.3.1 on page 15 for the system (2.3.19). As such, it can be viewed as a type of weak law of large numbers in *functional form*, for the system (2.4.52) and (2.4.53). We saw in Remark 2.3.7 that Theorem 2.3.6 on page 20 gives a "central limit theorem" which complements the weak law of large numbers provided by Theorem 2.3.1 on page 15 for the system (2.3.19). It is therefore natural to try to establish a central limit theorem which is an analogue of Theorem 2.3.6 on page 20, and which bears the same general relation to the weak law of large numbers in Theorem 2.4.1 on page 25 as the central limit theorem of Theorem 2.3.6 on page 20 does to the weak law of large numbers in Theorem 2.3.1 on page 15. Establishing a result of this kind is our main research problem in the present thesis, and is addressed in the next chapter.

The organization of the thesis is as follows: In Section 3.2 of Chapter 3 we shall declare basic conditions on the coefficients of the coupled system (2.4.52) and (2.4.53) which are sufficient to ensure that such a central limit theorem does indeed hold, and in Section 3.3 we shall state the main result of the thesis (see Theorem 3.3.3 on page 42). The conditions that we postulate will be in fairly abstract form, and in particular will entail solvability of certain Poisson-type partial differential equations associated with the linear second-order differential operator arising from the coefficients of the second of the coupled equations (2.4.53). In Section 3.4 of Chapter 3 we shall therefore formulate some simple sufficient conditions on the coefficients of (2.4.52), (2.4.53), which are enough to ensure verification of the rather abstract conditions in Section 3.4. The proof of the main result (Theorem 3.3.3) on page 42 is given in Appendix A, and all results stated in Section 3.4 are proved in Appendix B. Appendix C lists some standard miscellaneous technical results, needed for the proofs in Appendices A and B, that we have collected for easy reference. Appendices D and E summarize standard background information, on ergodicity and mixing, and on solvability of Poisson equations associated with discrete-parameter Markov chains, that are useful for the discussions in the thesis. Finally, Appendix F summarizes the main ideas due to Gihman and Skorohod [14] on so-called L_2-derivatives of solutions of SDE's; the ideas and results summarized here are needed for the proofs in Appendix B.

Chapter 3

Averaging for Coupled Itô Equations

3.1 Introduction

In this chapter we consider the coupled Itô equations (2.4.52) and (2.4.53), which we reproduce for convenience as follows

$$x^\epsilon(t) \triangleq x_0 + \epsilon \int_0^t F(x^\epsilon(s), y^\epsilon(s))ds + \epsilon^{1/2} \int_0^t G(x^\epsilon(s))dw(s), \tag{3.1.1}$$

$$y^\epsilon(t) \triangleq y_0 + \int_0^t b(x^\epsilon(s), y^\epsilon(s))ds + \int_0^t \sigma(x^\epsilon(s), y^\epsilon(s))d\beta(s). \tag{3.1.2}$$

Here $x_0 \in \mathbb{R}^d$ and $y_0 \in \mathbb{R}^D$ are nonrandom, and $\{w(t),\ t \in [0,\infty)\}$ and $\{\beta(t),\ t \in [0,\infty)\}$ are independent standard Wiener processes on the same probability space $(\Omega, \mathcal{F}, P)$, taking values in $\mathbb{R}^M$ and $\mathbb{R}^N$ respectively.

Suppose that the mappings $b(.,.)$ and $\sigma(.,.)$ in (3.1.2) are such that, for each $x \in \mathbb{R}^d$, there is a unique invariant probability measure π_x on $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ for the

Markov process $\{\xi(t,x),\ t \in [0,\infty)\}$ defined by

$$d\xi(t) = b(x,\xi(t))dt + \sigma(x,\xi(t))d\beta(t). \tag{3.1.3}$$

Then we can define an "averaged" drift

$$\overline{F}(x) \triangleq \int_{\mathbb{R}^D} F(x,\xi)d\pi_x(\xi), \qquad \forall\, x \in \mathbb{R}^d. \tag{3.1.4}$$

and use this to define an "averaged version" of (3.1.1) namely

$$\overline{x}^\epsilon(t) = x_0 + \epsilon \int_0^t \bar{F}(\overline{x}^\epsilon(s))ds + \epsilon^{1/2}\int_0^t G(\overline{x}^\epsilon(s))dw(s), \quad 0 \le t \le \epsilon^{-1}. \tag{3.1.5}$$

In Theorem 2.4.1 on page 25 conditions were introduced which are sufficient to ensure that

$$\sup_{0\le\tau\le1} \left|X^\epsilon(\tau) - \overline{X}^\epsilon(\tau)\right| \tag{3.1.6}$$

goes to zero in probability as $\epsilon \to 0$, where we have used the re-scalings

$$X^\epsilon(\tau) \triangleq x^\epsilon(\tau/\epsilon) \quad \text{and} \quad \overline{X}^\epsilon(\tau) \triangleq \overline{x}^\epsilon(\tau/\epsilon), \;\; \forall 0 \le \tau \le 1. \tag{3.1.7}$$

As was noted in Remark 2.3.7 and Section 2.5, this result is a type of "weak law of large numbers", and the question of an associated rate of convergence naturally arises. In this chapter we will look at the "normalized discrepancy process" $\{Z^\epsilon(\tau),\ 0 \le \tau \le 1\}$, defined by

$$Z^\epsilon(\tau) \triangleq \frac{X^\epsilon(\tau) - \overline{X}^\epsilon(\tau)}{\epsilon^{1/2}}, \quad 0 \le \tau \le 1, \tag{3.1.8}$$

and try to establish reasonable and natural conditions on the mappings $F(.,.), G(.)$, $b(.,.)$ and $\sigma(.,.)$ in (3.1.1) and (3.1.2) which are sufficient to ensure weak convergence of $\{Z^\epsilon(\tau),\ 0 \le \tau \le 1\}$ to a limiting process $\{\hat{Z}(\tau),\ 0 \le \tau \le 1\}$. We shall also fully characterize this limit process. The motivation for this is, of course, Theorem 2.3.6 on page 20 which provides a similar type of rate of convergence for the "weak law of large numbers" given by Theorem 2.3.1 on page 15 for averaging in the Itô stochastic differential equation (2.3.19). Thus our goal in this chapter is really to establish an analogue of Theorem 2.3.6 on page 20 but for the system (3.1.1), (3.1.2).

There is in fact a special case of such a result due to Skorohod (see §3 and Theorem 15 on page 163 of [33]) who studies the system (3.1.1) and (3.1.2) with $G \equiv 0$ (so that (3.1.1) is an ordinary differential equation) and it is *assumed that the averaged drift* $\bar{F}(.)$ in (3.1.4) is *identically zero*. This is the so - called "neutral case" in which the averaged equation (3.1.5) is trivial (since $G \equiv 0$ and $\bar{F} \equiv 0$ by assumption) and it therefore follows from (3.1.5) that

$$\bar{x}^{\epsilon}(t) \equiv x_0. \tag{3.1.9}$$

With these additional assumptions it is established in Theorem 15 on page 163 of [33] that

$$\frac{X^{\epsilon}(\tau) - x_0}{\sqrt{\epsilon}}, \quad 0 \le \tau \le 1 \tag{3.1.10}$$

converges weakly to a limiting process (which is Gaussian in this case). Extending this result to the "non-neutral case", in which we do *not* suppose that $G \equiv 0$ and $\bar{F} \equiv 0$, involves definite technical challenges, and is the main goal of this chapter. Our approach to this problem will be quite different from that found in Liptser and Stoyanov [25] and Skorohod [33], and will be based on a method of singularly perturbed martingale problems introduced by Papanicolaou, Stroock and Varadhan [29] for problems in the theory of turbulent flows.

3.2 Conditions

The following conditions will henceforth be supposed for the equations (3.1.1) and (3.1.2):

Condition 3.2.1. The mappings $F : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^d$, $b : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^D$, $\sigma : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}^{D\otimes N}$ in (3.1.1) and (3.1.2) are linearly bounded and locally Lipschitz continuous. That is, there is a constant $C_1 \in [0, \infty)$ such that

$$|F(x, y)| \leq C_1[1 + |x| + |y|], \quad \forall\, x \in \mathbb{R}^d, \quad \forall\, y \in \mathbb{R}^D, \tag{3.2.11}$$

and, for each $R \in (0, \infty)$, there is some constant $C(R) \in [0, \infty)$ such that

$$|F(x_1, y_1) - F(x_2, y_2)| \leq C(R)[|x_1 - x_2| + |y_1 - y_2|], \tag{3.2.12}$$

for all $(x_i, y_i) \in S_R^d \otimes S_R^D$, $i = 1, 2$, with similar bounds for $b(\cdot, \cdot)$, $\sigma(\cdot, \cdot)$.

Also, $G : \mathbb{R}^d \to \mathbb{R}^{d\otimes M}$ is linearly bounded and a globally Lipschitz continuous and C^2-function. In particular, there exists a constant $L_2 \in [0, \infty)$ such that

$$|G(x_1) - G(x_2)| \leq L_2|x_1 - x_2|, \quad \forall x_1, x_2 \in \mathbb{R}^d$$

Condition 3.2.2. In (3.1.1) and (3.1.2) the initial values $x_0 \in \mathbb{R}^d$ and $y_0 \in \mathbb{R}^D$ are nonrandom. Also, $\{(w(t), \mathcal{F}_t);\, t \in [0, \infty)\}$ and $\{(\beta(t), \mathcal{F}_t);\, t \in [0, \infty)\}$ are independent standard Wiener processes, $\mathbb{R}^M$ and $\mathbb{R}^N$-valued respectively, defined on the complete probability space $(\Omega, \mathcal{F}, P)$, and each $\mathcal{F}_t$ includes all P-null events in $\mathcal{F}$. Without loss of generality we shall take $\mathcal{F}_t$ to be given by

$$\mathcal{F}_t \triangleq \sigma\{w(s), \beta(s), 0 \leq s \leq t\} \vee \{\text{P-null events in } \mathcal{F}\}. \tag{3.2.13}$$

We next define the drift $\overline{F}(x)$ in the averaged equation (3.1.5) and formulate conditions on this drift:

Condition 3.2.3. For each $x \in \mathbb{R}^d$ there is a unique invariant probability measure π_x on $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ for the Markov process $\{\xi(t, x)\}$defined by (3.1.3). The integral in (3.1.4) exists for each $x \in \mathbb{R}^d$, and the mapping $\overline{F} : \mathbb{R}^d \to \mathbb{R}^d$ defines a globally Lipschitz continuous and C^2-function.

Remark 3.2.4. By Theorem 5.2.5 of Karatzas and Shreve [19], Conditions 3.2.1, 3.2.2, and 3.2.3 ensure that the coupled Itô equations (3.1.1), (3.1.2) have a pathwise unique strong solution $\{(x^\epsilon(t), y^\epsilon(t)),\ t \in [0, \infty)\}$. Likewise, we see that (3.1.5) with $\bar{F}(x)$ given by (3.1.4) has a pathwise unique strong solution $\{\bar{x}^\epsilon(t);\ \ t \in [0, \infty)\}$.

Remark 3.2.5. Condition 3.2.3 of course raises the question of recognizing when there is a unique invariant probability measure π_x for each $x \in \mathbb{R}^d$. In Section 3.4 we will give simple sufficient conditions on $b(.,.)$ and $\sigma(.,.)$ in (3.1.3) which ensure that this is the case.

Remark 3.2.6. Let $\Theta : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ be a mapping such that $y \to \Theta(x, y) : \mathbb{R}^D \to \mathbb{R}$ is a C^2-mapping for each $x \in \mathbb{R}^d$. Put

$$\mathcal{A}\Theta(x, y) \triangleq \sum_{i=1}^{D} b^i(x, y)(\partial_{y^i}\Theta)(x, y) + 1/2 \sum_{i,j=1}^{D} c^{ij}(x, y)(\partial_{y^i}\partial_{y^j}\Theta)(x, y), \quad (3.2.14)$$

where

$$c(x, y) \triangleq \sigma(x, y)\sigma^T(x, y), \quad \forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D. \quad (3.2.15)$$

Remark 3.2.7. Let $C^{2,2}(\mathbb{R}^d \otimes \mathbb{R}^D)$ denote the collection of all continuous functions $\Theta : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ such that (i) the partial derivative functions $(\partial_{x^l}\Theta)(x, y)$ and $(\partial_{x^l}\partial_{x^k}\Theta)(x, y)$, $l, k = 1, 2, \dots, d$, exist and are continuous at all $(x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$, and (ii) the partial derivative functions $(\partial_{y^l}\Theta)(x, y)$ and $(\partial_{y^l}\partial_{y^k}\Theta)(x, y)$, $l, k = 1, 2, \dots, D$, exist and are continuous at all $(x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Notice that there is no requirement that the mixed partial derivatives $(\partial_{x^l}\partial_{y^k}\Theta)(x, y)$, $l = 1, 2, \dots, d$, $k = 1, 2, \dots, D$, need exist.

Conditions 3.2.8 and 3.2.15 which follow postulate solvability of certain Poisson partial differential equation in terms of the operator $\mathcal{A}$ given by (3.2.14). The solutions of these Poisson equations will be essential for the asymptotic analysis of the coupled system (3.1.1) and (3.1.2).

Condition 3.2.8. For each $i = 1, 2, \dots, d$, there exist functions $(x, y) \mapsto \Phi^i(x, y) : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ in $C^{2,2}(\mathbb{R}^d \otimes \mathbb{R}^D)$, such that

1. The mappings $\Phi^i(x, y)$ are polynomially y-bounded of order q_2 locally in x for some constant $q_2 \in [0, \infty)$, i.e. for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$|\Phi^i(x, y)| \le C(R)[1 + |y|^{q_2}], \tag{3.2.16}$$

for each $(x, y) \in S_R^d \otimes \mathbb{R}^D$, $i = 1, 2, \dots, d$. The partial derivative functions $(\partial_{y^l}\Phi^i)(x, y)$, $(\partial_{y^l}\partial_{y^k}\Phi^i)(x, y)$, are uniformly y-bounded locally in x, namely for each $R \in [0, \infty)$ there is some constant $C(R) \in [0, \infty)$ such that

$$|(\partial_{y^l}\Phi^i)(x, y)| + |(\partial_{y^l}\partial_{y^k}\Phi^i)(x, y)| \le C(R), \quad \forall\, (x, y) \in S_R^d \otimes \mathbb{R}^D, \tag{3.2.17}$$

for all $i = 1, 2, \dots, d$, and $l, k \in 1, 2, \dots, D$. The mappings $(\partial_{x^l}\Phi^i)(x, y)$ and $(\partial_{x^l}\partial_{x^k}\Phi^i)(x, y)$ are polynomially y-bounded of order q_2 locally in x (where q_2 is the constant in (3.2.16)), i.e. for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$|(\partial_{x^l}\Phi^i)(x, y)| + |(\partial_{x^l}\partial_{x^k}\Phi^i)(x, y)| \le C(R)(1 + |y|^{q_2}), \tag{3.2.18}$$

for each $(x, y) \in S_R^d \otimes \mathbb{R}^D$.

2. $\Phi^i(x, y)$ satisfies the Poisson equation

$$\mathcal{A}\Phi^i(x, y) = \overline{F}^i(x) - F^i(x, y), \quad \forall\, (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D, \quad \forall i = 1, 2, \dots, d \tag{3.2.19}$$

(where we use $F^i(x, y)$ and $\overline{F}^i(x)$ to denote the i-th scalar entries of the $\mathbb{R}^d$-vectors $F(x, y)$ and $\overline{F}(x)$).

3. For the constant q_2 in Condition 3.2.8(1), we have

$$\int_{\mathbb{R}^D} |\xi|^{2+2q_2} d\pi_x(\xi) < \infty, \quad \forall\, x \in \mathbb{R}^d. \tag{3.2.20}$$

4. The solutions $\Phi^i(x,y)$ of (3.2.19) subject to (3.2.17) are unique in the following sense: if a mapping $\hat{\Phi}^i : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ is such that $y \to \hat{\Phi}^i(x,y)$ is C^2 - mapping for each $x \in \mathbb{R}^d$ with

$$\mathcal{A}\hat{\Phi}^i(x,y) = \overline{F}^i(x) - F^i(x,y), \quad \forall\, (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D, \quad \forall i = 1,2,\ldots,d \tag{3.2.21}$$

and the partial derivative functions $(\partial_{y^l}\hat{\Phi}^i)(x,y)$, $(\partial_{y^l}\partial_{y^k}\hat{\Phi}^i)(x,y)$, are uniformly y-bounded locally in x, then $\Phi^i(x,y) - \hat{\Phi}^i(x,y)$ is a function of x only.

Remark 3.2.9. The preceding Condition 3.2.8 may look somewhat strange. It turns out that there are strong connections between the asymptotic properties of a Markov process $\{\xi(t)\}$ and solvability of a corresponding Poisson equation. Indeed, let $\{\xi(t)\}$ be a strictly stationary Markov process (assumed to be $\mathbb{R}^D$-valued for the sake of argument) with unique invariant probability π. Then $\{\xi(t)\}$ is an *ergodic* Markov process (see Corollary D.0.19). In general, ergodicity by itself is not enough to establish second - order asymptotic properties of $\{\xi(t)\}$ such as central limit theorems, and must be supplemented by additional conditions on $\{\xi(t)\}$. For example, one could postulate that $\{\xi(t)\}$ is not just ergodic but also strong mixing (in the sense summarized in Appendix D.0.3) and this is typically enough to establish a central limit theorem. An alternative strengthening of ergodicity to get second-order properties, which avoids the strong mixing hypothesis and is particularly well-suited to the Markov case, can be formulated as follows: Let $\mathcal{G}$ be the infinitesimal generator of the Markov process $\{\xi(t)\}$ and suppose that $F : \mathbb{R}^D \to \mathbb{R}^d$ is some bounded Borel - measurable mapping such that $\bar{F}^i - F^i(.)$ is in the range of $\mathcal{G}$ (for $\bar{F}^i \triangleq \int_{\mathbb{R}^D} F^i(\xi)\, d\pi(\xi)$) for each $i = 1,2,\ldots d$. Equivalently, there are functions $\Phi^i \in$ domain $[\mathcal{G}]$ which solve the "Poisson equations"

$$\mathcal{G}\Phi^i(y) = \bar{F}^i - F^i(y), \quad \forall y \in \mathbb{R}^D, \;\; i = 1,2,\ldots d. \tag{3.2.22}$$

Ergodicity of $\{\xi(t)\}$ ensures that the functions Φ^i in (3.2.22) are unique to within a constant, that is, for any $\hat{\Phi}^i \in$ domain$[\mathcal{G}]$ such that

$$\mathcal{G}\hat{\Phi}^i(y) = \bar{F}^i - F^i(y), \quad \forall y \in \mathbb{R}^D, \ i = 1, 2, \dots, d, \tag{3.2.23}$$

it necessarily follows that

$$\Phi^i(y) - \hat{\Phi}^i(y) \equiv \alpha^i, \quad \forall y \in \mathbb{R}^D, \tag{3.2.24}$$

for some constants $\alpha^i, i = 1, 2, \dots, d$. (This follows e.g. from Theorem 1.3.7 of Kunita [21]). Then (3.2.22), together with uniqueness modulo constants of the Φ^i, may be used to establish second - order properties of $\{\xi(t)\}$ such as central limit theorems. Indeed this very approach (which originates in the work of Doeblin [8] and Doob (Section V.7 of [9])) is used in Jacod and Shiryayev ([16], Theorem 3.65, page 445) to show that the random process

$$Z^\epsilon(\tau) \triangleq \sqrt{\epsilon} \int_0^{\tau/\epsilon} \left[F(\xi(s)) - \bar{F}\right] ds, \quad 0 \leq \tau \leq 1, \tag{3.2.25}$$

converges to a Gaussian distribution on $C[0, 1]$ (the space of $\mathbb{R}^d$-valued continuous functions on the unit interval $[0, 1]$) as $\epsilon \to 0$. Notice that our problem reduces to exactly this case if, in place of the equations (3.1.1) and (3.1.2), we consider the simpler relation

$$x^\epsilon(t) = x_0 + \epsilon \int_0^t F(\xi(s))ds. \tag{3.2.26}$$

That is, we put $G \equiv 0$ and remove all dependence on $x^\epsilon(t)$ in the right hand side of (3.1.1), and take the fast perturbing process in (3.1.1) to be a stationary $\mathbb{R}^D$-valued Markov process $\{\xi(t)\}$ with infinitesimal generator $\mathcal{G}$, which evolves independently of $\{x^\epsilon(t)\}$ (unlike the perturbing process $\{y^\epsilon(t)\}$, which is conditioned by $x^\epsilon(t)$). This is a considerably simpler situation than the one represented by the coupled pair (3.1.1), (3.1.2), but it turns out that this basic methodology, built on the Poisson equation (3.2.22), nevertheless can be made to work for the system (3.1.1) and (3.1.2). The technical challenges that must be dealt with in doing this are, briefly, as follows:

(1) The perturbing process $\{y^\epsilon(t)\}$ in (3.1.1) is a component of the $\mathbb{R}^{d+D}$-valued Markov diffusion $\{(x^\epsilon(t), y^\epsilon(t))\}$. As is well known (see the comments on page 243 of Rogers and Williams V.1 [30]) it is generally impossible to fully characterize the infinitesimal generator of Markov diffusions, so we cannot use (3.2.22) itself. However, if $\{\xi(t,x)\}$ is an $\mathbb{R}^D$-valued Markov diffusion defined by the stochastic differential relation (3.1.3), then it is known that its second-order differential operator (which is defined by (3.2.14)) agrees with the infinitesimal generator of $\{\xi(t,x)\}$ on a large part of its domain (see Theorem III.13.3 of Rogers and Williams V.1 [30]) and hence our analogue of (3.2.22) is the Poisson equation (3.2.19). In fact, the SDE (3.1.3) really gives a *family* of $\mathbb{R}^D$-valued Markov processes $\{\xi(t,x),\ t \in [0,\infty)\}$, parametrized by $x \in \mathbb{R}^d$, and (3.2.19) is the corresponding *family* of Poisson equations in the variable y, parametrized by the variable $x \in \mathbb{R}^d$. The properties of the solution $\Phi^i(x,.)$ of (3.2.19) formulated by Condition 3.2.8, will be necessary when we exploit the relation (3.2.19).

(2) In place of the static system (3.2.26) we are considering the dynamically coupled pair (3.1.1) and (3.1.2). In order to deal with these dynamics we shall integrate the method of Poisson equations with the methodology of the Stroock-Varadhan martingale problem, in a way that is suggested by the work of Papanicolaou, Stroock and Varadhan [29] on limit theorems in hydrodynamics. In particular, we are motivated by the study on averaging in coupled *ordinary differential equations* (corresponding to putting $G \equiv 0$ and $\sigma \equiv 0$ in (3.1.1) and (3.1.2) due to Papanicolaou [28]). Indeed, our whole approach is essentially to generalize the methodology and ideas of [28] to deal with the diffusion terms in (3.1.1) and (3.1.2).

(3) In (3.1.1) and (3.1.2), the fast - perturbing process $\{y^\epsilon(t)\}$ is not stationary and ergodic, unlike the perturbation process $\{\xi(t)\}$ in (3.2.25).

Remark 3.2.10. There remains the question of recognizing when the Poisson equation (3.2.19) is solvable, and has solutions that satisfy the requirements of Condition 3.2.8. An important secondary goal of this thesis is to develop simple sufficient conditions on the mappings $b(.,.)$ and $\sigma(.,.)$ in (3.1.3) which ensure that Condition 3.2.8 holds. This issue is taken up in Section 3.4.

Remark 3.2.11. To state the remaining conditions define

$$\tilde{F}(x,y) \triangleq F(x,y) - \overline{F}(x), \qquad \forall\, (x,y) \in I\!R^d \otimes I\!R^D, \tag{3.2.27}$$

and the symmetric $d \times d$-matrices $a(x,y)$ and $\overline{a}(x)$ by

$$a^{i,j}(x,y) \triangleq \tilde{F}^i(x,y)\Phi^j(x,y) + \tilde{F}^j(x,y)\Phi^i(x,y), \quad \forall \quad (x,y) \in I\!R^d \otimes I\!R^D, \tag{3.2.28}$$

$$\overline{a}^{i,j}(x) \triangleq \int_{I\!R^D} a^{i,j}(x,\xi)d\pi_x(\xi), \qquad \forall\, x \in I\!R^d \tag{3.2.29}$$

From (3.2.20), (3.2.16), (3.2.11) we see that the integral in (3.2.29) certainly exists for each $x \in I\!R^d$.

Remark 3.2.12. Since the solutions $\Phi^i(x,y)$ of (3.2.19) are unique modulo functions of x only (see Condition 3.2.8 (4)), and we see from (3.1.4) and (3.2.27) that

$$\int_{I\!R^D} \tilde{F}^i(x,\xi)d\pi_x(\xi) = 0, \tag{3.2.30}$$

it follows that the functions $\overline{a}^{i,j}(x)$ in (3.2.29) are uniquely defined. Indeed, suppose $\hat{\Phi}^i$ is also a solution of (3.2.21), $i = 1, 2, \ldots, d$, and put

$$\hat{a}^{i,j}(x,y) \triangleq \tilde{F}^i(x,y)\hat{\Phi}^j(x,y) + \tilde{F}^j(x,y)\hat{\Phi}^i(x,y). \tag{3.2.31}$$

Then, from (3.2.31) and (3.2.29),

$$\begin{aligned}\int_{\mathbb{R}^d} \left[a^{i,j}(x,\xi) - \hat{a}^{i,j}(x,\xi)\right] \, d\pi_x(\xi) &= \int_{\mathbb{R}^d} \tilde{F}^i(x,\xi) \left[\Phi^j(x,\xi) - \hat{\Phi}^j(x,\xi)\right] \, d\pi_x(\xi) \\ &\quad + \int_{\mathbb{R}^d} \tilde{F}^j(x,\xi) \left[\Phi^i(x,\xi) - \hat{\Phi}^i(x,\xi)\right] \, d\pi_x(\xi).\end{aligned} \tag{3.2.32}$$

By Condition 3.2.8 (4) we know that $\Phi^i(x,\xi) - \hat{\Phi}^i(x,\xi)$ is a function of x only, say $\alpha^i(x), i = 1, 2, \ldots, d$. Then, from (3.2.32), (3.2.30),

$$\begin{aligned}\int_{\mathbb{R}^d} \left[a^{i,j}(x,\xi) - \hat{a}^{i,j}(x,\xi)\right] \, d\pi_x(\xi) &= \alpha^j(x) \int_{\mathbb{R}^d} \tilde{F}^i(x,\xi) d\pi_x(\xi) + \alpha^i(x) \int_{\mathbb{R}^d} \tilde{F}^j(x,\xi) d\pi_x(\xi) \\ &= 0.\end{aligned} \tag{3.2.33}$$

This shows that $\bar{a}^{i,j}(x)$ is uniquely defined, regardless of which solutions Φ^i of (3.2.19) we use in (3.2.28).

In Section A.2 of Appendix A, we will prove the following result:

Proposition 3.2.13. *Suppose Conditions 3.2.1, 3.2.2, 3.2.3 and 3.2.8 hold. Then the function $\overline{a}(x)$, defined by (3.2.29), is nonnegative-definite for each $x \in \mathbb{R}^d$.*

Remark 3.2.14. It follows from Proposition 3.2.13 on page 37 that there is a unique symmetric nonnegative semidefinite matrix $\overline{a}^{1/2}(x)$ such that

$$\overline{a}(x) = (\overline{a}^{1/2}(x))^2. \tag{3.2.34}$$

We shall see later that $\overline{a}(x)$ determines the weak limit of Z^ϵ (in (3.1.8)) in much the same way that $\nu(x)$ (see (2.3.42)) determines the weak limit of Z^ϵ in Theorem 2.3.6 on page 20. To this end, we must impose some technical restrictions on $\bar{a}(x)$ as follows:

Condition 3.2.15. 1. The mappings $x \mapsto \overline{a}^{i,j}(x)$ defined by (3.2.29), $i, j = 1, 2, \ldots, d$, are locally Lipschitz continuous on $\mathbb{R}^d$.

2. For each $i, j = 1, 2, \ldots, d$, there exists a function $(x, y) \mapsto \Psi^{i,j}(x, y) : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ in $C^{2,2}(\mathbb{R}^d \otimes \mathbb{R}^D)$ such that $\Psi^{i,j}(x, y)$, $(\partial_{x^l}\Psi^{i,j})(x, y)$, $(\partial_{y^l}\Psi^{i,j})(x, y)$, $(\partial_{x^l}\partial_{x^k}\Psi^{i,j})(x, y)$, and $(\partial_{y^l}\partial_{y^k}\Psi^{i,j})(x, y)$ are polynomially y- bounded of order q_3 locally in x for some constant $q_3 \in [0, \infty)$, i.e. for each $R \in [0, \infty)$, there is a constant $C(R) \in [0, \infty)$, such that

$$\begin{aligned} &|\Psi^{i,j}(x, y)| + |(\partial_{x^l}\Psi^{i,j})(x, y)| + |(\partial_{y^l}\Psi^{i,j})(x, y)| \\ &\quad + |(\partial_{x^l}\partial_{x^k}\Psi^{i,j})(x, y)| + |(\partial_{y^l}\partial_{y^k}\Psi^{i,j})(x, y)| \le C(R)(|1 + |y|^{q_3}), \end{aligned} \tag{3.2.35}$$

for each $(x, y) \in S^d_R \otimes \mathbb{R}^D$.

3. The mappings $\Psi^{i,j}(x, y)$ in (2) satisfy the Poisson equation

$$\mathcal{A}\Psi^{i,j}(x, y) = \overline{a}^{i,j}(x) - a^{i,j}(x, y), \quad \forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D. \tag{3.2.36}$$

Remark 3.2.16. Observe that, in contrast to Condition 3.2.8, we do not insist on the uniqueness of solutions $\Psi^{i,j}$ of (3.2.36) modulo functions of x only. This is because we do not use the $\Psi^{i,j}$ to define any quantities, analogous to (3.2.27) and (3.2.28) where the uniqueness of Φ^i was essential (see Remark 3.2.12). In fact, the function $\Psi^{i,j}$ furnished by (3.2.36) will be used only to cancel terms that depend on y when we utilize the so-called *near identity method* of singular perturbations, resulting in a function of (x, z) only, from which we can deduce the desired limit theorems. That said, in Section 3.4 we will develop a solvability theory for Poisson equations which includes (3.2.19) and (3.2.36), and which provides uniqueness of solutions, modulo functions of x only, for both of these equations.

Remark 3.2.17. Our final condition (see Condition 3.2.18 to follow), which is suggested by Condition (90) on page 143 of Skorohod [33], is really a type of stability condition for the stochastic differential equation (3.1.3). This condition will play a very important role in our asymptotic analysis. It will allow us to keep the moments of $y^\epsilon(t)$ in (3.1.1), (3.1.2), under control as $\epsilon \to 0$, and provide us with the means to deal with the lack of stationarity and ergodicity of $\{y^\epsilon(t)\}$ which was noted in Remark 3.2.9.

Condition 3.2.18. There is a constant $q_4 \in (2 + 2q_2, \infty) \cap (2 + 2q_3, \infty)$ (where q_2 and q_3 are specified in Conditions 3.2.8 and 3.2.15 respectively), such that, for

$$\varphi(y) = |y|^{q_4}, \quad \forall y \in I\!R^D, \tag{3.2.37}$$

we have the following: corresponding to each $R \in [0, \infty)$ there are constants $\lambda_R \in (0, \infty)$ and $\alpha_R \in [0, \infty)$ such that

$$\mathcal{A}\varphi(x, y) + \lambda_R \varphi(y) \leq \alpha_R, \qquad \forall\, (x, y) \in S_R^d \otimes I\!R^D. \tag{3.2.38}$$

Remark 3.2.19. In this section we have formulated Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8, 3.2.15 and 3.2.18. These will be the basic conditions that we need to establish convergence of the scaled discrepancy process $\{Z^\epsilon(\tau)\}$ in (3.1.8). Conditions 3.2.1 and 3.2.2 are quite elementary, and are needed for (3.1.1), (3.1.2), to have a pathwise unique strong solution. Condition 3.2.3 enables us to formulate an "averaged" stochastic differential equation (3.1.5). On the other hand Conditions 3.2.8, 3.2.15, and 3.2.18 represent "stability - type" conditions on the stochastic differential equation (3.1.3) which will ensure that the fast process $\{y^\epsilon(t)\}$ is well enough behaved for $\{Z^\epsilon(\tau)\}$ in (3.1.8) to converge to a limit. In Section 3.4 we will present some simple sufficient conditions on the coefficients in (3.1.1) and (3.1.2) which are enough to ensure that Conditions 3.2.3, 3.2.8, 3.2.15 and 3.2.18 hold.

3.3 Main Result

We first declare the basic notation in terms of which the main result of this chapter is going to be formulated. Some of this notation has already been introduced previously but we repeat it here so that it is all defined in one place. For each $\epsilon \in (0,1]$ and $\tau \in [0,1]$ put

$$X^\epsilon(\tau) \triangleq x^\epsilon(\tau/\epsilon), \quad \overline{X}^\epsilon(\tau) \triangleq \overline{x}^\epsilon(\tau/\epsilon), \tag{3.3.39}$$

$$Y^\epsilon(\tau) \triangleq y^\epsilon(\tau/\epsilon), \quad W^\epsilon(\tau) \triangleq \epsilon^{1/2}w(\tau/\epsilon), \quad B^\epsilon(\tau) \triangleq \epsilon^{1/2}\beta(\tau/\epsilon). \tag{3.3.40}$$

Recalling Condition 3.2.2, for each $\epsilon \in (0,1]$ define the filtration $\{\mathcal{G}^\epsilon_\tau, \tau \in [0,1]\}$ by

$$\mathcal{G}^\epsilon_\tau \triangleq \mathcal{F}_{\tau/\epsilon}, \quad \forall\, \tau \in [0,1]. \tag{3.3.41}$$

In view of Condition 3.2.2 we see that $\{(W^\epsilon(\tau)), \mathcal{G}^\epsilon_\tau), \ \tau \in [0,1]\}$ and $\{(B^\epsilon(\tau), \mathcal{G}^\epsilon_\tau), \tau \in [0,1]\}$ are independent standard Wiener processes on $(\Omega, \mathcal{F}, P)$ for each $\epsilon \in (0,1]$, and it follows from (3.1.1) and (3.1.2) that $\{(X^\epsilon(\tau), Y^\epsilon(\tau)), \tau \in [0,1]\}$ solves the re-scaled equations

$$X^\epsilon(\tau) = x_0 + \int_0^\tau F(X^\epsilon(s), Y^\epsilon(s))ds + \int_0^\tau G(X^\epsilon(s))dW^\epsilon(s), \tag{3.3.42}$$

$$\begin{aligned} Y^\epsilon(\tau) &= y_0 + \epsilon^{-1}\int_0^\tau b(X^\epsilon(s), Y^\epsilon(s))ds \\ &\quad + \epsilon^{-1/2}\int_0^\tau \sigma(X^\epsilon(s), Y^\epsilon(s))dB^\epsilon(s). \end{aligned} \tag{3.3.43}$$

Also, we see from (3.1.5) and (3.3.39), $\{\overline{X}^\epsilon(\tau), \ \tau \in [0,1]\}$ solves the re-scaled equation

$$\overline{X}^\epsilon(\tau) = x_0 + \int_0^\tau \overline{F}(\overline{X}^\epsilon(s))ds + \int_0^\tau G(\overline{X}^\epsilon(s))dW^\epsilon(s). \tag{3.3.44}$$

For each $\epsilon \in (0,1]$ define the scaled discrepancy process $\{Z^\epsilon(\tau),\ \tau \in [0,1]\}$ by

$$Z^\epsilon(\tau) \triangleq \epsilon^{-1/2}[X^\epsilon(\tau) - \overline{X}^\epsilon(\tau)], \qquad \forall\, \tau \in [0,1]. \tag{3.3.45}$$

We are going to show that, as $\epsilon \to 0$, the process $\{(X^\epsilon(\tau), Z^\epsilon(\tau)),\ \tau \in [0,1]\}$ converges weakly to a limiting process $\{(\hat{X}(\tau), \hat{Z}(\tau)),\ \tau \in [0,1]\}$ which solves the system of equations

$$\hat{X}(\tau) = x_0 + \int_0^\tau \overline{F}(\hat{X}(s))ds + \int_0^\tau G(\hat{X}(s))d\hat{W}_1(s), \tag{3.3.46}$$

$$\hat{Z}(\tau) = \int_0^\tau (\partial_x \overline{F})(\hat{X}(s))\hat{Z}(s)ds + \int_0^\tau (\partial_x G)(\hat{X}(s))[\hat{Z}(s)]d\hat{W}_1(s) + \int_0^\tau \overline{a}^{1/2}(\hat{X}(s))d\hat{W}_2(s), \tag{3.3.47}$$

where $\{\hat{W}_1(\tau),\ \tau \in [0,1]\}$ and $\{\hat{W}_2(\tau),\ \tau \in [0,1]\}$ are independent standard Wiener processes, $\mathbb{R}^M$ and $\mathbb{R}^d$-valued processes respectively, on some probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ (see "Basic Notation and Terminology IV" for the definition of the d by M matrix $(\partial G)(x)[z], x, z \in \mathbb{R}^d$).

Remark 3.3.1. Since $\{(W^\epsilon(\tau), \mathcal{G}^\epsilon_\tau),\ \tau \in [0,1]\}$ is a standard Wiener process for each $\epsilon \in (0,1]$ it follows from (3.3.44) that the solution $\{\overline{X}^\epsilon(\tau),\ \tau \in [0,1]\}$ has a distribution which is invariant with respect to $\epsilon \in (0,1]$. In fact, comparing (3.3.44) with (3.3.46) one sees that

$$\left\{\overline{X}^\epsilon(\tau),\ \tau \in [0,1]\right\} \stackrel{\mathcal{D}}{=} \left\{\hat{X}(\tau),\ \tau \in [0,1]\right\}, \tag{3.3.48}$$

$\forall \epsilon \in (0,1]$, where $\stackrel{\mathcal{D}}{=}$ denotes equality in probability law.

Remark 3.3.2. From Condition 3.2.1, Condition 3.2.3, and the bilinearity of (3.3.47), the system of equations (3.3.46), (3.3.47), has a pathwise unique strong solution, and hence a result of Yamada and Watanabe (see Proposition 5.3.20 of Karatzas and Shreve [19]) ensures that the law of the process $\{(\hat{X}(\tau), \hat{Z}(\tau)),\ \tau \in$

$[0, 1]\}$ is uniquely defined on the measure space $(\Omega^*, \mathcal{F}^*)$, where

$$\Omega^* \triangleq C[0,1] \otimes C[0,1], \tag{3.3.49}$$

$$\mathcal{F}^* \triangleq \mathcal{B}(C[0,1] \otimes C[0,1]), \tag{3.3.50}$$

regardless of which the square root function $\overline{a}(x)^{1/2}$ is used in the equation(3.3.47).

The main result of this chapter is

Theorem 3.3.3. *Suppose that Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8, 3.2.15 and 3.2.18 hold. Then we have*

$$\lim_{\epsilon\to 0} P\{\sup_{\tau\in[0,1]} |X^\epsilon(\tau) - \overline{X}^\epsilon(\tau)| \geq \delta\} = 0, \qquad \forall \delta \in (0, \infty), \tag{3.3.51}$$

and

$$\lim_{\epsilon\to 0} \mathcal{L}(X^\epsilon, Z^\epsilon) = \mathcal{L}(\hat{X}, \hat{Z}), \tag{3.3.52}$$

where the convergence is weak convergence of probability measures over $(\Omega^*, \mathcal{F}^*)$ defined by (3.3.49) and (3.3.50).

Remark 3.3.4. Theorem 3.3.3 on page 42 is proved in Appendix A.

Remark 3.3.5. The weak convergence in (3.3.52) is of course the result of primary interest in Theorem 3.3.3 on page 42. Notice however, from (3.3.51), that Theorem 3.3.3 on page 42 also establishes a convergence in probability like that of Theorem 2.4.1 on page 25, which also deals with essentially the coupled system (3.1.1) and (3.1.2). However, the hypotheses that we postulate in Theorem 3.3.3 on page 42 for this result are somewhat different from the hypotheses postulated for Theorem 2.4.1 on page 25. We believe that the conditions in Theorem 3.3.3 on page 42 are somewhat more natural than those in Theorem 2.4.1 on page 25. In particular, the uniformity of convergence in (2.4.67) that is required for Theorem 2.4.1 on page 25 looks rather stringent and difficult to verify. On the other hand, as we shall see in

Section 3.4, we can propose simple sufficient conditions on the coefficients in (3.1.1) and (3.1.2) which imply satisfaction of the conditions for Theorem 3.3.3 on page 42. Of course, the main motivation for our hypotheses in Theorem 3.3.3 on page 42 is not just to improve on the conditions in Theorem 2.4.1 on page 25, but rather to ensure that the weak convergence in (3.3.52) holds.

3.4 Sufficient Hypotheses for Conditions 3.2.3, 3.2.8, 3.2.15 and 3.2.18

In this section, we are going to formulate some simple sufficient conditions on the coefficients of the coupled Itô equations (3.1.1) and (3.1.2) which imply Conditions 3.2.3, 3.2.8, 3.2.15, and 3.2.18 of Section 3.2. The essential aspect of Conditions 3.2.3, 3.2.8 and 3.2.15 is existence and uniqueness of an invariant probability measure π_x for the Markov process $\{\xi(t,x)\}$ defined by the stochastic differential equation (3.1.3) together with solvability of the equations (3.2.19) and (3.2.36), which are "Poisson equations" in the variable y, parametrized by $x \in \mathbb{R}^d$ (see Remark 3.2.9). Our approach to establishing these sufficient conditions is motivated by the following result of Bhattacharya and Waymire [3], which essentially establishes conditions on the coefficients of an Itô stochastic differential equation of the form

$$d\xi(t) = b(\xi(t))dt + \sigma(\xi(t))d\beta(t), \tag{3.4.53}$$

to ensure that the corresponding Markov process has a unique invariant probability measure:

Theorem 3.4.1. *(see Theorem 4.2 on page 593 of Bhattacharya and Waymire [3]) Suppose that the following hold:*

(i) *The functions $b : \mathbb{R}^D \to \mathbb{R}^D$ and $\sigma : \mathbb{R}^D \to \mathbb{R}^{D\otimes N}$ are globally Lipschitz* *continuous with*

$$|\sigma(\xi_1) - \sigma(\xi_2)| \leq \Lambda_0 |\xi_1 - \xi_2| \qquad \forall \xi_1, \xi_2 \in \mathbb{R}^D, \tag{3.4.54}$$

for some constant $\Lambda_0 \in [0, \infty)$.

(ii) *The function* $b(\cdot)$ *is a* C^1*- mapping on* $\mathbb{R}^D$ *with* $D \times D$ *Jacobian matrix*

$$J(\xi) \triangleq (\partial_\xi b)(\xi), \tag{3.4.55}$$

(iii) *We have*

$$\Lambda_1 < -(1/2)\Lambda_0^2, \tag{3.4.56}$$

where

$$\Lambda_1 \triangleq \sup_{\xi \in \mathbb{R}^D} \lambda_{\max}(\xi) \tag{3.4.57}$$

and $\lambda_{\max}$ *denotes the largest eigenvalue of the* $D \times D$ *symmetric matrix*

$$(1/2)[J(\xi) + J^T(\xi)].$$

Then there exists a unique invariant probability measure π *on* $\mathbb{R}^D$ *for the Markov process* $\{\xi(t)\}$ *defined by*

$$d\xi(t) = b(\xi(t))dt + \sigma(\xi(t))d\beta(t), \tag{3.4.58}$$

where $\{\beta(t)\}$ *is a standard* $\mathbb{R}^N$*-valued Wiener process. Moreover, for each bounded continuous function* $\varphi : \mathbb{R}^D \to \mathbb{R}$, *we have*

$$\lim_{t \to \infty} E\varphi(\xi(t, y)) = \int_{\mathbb{R}^D} \varphi \, d\pi, \qquad \forall y \in \mathbb{R}^D, \tag{3.4.59}$$

where $\{\xi(t, y)\}$ *is the solution of the stochastic differential equation (3.4.58) subject to* $\xi(0, y) \triangleq y$.

Remark 3.4.2. 1. Our statement of Theorem 3.4.1 on page 43 differs slightly from that in [3] because we are using the Frobenius norm for matrices (see "Basic Notation and Terminology") to define the Lipschitz constant Λ_0 for $\sigma(.)$, whereas the operator norm is used in [3].

2. Notice that Theorem 3.4.1 on page 43 does not insist on nonsingularity of the matrix $\sigma\sigma^T$, and therefore includes the important case of "degenerate" diffusions.

3. The basic intuition in Theorem 3.4.1 on page 43 is that sharp local variations in $\sigma(\cdot)$ (reflected in a large value for the Lipschitz constant Λ_0) must be compensated by having all eigenvalues of $(1/2)[J(\xi) + J^T(\xi)]$ sufficiently negative, uniformly with respect to ξ (see (3.4.56)).

Remark 3.4.3. Theorem 3.4.1 on page 43 ensures that the transition probability function

$$P_t(y,\Gamma) \triangleq E\left[I_\Gamma(\xi(t,y))\right], \quad t \in [0,\infty), \quad y \in \mathbb{R}^D, \quad \Gamma \in \mathcal{B}(\mathbb{R}^D), \tag{3.4.60}$$

for the Markov process given by (3.4.58) has a *unique* invariant probability measure π. Thus if $\{\xi(t)\}$ is a solution of (3.4.58) with $\xi(0)$ independent of $\{\beta(t),\ t \in [0,\infty)\}$ and distribution given by π, then $\{\xi(t)\}$ is a strictly stationary Markov process (as follows from the standard theory of Itô stochastic differential equations - see for example Theorem 10.11 of Chung and Williams [7]). Moreover it follows from Corollary D.0.19 that $\{\xi(t)\}$ is ergodic.

Theorem 3.4.1 on page 43 does not address the solvability of Poisson equations associated with the second-order linear differential operator for the Markov diffusion $\{\xi(t)\}$. We shall therefore introduce a strengthening of its basic hypothesis to establish solvability of these Poisson equations. We will then "transfer" these conditions to the parametrized stochastic differential equation (3.1.3) and in this way get sufficient conditions on the coefficients in (3.1.3) which ensure solvability of the Poisson equations (3.2.19) and (3.2.36). Here our approach is influenced by a result due to Benveniste, Metivier and Priouret ([1], page 255, Proposition 3) which we reproduce as Proposition E.0.27 on page 216 in Appendix E. Among other things, this result ensures solvability of the Poisson - type operator equation

$$u - \Pi u = g - \int g \, dm \tag{3.4.61}$$

for a discrete - parameter Markov chain in $\mathbb{R}^D$ with one - step transition probability function $\Pi(x, A), x \in \mathbb{R}^D, A \in \mathcal{B}(\mathbb{R}^D)$, having a unique invariant probability measure m.

In brief we extend Theorem 3.4.1 on page 43, due to Bhattacharya and Waymire [3], to cover the case of the stochastic differential equation (3.1.3) which is parametrized by $x \in \mathbb{R}^d$. Also, motivated by Proposition E.0.27 on page 216 we also extend Theorem 3.4.1 on page 43 to ensure not just a unique invariant probability measure, but also solvability of the Poisson equations (3.2.19) and (3.2.36). In contrast to the Poisson-type operator $(1 - \Pi)$ which features in (3.4.61) and which is the generator of a discrete-parameter Markov chain with transition probability Π, we are dealing here with Poisson equations that arise from the parametrized second order differential operator $\mathcal{A}$ in (3.2.14). The technical problems of verifying solvability in this case are a good deal more delicate than in the case of the operator equation (3.4.61), mainly because the second-order linear differential operator is unbounded and non-closed. In the course of handling these problems, we shall frequently use the theory of L_2-derivatives for the solution of stochastic differential equations due to Gihman and Skorohod [14] which is summarized for easy reference in Appendix F.

We now begin to formulate our sufficient conditions on the coefficients in (3.1.3):

Condition 3.4.4. The mappings $(x, \xi) \mapsto b^i(x, \xi)$ and $(x, \xi) \mapsto \sigma^{i,j}(x, \xi)$ are C^3-

functions, $\forall i = 1, 2, \ldots, D$, $j = 1, 2, \ldots, N$, and the first, second and third (x, ξ)-derivative functions of $b^i(x, \xi)$ and $\sigma^{ij}(x, \xi)$ are uniformly ξ-bounded locally in x. That is, for every multi-index $(\alpha_1, \ldots, \alpha_d, \eta_1, \ldots, \eta_D)$, in which the α_p and η_q are non-negative integers such that

$$\sum_1^d \alpha_p + \sum_1^D \eta_q \leq 3,$$

we have the following: for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$|(\partial_{x^1}^{\alpha_1} \partial_{x^2}^{\alpha_2} \ldots \partial_{x^d}^{\alpha_d} \partial_{\xi^1}^{\eta_1} \partial_{\xi^2}^{\eta_2} \ldots \partial_{\xi^D}^{\eta_D} b^i)(x, \xi)| \leq C(R), \tag{3.4.62}$$

for all $(x, \xi) \in S_R^d \times \mathbb{R}^D$, with identical bounds holding for $\sigma^{ij}(x, \xi)$ in place of $b^i(x, \xi)$.

The following simple Lemma establishes a global Lipschitz constant for the mapping $\xi \to \sigma(x, \xi) : \mathbb{R}^D \to \mathbb{R}^{D \otimes N}$.

Lemma 3.4.5. *Suppose that Condition 3.4.4 holds. Then, for each $R \in [0, \infty)$, we have*

$$|\sigma(x, \xi_1) - \sigma(x, \xi_2)| \leq \Lambda_0(R)|\xi_1 - \xi_2|, \qquad \forall x \in S_R^d, \quad \forall \xi_1, \xi_2 \in \mathbb{R}^D, \tag{3.4.63}$$

where $\Lambda_0(R)$ is given by

$$\Lambda_0(R) \triangleq \left[\sup_{(x,\xi) \in S_R^d \otimes \mathbb{R}^D} \left\{ \sum_{l=1}^D |(\partial_{\xi^l} \sigma)(x, \xi)|^2 \right\} \right]^{1/2}, \tag{3.4.64}$$

and $\Lambda_0(R) < \infty, \forall R \in [0, \infty)$.

Remark 3.4.6. For each $(x, \xi) \in \mathbb{R}^d \otimes \mathbb{R}^D$, let $\lambda_{max}(x, \xi)$ denote the *largest* eigenvalue of the $D \times D$ symmetric matrix

$$(1/2)[J(x, \xi) + J(x, \xi)^T],$$

where

$$J(x,\xi)^{ij} \triangleq (\partial_{\xi^j} b^i)(x,\xi), \quad i,j = 1,2,\dots,D, \tag{3.4.65}$$

and, for each $R \in [0,\infty)$, put

$$\Lambda_1(R) \triangleq \sup_{(x,\xi)\in S_R^d \otimes \mathbb{R}^D} [\lambda_{max}(x,\xi)]. \tag{3.4.66}$$

The next condition is motivated by (3.4.56) of Theorem 3.4.1 on page 43, but applies to the stochastic differential equation (3.1.3) which is parametrized by $x \in \mathbb{R}^d$:

Condition 3.4.7. There is a constant $q \in (8,\infty)$ such that

$$\Lambda_1(R) < \frac{(1-q)}{2}[\Lambda_0(R)]^2, \qquad \forall R \in [0,\infty), \tag{3.4.67}$$

(for $\Lambda_0(R)$ given by (3.4.64)).

Remark 3.4.8. Suppose that there is no dependence on x in (3.1.3), i.e. (3.1.3) is just (3.4.58). Then condition (3.4.67) is still non-trivially stronger than (3.4.56), since we are supposing $q > 8$ in (3.4.67), and (3.4.56) follows from (3.4.67) with $q = 2$. The reason that we insist on taking $q > 8$ in Condition 3.4.7, is because we want enough "stability" in (3.1.3) to get not only a unique invariant probability measure π_x but also solvability of the Poisson equations (3.2.19) and (3.2.36) which are associated with the second order linear differential operator (3.2.14) for the diffusion given by (3.1.3). As will be seen in due course (see Remark 3.4.23) a value of $q > 8$ turns out to be enough to ensure this.

Remark 3.4.9. Motivated by Remark E.0.26, for later use we define the following spaces of locally Lipschitz continuous functions from $\mathbb{R}^D \to \mathbb{R}$: For a Borel measurable function $h : \mathbb{R}^D \to \mathbb{R}$ and some $r \in [0,\infty)$, put

$$[h]_r \triangleq \sup_{y_1\neq y_2} \frac{|h(y_1)-h(y_2)|}{|y_1-y_2|[1+|y_1|^r+|y_2|^r]}, \qquad \| h \|_{r+1} \triangleq \sup_y \frac{|h(y)|}{1+|y|^{r+1}},$$

and

$$M_r(h) \triangleq \max\{[h]_r, \| h \|_{r+1}\}.$$

For each $r \in [0, \infty)$, define the function spaces

$$Li(r) \triangleq \{h : \mathbb{R}^D \to \mathbb{R} \mid [h]_r < \infty\}. \tag{3.4.68}$$

Clearly, $h \in Li(0)$ if and only if $h(.)$ is globally Lipschitz continuous on $\mathbb{R}^D$.

Definition 3.4.10. *Suppose that $r \in [0, \infty)$. Then $Li(r)_{loc}$ is the set of Borel-measurable functions $g : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ with the following property: for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that*

$$M_r(g(x, \cdot)) \leq C(R), \quad \forall x \in S_R^d.$$

Remark 3.4.11. Thus $g \in Li(r)_{loc}$ implies that $g(x, .) \in Li(r)$ with

$$\sup_{x \in S_R^d} M_r(g(x, .)) < \infty, \quad \forall R \in [0, \infty). \tag{3.4.69}$$

Clearly, $g \in Li(r)_{loc}$ for some $r \in [0, \infty)$ implies that g is polynomially y-bounded of order $r + 1$ locally in x, namely, for each $R \in [0, \infty)$, there exists a constant $C(R)$ such that

$$|g(x, y)| \leq C(R)(1 + |y|^{r+1}), \quad \forall x \in S_R^d, \ \forall y \in \mathbb{R}^D, \tag{3.4.70}$$

and

$$|g(x, y_1) - g(x, y_2)| \leq C(R)\,|y_1 - y_2|\,[1 + |y_1|^r + |y_2|^r], \quad \forall x \in S_R^d, \ \forall y_1, y_2 \in \mathbb{R}^D. \tag{3.4.71}$$

The next Proposition will be needed for the main result of this section, namely Proposition 3.4.16 on page 51 on solvability of Poisson equations, and will also provide us with the means for verifying Condition 3.2.18 later in this section (see Remark 3.4.23):

Proposition 3.4.12. *Suppose that Conditions 3.4.4 and 3.4.7 hold, and put*

$$\varphi(y) \triangleq |y|^q, \quad \forall y \in \mathbb{R}^D \tag{3.4.72}$$

where q is the constant given by Condition 3.4.7. Then, for each $R \in [0, \infty)$, there are constants $\alpha_R \in [0, \infty)$, $\lambda_R \in (0, \infty)$, such that

$$\mathcal{A}\varphi(x, y) + \lambda_R \varphi(y) \leq \alpha_R, \qquad \forall (x, y) \in S_R^d \otimes \mathbb{R}^D. \tag{3.4.73}$$

Remark 3.4.13. from now on write $\{\xi(t, x, y),\ t \in [0, \infty)\}$ for the (pathwise unique) solution of (3.1.3) with initial value y at $t = 0$, namely

$$\xi(t, x, y) = y + \int_0^t b(x, \xi(s, x, y))ds + \int_0^t \sigma(x, \xi(s, x, y))d\beta(s), \tag{3.4.74}$$

Proposition 3.4.14. *Suppose that $b(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ in (3.1.3) are subject to Conditions 3.2.1, 3.4.4 and 3.4.7. Then, for each $x \in \mathbb{R}^d$, there exists a unique invariant probability measure π_x on $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ for the Markov process $\{\xi(t, x)\}$ defined by (3.1.3), and*

$$\int_{\mathbb{R}^D} |\xi|^q d\pi_x(\xi) < \infty, \qquad \forall x \in \mathbb{R}^d, \tag{3.4.75}$$

(where q is given by Condition 3.4.7). Furthermore, if $f \in Li(r_1)_{loc}$ for some $r_1 \in [0, q-1]$, then, for each $R \in [0, \infty)$, there are some constants $C(R) \in [0, \infty)$, $\gamma(R) \in (0, \infty)$, such that

$$|Ef(x, \xi(t, x, y)) - \int_{\mathbb{R}^D} f(x, \xi)\, d\pi_x(\xi)| \leq C(R)e^{-\gamma(R)t}[1 + |y|^{1+r_1}], \tag{3.4.76}$$

for each $(t, x, y) \in [0, \infty) \otimes S_R^d \otimes \mathbb{R}^D$.

Remark 3.4.15. If Conditions 3.2.1, 3.4.4 and 3.4.7 hold, and $f \in Li(r)_{loc}$ for some $r \in [0, q-1]$, where q is given by Condition 3.4.7, then Proposition 3.4.14 on page 50 ensures that the functions

$$\overline{f}(x) \triangleq \int_{\mathbb{R}^D} f(x, \xi)d\pi_x(\xi), \qquad \forall x \in \mathbb{R}^d, \tag{3.4.77}$$

$$\Theta(x, y) \triangleq \int_0^\infty E[f(x, \xi(t, x, y)) - \overline{f}(x)]dt, \qquad \forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D, \tag{3.4.78}$$

are well defined. This will be important for the next result, which is the main result of this section, and which essentially deals with solvability of the Poisson equations that occur in Condition 3.2.8 and Condition 3.2.15:

Proposition 3.4.16. *Suppose that*

(i) $b(\cdot,\cdot)$ *and* $\sigma(\cdot,\cdot)$ *in (3.1.3) are subject to Conditions 3.2.1, 3.4.4 and 3.4.7;*

(ii) $f \in Li(r)_{loc}$ *for some* $r \in [0, q/2)$*, where* q *is given by Condition 3.4.7, and suppose in addition that* $y \to f(x,y)$ *is a* C^2*-mapping for each* $x \in \mathbb{R}^d$ *such that the partial derivative functions* $(\partial_{y^l} f)(x,y)$*, and* $(\partial_{y^l}\partial_{y^k} f)(x,y)$ *are continuous in* (x,y)*, and are polynomially* y*-bounded of order* r *locally in* x*; namely, for each* $R \in [0,\infty)$ *there is a constant* $C_1(R) \in [0,\infty)$*, such that*

$$\left|(\partial_{y^l} f)(x,y)\right| + \left|(\partial_{y^l}\partial_{y^k} f)(x,y)\right| \le C_1(R)\left[1+|y|^r\right], \quad \forall x \in S_R^d, \ \forall y \in \mathbb{R}^D. \tag{3.4.79}$$

Then for the functions $\overline{f}(x)$ *and* $\Theta(x,y)$ *defined by (3.4.77) and (3.4.78), we have*

1. $\Theta \in Li(r)_{loc}$;

2. $y \to \Theta(x,y)$ *is a* C^2*-mapping for each* $x \in \mathbb{R}^d$*, and the partial derivative functions* $(\partial_{y^l}\Theta)(x,y)$ *and* $(\partial_{y^l}\partial_{y^k}\Theta)(x,y)$ *are continuous in* (x,y)*, and are polynomially* y*-bounded of order* r *locally in* x*, i.e., for each* $R \in [0,\infty)$*, there is a constant* $C(R) \in [0,\infty)$ *such that*

$$|(\partial_{y^l}\Theta)(x,y)| + |(\partial_{y^l}\partial_{y^k}\Theta)(x,y)| \le C(R)[1+|y|^r], \quad \forall (x,y) \in S_R^d \otimes \mathbb{R}^D; \tag{3.4.80}$$

3. *The following relation holds:*

$$\mathcal{A}\Theta(x,y) = \overline{f}(x) - f(x,y), \qquad \forall (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D, \tag{3.4.81}$$

where $\mathcal{A}$ *is defined by (3.2.14).*

Proposition 3.4.16 on page 51 basically postulates conditions on the coefficients $b(.,.)$ and $\sigma(.,.)$ in (3.1.3) and on the function $f(\cdot,\cdot)$ which are sufficient to ensure solvability of the Poisson equation (3.4.81). It remains to settle the sense in which solutions of this equation are unique. This is dealt with by the next result, which essentially says that if $\Theta_i(x,y), i = 1,2$, satisfy (3.4.81) then $\Theta_1(x,y) - \Theta_2(x,y)$ is a function of x only:

Proposition 3.4.17. *Suppose that*

(i) *of Proposition 3.4.16 on page 51 holds;*

(ii) *the mappings $f, \tilde{\Theta} : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ belong to $Li(r)_{loc}$ for some $r \in [0, q/2)$, the mappings $y \mapsto f(x,y) : \mathbb{R}^D \to \mathbb{R}$ and $y \mapsto \tilde{\Theta}(x,y) : \mathbb{R}^D \to \mathbb{R}$ are C^2-functions for each $x \in \mathbb{R}^d$, and the partial derivatives $(\partial_{y^l}\tilde{\Theta})(x,y)$ and $(\partial_{y^l}\partial_{y^k}\tilde{\Theta})(x,y)$ are polynomially y-bounded of order r locally in x;*

(iii) *the following relation holds:*

$$\mathcal{A}\tilde{\Theta}(x,y) = \overline{f}(x) - f(x,y), \qquad \forall (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D, \tag{3.4.82}$$

for $\overline{f}(x)$ defined by (3.4.77).

Then the solution $\tilde{\Theta}(x,y)$ of (3.4.82) satisfies the relation

$$\tilde{\Theta}(x,y) = \bar{\Theta}(x) + \Theta(x,y), \qquad \forall (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D, \tag{3.4.83}$$

where $\Theta(x,y)$ is given by (3.4.78), and

$$\bar{\Theta}(x) \triangleq \int_{\mathbb{R}^D} \tilde{\Theta}(x,\xi)\, d\pi_x(\xi). \tag{3.4.84}$$

In particular, the mapping

$$(x,y) \to \tilde{\Theta}(x,y) - \Theta(x,y) = \bar{\Theta}(x)$$

is a function of x only.

The following proposition deals with the dependence on x of the mappings $\bar{f}(x)$ in (3.4.77) and $\Theta(x, y)$ in (3.4.78):

Proposition 3.4.18. *Suppose that*

(i) *of Proposition 3.4.16 on page 51 holds.*

(ii) *of Proposition 3.4.16 on page 51 holds. In addition, suppose that $x \to f(x, y)$ is a C^2 - mapping for each $y \in \mathbb{R}^D$, such that the partial derivative functions $(\partial_{x^l} f)(x, y)$ and $(\partial_{x^l}\partial_{x^k} f)(x, y)$ are continuous in (x, y), and are polynomially y - bounded of order r locally in x; namely for each $R \in [0, \infty)$ there is a constant $C_2(R) \in [0, \infty)$, such that*

$$|(\partial_{x^l} f)(x, y)| + |(\partial_{x^l}\partial_{x^k} f)(x, y)| \leq C_2(R)\,[1 + |y|^r], \quad \forall x \in S^d_R,\ \forall y \in \mathbb{R}^D. \tag{3.4.85}$$

Then :

(a) *The mapping $\bar{f}(x)$ defined by (3.4.77) is a C^2 - function on $\mathbb{R}^d$.*

(b) *For the mapping $\Theta(x, y)$ defined by (3.4.78), the partial derivative functions $(\partial_{x^l}\Theta)(x, y)$ and $(\partial_{x^l}\partial_{x^k}\Theta)(x, y)$ exist, are continuous in (x, y), and are polynomially y-bounded of order $(r + 1)$ locally in x.*

We are now going to use Propositions 3.4.14 on page 50, 3.4.16, 3.4.17, and 3.4.18 to check Conditions 3.2.3, 3.2.8, 3.2.15 and 3.2.18. To this end, we will suppose that $b(.,.)$ and $\sigma(.,.)$ in (3.1.2) satisfy the Conditions 3.2.1, 3.4.4 and 3.4.7, which are the basic hypotheses on these mappings that are postulated by each of these Propositions. We will also suppose that $F(.,.)$ in (3.1.1) is subject to the following condition (which supplements the requirements imposed on $F(.,.)$ by Condition 3.2.1).

Condition 3.4.19. The mappings $F^i : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$ are members of $C^{2,2}(\mathbb{R}^d \otimes \mathbb{R}^D)$ for each $i = 1, 2, \ldots, d$, (recall Remark 3.2.7), and the partial derivative functions $(\partial_{x^l} F^i)(x,y), (\partial_{x^l}\partial_{x^k} F^i)(x,y), (\partial_{y^l} F^i)(x,y)$, and $(\partial_{y^l}\partial_{y^k} F^i)(x,y)$ are uniformly y - bounded locally in x. That is, for each $R \in [0,\infty)$ there exists a constant $C(R) \in [0,\infty)$ such that

$$|(\partial_{x^l} F^i)(x,y)| + |(\partial_{x^l}\partial_{x^k} F^i)(x,y)| + |(\partial_{y^l} F^i)(x,y)| + |(\partial_{y^l}\partial_{y^k} F^i)(x,y)| \leq C(R), \tag{3.4.86}$$

for all $(x,y) \in S_R^d \otimes \mathbb{R}^D$.

Remark 3.4.20. Notice that Condition 3.4.19 certainly ensures that $F^i \in Li(0)_{loc}$, $i = 1, 2, \ldots, d$. In fact, from Condition 3.4.19 and the mean-value theorem, we see that for each $R \in [0,\infty)$ there is a constant $C(R) \in [0,\infty)$ such that

$$|F^i(x,\xi)| \leq C(R)[1 + |\xi|], \qquad \forall (x,\xi) \in S_R^d \otimes \mathbb{R}^D. \tag{3.4.87}$$

Again by Condition 3.4.19 and the mean-value theorem, we see that for each $R \in [0,\infty)$ there exists a constant $C(R) \in [0,\infty)$ such that

$$|F^i(x,\xi_1) - F^i(x,\xi_2)| \leq C(R)|\xi_1 - \xi_2|, \qquad \forall x \in S_R^d, \quad \forall \xi_1, \ \xi_2 \in \mathbb{R}^D. \tag{3.4.88}$$

Now it follows from (3.4.87) and (3.4.88) that $F^i \in Li(0)_{loc}$ (recall Definition 3.4.10).

Remark 3.4.21. From now on we suppose that Conditions 3.2.1, 3.4.4, 3.4.7 and 3.4.19 hold. We will show that these conditions are sufficient to ensure that the Conditions 3.2.3, 3.2.8, 3.2.15 and 3.2.18 hold (recall Remark 3.2.19):

Partial Check of Condition 3.2.3: One sees from Proposition 3.4.14 on page 50 that, for each $x \in \mathbb{R}^d$, there is a unique invariant probability measure π_x on $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ for the Markov process $\{\xi(t,x)\}$ defined by (3.1.3). By Condition 3.4.7 we have $q \in (8,\infty)$, and thus

$$\int_{\mathbb{R}^D} |\xi|^6 \, \pi_x(d\xi) < \infty, \tag{3.4.89}$$

for each $x \in I\!R^d$ (see (3.4.75). It follows from (3.4.87) and (3.4.89) that the integral in (3.1.4) certainly exits for each $x \in I\!R^d$, and Proposition 3.4.18 on page 53 ensures that $\overline{F}(\cdot)$ is a C^2-mapping on $I\!R^d$. We have thus checked Condition 3.2.3. except for the required global Lipschitz continuity of $\overline{F}(\cdot)$ (see Remark 3.4.23), and turn next to Condition 3.2.8.

Remark 3.4.22. The main tool for verifying Condition 3.2.8(1) and (2) will be Proposition 3.4.16 on page 51. Observe that (i) of Proposition 3.4.16 on page 51 holds by virtue of our hypotheses (recall Remark 3.4.21). As for (ii) of Proposition 3.4.16 on page 51, we know from Remark 3.4.20 that $F^i \in Li(0)_{loc}$, $i = 1, 2, \ldots, d$. Moreover, from Condition 3.4.19, we see that the functions $(\partial_{\xi^l} F^i)(x, \xi)$ and $(\partial_{\xi^l}\partial_{\xi^k} F^i)(x, \xi)$ are polynomially ξ-bounded of order $r \triangleq 0$ locally in x. Thus (ii) of Proposition 3.4.16 on page 51 holds when we take $r \triangleq 0$ and $f \triangleq F^i$. Now (compare (3.1.4), (3.4.77) and (3.4.78)) we define

$$\Phi^i(x, y) \triangleq \int_0^\infty E[F^i(x, \xi(t, x, y) - \overline{F}^i(x)]dt, \qquad \forall (x, y) \in I\!R^d \otimes I\!R^D. \quad (3.4.90)$$

We are now able to apply Proposition 3.4.16 on page 51 when we take $r \triangleq 0$, $f \triangleq F^i$, and $\Theta \triangleq \Phi^i$, for each $i = 1, 2, \ldots, d$, and establish the following

Check of Condition 3.2.8(1): Proposition 3.4.16(1) on page 51 ensures that $\Phi^i \in Li(0)_{loc}$, and hence that $\Phi^i(x, y)$ is polynomially y-bounded of order $q_2 \triangleq 1$ locally in x (see Remark 3.4.11). By Proposition 3.4.16(2) on page 51 (with $r \triangleq 0$, $f \triangleq F^i$, $\Theta \triangleq \Phi^i$), it follows that $y \mapsto \Phi^i(x, y)$ is a C^2-mapping for each $x \in I\!R^d$, and that the partial derivative functions $(\partial_{y^l}\Phi^i)(x, y)$ and $(\partial_{y^l}\partial_{y^k}\Phi^i)(x, y)$ are continuous in (x, y) and are uniformly y-bounded locally in x, that is, for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that (3.2.17) holds. As for the x-derivative functions $(\partial_{x^l}\Phi^i)(x, y)$ and $(\partial_{x^l}\partial_{x^k}\Phi^i)(x, y)$, we see from Proposition 3.4.18 on page 53 (with $r \triangleq 0$, $f \triangleq F^i$, $\Theta \triangleq \Phi^i$), that these partial derivatives exist, are continuous in (x, y), and are polynomially y-bounded of unit order locally in x, that is, for each

$R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$|(\partial_{x^l}\Phi^i)(x, y)| + |(\partial_{x^l}\partial_{x^k}\Phi^i)(x, y)| \le C(R)(1 + |y|), \tag{3.4.91}$$

for each $(x, y) \in S^d_R \otimes I\!R^D$, so that (3.2.18) holds with $q_2 \stackrel{\triangle}{=} 1$. This verifies Condition 3.2.8(1).

Check of Condition 3.2.8(2): By Proposition 3.4.16(3) on page 51 (with $r \stackrel{\triangle}{=} 0$, $f \stackrel{\triangle}{=} F^i$, $\Theta \stackrel{\triangle}{=} \Phi^i$), we see that (3.2.19) holds, hence we have checked Condition 3.2.8(2).

Check of Condition 3.2.8(3): This is an immediate consequence of (3.4.89) and the fact that $q_2 \stackrel{\triangle}{=} 1$.

Check of Condition 3.2.8(4): We shall use Proposition 3.4.17 on page 52 as the tool for checking Condition 3.2.8(4). Suppose that a mapping $\hat{\Phi}^i : I\!R^d \otimes I\!R^D \to I\!R$ has the properties stipulated in Condition 3.2.8(4). That is, $y \to \hat{\Phi}(x, y) : I\!R^D \to I\!R$ is a C^2-mapping for each $x \in I\!R^d$ which satisfies the Poisson equation (3.2.21), and the partial derivative functions $(\partial_{y^l}\hat{\Phi}^i)(x, y)$, $(\partial_{y^l}\partial_{y^k}\hat{\Phi}^i)(x, y)$, are uniformly y-bounded locally in x, namely for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$|(\partial_{y^l}\hat{\Phi}^i)(x, y)| + |(\partial_{y^l}\partial_{y^k}\hat{\Phi}^i)(x, y)| \le C(R), \quad \forall\, (x, y) \in S^d_R \otimes I\!R^D, \tag{3.4.92}$$

for all $i = 1, 2, \dots, d$, and $l, k \in 1, 2, \dots, D$. Clearly (i) of Proposition 3.4.17 on page 52 holds (see Remark 3.4.21). Moreover, we have $\hat{\Phi}^i \in Li(0)_{loc}$ (by (3.4.92) and the mean-value theorem), and we have already seen (in Remark 3.4.22) that $F^i \in Li(0)_{loc}$. It follows that (i), (ii), and (iii) of Proposition 3.4.17 on page 52 hold when we take $r \stackrel{\triangle}{=} 0$, $f \stackrel{\triangle}{=} F^i$, $\tilde{\Theta} \stackrel{\triangle}{=} \hat{\Phi}^i$, and thus, in view of (3.4.83) and (3.4.90), we have

$$\begin{aligned}\hat{\Phi}^i(x,y) &= \int_{\mathbb{R}^D} \hat{\Phi}^i(x,z)d\pi_x(z) + \int_0^\infty E[F^i(x,\xi(t,x,y)) - \overline{F}^i(x)]dt. \\ &= \int_{\mathbb{R}^D} \hat{\Phi}^i(x,z)d\pi_x(z) + \Phi^i(x,y),\end{aligned}$$

for all $(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Thus $\Phi^i(x,y) - \hat{\Phi}^i(x,y)$ is a function of x only, as required by Condition 3.2.8(4).

Check of Condition 3.2.15(1): We will use Proposition 3.4.18 on page 53 to check Condition 3.2.15(1). From (3.4.87), (3.4.88), and (3.2.27), we see that $\tilde{F}^i \in Li(0)_{loc}$, thus for each $R \in [0,\infty)$ there is a constant $C(R) \in [0,\infty)$ such that

$$|\tilde{F}^i(x,y)| \leq C(R)[1+|y|], \quad \forall\, (x,y) \in S_R^d \otimes \mathbb{R}^D, \quad i = 1,2,\ldots,d. \qquad (3.4.93)$$

Also, we have seen that $\Phi^i \in Li(0)_{loc}$ (see Check of Condition 3.2.8(1)), hence $\Phi^i(x,y)$ is polynomially y-bounded of order $q_2 \triangleq 1$ locally in x, that is, for each $R \in [0,\infty)$ there is a constant $C(R) \in [0,\infty)$ such that

$$|\Phi^i(x,y)| \leq C(R)[1+|y|], \quad \forall\, (x,y) \in S_R^d \otimes \mathbb{R}^D, \quad i = 1,2,\ldots,d. \qquad (3.4.94)$$

Thus, from (3.4.94), (3.4.93) and (3.2.28) we see the following: for each $R \in [0,\infty)$ there is a constant $C(R) \in [0,\infty)$ such that

$$\left|a^{i,j}(x,y)\right| \leq C(R)[1+|y|^2], \quad \forall\, (x,y) \in S_R^d \otimes \mathbb{R}^D, \quad i,j = 1,2,\ldots,d. \qquad (3.4.95)$$

Moreover, we have

$$\begin{aligned}
&\left|\tilde{F}^i(x,y_1)\Phi^j(x,y_1) - \tilde{F}^i(x,y_2)\Phi^j(x,y_2)\right| \\
&\quad = \Big|\tilde{F}^i(x,y_1)\Phi^j(x,y_1) - \tilde{F}^i(x,y_1)\Phi^j(x,y_2) \\
&\qquad\quad + \tilde{F}^i(x,y_1)\Phi^j(x,y_2) - \tilde{F}^i(x,y_2)\Phi^j(x,y_2)\Big| \\
&\quad \le \left|\tilde{F}^i(x,y_1)\right| \left|\Phi^j(x,y_1) - \Phi^j(x,y_2)\right| \\
&\qquad\quad + \left|\tilde{F}^i(x,y_1) - \tilde{F}^i(x,y_2)\right| \left|\Phi^j(x,y_2)\right|
\end{aligned} \tag{3.4.96}$$

We have seen that $\Phi^i \in Li(0)_{loc}$ and $\tilde{F}^i \in Li(0)_{loc}$, hence, for each $R \in [0,\infty)$ there is a constant $C(R) \in [0,\infty)$ such that

$$\left|\Phi^j(x,y_1) - \Phi^j(x,y_2)\right| \le C(R)\,|y_1 - y_2| \tag{3.4.97}$$

$$\left|\tilde{F}^i(x,y_1) - \tilde{F}^i(x,y_2)\right| \le C(R)\,|y_1 - y_2|\,, \tag{3.4.98}$$

for each $x \in S_R^d$, $\forall y_1, y_2 \in I\!R^D$. Now, from (3.4.93), (3.4.94), (3.4.96), (3.4.97), (3.4.98), and (3.2.28), we see the following: for each $R \in [0,\infty)$ there is a constant $C(R) \in [0,\infty)$ such that

$$\left|a^{i,j}(x,y_1) - a^{i,j}(x,y_2)\right| \le C(R)\,|y_1 - y_2|\,[1 + |y_1| + |y_2|]\,, \tag{3.4.99}$$

$\forall x \in S_R^d, \forall y_1, y_2 \in I\!R^D$. Thus from (3.4.99), (3.4.95) and Definition 3.4.10 we see that $a^{i,j} \in Li(1)_{loc}$. We next look at the x-derivative functions $(\partial_{x^l} a^{i,j})(x,y)$ and $(\partial_{x^l}\partial_{x^k} a^{i,j})(x,y)$, and will see that these mappings exist, are continuous in (x,y), and are polynomially y-bounded of order $r \triangleq 2$ locally in x: We have

$$(\partial_{x^l}\tilde{F}^i\Phi^j)(x,y) = \tilde{F}^i(x,y)(\partial_{x^l}\Phi^j)(x,y) + (\partial_{x^l}\tilde{F}^i)(x,y)\Phi^j(x,y). \tag{3.4.100}$$

From Condition 3.4.19, we know that for each $R \in [0,\infty)$ there is a constant

$C(R) \in [0, \infty)$ such that

$$|(\partial_{x^l} \tilde{F}^i)(x,y)| + |(\partial_{x^l}\partial_{x^k} \tilde{F}^i)(x,y)| \leq C(R), \qquad \forall (x,y) \in S_R^d \otimes \mathbb{R}^D. \tag{3.4.101}$$

In view of (3.4.101), (3.4.100), (3.4.91), (3.4.94) and (3.4.93), we see that for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$\left|(\partial_{x^l} \tilde{F}^i \Phi^j)(x,y)\right| \leq C(R)\left[1 + |y|^2\right], \tag{3.4.102}$$

for all $(x,y) \in S_R^d \otimes \mathbb{R}^D$. Thus, in view of (3.2.28), (3.4.100), and (3.4.102), we see that the partial derivative function $(\partial_{x^l} a^{i,j})(x,y)$ exists, is continuous in (x,y), and is polynomially y-bounded of order $r \triangleq 2$ locally in x. As for the second-derivative function $(\partial_{x^l}\partial_{x^k} a^{i,j})(x,y)$, from (3.4.100) we have

$$\begin{aligned}(\partial_{x^l}\partial_{x^k} \tilde{F}^i \Phi^j)(x,y) &= (\partial_{x^k} \tilde{F}^i)(x,y)(\partial_{x^l} \Phi^j)(x,y) \\ &+ \tilde{F}^i(x,y)(\partial_{x^k}\partial_{x^l} \Phi^j)(x,y) \\ &+ (\partial_{x^k}\partial_{x^l} \tilde{F}^i)(x,y)\Phi^j(x,y) \\ &+ (\partial_{x^l} \tilde{F}^i)(x,y)(\partial_{x^k} \Phi^j)(x,y).\end{aligned} \tag{3.4.103}$$

From (3.4.103), (3.4.101), (3.4.91), (3.4.94), and (3.4.93), we see that for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$\left|(\partial_{x^l}\partial_{x^k} \tilde{F}^i \Phi^j)(x,y)\right| \leq C(R)\left[1 + |y|^2\right], \tag{3.4.104}$$

for all $(x,y) \in S_R^d \otimes \mathbb{R}^D$. Thus, in view of (3.2.28), (3.4.103), and (3.4.104), we see that the partial derivative function $(\partial_{x^l}\partial_{x^k} a^{i,j})(x,y)$ exists, is continuous in (x,y), and is polynomially y-bounded of order $r \triangleq 2$ locally in x. Now we have shown that $a^{i,j} \in Li(1)_{loc} \subset Li(2)_{loc}$, so that we can use Proposition 3.4.18(a) on page 53 (with $r \triangleq 2$ and $f \triangleq a^{i,j}$) to conclude that $\bar{a}^{i,j}(\cdot)$ is a C^2-mapping, and therefore it is locally Lipschitz continuous on $\mathbb{R}^d$. This checks Condition 3.2.15(1).

Check of Condition 3.2.15(2): From Condition 3.4.19, the fact that $y \to \Phi^i(x,y)$

are C^2-mappings for each $x \in I\!R^d$ (see Partial Check of Condition 3.2.8(1)), and (3.2.28), one sees that $y \to a^{ij}(x,y)$ is a C^2-mapping for each $x \in I\!R^d$. In the following we check that the partial derivative functions $(\partial_{y^l} a^{ij})(x,y)$ and $(\partial_{y^l}\partial_{y^k} a^{ij})(x,y)$ are polynomially y-bounded of order $r \triangleq 1$ locally in x: We have

$$(\partial_{y^l}\tilde{F}^i\Phi^j)(x,y) = \tilde{F}^i(x,y)(\partial_{y^l}\Phi^j)(x,y) + (\partial_{y^l}\tilde{F}^i)(x,y)\Phi^j(x,y). \quad (3.4.105)$$

Moreover, from Condition 3.4.19, we know that for each $R \in [0,\infty)$ there is a constant $C(R) \in [0,\infty)$ such that

$$|(\partial_{y^l}\tilde{F}^i)(x,\xi)| + |(\partial_{y^l}\partial_{y^k}\tilde{F}^i)(x,y)| \leq C(R), \qquad \forall (x,y) \in S_R^d \otimes I\!R^D. \quad (3.4.106)$$

Thus, from (3.4.105), and the bounds given by (3.4.93), (3.2.17), (3.4.106), and (3.4.94), we see that for each $R \in [0,\infty)$ there is a constant $C(R) \in [0,\infty)$ such that

$$\left|(\partial_{y^l}\tilde{F}^i\Phi^j)(x,y)\right| \leq C(R)\left[1+|y|\right], \quad (3.4.107)$$

for all $(x,y) \in S_R^d \otimes I\!R^D$. Thus, in view of (3.2.28), (3.4.105) and (3.4.107) we see that the partial derivative function $(\partial_{y^l} a^{ij})(x,y)$ exits, is continuous in (x,y), and is polynomially y-bounded of order $r \triangleq 1$ locally in x. As for the second derivative function, from (3.4.105) we find

$$\begin{aligned}(\partial_{y^l}\partial_{y^k}\tilde{F}^i\Phi^j)(x,y) &= (\partial_{y^k}\tilde{F}^i)(x,y)(\partial_{y^l}\Phi^j)(x,y) \\ &+ \tilde{F}^i(x,y)(\partial_{y^k}\partial_{y^l}\Phi^j)(x,y) \\ &+ (\partial_{y^k}\partial_{y^l}\tilde{F}^i)(x,y)\Phi^j(x,y) \\ &+ (\partial_{y^l}\tilde{F}^i)(x,y)(\partial_{y^k}\Phi^j)(x,y). \end{aligned} \quad (3.4.108)$$

Thus, from (3.4.108), together with the bounds given by (3.4.106), (3.2.17), (3.4.93), and (3.4.94), we see the following: for each $R \in [0,\infty)$ there is a constant $C(R) \in [0,\infty)$ such that

$$\left|(\partial_{y^l}\partial_{y^k}\tilde{F}^i\Phi^j)(x,y)\right| \leq C(R)\left[1+|y|\right], \quad (3.4.109)$$

for all $(x, y) \in S_R^d \otimes I\!R^D$. Hence, from (3.2.28), (3.4.108), and (3.4.109), it follows that the second derivative functions $(\partial_{y^l}\partial_{y^k} a^{i,j})(x, y)$ exist, are continuous in (x, y), and are polynomially y-bounded of order $r \triangleq 1$ locally in x. This checks (3.4.79) when $f \triangleq a^{i,j}$ and $r \triangleq 1$ (note that $r \in [0, q/2)$ since $q > 8$ in Condition 3.4.7). Now define

$$\Psi^{i,j}(x, y) \triangleq \int_0^\infty E[a^{i,j}(x, \xi(t, x, y) - \overline{a}^{i,j}(x)]dt \qquad \forall (x, y) \in I\!R^d \otimes I\!R^D. \quad (3.4.110)$$

We have already seen that $a^{i,j} \in Li(1)_{loc}$ (see Check of Condition 3.2.15(1)), so that we have checked all hypotheses for Proposition 3.4.16 on page 51 when $r \triangleq 1$, $f \triangleq a^{i,j}$ and $\Theta \triangleq \Psi^{i,j}$. Then Proposition 3.4.16(1) on page 51 ensures that $\Psi^{i,j} \in Li(1)_{loc}$. Thus from Remark 3.4.11, for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$|\Psi^{i,j}(x, y)| \leq C(R)[1 + |y|^2], \qquad \forall\, (x, y) \in S_R^d \times I\!R^D. \quad (3.4.111)$$

Proposition 3.4.16(2) on page 51 ensures that $y \mapsto \Psi^{i,j}(x, y)$ is a C^2-mapping for each $x \in I\!R^d$, and that the partial derivative functions $(\partial_{y^l}\Psi^{i,j})(x, y)$ and $(\partial_{y^l}\partial_{y^k}\Psi^{i,j})(x, y)$ are continuous in (x, y) and are polynomially y-bounded of order $r \triangleq 1$ locally in x; that is, for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$|(\partial_{y^l}\Psi^{i,j})(x, y)| + |(\partial_{y^l}\partial_{y^k}\Psi^{i,j})(x, y)| \leq C(R)[1 + |y|], \qquad \forall\, (x, y) \in S_R^d \times I\!R^D. \quad (3.4.112)$$

As for the derivative functions $(\partial_{x^l}\Psi^{i,j})(x, y)$ and $(\partial_{x^l}\partial_{x^k}\Psi^{i,j})(x, y)$, we have already seen that the hypotheses of Proposition 3.4.18 on page 53 are verified for $r \triangleq 2$ and $f \triangleq a^{i,j}$ (see Check of Condition 3.2.15(1)). Thus, from Proposition 3.4.18(b) on page 53, with $\Theta \triangleq \Psi^{i,j}$, we conclude that the partial derivatives $(\partial_{x^l}\Psi^{i,j})(x, y)$ and $(\partial_{x^l}\partial_{x^k}\Psi^{i,j})(x, y)$ exist, are continuous in (x, y), and are polynomially y-bounded of order $r + 1 = 3$ locally in x. That is, for each $R \in [0, \infty)$ there is a constant

$C(R) \in [0, \infty)$ such that

$$|(\partial_{x^l}\Psi^{i,j})(x,y)| + |(\partial_{x^l}\partial_{x^k}\Psi^{i,j})(x,y)| \le C(R)[1 + |y|^3], \qquad \forall\ (x,y) \in S^d_R \times \mathbb{R}^D. \tag{3.4.113}$$

Putting together (3.4.111), (3.4.112), and (3.4.113), we see that for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that

$$\begin{aligned} &|\Psi^{i,j}(x,y)| + |(\partial_{x^l}\Psi^{i,j})(x,y)| + |(\partial_{y^l}\Psi^{i,j})(x,y)| \\ &\quad + |(\partial_{x^l}\partial_{x^k}\Psi^{i,j})(x,y)| + |(\partial_{y^l}\partial_{y^k}\Psi^{i,j})(x,y)| \le C(R)(|1 + |y|^3), \end{aligned}$$

for each $(x, y) \in S^d_R \otimes \mathbb{R}^D$. This checks (3.2.35) with $q_3 \triangleq 3$, and so we have verified Condition 3.2.15(2)).

Check of Condition 3.2.15(3): Proposition 3.4.16(3) on page 51 (with $f \triangleq a^{i,j}$, $r \triangleq 1$, $\Theta \triangleq \Psi^{i,j}$) shows that the Poisson relation (3.2.36) holds, which checks Condition 3.2.15(3).

Check of Condition 3.2.18: We have now shown that Condition 3.2.8 holds with $q_2 = 1$, and Condition 3.2.15 holds with $q_3 = 3$. We now put $q_4 \triangleq q$, where q is the constant in Condition 3.4.7. Then, since $q > 8$, we have $q_4 \in (2 + 2q_2, \infty) \cap (2 + 2q_3, \infty) \equiv (8, \infty)$. Moreover, from Proposition 3.4.12 on page 50, we see that for each $R \in [0, \infty)$ there are constants $\alpha_R \in [0, \infty)$ and $\lambda_R \in (0, \infty)$ such that (3.2.38) holds. This checks Condition 3.2.18.

Remark 3.4.23. We have now shown that, if Conditions 3.2.1, 3.4.4, 3.4.7 and 3.4.19 hold, then Condition 3.2.8 holds with $q_2 = 1$, Condition 3.2.15 holds with $q_3 = 3$, and Condition 3.2.18 holds with $q_4 = q$ in Condition 3.4.7. We also showed that Condition 3.2.3 holds, except for the global Lipschitz continuity of the mapping $\bar{F}$. This last requirement appears to be rather difficult to verify in general terms, mainly because the dependence on x of the unique invariant probability measure π_x in (3.1.4) is difficult to characterize. In particular examples, such as when (3.1.3) has no dependence on x so that the invariant probability measure π is fixed, this global Lipschitz continuity may be clear.

Remark 3.4.24. The results of this section do not require that $\sigma\sigma^T(x,y)$ be non-singular. Condition 3.4.7 provides enough "stability" in (3.1.3) to ensure that there is a unique invariant probability measure π_x for the Markov process defined by (3.1.3) and to ensure that Conditions 3.2.8, 3.2.15, and 3.2.18 hold. Verification of Condition 3.4.7 can be facilitated when there is additional structure in the system (3.1.1) and (3.1.2). For example:

(a) if $\sigma(x,y) = \sigma(x)$, $\forall(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$, then $\Lambda_0(R) = 0$, $R \in [0,\infty)$, in Condition 3.4.7.

(b) if $\sigma(x,y) = \sigma_1(x) + \sigma_2(y)$ and $b(x,y) = b_1(x) + b_2(y)$ on $\mathbb{R}^d \otimes \mathbb{R}^D$, then Λ_0 in (3.4.64) is determined by $\sigma_2(.)$ and no longer depends on R, while λ_{max} (see Remark 3.4.6)) is determined by $b_2(.)$ and is a function of y only. Thus Condition 3.4.7 follows when

$$\sup_{\xi \in \mathbb{R}^D} [\lambda_{max}(\xi)] < \frac{(1-q)}{2}\Lambda_0^2.$$

A special case of this occurs when $\sigma(x,y)$ and $b(x,y)$ are functions of y only.

(c) When $D = N = 1$ then Condition 3.4.7 reduces to a simple condition on the scalar function $(\partial_\xi \sigma)(x,\xi)$ and $(\partial_\xi b)(x,\xi)$.

In Proposition 3.2.13 on page 37 we saw that the matrix $\overline{a}(x)$ defined by the relation (3.2.29) is positive semi-definite for each $x \in \mathbb{R}^d$ when Conditions 3.2.1, 3.2.2, 3.2.3 and 3.2.8 hold. In the next result we establish a more intuitively appealing expression for $\overline{a}(x)$ when the conditions of this Section are assumed:

Proposition 3.4.25. *Suppose that Conditions 3.2.1, 3.2.2, 3.4.4, 3.4.7 and 3.4.19 hold. Then the integrals in (3.4.114) are well-defined, and* $\overline{a}(x)$ *defined by (3.2.29) is alternatively given by*

$$\overline{a}(x) \triangleq \int_0^\infty \left\{ \bar{E}\{\bar{F}(x,\overline{\xi}(s,x))\bar{F}^T(x,\overline{\xi}(0;x))\} + \bar{E}\{\bar{F}(x,\overline{\xi}(0;x))\bar{F}^T(x,\overline{\xi}(s;x))\} \right\} ds, \tag{3.4.114}$$

for each $x \in \mathbb{R}^d$, *where* $\{\overline{\xi}(t;x),\ t \in [0,\infty)\}$ *is some* $\mathbb{R}^D$*-valued stationary Markov process defined by (3.1.3) (on some probability space* $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$*) with marginal distribution given by the invariant probability measure* π_x *of Proposition 3.4.14 on page 50.*

Appendix A

Proofs for Section 3.3

Remark A.0.26. For the proofs of this section, define the stopping times:

$$t_R^\epsilon \triangleq \inf\{t \in [0, \epsilon^{-1}] : |x^\epsilon(t)| \geq R\} \tag{A.0.1}$$

$$T_R^\epsilon \triangleq \inf\{\tau \in [0, 1] : |X^\epsilon(\tau)| \geq R\}, \tag{A.0.2}$$

for each $R \in (0, \infty)$ and $\epsilon \in (0, 1]$. From (3.1.1) and (3.1.2), and Condition 3.2.2 we see that t_R^ϵ is a $\{\mathcal{F}_t\}$-stopping time, and, from Section 3.3, it follows that T_R^ϵ is a $\mathcal{G}_\tau^\epsilon$-stopping time for each $\epsilon \in (0, 1]$ (recall (3.3.41)); and

$$T_R^\epsilon \triangleq \epsilon t_R^\epsilon, \qquad \forall \epsilon \in (0, 1], \ \forall R \in (0, \infty). \tag{A.0.3}$$

A.1 The Main Result

Proof of Theorem 3.3.3 on page 42: Let $\mathcal{C}$ denote the linear second order differential operator for the diffusion process defined in the equations (3.3.46) and

(3.3.47). Thus, for each $g(x,z) \in C^2(\mathbb{R}^d \otimes \mathbb{R}^d)$ we have

$$
\begin{aligned}
&\mathcal{C}g(x,z) \\
&\triangleq \sum_{i=1}^{d} \overline{F}^i(x)\,(\partial_{x^i} g)(x,z) + \sum_{i=1}^{d}\sum_{j=1}^{d} z^j \left(\partial_{x^j}\overline{F}^i\right)(x)\,(\partial_{z^i} g)(x,z) \\
&+\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M} G^{i,l}(x)G^{j,l}(x)\,(\partial_{x^i}\partial_{x^j} g)(x,z) \\
&+\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M} (\partial_{x^i}\partial_{z^j} g)(x,z)G^{i,l}(x)\left[(\partial_x G^{j,l})(x)z\right] \\
&+\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\left\{\left(\overline{a}^{i,j}(x) + \sum_{l=1}^{M}\left[(\partial_x G^{i,l})(x)z\right]\left[(\partial_x G^{j,l})(x)z\right]\right)(\partial_{z^i}\partial_{z^j} g)(x,z)\right\}.
\end{aligned}
\tag{A.1.4}
$$

Remark A.1.1. In proving (3.3.52) we use the following method: first establish tightness of the family of measures $\{\mathcal{L}(X^\epsilon, Z^\epsilon),\ \epsilon \in (0,1]\}$ and then show that each weak accumulation point of the family $\{\mathcal{L}(X^\epsilon, Z^\epsilon),\ \epsilon \in (0,1]\}$ coincides with $\mathcal{L}(\hat{X},\hat{Z})$ for $(\hat{X},\hat{Z})$ given by (3.3.46) and (3.3.47). Now (3.3.52) follows from Remark 3.3.2. In following this plan we use the Stroock and Varadhan martingale problem, and the so-called *perturbed near-identity transformation method.* That is, if $\mathcal{G}^\epsilon$ is the usual linear second order differential operator for $(X^\epsilon, Y^\epsilon, Z^\epsilon)$ (given by (3.3.42), (3.3.43), (3.3.45)), then for each $g \in C_c^\infty(\mathbb{R}^d \otimes \mathbb{R}^d)$ we will construct a C^2-mapping $\Psi_g^\epsilon : \mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d \to \mathbb{R}$ having the form

$$\Psi_g^\epsilon(x,y,z) \triangleq g(x,z) + \epsilon^{1/2}\phi_g^1(x,y,z) + \epsilon\phi_g^2(x,y,z), \tag{A.1.5}$$

where $\phi_g^1(x,y,z)$ and $\phi_g^2(x,y,z)$ will be chosen such that $\mathcal{G}^\epsilon\Psi_g^\epsilon$ is of the form

$$\mathcal{G}^\epsilon\Psi_g^\epsilon(x,y,z) = \mathcal{C}g(x,z) + \epsilon\Delta(\epsilon,x,y,z). \tag{A.1.6}$$

Here the linear operator $\mathcal{C}$, defined in (A.1.4), is a consequence of the averaging procedure and acts on functions depending on (x,z) only, and the mapping Δ is a bounded continuous function. The relation (A.1.6) implies that as ϵ goes to zero the linear second order differential operator $\mathcal{G}^\epsilon$ "converges" to the linear second order differential operator $\mathcal{C}$ in (A.1.4), which corresponds to the limiting diffusion process $(\hat{X}, \hat{Z})$ defined by (3.3.46) and (3.3.47). Thus, using the martingale problem method and (A.1.6), one can prove the weak convergence $\mathcal{L}(X^\epsilon, Z^\epsilon) \to \mathcal{L}(\hat{X}, \hat{Z})$, as $\epsilon \to 0$.

The perturbed near-identity transformation method introduced above is broadly used in singular perturbation theory when dealing with averaging problems. This method, sometimes in the literature called the *perturbed test function method* , is used for proving the weak convergence theorem, as the functions Ψ_g^ϵ in (A.1.5) are perturbed by the factor depending on ϵ. The general idea of applying this method was introduced in Kurtz [22], Blankenship and Papanicolau [5], Papanicolau, Stroock, and Varadhan [29] and subsequently extended in Kushner(e.g see [23]).

Following the procedure outlined above we are now ready to establish (3.3.52) in three steps:

STEP I: Let $(\Omega^*, \mathcal{F}^*)$ be the measurable space defined by (3.3.49) and (3.3.50). Also, let $\{\mathcal{B}(\tau),\ \tau \in [0,1]\}$ be a filtration defined in $(\Omega^*, \mathcal{F}^*)$ by

$$\mathcal{B}(\tau) \stackrel{\triangle}{=} \sigma\{X(s), Z(s),\ s \in [0,\tau]\}. \tag{A.1.7}$$

To introduce the martingale problem, for each $g(x,z) \in C_c^\infty(\mathbb{R}^d \otimes \mathbb{R}^d)$, define

$$\Lambda_g(X,Z)(\tau) \stackrel{\triangle}{=} g(X(\tau), Z(\tau)) - g(X(0), Z(0)) - \int_0^\tau \mathcal{C}g(X(s), Z(s))ds, \tag{A.1.8}$$

for every $(X, Z) \in \Omega^*$ and $\tau \in [0,1]$. For each $\epsilon \in (0,1]$, put

$$P_\epsilon^* \stackrel{\triangle}{=} \mathcal{L}(X^\epsilon, Z^\epsilon). \tag{A.1.9}$$

Now, fix an arbitrary sequence $\{\epsilon_n,\ n \in \mathbb{N}\} \subset (0,1]$ such that $\epsilon_n \to 0$ as $n \to \infty$.

To get (3.3.52) it is enough to prove

$$\lim_{n\to\infty} P^\bullet_{\epsilon_n} = \mathcal{L}(\hat{X}, \hat{Z}). \tag{A.1.10}$$

To establish this, fix an arbitrary subsequence $\{P^\bullet_{\epsilon_{n(k)}}\}_k$ of $\{P^\bullet_{\epsilon_n}\}$. From Proposition A.3.1 on page 98 we know that $\{P^\bullet_{\epsilon_n}\}$ is weakly relatively compact, hence the subsequence $\{P^\bullet_{\epsilon_{n(k)}}\}_k$ is also weakly relatively compact. Hence, by definition of weak relative compactness, there is a further subsequence $\{P^\bullet_{\epsilon_{n(k(m))}}\}_m$ with some probability measure $P^\bullet$ on $(\Omega^\bullet, \mathcal{F}^\bullet)$, such that

$$\lim_{m\to\infty} \{P^\bullet_{\epsilon_{n(k(m))}}\}_m = P^\bullet. \tag{A.1.11}$$

It remains to show that

$$P^\bullet = \mathcal{L}(\hat{X}, \hat{Z}) \tag{A.1.12}$$

for then (A.1.10) follows by Fact C.0.5 on page 203. To simplify the notation we will write $P^\bullet_m$ for $P^\bullet_{\epsilon_{n(k(m))}}$, so that (A.1.11) becomes

$$\lim_{m\to\infty} P^\bullet_m = P^\bullet. \tag{A.1.13}$$

In view of Remark 3.3.2, the probability law of each solution $(\hat{X}, \hat{Z})$ of (3.3.46), (3.3.47), is uniquely defined. Thus, by Corollary 5.4.9 of Karatzas and Shreve [19] there is a unique probability measure P on $(\Omega^\bullet, \mathcal{F}^\bullet)$ with

$$P\left[(X, Z) \in \Omega^\bullet : \ (X(0), Z(0)) \in \Gamma\right] = \delta(x_0, z_0), \qquad \forall \Gamma \in \mathcal{B}(\mathbb{R}^d, \mathbb{R}^d), \tag{A.1.14}$$

such that $\{(\Lambda_g(\tau), \mathcal{B}(\tau)),\ \tau \in [0,1]\}$ is a martingale on $(\Omega^\bullet, \mathcal{F}^\bullet, P)$, $\forall g \in C_c^\infty(\mathbb{R}^d \otimes \mathbb{R}^d)$, and this probability is the (unique) law $\mathcal{L}(\hat{X}, \hat{Z})$. Thus, it is enough to show that $\{(\Lambda_g(\tau), \mathcal{B}(\tau)),\ \tau \in [0,1]\}$ is a martingale on $(\Omega^\bullet, \mathcal{F}^\bullet, P^\bullet)$ (where $P^\bullet$ is the limit in (A.1.13)) for each $g \in C_c^\infty(\mathbb{R}^d \otimes \mathbb{R}^d)$, for then (A.1.12) follows, as required to establish the Theorem.

From now on fix $g(x, z) \in C_c^\infty(\mathbb{R}^d \otimes \mathbb{R}^d)$. To see that $\{(\Lambda_g(\tau), \mathcal{B}(\tau),\ \tau \in [0,1]\}$ is a martingale on $(\Omega^\bullet, \mathcal{F}^\bullet, P^\bullet)$, it is enough to establish that, for arbitrary $0 \leq$

$\tau_1 < \tau_2 \leq 1$ and mappings $\Gamma : \Omega^* \to \mathbb{R}$ having the form

$$\Gamma(X, Z) \triangleq \prod_{i=1}^{n} h_i(X(s_i), Z(s_i)), \qquad \forall (X, Z) \in \Omega^*, \tag{A.1.15}$$

with $0 \leq s_0 < s_1 < \ldots < s_n \leq \tau_1$ and continuous uniformly bounded functions $h_i : \mathbb{R}^d \otimes \mathbb{R}^d \to \mathbb{R}$, one has

$$E^{P^*}\left[(\Lambda_g(X, Z)(\tau_2) - \Lambda_g(X, Z)(\tau_1))\, \Gamma(X, Z)\right] = 0. \tag{A.1.16}$$

It then follows from Fact C.0.4 on page 202 that (A.1.16) continues to hold for all $\mathcal{B}(\tau_1)$-measurable and uniformly bounded mappings $\Gamma : \Omega^* \to \mathbb{R}$, which proves that $\{(\Lambda_g(\tau), \mathcal{B}(\tau)),\ \tau \in [0, 1]\}$ is a martingale. Thus fix arbitrary partition $0 \leq s_0 < s_1 < \ldots < s_n \leq \tau_1$ and fix continuous uniformly bounded functions $h_i : \mathbb{R}^d \otimes \mathbb{R}^d \to \mathbb{R}$, $i = 1, 2, \ldots, n$. It remains to show that (A.1.16) holds when Γ is given by (A.1.15).

Since $(X, Z) \to [(\Lambda_g(\tau_2) - \Lambda_g(\tau_1))\Gamma] : \Omega^* \longrightarrow \mathbb{R}$ is clearly a uniformly bounded and continuous mapping, it follows from (A.1.13) that

$$\begin{aligned} \lim_{m\to\infty} E^{P^*_m}\left[(\Lambda_g(X, Z)(\tau_2) - \Lambda_g(X, Z)(\tau_1))\, \Gamma(X, Z)\right] \\ = E^{P^*}\left[(\Lambda_g(X, Z)(\tau_2) - \Lambda_g(X, Z)(\tau_1))\, \Gamma(X, Z)\right]. \end{aligned} \tag{A.1.17}$$

Now one sees from (A.1.9) that

$$\begin{aligned} E^{P^*_\epsilon}\left[(\Lambda_g(X, Z)(\tau_2) - \Lambda_g(X, Z)(\tau_1))\, \Gamma(X, Z)\right] \\ = E\left[(\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1))\, \Gamma(X^\epsilon, Z^\epsilon)\right]. \end{aligned} \tag{A.1.18}$$

Thus, the relation (A.1.16)(hence Theorem 3.3.3) on page 42 follows from (A.1.17) and (A.1.18) once it is shown that

$$\lim_{\epsilon\to 0} E\left[(\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1))\, \Gamma(X^\epsilon, Z^\epsilon)\right] = 0. \tag{A.1.19}$$

In order to get (A.1.19), we shall establish in Step II the following

$$\lim_{\epsilon\to 0} E\left[\{\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)\}\,\Gamma(X^\epsilon, Z^\epsilon)\right] = 0, \quad \forall R \in (R_0, \infty). \tag{A.1.20}$$

for a fixed $R_0 \in (0, \infty)$ being large enough that support of $g(x, z)$ is within the interior of $S^d_{R_0} \otimes \mathbb{R}^d$ (T_R^ϵ is defined in (A.0.2)). To see that (A.1.19) follows from (A.1.20), note from (A.1.4), (A.1.8), the uniform bounds on the h_i in (A.1.15), and $g \in C_c^\infty(\mathbb{R}^d \otimes \mathbb{R}^d)$, that

$$|\Lambda_g(X, Z)(\tau)\Gamma(X, Z)| \le C_1, \tag{A.1.21}$$

for all $(X, Z) \in \Omega^*, \tau \in [0, 1]$, and some constant $C_1 \in [0, \infty)$. Now we can write

$$\begin{aligned}
&E\left([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1)]\,\Gamma(X^\epsilon, Z^\epsilon)\right)\\
&\quad = E\left([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon \ge 1\}\right)\\
&\quad + E\left([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon < 1\}\right)\\
&\quad = E\left([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon \ge 1\}\right)\\
&\quad + E\left([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon < 1\}\right)\\
&\quad = E\left([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)]\,\Gamma(X^\epsilon, Z^\epsilon)\right)\\
&\quad - E\left([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon < 1\}\right)\\
&\quad + E\left([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon < 1\}\right)
\end{aligned} \tag{A.1.22}$$

Thus, from (A.1.21) and (A.1.22), there is some constant $C \in [0, \infty)$ such that

$$
\begin{aligned}
&|E([(\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1))\,\Gamma(X^\epsilon, Z^\epsilon)])| \\
&\quad \le |E([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)]\,\Gamma(X^\epsilon, Z^\epsilon))| \\
&\quad + |E([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon < 1\})| \\
&\quad + |E([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon < 1\})| \\
&\quad \le |E([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)]\,\Gamma(X^\epsilon, Z^\epsilon))| \\
&\quad + E|[\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon < 1\}| \\
&\quad + E|[\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1)]\,\Gamma(X^\epsilon, Z^\epsilon) I\{T_R^\epsilon < 1\}| \\
&\quad \le |E([\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)]\,\Gamma(X^\epsilon, Z^\epsilon))| \\
&\qquad + 2CP[T_R^\epsilon < 1]. \qquad (A.1.23)
\end{aligned}
$$

$\forall \epsilon \in (0, 1]$, $\forall R \in [0, \infty)$. Now fix some small $\eta \in (0, \infty)$. By Proposition A.4.4 on page 111 there is some $R(\eta) \in (R_0, \infty)$ and $\epsilon_1(\eta) \in (0, 1]$ such that

$$CP[T_{R(\eta)}^\epsilon < 1] < \frac{\eta}{2}, \qquad \forall \epsilon \in (0, \epsilon_1(\eta)). \qquad (A.1.24)$$

Moreover, by (A.1.20), there is some $\epsilon_2(\eta) \in (0, 1]$ such that

$$|E\left[(\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_{R(\eta)}^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_{R(\eta)}^\epsilon))\,\Gamma(X^\epsilon, Z^\epsilon)\right]| < \frac{\eta}{2}, \quad (A.1.25)$$

$\forall \epsilon \in (0, \epsilon_2(\eta))$. Using (A.1.23), (A.1.24) and (A.1.25) we get

$$|E\left[(\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1))\,\Gamma(X^\epsilon, Z^\epsilon)\right]| < \eta, \qquad (A.1.26)$$

$\forall \epsilon \in (0, \epsilon_1(\eta) \wedge \epsilon_2(\eta))$, which proves (A.1.19), and hence Theorem 3.3.3 on page 42 follows.

Remark A.1.2. Thus, it remains to prove (A.1.20) in order to establish Theorem 3.3.3 on page 42. We shall do this in Steps II to IV using the near-identity method summerized in Remark A.1.1.

STEP II: Since $\overline{F}^i(\cdot)$ is a C^2-function(by Condition 3.2.3) we have by the mean value theorem:

$$\overline{F}^i(x - \epsilon^{1/2}z) = \overline{F}^i(x) - \epsilon^{1/2}\left[(\partial_x \overline{F}^i)(x)z\right] + \epsilon I_1^i(\epsilon, x, z), \tag{A.1.27}$$

(recall that $\partial_x \overline{F}^i(x)$ is a row vector of length d), where

$$I_1^i(\epsilon, x, z) \triangleq \int_0^1 \left(\sum_{k=1}^d z_k(1-\zeta)\sum_{n=1}^d z_n \left(\partial_{x_n}\partial_{x_k} F^i\right)(x - \epsilon^{1/2}\zeta z)\right) d\zeta, \tag{A.1.28}$$

$\forall (x, z) \in \mathbb{R}^d \otimes \mathbb{R}^d$, $\forall \epsilon \in (0, 1]$, $i = 1, 2, \ldots, d$. Similarly, since $G^{i,j}(.)$ is a C^2-function (see Condition 3.2.1) we have

$$G^{i,l}(x - \epsilon^{1/2}z) = G^{i,l}(x) - \epsilon^{1/2}\left[(\partial_x G^{i,l})(x)z\right] + \epsilon I_2^{i,l}(\epsilon, x, z), \tag{A.1.29}$$

(recall that $\partial_x G^{i,l}(x)$ is a row vector of length d), where

$$I_2^{i,l}(\epsilon, x, z) \triangleq \int_0^1 \left(\sum_{k=1}^d z_k(1-\zeta)\sum_{n=1}^d z_n \left(\partial_{x_n}\partial_{x_k} G^{i,l}\right)(x - \epsilon^{1/2}\zeta z)\right) d\zeta, \tag{A.1.30}$$

$\forall (x, z) \in \mathbb{R}^d \otimes \mathbb{R}^d$, $\forall \epsilon \in (0, 1]$, $i = 1, 2, \ldots, d, l = 1, 2, \ldots, M$. Moreover, from (3.3.45), (A.1.27), and (A.1.29) one has

$$\overline{F}^i(\overline{X}^\epsilon(\tau)) = \overline{F}^i(X^\epsilon(\tau)) - \epsilon^{1/2}\left[(\partial_x \overline{F}^i)(X^\epsilon(\tau))Z^\epsilon(\tau)\right] + \epsilon I_1^i(\epsilon, X^\epsilon(\tau), Z^\epsilon(\tau)), \tag{A.1.31}$$

$\forall \tau \in [0, 1]$, $\forall i = 1, 2, \ldots, d$, and

$$G^{i,l}(X^\epsilon(\tau)) - G^{i,l}(\overline{X}^\epsilon(\tau)) = \epsilon^{1/2}\left[(\partial_x G^{i,l})(X^\epsilon(\tau))Z^\epsilon(\tau)\right] - \epsilon I_2^{i,l}(\epsilon, X^\epsilon(\tau), Z^\epsilon(\tau)), \tag{A.1.32}$$

$\forall \tau \in [0, 1]$, $\forall i = 1, 2, \ldots, d$, and $l = 1, 2, \ldots, M$. Now, from (3.3.45), (3.3.42), (3.3.44) we get

$$\begin{aligned} Z_i^\epsilon(\tau) &= \epsilon^{-1/2}\left\{\int_0^\tau \left(F^i(X^\epsilon(s), Y^\epsilon(s)) - \overline{F}^i(\overline{X}^\epsilon(s)\right) ds \right. \\ &\left. \qquad + \sum_{k=1}^M \int_0^\tau \left(G^{i,k}(X^\epsilon(s)) - G^{i,k}(\overline{X}^\epsilon(s))\right) dW_k^\epsilon(s)\right\}. \end{aligned} \tag{A.1.33}$$

Then, using (A.1.31) and (A.1.32) in (A.1.33), we can write:

$$\begin{aligned} Z_i^\epsilon(\tau) = & \left\{ \epsilon^{-1/2} \int_0^\tau \left[F^i(X^\epsilon(s), Y^\epsilon(s)) - \overline{F}^i(X^\epsilon(s)) \right] ds \right. \\ & \left. + \int_0^\tau \left(\partial_z \overline{F}^i \right) (X^\epsilon(s)) Z^\epsilon(s) ds - \epsilon^{1/2} \int_0^\tau I_1^i(\epsilon, X^\epsilon(s), Z^\epsilon(s)) ds \right\} \\ & + \sum_{k=1}^M \int_0^\tau \left(\partial_z G^{i,k} \right) (X^\epsilon(s)) Z^\epsilon(s) dW_k^\epsilon(s) \\ & - \epsilon^{1/2} \sum_{k=1}^M \int_0^\tau I_2^{i,k}(\epsilon, X^\epsilon(s), Z^\epsilon(s)) dW_k^\epsilon(s). \end{aligned} \tag{A.1.34}$$

We next calculate the variation and cross-variation processes of $\{X^\epsilon(\tau),\ \tau \in [0,1]\}$, $\{Y^\epsilon(\tau),\ \tau \in [0,1]\}$, and $\{Z^\epsilon(\tau),\ \tau \in [0,1]\}$, using (3.3.42), (3.3.43), (A.1.33) and independence of $\{W^\epsilon(\tau),\ \tau \in [0,1]\}$ and $\{B^\epsilon(\tau),\ \tau \in [0,1]\}$(as follows from (3.3.40) and Condition 3.2.2):

$$\begin{aligned}
\left[X_i^\epsilon, X_j^\epsilon\right](\tau) &= \int_0^\tau \sum_{l=1}^{M} G^{i,l}(X^\epsilon(s))G^{j,l}(X^\epsilon(s))ds, \\
\left[X_i^\epsilon, Y_j^\epsilon\right](\tau) &= \epsilon^{-1/2}\int_0^\tau \sum_{l=1}^{M}\sum_{k=1}^{N} G^{i,l}(X^\epsilon(s))\sigma^{j,k}(X^\epsilon(s), Y^\epsilon(s))d[W_i^\epsilon, B_k^\epsilon] \\
&\equiv 0, \\
\left[X_i^\epsilon, Z_j^\epsilon\right](\tau) &= \int_0^\tau \sum_{l=1}^{M} G^{i,l}(X^\epsilon(s))\left[(\partial_x G^{j,l})(X^\epsilon(s))Z^\epsilon(s)\right] ds \\
&\quad -\epsilon^{1/2}\int_0^\tau \sum_{l=1}^{M} G^{i,l}(X^\epsilon(s))I_2^{j,l}\left(\epsilon, X^\epsilon(s), Z^\epsilon(s)\right) ds, \\
\left[Y_i^\epsilon, Y_j^\epsilon\right](\tau) &= \epsilon^{-1}\int_0^\tau \sum_{k=1}^{N} \left[\sigma^{i,k}(X^\epsilon(s), Y^\epsilon(s))\sigma^{j,k}(X^\epsilon(s), Y^\epsilon(s))\right] ds, \\
\left[Y_i^\epsilon, Z_j^\epsilon\right](\tau) &= \epsilon^{-1/2}\int_0^\tau \sum_{l=1}^{M}\sum_{k=1}^{N} \sigma^{i,k}(X^\epsilon(s), Y^\epsilon(s))\left\{G^{j,l}(X^\epsilon(s))\right. \\
&\qquad\qquad \left. - G^{j,l}(\overline{X}^\epsilon(s))\right\} d[W_i^\epsilon, B_k^\epsilon] \\
&\equiv 0, \\
\left[Z_i^\epsilon, Z_j^\epsilon\right](\tau) &= \int_0^\tau \sum_{l=1}^{M} \left[(\partial_x G^{i,l})(X^\epsilon(s))Z^\epsilon(s)\right]\left[(\partial_x G^{j,l})(X^\epsilon(s))Z^\epsilon(s)\right] ds \\
&\quad -\epsilon^{1/2}\int_0^\tau \sum_{l=1}^{M} \left[(\partial_x G^{i,l})(X^\epsilon(s))Z^\epsilon(s)\right] I_2^{j,l}(\epsilon, X^\epsilon(s), Z^\epsilon(s))ds \\
&\quad -\epsilon^{1/2}\int_0^\tau \sum_{l=1}^{M} \left[(\partial_x G^{j,l})(X^\epsilon(s))Z^\epsilon(s)\right] I_2^{i,l}(\epsilon, X^\epsilon(s), Z^\epsilon(s))ds \\
&\quad +\epsilon\int_0^\tau \sum_{l=1}^{M} I_2^{i,l}(\epsilon, X^\epsilon(s), Z^\epsilon(s))I_2^{j,l}(\epsilon, X^\epsilon(s), Z^\epsilon(s))ds.
\end{aligned} \tag{A.1.35}$$

Remark A.1.3. Fix some arbitrary continuous mapping $(x, y, z) \to \Psi(x, y, z)$: $(\mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d) \longrightarrow \mathbb{R}$, whose first partial derivative functions $(\partial_{x^i}\Psi)$, $i = 1, 2, \ldots, d$, $(\partial_{y^i}\Psi)$, $i = 1, 2, \ldots, D$, $(\partial_{z^i}\Psi)$, $i = 1, 2, \ldots, d$, exist and are continuous on $\mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d$, and whose second partial derivative functions $(\partial_{x^i}\partial_{x^j}\Psi)$, $i,j = 1,2,\dots,d$, $(\partial_{y^i}\partial_{y^j}\Psi)$, $i,j = 1,2,\dots,D$, $(\partial_{z^i}\partial_{z^j}\Psi)$, $i,j = 1,2,\dots,d$, $(\partial_{x^i}\partial_{z^j}\Psi)$, $i,j = 1,2,\dots,d$, exist and are continuous on $\mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d$. Notice that we do not require the mixed derivative functions $\partial_{x^i}\partial_{y^j}$ and $\partial_{y^j}\partial_{z^i}$ to exist. Since we have seen that the cross-variations $[X_i^\epsilon, Y_j^\epsilon]$ and $[Y_i^\epsilon, Z_j^\epsilon]$ are identically zero, we will not need these derivatives when we make an expansion of $\Psi(X^\epsilon, Y^\epsilon, Z^\epsilon)$, as we do next.

For a mapping Ψ as in Remark A.1.3, Itô's formula gives

$$\begin{aligned}\Psi(X^\epsilon(\tau), Y^\epsilon(\tau), Z^\epsilon(\tau)) &= \Psi(x_0, y_0, z_0) + \sum_{i=1}^{d} \int_0^\tau (\partial_{x^i}\Psi)(X^\epsilon, Y^\epsilon, Z^\epsilon)(s)\,dX_i^\epsilon(s)\\ &+ \sum_{i=1}^{D} \int_0^\tau (\partial_{y^i}\Psi)(X^\epsilon, Y^\epsilon, Z^\epsilon)(s)\,dY_i^\epsilon(s)\\ &+ \sum_{i=1}^{d} \int_0^\tau (\partial_{z^i}\Psi)(X^\epsilon, Y^\epsilon, Z^\epsilon)(s)\,dZ_i^\epsilon(s)\\ &+ \frac{1}{2}\left\{\sum_{i=1}^{d}\sum_{j=1}^{d} \int_0^\tau (\partial_{x^i}\partial_{x^j}\Psi)(X^\epsilon, Y^\epsilon, Z^\epsilon)(s)\,d[X_i^\epsilon, X_j^\epsilon](s)\right.\\ &+ 2\sum_{i=1}^{d}\sum_{j=1}^{d} \int_0^\tau (\partial_{x^i}\partial_{z^j}\Psi)(X^\epsilon, Y^\epsilon, Z^\epsilon)(s)\,d[X_i^\epsilon, Z_j^\epsilon](s)\\ &+ \sum_{i=1}^{D}\sum_{j=1}^{D} \int_0^\tau (\partial_{y^i}\partial_{y^j}\Psi)(X^\epsilon, Y^\epsilon, Z^\epsilon)(s)\,d[Y_i^\epsilon, Y_j^\epsilon](s)\\ &+ \left.\sum_{i=1}^{d}\sum_{j=1}^{d} \int_0^\tau (\partial_{z^i}\partial_{z^j}\Psi)(X^\epsilon, Y^\epsilon, Z^\epsilon)(s)\,d[Z_i^\epsilon, Z_j^\epsilon](s)\right\}\end{aligned} \tag{A.1.36}$$

$\forall \tau \in [0,1]$. Putting together (A.1.32), (A.1.33), (A.1.35), (A.1.36), (3.2.14), (3.3.42), and (3.3.43), one sees that

$$\Psi(X^\epsilon, Y^\epsilon, Z^\epsilon)(\tau) = \Psi(x_0, y_0, z_0) + \int_0^\tau \mathcal{G}^\epsilon\Psi(X^\epsilon, Y^\epsilon, Z^\epsilon)(s)\,ds + M_\Psi^\epsilon(\tau),\ \forall \tau \in [0,1], \tag{A.1.37}$$

where

$$\begin{aligned}
M_\Psi^\epsilon(\tau) \triangleq\ & \sum_{i=1}^{d}\sum_{l=1}^{M}\int_0^\tau (\partial_{z^i}\Psi)\,(X^\epsilon,Y^\epsilon,Z^\epsilon)(s)G^{i,l}(X^\epsilon(s))dW_l^\epsilon(s) \\
& +\epsilon^{-1/2}\sum_{i=1}^{D}\sum_{k=1}^{N}\int_0^\tau \left(\partial_{y^i}\Psi\right)(X^\epsilon,Y^\epsilon,Z^\epsilon)(s)\sigma^{i,k}(X^\epsilon(s),Y^\epsilon(s))dB_k^\epsilon(s) \\
& +\sum_{i=1}^{d}\sum_{l=1}^{M}\int_0^\tau (\partial_{z^i}\Psi)\,(X^\epsilon,Y^\epsilon,Z^\epsilon)(s)\left[(\partial_x G^{i,l})(X^\epsilon(s))Z^\epsilon(s)\right]dW_l^\epsilon(s) \\
& -\epsilon^{1/2}\sum_{i=1}^{d}\sum_{l=1}^{M}\int_0^\tau (\partial_{z^i}\Psi)(X^\epsilon,Y^\epsilon,Z^\epsilon)(s)I_2^{i,l}(\epsilon,X^\epsilon,Z^\epsilon)dW_l^\epsilon(s)
\end{aligned} \tag{A.1.38}$$

and we have put

$$
\begin{aligned}
\mathcal{G}^{\epsilon}\Psi(x,y,z) \;&\triangleq\; \epsilon^{-1}\mathcal{A}\Psi(x,y,z) \\
&+\; \epsilon^{-1/2}\sum_{i=1}^{d}\left(\partial_{z^i}\Psi\right)(x,y,z)\left[F^i(x,y)-\overline{F}^i(x)\right] \\
&+\; \left\{\sum_{i=1}^{d}\left(\partial_{x^i}\Psi\right)(x,y,z)F^i(x,y)\right. \\
&+\; \sum_{i=1}^{d}\sum_{j=1}^{d}\left(\partial_{z^i}\Psi\right)(x,y,z)z^j\partial_{x^j}\overline{F}^i(x) \\
&+\; \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{x^j}\Psi\right)(x,y,z)G^{i,l}(x)G^{j,l}(x) \\
&+\; \sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{z^j}\Psi\right)(x,y,z)G^{i,l}(x)\left[(\partial_x G^{j,l})(x)z\right] \\
&+\; \left.\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{z^i}\partial_{z^j}\Psi\right)(x,y,z)\left[(\partial_x G^{i,l})(x)z\right]\left[(\partial_x G^{j,l})(x)z\right]\right\} \\
&-\; \epsilon^{1/2}\left\{\sum_{i=1}^{d}\left(\partial_{z^i}\Psi\right)(x,y,z)\Gamma_1^i(\epsilon,x,z)\right. \\
&+\; \sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{z^j}\Psi\right)(x,y,z)G^{i,l}(x)\Gamma_2^{j,l}(\epsilon,x,z) \\
&+\; \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{z^i}\partial_{z^j}\Psi\right)(x,y,z)\left[(\partial_x G^{i,l})(x)z\right]\Gamma_2^{j,l}(\epsilon,x,z) \\
&+\; \left.\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{z^i}\partial_{z^j}\Psi\right)(x,y,z)\left[(\partial_x G^{j,l})(x)z\right]\Gamma_2^{i,l}(\epsilon,x,z)\right\} \\
&+\; \frac{\epsilon}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{z^i}\partial_{z^j}\Psi\right)(x,y,z)\Gamma_2^{i,l}(\epsilon,x,z)\Gamma_2^{j,l}(\epsilon,x,z) \qquad \text{(A.1.39)}
\end{aligned}
$$

Next, we introduce the class of perturbed near-identity functions: for the fixed function $(x,z)\to g(x,z)\in C_c^{\infty}(\mathbb{R}^d\times\mathbb{R}^d)$ and for each $\epsilon\in(0,1]$, define the function

$(x, y, z) \to \Psi_g^\epsilon(x, y, z) : \mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d \to \mathbb{R}$ by

$$\Psi_g^\epsilon(x, y, z) \triangleq g(x, z) + \epsilon^{1/2}\Xi_1(x, y, z) + \epsilon\Xi_2(x, y, z) + \epsilon\Xi_3(x, y, z), \tag{A.1.40}$$

$\forall (x, y, z) \in \mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d$, where

$$\begin{aligned} \Xi_1(x, y, z) &\triangleq \sum_{i=1}^{d} \Phi^i(x, y) \, (\partial_{z^i} g) \, (x, z), \\ \Xi_2(x, y, z) &\triangleq \sum_{i=1}^{d} \Phi^i(x, y) \, (\partial_{x^i} g) \, (x, z), \\ \Xi_3(x, y, z) &\triangleq \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \Psi^{i,j}(x, y) \, (\partial_{z^i} \partial_{z^j} g) \, (x, z), \end{aligned} \tag{A.1.41}$$

and the mappings $\Phi^i(x, y)$ and $\Psi^{i,j}(x, y)$ are defined in Conditions 3.2.8 and 3.2.15 respectively.

We have defined the functions Ξ_1, Ξ_2 and Ξ_3 in (A.1.41) such that, upon taking $\Psi \triangleq \Psi_g^\epsilon$ in (A.1.39), we will be able to use

(a) the fact that $\mathcal{A}g(x, z) \equiv 0$ (by (3.2.14) to remove the term involving ϵ^{-1},

(b) the relation (3.2.19) in Condition 3.2.8 to cancel the terms involving $\epsilon^{-1/2}$,

(c) the equations (3.2.28), (3.2.19) and (3.2.36) to get the terms of power ϵ^0 independent of y.

To this end, we first evaluate each term on the right side of (A.1.39) with $\Psi \triangleq \Psi_g^\epsilon$. Before doing this, notice from (A.1.41) and the fact that $\Phi^i, \Psi^{i,j} \in C^{2,2}(\mathbb{R}^d \otimes \mathbb{R}^D)$ (recall Remark 3.2.7), that Ψ_g^ϵ has all of the derivatives required by Remark A.1.3, so that $\mathcal{G}^\epsilon \Psi_g^\epsilon$ is indeed defined. From Remark 3.2.6, Condition 3.2.8, Condition 3.2.15, and (A.1.41) we see that:

$$\mathcal{A}g(x, y, z) \equiv 0, \tag{A.1.42}$$

and

$$
\begin{aligned}
&\epsilon^{-1}\mathcal{A}\Psi_g^\epsilon(x,y,z) \\
&\quad = \epsilon^{-1/2}\left\{\sum_{i=1}^{d}\left[\overline{F}^i(x) - F^i(x,y)\right](\partial_{z^i} g)(x,z)\right\} \\
&\quad + \sum_{i=1}^{d}\left[\overline{F}^i(x) - F^i(x,y)\right](\partial_{x^i} g)(x,z) \\
&\quad + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\left[\overline{a}^{i,j}(x) - a^{i,j}(x,y)\right](\partial_{z^i}\partial_{z^j} g)(x,z). \qquad \text{(A.1.43)}
\end{aligned}
$$

Also, from (A.1.40) and (A.1.41):

$$
\begin{aligned}
&\epsilon^{-1/2}\sum_{i=1}^{d}\left(\partial_{z^i}\Psi_g^\epsilon\right)(x,y,z)\left[F^i(x,y) - \overline{F}^i(x)\right] \\
&\quad = \epsilon^{-1/2}\sum_{i=1}^{d}\left[F^i(x,y) - \overline{F}^i(x)\right](\partial_{z^i} g)(x,z) \\
&\quad + \sum_{i=1}^{d}\sum_{j=1}^{d}\Phi^j(x,y)\left[F^i(x,y) - \overline{F}^i(x)\right](\partial_{z^i}\partial_{z^j} g)(x,z) \\
&\quad + \epsilon^{1/2}\sum_{i=1}^{d}(\partial_{z^i}\Xi_2)(x,y,z)\left[F^i(x,y) - \overline{F}^i(x)\right] \\
&\quad + \epsilon^{1/2}\sum_{i=1}^{d}(\partial_{z^i}\Xi_3)(x,y,z)\left[F^i(x,y) - \overline{F}^i(x)\right]. \qquad \text{(A.1.44)}
\end{aligned}
$$

From (A.1.40) and (A.1.41):

$$\begin{aligned}\sum_{i=1}^{d} \left(\partial_{x^i}\Psi_g^\epsilon\right)(x,y,z)F^i(x,y) &= \sum_{i=1}^{d} \left(\partial_{x^i}g\right)(x,z)F^i(x,y) \\ &+\epsilon^{1/2}\sum_{i=1}^{d} \left(\partial_{x^i}\Xi_1\right)(x,y,z)F^i(x,y) \\ &+\epsilon\sum_{i=1}^{d} \left(\partial_{x^i}\Xi_2\right)(x,y,z)F^i(x,y) \\ &+\epsilon\sum_{i=1}^{d} \left(\partial_{x^i}\Xi_3\right)(x,y,z)F^i(x,y) \end{aligned} \quad (A.1.45)$$

From (A.1.40) and (A.1.41):

$$\begin{aligned}\sum_{i=1}^{d}\sum_{j=1}^{d} \left(\partial_{z^i}\Psi_g^\epsilon\right)(x,y,z)z^j\left(\partial_{x^j}\overline{F}^i\right)(x) &= \sum_{i=1}^{d}\sum_{j=1}^{d} \left(\partial_{z^i}g\right)(x,z)z^j\left(\partial_{x^j}\overline{F}^i\right)(x) \\ &+\epsilon^{1/2}\sum_{i=1}^{d}\sum_{j=1}^{d} \left(\partial_{z^i}\Xi_1\right)(x,y,z)z^j\left(\partial_{x^j}\overline{F}^i\right)(x) \\ &+\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d} \left(\partial_{z^i}\Xi_2\right)(x,y,z)z^j\left(\partial_{x^j}\overline{F}^i\right)(x) \\ &+\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d} \left(\partial_{z^i}\Xi_3\right)(x,y,z)z^j\left(\partial_{x^j}\overline{F}^i\right)(x) \end{aligned} \quad (A.1.46)$$

$$
\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{x^j}\Psi_g^{\epsilon}\right)(x,y,z)G^{i,l}(x)G^{j,l}(x)\\
&\qquad=\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{x^j}g\right)(x,z)G^{i,l}(x)G^{j,l}(x)\\
&\qquad+\frac{1}{2}\epsilon^{1/2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{x^j}\Xi_1\right)(x,y,z)G^{i,l}(x)G^{j,l}(x)\\
&\qquad+\frac{1}{2}\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{x^j}\Xi_2\right)(x,y,z)G^{i,l}(x)G^{j,l}(x)\\
&\qquad+\frac{1}{2}\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{x^j}\Xi_3\right)(x,y,z)G^{i,l}(x)G^{j,l}(x) \qquad \text{(A.1.47)}
\end{aligned}
$$

From (A.1.40) and (A.1.41):

$$
\begin{aligned}
&\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{z^j}\Psi_g^{\epsilon}\right)(x,y,z)G^{i,l}(x)\left[(\partial_x G^{j,l})(x)z\right]\\
&\qquad=\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{z^j}g\right)(x,z)G^{i,l}(x)\left[(\partial_x G^{j,l})(x)z\right]\\
&\qquad+\epsilon^{1/2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{z^j}\Xi_1\right)(x,y,z)G^{i,l}(x)\left[(\partial_x G^{j,l})(x)z\right]\\
&\qquad+\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{z^j}\Xi_2\right)(x,y,z)G^{i,l}(x)\left[(\partial_x G^{j,l})(x)z\right]\\
&\qquad+\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{z^j}\Xi_3\right)(x,y,z)G^{i,l}(x)\left[(\partial_x G^{j,l})(x)z\right] \quad \text{(A.1.48)}
\end{aligned}
$$

And, also from (A.1.40) and (A.1.41):

$$
\begin{aligned}
\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{z^i}\partial_{z^j}\Psi_g^{\epsilon}\right)(x,y,z)\left[(\partial_x G^{i,l})(x)z\right]\left[(\partial_x G^{j,l})(x)z\right]& \\
= \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{z^i}\partial_{z^j}g\right)(x,z)\left[(\partial_x G^{i,l})(x)z\right]\left[(\partial_x G^{j,l})(x)z\right]& \\
+\frac{1}{2}\epsilon^{1/2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{z^i}\partial_{z^j}\Xi_1\right)(x,y,z)\left[(\partial_x G^{i,l})(x)z\right]\left[(\partial_x G^{j,l})(x)z\right]& \\
+\frac{1}{2}\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{z^i}\partial_{z^j}\Xi_2\right)(x,y,z)\left[(\partial_x G^{i,l})(x)z\right]\left[(\partial_x G^{j,l})(x)z\right]& \\
+\frac{1}{2}\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{z^i}\partial_{z^j}\Xi_3\right)(x,y,z)\left[(\partial_x G^{i,l})(x)z\right]\left[(\partial_x G^{j,l})(x)z\right]&
\end{aligned}
\tag{A.1.49}
$$

We substitute (A.1.43) to (A.1.49) in (A.1.39) and simplify to get (A.1.50) which follows. Note that, in (A.1.50), we collect all terms multiplied by $\epsilon^{k/2}$ into the term $\Delta_k(\epsilon, x, y, z)$, for each $k = 1, 2, 3, 4$.

$$\begin{aligned}
\mathcal{G}^\epsilon \Psi_g^\epsilon(x, y, z) \;=\; & \left\{ \sum_{i=1}^{d} \left[\overline{F}^i(x) - F^i(x, y)\right] (\partial_{x^i} g)(x, z) \right. \\
& + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \left[\overline{a}^{ij}(x) - a^{ij}(x, y)\right] (\partial_{z^i} \partial_{z^j} g)(x, z) \\
& + \sum_{i=1}^{d} \sum_{j=1}^{d} \Phi^j(x, y) \left[F^i(x, y) - \overline{F}^i(x)\right] (\partial_{z^i} \partial_{z^j} g)(x, z) \\
& + \sum_{i=1}^{d} (\partial_{x^i} g)(x, z) F^i(x, y) \\
& + \sum_{i=1}^{d} \sum_{j=1}^{d} (\partial_{z^i} g)(x, z) z^j \left(\partial_{x^j} \overline{F}^i\right)(x) \\
& + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \sum_{l=1}^{M} (\partial_{x^i} \partial_{x^j} g)(x, z) G^{i,l}(x) G^{j,l}(x) \\
& + \sum_{i=1}^{d} \sum_{j=1}^{d} \sum_{l=1}^{M} (\partial_{x^i} \partial_{z^j} g)(x, z) G^{i,l}(x) \left[(\partial_x G^{j,l})(x) z\right] \\
& \left. + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \sum_{l=1}^{M} (\partial_{z^i} \partial_{z^j} g)(x, z) \left[(\partial_x G^{i,l})(x) z\right] \left[(\partial_x G^{j,l})(x) z\right] \right\} \\
& + \epsilon^{1/2} \Delta_1(\epsilon, x, y, z) + \epsilon \Delta_2(\epsilon, x, y, z) + \epsilon^{3/2} \Delta_3(\epsilon, x, y, z) \\
& \qquad + \epsilon^2 \Delta_4(\epsilon, x, y, z)
\end{aligned} \tag{A.1.50}$$

Moreover, from (3.2.28):

$$\sum_{i=1}^{d}\sum_{j=1}^{d}\Phi^j(x,y)\left[F^i(x,y)-\overline{F}^i(x)\right](\partial_{z^i}\partial_{z^j}g)(x,z)$$
$$=\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\Phi^j(x,y)\left[F^i(x,y)-\overline{F}^i(x)\right](\partial_{z^i}\partial_{z^j}g)(x,z)$$
$$+\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\Phi^i(x,y)\left[F^j(x,y)-\overline{F}^j(x)\right](\partial_{z^j}\partial_{z^i}g)(x,z)$$
$$=\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}a^{i,j}(x,y)(\partial_{z^i}\partial_{z^j}g)(x,z). \tag{A.1.51}$$

Then, substituting (A.1.51) in (A.1.50), and simplifing, we get

$$\begin{aligned}\mathcal{G}^\epsilon\Psi_g^\epsilon(x,y,z) &= \sum_{i=1}^{d}\overline{F}^i(x)\left(\partial_{z^i}g\right)(x,z)+\sum_{i=1}^{d}\sum_{j=1}^{d}z^j\left(\partial_{x^j}\overline{F}^i\right)(x)\left(\partial_{z^i}g\right)(x,z)\\ &+ \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}G^{i,l}(x)G^{j,l}(x)\left(\partial_{x^i}\partial_{x^j}g\right)(x,z)\\ &+ \sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{l=1}^{M}\left(\partial_{x^i}\partial_{z^j}g\right)(x,z)G^{i,l}(x)\left[(\partial_x G^{j,l})(x)z\right]\\ &+ \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\Big\{\Big(\overline{a}^{i,j}(x)\\ &\qquad + \sum_{l=1}^{M}\left[(\partial_x G^{i,l})(x)z\right]\left[(\partial_x G^{j,l})(x)z\right]\Big)\left(\partial_{z^i}\partial_{z^j}g\right)(x,z)\Big\}\\ &+ \epsilon^{1/2}\Delta_1(\epsilon,x,y,z)+\epsilon\Delta_2(\epsilon,x,y,z)+\epsilon^{3/2}\Delta_3(\epsilon,x,y,z)\\ &\qquad +\epsilon^2\Delta_4(\epsilon,x,y,z).\end{aligned} \tag{A.1.52}$$

Thus, in view of (A.1.4) and (A.1.52), we establish

$$\mathcal{G}^\epsilon\Psi_g^\epsilon(x,y,z)=\mathcal{C}g(x,z)+\sum_{K=1}^{4}\epsilon^{K/2}\Delta_K(\epsilon,x,y,z), \tag{A.1.53}$$

(compare (A.1.40) with (A.1.5), and (A.1.53) with (A.1.6)). In this equation we have not given the explicit forms for the functions $\Delta_K(\epsilon, x, y, z)$(in terms of $\Phi(x, y)$, $\Psi(x, y)$, $F(x, y)$, etc.) since these are not important(and quite lengthy). Indeed, the only significant thing about these functions is that $(x, y, z) \to \Delta_K(\epsilon, x, y, z)$ is continuous on $(\mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d)$ for each $\epsilon \in (0, 1]$, and, corresponding to each $R \in [0, \infty)$, there is a constant $C_1(R) \in [0, \infty)$ such that, for each $K = 1, 2, 3, 4$, one has

$$|\Delta_K(\epsilon, x, y, z)| \leq C_1(R)[1 + |y|^{1+(q_2 \vee q_3)}], \qquad \text{(A.1.54)}$$

$\forall \epsilon \in (0, 1]$, and, $\forall (x, y, z) \in S_R^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d$. This follows easily by checking all terms for the individual $\Delta_K(x, y, z)$, and using the compact support of $g(x, z)$ and its partial derivatives, together with Condition 3.2.8 (which ensures that $\Phi^i(x, y)$, $\partial_{x^l}\Phi^i(x, y)$, and $\partial_{x^l}\partial_{x^k}\Phi^i(x, y)$ are polynomially y-bounded of order q_2 locally in x), and Condition 3.2.15 (which ensures that $\Psi^{i,j}(x, y)$, $\partial_{x^l}\Psi^{i,j}(x, y)$, and $\partial_{x^l}\partial_{x^k}\Psi^{i,j}(x, y)$ are polynomially y-bounded of order q_3 locally in x), and Condition 3.2.1 (which ensures that $F^i(x, y)$ is polynomially y-bounded of unit order, locally in x).

Now let $M_g^\epsilon(\tau)$ denote $M_\Psi^\epsilon(\tau)$ in (A.1.38) when $\Psi \triangleq \Psi_g^\epsilon$, and combine (A.1.37), (A.1.40), (A.1.41), (A.1.53) to get

$$\begin{aligned}
M_g^\epsilon(\tau) &= g(X^\epsilon, Z^\epsilon)(\tau) + \epsilon^{1/2}\sum_{i=1}^{d}\Phi^i(X^\epsilon, Y^\epsilon)(\tau)\,(\partial_{z^i}g)\,(X^\epsilon, Z^\epsilon)(\tau) \\
&\quad +\epsilon\sum_{i=1}^{d}\Phi^i(X^\epsilon, Y^\epsilon)(\tau)\,(\partial_{x^i}g)\,(X^\epsilon, Z^\epsilon)(\tau) \\
&\quad +\frac{1}{2}\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\Psi^{i,j}(X^\epsilon, Y^\epsilon)(\tau)\,(\partial_{z^i}\partial_{z^j}g)\,(X^\epsilon, Z^\epsilon)(\tau) \\
&\quad -g(x_0, z_0) - \epsilon^{1/2}\sum_{i=1}^{d}\Phi^i(x_0, y_0)\,(\partial_{z^i}g)\,(x_0, z_0) \\
&\quad -\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\Phi^i(x_0, y_0)\,(\partial_{x^i}g)\,(x_0, z_0) \\
&\quad - \frac{1}{2}\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\Psi^{i,j}(x_0, y_0)\,(\partial_{z^i}\partial_{z^j}g)\,(x_0, z_0) \\
&\quad -\int_0^\tau \mathcal{C}g(X^\epsilon, Z^\epsilon)(s)ds - \sum_{K=1}^{4}\int_0^\tau \epsilon^{K/2}\Delta_K(\epsilon, X^\epsilon, Y^\epsilon, Z^\epsilon)(s)ds.
\end{aligned} \tag{A.1.55}$$

Rearranging (A.1.55):

$$\begin{aligned}
M_g^\epsilon(\tau) &= g(X^\epsilon, Z^\epsilon)(\tau) - g(x_0, z_0) - \int_0^\tau \mathcal{C}g(X^\epsilon, Z^\epsilon)(s)ds \\
&\quad +\epsilon^{1/2}\sum_{i=1}^{d}\left\{\Phi^i(X^\epsilon, Y^\epsilon)(\tau)\,(\partial_{z^i}g)\,(X^\epsilon, Z^\epsilon)(\tau) - \Phi^i(x_0, y_0)\,(\partial_{z^i}g)\,(x_0, z_0)\right\} \\
&\quad +\epsilon\sum_{i=1}^{d}\left\{\Phi^i(X^\epsilon, Y^\epsilon)(\tau)\,(\partial_{x^i}g)\,(X^\epsilon, Z^\epsilon)(\tau) - \Phi^i(x_0, y_0)\,(\partial_{x^i}g)\,(x_0, z_0)\right\} \\
&\quad +\frac{1}{2}\epsilon\sum_{i=1}^{d}\sum_{j=1}^{d}\left\{\Psi^{i,j}(X^\epsilon, Y^\epsilon)\,(\partial_{z^i}\partial_{z^j}g)\,(X^\epsilon, Z^\epsilon)(\tau)\right. \\
&\qquad\qquad \left. - \Psi^{i,j}(x_0, y_0)\,(\partial_{z^i}\partial_{z^j}g)\,(x_0, z_0)\right\} \\
&\quad -\sum_{K=1}^{4}\epsilon^{K/2}\int_0^\tau \Delta_K(\epsilon, X^\epsilon, Y^\epsilon, Z^\epsilon)(s)ds, \qquad \forall\tau\in[0,1].
\end{aligned} \tag{A.1.56}$$

In view of (A.1.8) and (A.1.56) we have

$$\begin{aligned} M_g^\epsilon(\tau) &= \Lambda_g(X^\epsilon, Z^\epsilon)(\tau) \\ &\quad + \epsilon^{1/2} \sum_{i=1}^{d} \left\{ \Phi^i(X^\epsilon, Y^\epsilon)(\tau)\, (\partial_{z^i} g)\,(X^\epsilon, Z^\epsilon)(\tau) - \Phi^i(x_0, y_0)\,(\partial_{z^i} g)\,(x_0, z_0) \right\} \\ &\quad + \epsilon \sum_{i=1}^{d} \left\{ \Phi^i(X^\epsilon, Y^\epsilon)(\tau)\, (\partial_{x^i} g)\,(X^\epsilon, Z^\epsilon)(\tau) - \Phi^i(x_0, y_0)\,(\partial_{x^i} g)\,(x_0, z_0) \right\} \\ &\quad + \frac{1}{2}\epsilon \sum_{i=1}^{d}\sum_{j=1}^{d} \left\{ \Psi^{i,j}(X^\epsilon, Y^\epsilon)\,(\partial_{z^i}\partial_{z^j} g)\,(X^\epsilon, Z^\epsilon)(\tau) \right. \\ &\qquad\qquad \left. - \Psi^{i,j}(x_0, y_0)\,(\partial_{z^i}\partial_{z^j} g)\,(x_0, z_0) \right\} \\ &\quad - \sum_{K=1}^{4} \epsilon^{K/2} \int_0^\tau \Delta_K(\epsilon, X^\epsilon, Y^\epsilon, Z^\epsilon)(s)\,ds, \qquad \forall \tau \in [0,1]. \end{aligned} \tag{A.1.57}$$

In order to establish (A.1.20) from (A.1.57), which will follow in Step III, we now upper-bound the expectations of terms on the right hand side of $M_g^\epsilon(\tau \wedge T_R^\epsilon)$ from (A.1.57) starting with the last term. From (A.0.2) we have $X^\epsilon(s) \in S_R^d$, $\forall 0 \le s \le T_R^\epsilon$, thus, from (A.1.54), and $q_4 > (1 + q_2 \vee q_3)$ (see Condition 3.2.18) we get

$$\begin{aligned} &|\Delta_K(\epsilon, X^\epsilon, Y^\epsilon, Z^\epsilon)(s)| I\{s \le T_R^\epsilon\} \\ &\qquad \le C_1(R)\left[1 + |Y^\epsilon(s)|^{(1+q_2 \vee q_3)}\right] I\{s \le T_R^\epsilon\} \\ &\qquad \le C_1(R)\left[1 + |Y^\epsilon(s)|^{q_4}\right] I\{s \le T_R^\epsilon\}. \end{aligned} \tag{A.1.58}$$

Taking expectation in (A.1.58) and using Lemma A.4.7 on page 123 we get a constant $C_2(R) \in [0, \infty)$ such that

$$E\left[|\Delta_K(\epsilon, X^\epsilon, Y^\epsilon, Z^\epsilon)(s)| I\{s \le T_R^\epsilon\}\right] \le C_2(R). \tag{A.1.59}$$

$\forall \epsilon \in (0,1]$, $\forall s \in [0,1]$, $K = 1,2,3,4$, and thus by Fubini's theorem one has

$$\begin{aligned} E\int_0^{\tau \wedge T_R^\epsilon} |\Delta_K(\epsilon, X^\epsilon, Y^\epsilon, Z^\epsilon)(s)|ds &= \int_0^\tau E[|\Delta_K(\epsilon, X^\epsilon, Y^\epsilon, Z^\epsilon)(s)| I\{s \le T_R^\epsilon\}]ds \\ &\le C_2(R) \end{aligned} \tag{A.1.60}$$

$\forall \epsilon \in (0,1]$, $\forall \tau \in [0,1]$, $K = 1,2,3,4$.

As for the second term on the right side of (A.1.57), fix some $R \in (R_0, \infty)$ where R_0 is specified at (A.1.20). Then we have $|X^\epsilon(T_R^\epsilon)| = R$, hence on the event $\{T_R^\epsilon < \tau\}$, $(X^\epsilon(\tau \wedge T_R^\epsilon), Z^\epsilon(\tau \wedge T_R^\epsilon)) \notin$ support $\{(\partial_{z^i} g)(.,.)\}$, and thus

$$\left\{\Phi^i(X^\epsilon(\tau \wedge T_R^\epsilon), Y^\epsilon(\tau \wedge T_R^\epsilon))\,(\partial_{z^i} g)\,(X^\epsilon(\tau \wedge T_R^\epsilon), Z^\epsilon(\tau \wedge T_R^\epsilon))\right\} I\left\{T_R^\epsilon < \tau\right\} = 0. \tag{A.1.61}$$

$\forall \tau \in [0,1]$, $\forall \epsilon \in (0,1]$. In view of Condition 3.2.8, which ensures that $\Phi^i(x,y)$ is polynomially y-bounded of order q_2 locally in x, and also using the fact that $(\partial_{z^i} g)(x,z)$ is uniformly bounded, there is a constant $C_3(R) \in [0,\infty)$ such that

$$\begin{aligned}|\Phi^i(X^\epsilon(\tau \wedge T_R^\epsilon), Y^\epsilon(\tau \wedge T_R^\epsilon))\,(\partial_{z^i} g)\,(X^\epsilon(\tau \wedge T_R^\epsilon), Z^\epsilon(\tau \wedge T_R^\epsilon))| I\left\{\tau \le T_R^\epsilon\right\}\\ \le C_3(R)[1 + |Y^\epsilon(\tau \wedge T_R^\epsilon)|^{q_2}] I\left\{\tau \le T_R^\epsilon\right\}.\end{aligned} \tag{A.1.62}$$

Now, from (A.1.61) and (A.1.62), we can write

$$\begin{aligned}|\Phi^i(X^\epsilon(\tau \wedge T_R^\epsilon), Y^\epsilon(\tau \wedge T_R^\epsilon))\,(\partial_{z^i} g)\,(X^\epsilon(\tau \wedge T_R^\epsilon), Z^\epsilon(\tau \wedge T_R^\epsilon))|\\ \le C_3(R)[1 + |Y^\epsilon(\tau \wedge T_R^\epsilon)|^{q_2}] I\left\{\tau \le T_R^\epsilon\right\},\end{aligned} \tag{A.1.63}$$

Thus, taking expectation in (A.1.63) and using Lemma A.4.7 on page 123 there is a constant $C_4(R) \in [0,\infty)$ such that

$$\begin{aligned}E|\Phi^i(X^\epsilon(\tau \wedge T_R^\epsilon), Y^\epsilon(\tau \wedge T_R^\epsilon))\,(\partial_{z^i} g)\,(X^\epsilon(\tau \wedge T_R^\epsilon), Z^\epsilon(\tau \wedge T_R^\epsilon))|\\ \le E\left(C_3(R)\,[1 + |Y^\epsilon(\tau \wedge T_R^\epsilon)|^{q_2}]\, I\left\{\tau \le T_R^\epsilon\right\}\right)\\ \le C_4(R),\end{aligned} \tag{A.1.64}$$

for all $\epsilon \in (0,1]$, $\tau \in [0,1]$. Considering the third term on the right hand side of (A.1.57), a similar argument shows that there is a constant $C_5(R)$ such that

$$E|\Phi^i(X^\epsilon(\tau \wedge T_R^\epsilon), Y^\epsilon(\tau \wedge T_R^\epsilon))\,(\partial_{x^i} g)\,(X^\epsilon(\tau \wedge T_R^\epsilon), Z^\epsilon(\tau \wedge T_R^\epsilon))| \le C_5(R), \tag{A.1.65}$$

for all $\epsilon \in (0,1]$, $\tau \in [0,1]$. For the fourth term on the right hand side of (A.1.57), an argument which is similar to that used for (A.1.64) together with (3.2.35) establishes; for each $R \in [0,\infty)$, there is some constant $C_6(R)$ such that

$$E|\Psi^{i,j}(X^\epsilon(\tau \wedge T_R^\epsilon), Y^\epsilon(\tau \wedge T_R^\epsilon))\,(\partial_{z^i}\partial_{z^j} g)\,(X^\epsilon(\tau \wedge T_R^\epsilon), Z^\epsilon(\tau \wedge T_R^\epsilon))| \leq C_6(R), \tag{A.1.66}$$

for all $\epsilon \in (0,1]$, $\tau \in [0,1]$. Let $J_g^\epsilon(\tau)$ denote the sum of all terms on the right side of (A.1.57), except for the first term $\Lambda_g(X^\epsilon, Z^\epsilon)(\tau)$, thus

$$M_g^\epsilon(\tau) = \Lambda_g(X^\epsilon, Z^\epsilon)(\tau) + J_g^\epsilon(\tau), \tag{A.1.67}$$

for all $\epsilon \in (0,1], \tau \in [0,1]$. From (A.1.60), (A.1.64), (A.1.65) and (A.1.66) we have seen that for each $R \in (R_0, \infty)$ there is a constant $C_7(R) \in [0,\infty)$ such that

$$E|J_g^\epsilon(\tau \wedge T_R^\epsilon)| \leq \epsilon^{1/2} C_7(R), \tag{A.1.68}$$

for all $\epsilon \in (0,1], \tau \in [0,1]$.

STEP III: In this step we show that $\left\{\left(M_g^\epsilon(\tau \wedge T_R^\epsilon), \mathcal{G}_\tau^\epsilon\right), \tau \in [0,1]\right\}$ is a martingale on $(\Omega, \mathcal{F}, P)$ for each $\epsilon \in (0,1]$, $R \in (0,\infty)$ and $g \in C_c^\infty(I\!R^d \otimes I\!R^d)$, recalling that $M_g^\epsilon(\tau)$ is given by (A.1.38) with $\Psi \triangleq \Psi_g^\epsilon$ from (A.1.40). We do this by showing that the stochastic integrals on the right side of (A.1.38) (which are clearly $\{\mathcal{G}_\tau^\epsilon\}$-local martingales) in fact become genuine $\{\mathcal{G}_\tau^\epsilon\}$-martingales when stopped at T_R^ϵ. Beginning with the second term on the right hand side of (A.1.38), take y^l-derivatives of Ψ_g^ϵ in (A.1.40) and use (A.1.41) to get

$$\begin{aligned} \partial_{y^l}\Psi_g^\epsilon(x,y,z) &= \epsilon^{1/2}\sum_{i=1}^d \left(\partial_{y^l}\Phi^i\right)(x,y)\,(\partial_{z^i} g)\,(x,z) \\ &+ \epsilon\sum_{i=1}^d \left(\partial_{y^l}\Phi^i\right)(x,y)\,(\partial_{z^i} g)\,(x,z) \\ &+ \frac{1}{2}\epsilon\sum_{i=1}^d\sum_{j=1}^d \left(\partial_{y^l}\Psi^{i,j}\right)(x,y)\,(\partial_{z^i}\partial_{z^j} g)\,(x,z). \end{aligned} \tag{A.1.69}$$

Since $\partial_{z^i} g(x,z)$ and $\partial_{x^i} g(x,z)$ (recall $g \in C_c^\infty(\mathbb{R}^d \otimes \mathbb{R}^d)$ and $\partial_{z^i}\partial_{z^i} g(x,z)$ are uniformly bounded, it follows from (A.1.69) and Conditions 3.2.8(see (3.2.17)) and 3.2.15 that, for the fixed $R \in (0,\infty)$, there is a constant $C_8(R) \in [0,\infty)$ such that

$$| \left(\partial_{y^l} \Psi_g^\epsilon\right)(x,y,z)| \le \epsilon^{1/2} C_8(R)\left[1 + |y|^{(q_2 \vee q_3)}\right], \tag{A.1.70}$$

$\forall \epsilon \in (0,1]$, and, $\forall (x,y,z) \in S_R^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d$. Thus, from (A.1.70) and Condition 3.2.1, there is a constant $C_9(R) \in [0,\infty)$ such that

$$| \left(\partial_{y^l} \Psi_g^\epsilon\right)(x,y,z)\sigma^{l,k}(x,y)|^2 \le \epsilon C_9(R)\left[1 + |y|^{2(1+q_2 \vee q_3)}\right], \tag{A.1.71}$$

$\forall \epsilon \in (0,1]$, and, $\forall (x,y,z) \in S_R^d \otimes \mathbb{R}^D \otimes \mathbb{R}^d$. Then from (A.1.71) we can write

$$\begin{aligned} &| \left(\partial_{y^l} \Psi_g^\epsilon\right)(X^\epsilon(s),Y^\epsilon(s),Z^\epsilon(s))\sigma^{l,k}(X^\epsilon(s),Y^\epsilon(s))|^2 I\left\{s \le T_R^\epsilon\right\} \\ &\qquad \le \epsilon C_8(R)\left[1 + |Y^\epsilon(s)|^{2(1+q_2 \vee q_3)}\right] I\left\{s \le T_R^\epsilon\right\}, \end{aligned} \tag{A.1.72}$$

$\forall \epsilon \in (0,1]$, $\forall s \in [0,1]$. Taking expectation in (A.1.72), applying Lemma A.4.7 on page 123 and recalling that $q_4 > 2(1 + q_2 \vee q_3)$ there is a constant $C_9(R) \in [0,\infty)$ such that

$$\begin{aligned} &E\left(| \left(\partial_{y^l} \Psi_g^\epsilon\right)(X^\epsilon(s),Y^\epsilon(s),Z^\epsilon(s))\sigma^{l,k}(X^\epsilon(s),Y^\epsilon(s))|^2 I\left\{s \le T_R^\epsilon\right\}\right) \\ &\qquad \le E\left(\epsilon C_9(R)\left[1 + |Y^\epsilon(s)|^{2(1+q_2 \vee q_3)}\right] I\left\{s \le T_R^\epsilon\right\}\right) \\ &\qquad \le \epsilon C_{10}(R), \end{aligned} \tag{A.1.73}$$

for all $s \in [0,1]$. Now from (A.1.73), Fubini's theorem, the Itô isometry, and the fact that $q_4 > 2(1 + q_2 \vee q_3)$ (see Condition 3.2.18) it follows that

$$
\begin{aligned}
&E \left| \int_0^{\tau \wedge T_R^\epsilon} \left(\partial_{y^i} \Psi_g^\epsilon\right) (X^\epsilon(s), Y^\epsilon(s), Z^\epsilon(s)) \sigma^{i,k}(X^\epsilon(s), Y^\epsilon(s)) dB_k^\epsilon(s) \right|^2 \\
&= E \left| \int_0^{\tau} \partial_{y^i} \Psi_g^\epsilon(X^\epsilon(s), Y^\epsilon(s), Z^\epsilon(s)) \sigma^{i,k}(X^\epsilon(s), Y^\epsilon(s)) I\left\{s \le T_R^\epsilon\right\} dB_k^\epsilon(s) \right|^2 \\
&= E \int_0^{\tau} |\partial_{y^i} \Psi_g^\epsilon(X^\epsilon(s), Y^\epsilon(s), Z^\epsilon(s)) \sigma^{i,k}(X^\epsilon(s), Y^\epsilon(s))|^2 I\left\{s \le T_R^\epsilon\right\} d(s) \\
&= \int_0^{\tau} E\left(|\partial_{y^i} \Psi_g^\epsilon(X^\epsilon(s), Y^\epsilon(s), Z^\epsilon(s)) \sigma^{i,k}(X^\epsilon(s), Y^\epsilon(s))|^2 I\left\{s \le T_R^\epsilon\right\}\right) d(s) \\
&\le \int_0^{\tau} \epsilon C_{10}(R) d(s) \\
&\le \epsilon C_{11}(R), \qquad \forall \tau \in [0,1],
\end{aligned} \tag{A.1.74}
$$

where $C_{11}(R) \in [0, \infty]$ is a constant. Thus, from (A.1.74), it follows that

$$
H(\tau) \triangleq \int_0^{\tau \wedge T_R^\epsilon} \left(\partial_{y^i} \Psi_g^\epsilon\right) (X^\epsilon(s), Y^\epsilon(s), Z^\epsilon(s)) \sigma^{i,k}(X^\epsilon(s), Y^\epsilon(s)) dB_k^\epsilon(s) \tag{A.1.75}
$$

is an L_2-continuous martingale with respect to $\{\mathcal{G}_\tau^\epsilon,\ \tau \in [0,1]\}$. Similar analysis for all terms in the right hand side of (A.1.38) shows that $\{(M_g^\epsilon(\tau \wedge T_R^\epsilon), \mathcal{G}_\tau^\epsilon),\ \tau \in [0,1]\}$ is a martingale on $(\Omega, \mathcal{F}, P)$, for each $\epsilon \in (0,1]$, $R \in (0, \infty)$, and $g \in C_c^\infty(\mathbb{R}^d \otimes \mathbb{R}^d)$.

STEP IV: In this step, we show that (A.1.20) holds, which establishes the result (see Remark A.1.2). Since $\Gamma(X^\epsilon, Z^\epsilon)$ is uniformly bounded and $\mathcal{G}_{\tau_1}^\epsilon$-measurable (recall (A.1.15)), it follows from Step III that

$$
\begin{aligned}
&E\left[\left(M_g^\epsilon(\tau_2 \wedge T_R^\epsilon) - M_g^\epsilon(\tau_1 \wedge T_R^\epsilon)\right) \Gamma(X^\epsilon, Z^\epsilon)\right] \\
&\qquad = E\left[E\left[M_g^\epsilon(\tau_2 \wedge T_R^\epsilon) - M_g^\epsilon(\tau_1 \wedge T_R^\epsilon) | \mathcal{G}_{\tau_1}^\epsilon\right] \Gamma(X^\epsilon, Z^\epsilon)\right] \\
&\qquad = 0,
\end{aligned} \tag{A.1.76}
$$

for each $\epsilon \in (0,1]$ and $0 \le \tau_1 < \tau_2 \le 1$. From (A.1.76) and (A.1.67):

$$
\begin{aligned}
&E\left[\left\{\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)\right\} \Gamma(X^\epsilon Z^\epsilon)\right] \\
&\qquad = E\left[\left\{M_g^\epsilon(\tau_2 \wedge T_R^\epsilon) - M_g^\epsilon(\tau_1 \wedge T_R^\epsilon)\right\} \Gamma(X^\epsilon, Z^\epsilon)\right] \\
&\qquad\quad - E\left[\left\{J_g^\epsilon(\tau_2 \wedge T_R^\epsilon) - J_g^\epsilon(\tau_1 \wedge T_R^\epsilon)\right\} \Gamma(X^\epsilon, Z^\epsilon)\right] \\
&\qquad = -E\left[\left\{J_g^\epsilon(\tau_2 \wedge T_R^\epsilon) - J_g^\epsilon(\tau_1 \wedge T_R^\epsilon)\right\} \Gamma(X^\epsilon, Z^\epsilon)\right].
\end{aligned} \tag{A.1.77}
$$

Moreover, from (A.1.68), and the uniform boundedness of $\Gamma(X^\epsilon, Z^\epsilon)$ we see that

$$\lim_{\epsilon\to 0} E\left[\left\{J_g^\epsilon(\tau_2 \wedge T_R^\epsilon) - J_g^\epsilon(\tau_1 \wedge T_R^\epsilon)\right\}\Gamma(X^\epsilon, Z^\epsilon)\right] = 0, \qquad \text{(A.1.78)}$$

for each $R \in (R_0, \infty)$. Now from (A.1.77) and (A.1.78) we get

$$\lim_{\epsilon\to 0} E\left[(\Lambda_g(X^\epsilon, Z^\epsilon)(\tau_2 \wedge T_R^\epsilon) - \Lambda_g(X^\epsilon, Z^\epsilon)(\tau_1 \wedge T_R^\epsilon)\Gamma(X^\epsilon, Z^\epsilon)\right] = 0, \qquad \text{(A.1.79)}$$

as required.

□

A.2 Semi-definiteness

Proof of Proposition 3.2.13 on page 37: Fix $x \in \mathbb{R}^d$, and fix $i, j \in \{1, 2, \ldots, d\}$ throughout the proof. Observe from (3.2.19), (3.2.27), (3.2.28) and (3.2.29),

$$\begin{aligned}\overline{a}^{i,j}(x) &= \int_{\mathbb{R}^D} a^{i,j}(x,y)\pi_x(dy) \\ &= -\int_{\mathbb{R}^D} \Phi^i(x,y)\mathcal{A}\Phi^j(x,y)\pi_x(dy) \\ &\quad -\int_{\mathbb{R}^D} \Phi^j(x,y)\mathcal{A}\Phi^i(x,y)\pi_x(dy)\end{aligned} \qquad \text{(A.2.80)}$$

We now study the terms on the right hand side of (A.2.80). Since i and j can be interchanged we need to study only the first term. Define

$$\theta(t,x,y) \triangleq E\left[\Phi^j\left(x, \xi(t,x,y)\right)\right], \qquad \forall (t,y) \in [0,\infty) \otimes \mathbb{R}^D, \qquad \text{(A.2.81)}$$

where $\{\xi(t,x,y)\}$ is given by (3.1.3) with $\xi(0,x,y) \triangleq y$, namely

$$\xi(t,x,y) = y + \int_0^t b(x, \xi(\tau,x,y))d\tau + \int_0^t \sigma(x, \xi(\tau,x,y))d\beta(\tau), \qquad \text{(A.2.82)}$$

and $\Phi^j(x,\xi)$ is subject to Condition 3.2.8. Expanding $\Phi^j(x,\xi(t,x,y))$ using (3.2.14) and Itô's formula gives

$$\begin{aligned}\Phi^j(x,\xi(t,x,y)) &= \Phi^j(x,y) + \int_0^t \mathcal{A}\Phi^j(x,\xi(s,x,y))\,ds \\ &+ \sum_{n=1}^{N}\sum_{k=1}^{D}\int_0^t \partial_{\xi^k}\Phi^j(x,\xi(s,x,y))\sigma^{k,n}(x,\xi(s,x,y))d\beta^n(s).\end{aligned} \tag{A.2.83}$$

Now fix $R \in [0,\infty)$ such that $x \in S_R^d$. From Condition 3.2.8, there is a constant $C_1(R) \in [0,\infty)$ such that

$$|\partial_{\xi^k}\Phi^j(x,\xi(s,x,y))| \le C_1(R)\left[1 + |\xi(s,x,y)|^{q_2}\right], \tag{A.2.84}$$

for each $(s,y) \in [0,\infty) \otimes \mathbb{R}^D$. Since (3.2.11) holds with $\sigma(\cdot,\cdot)$ in place of $F(\cdot,\cdot)$ (see Condition 3.2.1) there is a constant $C_2(R) \in [0,\infty)$ such that

$$|\sigma^{k,n}(x,\xi(s,x,y))| \le C_2(R)\left[1 + |\xi(s,x,y)|\right], \tag{A.2.85}$$

for each $(s,y) \in [0,\infty) \otimes \mathbb{R}^D$. From (A.2.84), (A.2.85) and using this fact in conjunction with (C.0.7) in Lemma C.0.6 on page 203, one checks that

$$\begin{aligned} E\int_0^t & |\partial_{\xi^k}\Phi^j(x,\xi(s,x,y))\sigma^{k,n}(x,\xi(s,x,y))|^2 ds \\ &\le C_3(R)\int_0^t \left[1 + E|\xi(s,x,y)|^{2(1+q_2)}\right] ds \\ &\le C_3(R)\cdot t\left[1 + E\left[\max_{0\le s\le t}|\xi(s,x,y)|^{2(1+q_2)}\right]\right] ds \\ &< \infty, \qquad \forall t \in [0,\infty), \end{aligned} \tag{A.2.86}$$

where $C_3(R) \in [0,\infty)$ is a constant. As a result, the stochastic integrals in (A.2.83) are L_2-martingales and null at the origin.

Now one easily sees from Conditions 3.2.1, 3.2.8, and (3.2.14), that $\mathcal{A}\Phi^j(x,y)$ is polynomially y-bounded of order $(2+q_2)$ locally in x. Since $x \in S_R^d$, there is a constant $C_3(R) \in [0,\infty)$ such that

$$|\mathcal{A}\Phi^j(x,\xi(s,x,y)\,| \quad \leq \quad C_3(R)\left[1+|\xi(s,x,y)|^{(2+q_2)}\right], \qquad \forall (s,y) \in [0,\infty) \otimes I\!R^D. \tag{A.2.87}$$

Hence, we can write

$$E \max_{s\in[0,t]} |\mathcal{A}\Phi^j(x,\xi(s,x,y)| \quad \leq \quad C_3(R)\left[1+E\max_{s\in[0,t]}|\xi(s,x,y)|^{(2+q_2)}\right], \tag{A.2.88}$$

$\forall (t,y) \in [0,\infty) \otimes I\!R^D$. Thus, by (C.0.7) in Lemma C.0.6 on page 203 and using (A.2.88), one sees that for each $T \in [0,\infty)$ there is a constant $C_5(R,T) \in [0,\infty)$ such that

$$E\left[\max_{0\leq s\leq T}|\mathcal{A}\Phi^j(x,\xi(s,x,y))|\right] \quad \leq \quad C_3(R)\left[1+C_5(R,T)\left[1+|y|^{(2+q_2)}\right]e^{C_5(R,T)T}\right] \quad < \quad \infty, \qquad \forall y \in I\!R^D, \tag{A.2.89}$$

thus

$$\int_0^t E|\mathcal{A}\Phi^j(x,\xi(s,x,y))|ds \quad \leq \quad t\cdot E\left[\max_{0\leq s\leq T}|\mathcal{A}\Phi^j(x,\xi(s,x,y))|\right] \quad < \quad \infty, \tag{A.2.90}$$

$\forall (t,y) \in [0,\infty) \otimes I\!R^D$. Thus taking expectation in (A.2.83) and using Fubini's theorem with (A.2.90) we have

$$E(\Phi^j(x,\xi(t,x,y)) = \Phi^j(x,y) + \int_0^t E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right]ds. \tag{A.2.91}$$

In view of (A.2.80) and (A.2.91):

$$\theta(t,x,y) = \Phi^j(x,y) + \int_0^t E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right]ds. \tag{A.2.92}$$

Now let $\{\bar{\xi}(t,x),\ t \in [0,\infty)\}$ be some $I\!R^D$-valued stationary Markov process defined by (3.1.3) with marginal distribution given by the invariant probability measure π_x of Condition 3.2.3. From (A.2.80) and the Markov property of $\{\xi(t,x,y)\}$ we have

$$E\left[\Phi^j(x,\bar{\xi}(t,x))|\bar{\xi}(0,x)\right] = \theta(t,x,\bar{\xi}(0,x)), \qquad a.s. \tag{A.2.93}$$

Thus, by (A.2.93) and the composition rule for conditional expectations,

$$\begin{aligned}
&E\left[\Phi^i(x,\bar{\xi}(0,x))\Phi^j(x,\bar{\xi}(t,x))\right] \\
&\quad = E\left[\Phi^i(x,\bar{\xi}(0,x))E\left[\Phi^j(x,\bar{\xi}(t,x))|\bar{\xi}(0,x)\right]\right] \\
&\quad = \int_{\mathbb{R}^d} \Phi^i(x,y)\theta(t,x,y)d\pi_x(y) \\
&\quad = \int_{\mathbb{R}^d} \Phi^i(x,y)E\left[\Phi^j(x,\xi(t,x,y))\right]d\pi_x(y). \qquad \text{(A.2.94)}
\end{aligned}$$

and so, from (A.2.91) and (A.2.94) we get

$$\begin{aligned}
&E\left\{\Phi^i(x,\bar{\xi}(0,x))\left[\Phi^j(x,\bar{\xi}(0,x)) - \Phi^j(x,\bar{\xi}(t,x))\right]\right\} \\
&\quad = E\left[\Phi^i(x,\bar{\xi}(0,x))\Phi^j(x,\bar{\xi}(0,x)) - \Phi^i(x,\bar{\xi}(0,x))\Phi^j(x,\bar{\xi}(t,x))\right] \\
&\quad = \int_{\mathbb{R}^d} \Phi^i(x,y)\Phi^j(x,y)d\pi_x(y) - \int_{\mathbb{R}^d} \Phi^i(x,y)E\left[\Phi^j(x,\xi(t,x,y))\right]d\pi_x(y) \\
&\quad = -\int_{\mathbb{R}^d} \Phi^i(x,y)\left\{E\left[\Phi^j(x,\xi(t,x,y))\right] - \Phi^j(x,y)\right\}d\pi_x(y) \\
&\quad = -\int_{\mathbb{R}^d} \Phi^i(x,y)\left\{\int_0^t E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right]ds\right\}d\pi_x(y). \qquad \text{(A.2.95)}
\end{aligned}$$

We now show that $s \to E\left[\mathcal{A}\Phi^j(\xi(s,x,y)\right], s \in [0,\infty)$, is continuous. Fix some sequence $\{s_n,\ n \in \mathbb{N}\} \subset [0,\infty)$ such that $s_n \to s$. Since $\xi(\cdot,x,y)$ is continuous and $\mathcal{A}\Phi^j(x,\cdot)$ is continuous, we have

$$\lim_{n\to\infty} \mathcal{A}\Phi^j(x,\xi(s_n,x,y)) = \mathcal{A}\Phi^j(x,\xi(x,\xi(s,x,y)), \qquad \text{(A.2.96)}$$

$\forall(x,y,\omega) \in \mathbb{R}^d \otimes \mathbb{R}^D \otimes \Omega$. Now, suppose that $0 \le s_n \le T < \infty, \forall n \in \mathbb{N}$, then, using (A.2.89), we get

$$\begin{aligned}
E|\mathcal{A}\Phi^j(x,\xi(s_n,x,y))| &\le E\left[\max_{s\in[0,T]} |\mathcal{A}\Phi^j(x,\xi(s,x,y))|\right] \\
&< \infty, \qquad \text{(A.2.97)}
\end{aligned}$$

$\forall(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Thus, using (A.2.97), (A.2.96) and the Dominated Convergence Theorem, it follows that

$$\lim_{n\to\infty} E\left[\mathcal{A}\Phi^j(x,\xi(s_n,x,y))\right] = E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right], \qquad \text{(A.2.98)}$$

$\forall(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$, which proves that $s \to E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right]$ is continuous over $[0,\infty)$. Hence fixing a sequence $\{t_n\}$ such that $0 < t_n \leq 1$ and $\lim_{n\to\infty} t_n = 0$, by the Fundamental Theorem of calculus we get

$$\begin{aligned}\lim_{n\to\infty} &\left\{\frac{1}{t_n}\int_0^{t_n} E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right] ds\right\} \\ &= E\left[\mathcal{A}\Phi^j(x,\xi(0,x,y))\right] \\ &= \mathcal{A}\Phi^j(x,y). \end{aligned} \tag{A.2.99}$$

Thus, clearly

$$\lim_{n\to\infty} \Phi^i(x,y)\left\{\frac{1}{t_n}\int_0^{t_n} E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right] ds\right\} = \Phi^i(x,y)\mathcal{A}\Phi^j(x,y), \tag{A.2.100}$$

for $\Phi^j(x,y)$ defined in Condition 3.2.8. But, from (A.2.89) and Condition 3.2.8(1), one sees that the mapping $(x,y) \to \Phi^i(x,y)\left\{\frac{1}{t_n}\int_0^{t_n} E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right] ds\right\}$ is polynomially y-bounded of order $(2+2q_2)$ locally in x; namely

$$\begin{aligned}&\left|\Phi^i(x,y)\left\{\frac{1}{t_n}\int_0^{t_n} E\left[\mathcal{A}\Phi^j(x,\xi(u,x,y))\right] du\right\}\right| \\ &\quad\leq \left|\Phi^i(x,y)\right|\left|E\left[\max_{0\leq s\leq 1}\left|\mathcal{A}\Phi^j(x,\xi(s,x,y))\right|\right]\right| \\ &\quad\leq C_6(R)\left[1+|y|^{(2+2q_2)}\right], \qquad \forall y \in \mathbb{R}^D, \quad n = 1,2,3,\dots, \end{aligned} \tag{A.2.101}$$

for some constant $C_6(R) \in [0,\infty)$. From (3.2.20) the dominating function $y \to C_6(R)\left[1+|y|^{(2+2q_2)}\right] : \mathbb{R}^D \to \mathbb{R}$ in (A.2.101) is π_x-integrable. Thus (A.2.100) and the Dominated Convergence Theorem give

$$\begin{aligned}\lim_{n\to\infty} \int_{\mathbb{R}^D} &\Phi^i(x,y)\left\{\frac{1}{t_n}\int_0^{t_n} E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right] ds\right\} d\pi_x(y) \\ &= \int_{\mathbb{R}^D} \Phi^i(x,y)\mathcal{A}\Phi^j(x,y)\,d\pi_x(y), \end{aligned} \tag{A.2.102}$$

and therefore, from (A.2.95) and (A.2.102):

$$\lim_{n\to\infty}\frac{1}{t_n}E\left[\Phi^i(x,\bar{\xi}(0,x))\left\{\Phi^j(x,\bar{\xi}(0,x))-\Phi^j(x,\bar{\xi}^j(x,\bar{\xi}(t_n,x))\right\}\right]$$

$$= -\lim_{n\to\infty}\int_{\mathbb{R}^D}\Phi^i(x,y)\left\{\frac{1}{t_n}\int_0^{t_n}E\left[\mathcal{A}\Phi^j(x,\xi(s,x,y))\right]ds\right\}d\pi_x(y)$$

$$= -\int_{\mathbb{R}^D}\Phi^i(x,y)\mathcal{A}\Phi^j(x,y)d\pi_x(y). \tag{A.2.103}$$

Since $\{\bar{\xi}(t,x)\}$ is stationary we have

$$E\left[\Phi^i(x,\bar{\xi}(t_n,x))\Phi^j(x,\bar{\xi}(t_n,x))\right] = E\left[\Phi^i(x,\bar{\xi}(0,x))\Phi^j(x,\bar{\xi}(0,x))\right], \tag{A.2.104}$$

thus

$$E\left[\left\{\Phi^i(x,\bar{\xi}(0,x))-\Phi^i(x,\bar{\xi}(t_n,x))\right\}\left\{\Phi^j(x,\bar{\xi}(0,x))-\Phi^j(x,\bar{\xi}(t_n,x))\right\}\right]$$

$$= E\left[\Phi^i(x,\bar{\xi}(0,x))\left\{\Phi^j(x,\bar{\xi}(0,x))-\Phi^j(x,\bar{\xi}(t_n,x))\right\}\right]$$

$$-E\left[\Phi^i(x,\bar{\xi}(t_n,x))\Phi^j(x,\bar{\xi}(0,x))\right]$$

$$+E\left[\Phi^i(x,\bar{\xi}(t_n,x))\Phi^j(x,\bar{\xi}(t_n,x))\right]$$

$$= E\left[\Phi^i(x,\bar{\xi}(0,x))\left\{\Phi^j(x,\bar{\xi}(0,x))-\Phi^j(x,\bar{\xi}(t_n,x))\right\}\right]$$

$$-E\left[\Phi^i(x,\bar{\xi}(t_n,x))\Phi^j(x,\bar{\xi}(0,x))\right]$$

$$+E\left[\Phi^i(x,\bar{\xi}(0,x))\Phi^j(x,\bar{\xi}(0,x))\right]$$

$$= E\left[\Phi^i(x,\bar{\xi}(0,x))\left\{\Phi^j(x,\bar{\xi}(0,x))-\Phi^j(x,\bar{\xi}(t_n,x))\right\}\right]$$

$$+E\left[\Phi^j(x,\bar{\xi}(0,x))\left\{\Phi^i(x,\bar{\xi}(0,x))-\Phi^i(x,\bar{\xi}(t_n,x))\right\}\right]. \tag{A.2.105}$$

Now combine (A.2.80), (A.2.103) and (A.2.105) to get

$$\begin{aligned}\overline{a}^{ij}(x) &= \lim_{n\to\infty}\frac{1}{t_n}E\left[\Phi^i(x,\overline{\xi}(0,x))\left\{\Phi^j(x,\overline{\xi}(0,x))-\Phi^j(x,\overline{\xi}(t_n,x))\right\}\right]\\ &\quad+\lim_{n\to\infty}\frac{1}{t_n}E\left[\Phi^j(x,\overline{\xi}(0,x))\left\{\Phi^i(x,\overline{\xi}(0,x))-\Phi^i(x,\overline{\xi}(t_n,x))\right\}\right]\\ &= \lim_{n\to\infty}\frac{1}{t_n}E\left[\left\{\Phi^i(x,\overline{\xi}(0,x))-\Phi^i(x,\overline{\xi}(t_n,x))\right\}\left\{\Phi^j(x,\xi(0,x))\right.\right.\\ &\qquad\qquad\left.\left.-\Phi(x,\xi(t_n,x))\right\}\right]\end{aligned} \tag{A.2.106}$$

Since the right hand side of (A.2.106) defines a nonnegative definite matrix, we see that $\overline{a}(x)$ is nonnegative definite.

□

A.3 Compactness

In this section we give the proof of the result on compactness which is needed to establish Theorem 3.3.3 on page 42.

Proposition A.3.1. *Suppose that Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8, and 3.2.18 hold, and* $\{\epsilon_n,\ n\in\mathbb{N}\}$ *is an arbitrary sequence in* $(0,1]$ *with* $\epsilon_n\to 0$, *as* $n\to\infty$. *Then the sequence of probability measures* $\{\mathcal{L}(X^{\epsilon_n},Z^{\epsilon_n}),\ n\in\mathbb{N}\}$ *over* $(\Omega^*,\mathcal{F}^*)$ *(defined by (3.3.49) and (3.3.50)) is weakly relatively compact.*

Proof: Fix some $\delta\in(0,1]$. From Proposition A.4.4 on page 111 there is some $R(\delta)\in(0,\infty)$ and positive integer $n_1(\delta)$ such that

$$P\left[T^{\epsilon_n}_{R(\delta)}<1\right]<\frac{\delta}{2},\qquad \forall n\geq n_1(\delta). \tag{A.3.107}$$

Let $\zeta \in (0, \infty)$, and choosing $0 < \gamma < 1$, write

$$\begin{aligned}
&\left\{ \max_{\substack{|\tau_2-\tau_1|\leq\gamma \\ \tau_1,\tau_2\in[0,1]}} |Z^{\epsilon_n}(\tau_2) - Z^{\epsilon_n}(\tau_1)| \geq \zeta \right\} \\
&\quad = \left[\left\{ \max_{\substack{|\tau_2-\tau_1|\leq\gamma \\ \tau_1,\tau_2\in[0,1]}} |Z^{\epsilon_n}(\tau_2) - Z^{\epsilon_n}(\tau_1)| \geq \zeta \right\} \cap \left\{ T^{\epsilon_n}_{R(\delta)} = 1 \right\} \right] \\
&\qquad \bigcup \left[\left\{ \max_{\substack{|\tau_2-\tau_1|\leq\gamma \\ \tau_1,\tau_2\in[0,1]}} |Z^{\epsilon_n}(\tau_2) - Z^{\epsilon_n}(\tau_1)| \geq \zeta \right\} \cap \left\{ T^{\epsilon_n}_{R(\delta)} < 1 \right\} \right] \\
&\quad = \left[\left\{ \max_{\substack{|\tau_2-\tau_1|\leq\gamma \\ \tau_1,\tau_2\in[0,1]}} |Z^{\epsilon_n}(\tau_2 \wedge T^{\epsilon_n}_{R(\delta)}) - Z^{\epsilon_n}(\tau_1 \wedge T^{\epsilon_n}_{R(\delta)})| \geq \zeta \right\} \cap \left\{ T^{\epsilon_n}_{R(\delta)} = 1 \right\} \right] \\
&\qquad \bigcup \left[\left\{ \max_{\substack{|\tau_2-\tau_1|\leq\gamma \\ \tau_1,\tau_2\in[0,1]}} |Z^{\epsilon_n}(\tau_2) - Z^{\epsilon_n}(\tau_1)| \geq \zeta \right\} \cap \left\{ T^{\epsilon_n}_{R(\delta)} < 1 \right\} \right] \\
&\quad \subset \left[\left\{ \max_{\substack{|\tau_2-\tau_1|\leq\gamma \\ \tau_1,\tau_2\in[0,1]}} |Z^{\epsilon_n}(\tau_2 \wedge T^{\epsilon_n}_{R(\delta)}) - Z^{\epsilon_n}(\tau_1 \wedge T^{\epsilon_n}_{R(\delta)})| \geq \zeta \right\} \cup \left\{ T^{\epsilon_n}_{R(\delta)} < 1 \right\} \right]
\end{aligned} \tag{A.3.108}$$

Thus, using (A.3.108), one sees that

$$\begin{aligned}
&P\left\{ \max_{\substack{|\tau_2-\tau_1|\leq\gamma \\ \tau_1,\tau_2\in[0,1]}} |Z^{\epsilon_n}(\tau_2) - Z^{\epsilon_n}(\tau_1)| \geq \zeta \right\} \\
&\quad \leq P\left\{ \max_{\substack{|\tau_2-\tau_1|\leq\gamma \\ \tau_1,\tau_2\in[0,1]}} |Z^{\epsilon_n}(\tau_2 \wedge T^{\epsilon_n}_{R(\delta)}) - Z^{\epsilon_n}(\tau_1 \wedge T^{\epsilon_n}_{R(\delta)})| \geq \zeta \right\} + P\left\{ T^{\epsilon_n}_{R(\delta)} < 1 \right\}
\end{aligned} \tag{A.3.109}$$

Now from Proposition A.4.2 on page 101, Chebyshev's inequality, and the fact that $p \in (2, \infty)$

(see Remark A.4.1) we can write

$$
\begin{aligned}
P&\left[\max_{\tau\in[\tau_1,\tau_1+\gamma]}|Z^{\epsilon_n}(\tau\wedge T^{\epsilon_n}_{R(\delta)})-Z^{\epsilon_n}(\tau_1\wedge T^{\epsilon_n}_{R(\delta)})|\geq\zeta\right]\\
&\leq\frac{E\left[\max_{\tau\in[\tau_1,\tau_1+\gamma]}|Z^{\epsilon_n}(\tau\wedge T^{\epsilon_n}_{R(\delta)})-Z^{\epsilon_n}(\tau_1\wedge T^{\epsilon_n}_{R(\delta)})|^p\right]}{\zeta^p}\\
&\leq\frac{C(R(\delta))}{\zeta^p}\left[(\epsilon_n)^{\frac{p-2}{2}}+\gamma^{\frac{p}{2}}\right],\quad\forall\tau_1\in[0,1],
\end{aligned}
\tag{A.3.110}
$$

for each $\gamma\in(0,1)$ and each $\zeta\in(0,\infty)$. Clearly, from (A.3.110), one has

$$
\begin{aligned}
\frac{1}{\gamma}P&\left[\max_{\tau\in[\tau_1,\tau_1+\gamma]}|Z^{\epsilon_n}(\tau\wedge T^{\epsilon_n}_{R(\delta)})-Z^{\epsilon_n}(\tau_1\wedge T^{\epsilon_n}_{R(\delta)})|\geq\zeta\right]\\
&\leq\frac{C(R(\delta))}{\zeta^p}\left[\gamma^{-1}(\epsilon_n)^{\frac{p-2}{2}}+\gamma^{\frac{p-2}{2}}\right],\quad\forall\tau_1\in[0,1],
\end{aligned}
\tag{A.3.111}
$$

for each $\gamma\in(0,1)$ and each $\zeta\in(0,\infty)$. Fix $\zeta,\beta\in(0,\infty)$, then there is a constant $\gamma_1\triangleq\gamma_1(\zeta,\beta)\in(0,\infty)$ and a positive integer $n_2\triangleq n_2(\zeta,\beta)$ such that

$$
\begin{aligned}
\frac{C(R(\delta))}{\zeta^p}\gamma_1^{\frac{(p-2)}{2}}&<\frac{\beta}{2},\\
\frac{C(R(\delta))}{\zeta^p}\gamma_1^{-1}(\epsilon_n)^{\frac{(p-2)}{2}}&<\frac{\beta}{2},\quad\forall n\geq n_2.
\end{aligned}
\tag{A.3.112}
$$

Hence, from (A.3.111) and (A.3.112), we get

$$
\frac{1}{\gamma_1}P\left[\max_{\tau\in[\tau_1,\tau_1+\gamma_1]}|Z^{\epsilon_n}(\tau\wedge T^{\epsilon_n}_{R(\delta)})-Z^{\epsilon_n}(\tau_1\wedge T^{\epsilon_n}_{R(\delta)})|\geq\zeta\right]<\beta\quad\forall n\geq n_2,\ \tau_1\in[0,1].
\tag{A.3.113}
$$

Now, by Theorem C.0.10 on page 204, it follows easily that the sequence of probability measures $\{\mathcal{L}(Z^{\epsilon_n}(.\wedge T^{\epsilon_n}_{R(\delta)})),\ n\in\mathbb{N}\}$ on $\left(C[0,1],\mathcal{B}_{C[0,1]}\right)$ is tight. Thus, by Theorem C.0.8 on page 204, there exists some $\gamma_2(\zeta,\delta)\in(0,1]$ and some positive integer $n_3(\zeta,\delta)$ such that

$$
P\left[\max_{\substack{|\tau_2-\tau_1|\leq\gamma_2\\ \tau_1,\tau_2\in[0,1]}}|Z^{\epsilon_n}(\tau_2\wedge T^{\epsilon_n}_{R(\delta)})-Z^{\epsilon_n}(\tau_1\wedge T^{\epsilon_n}_{R(\delta)})|\geq\zeta\right]<\frac{\delta}{2},\quad\forall n\geq n_3(\delta,\zeta).
\tag{A.3.114}
$$

Combining (A.3.107), (A.3.109) and (A.3.114), we get

$$P\left[\max_{\substack{|\tau_2-\tau_1|\leq\tau_2\\ \tau_1,\tau_2\in[0,1]}} |Z^{\epsilon_n}(\tau_2) - Z^{\epsilon_n}(\tau_1)| \geq \zeta\right] < \delta, \qquad \forall n \geq n_3(\delta,\zeta). \tag{A.3.115}$$

Thus, using Theorem C.0.8 on page 204, one sees that the sequence $\{\mathcal{L}\ (Z^{\epsilon_n}),\ n \in \mathbb{N}\}$ is tight, hence weakly relatively compact (by Theorem 6.1 on page 37 of Billingsley [4]). A similar argument using Propositions A.4.4 on page 111 and A.4.3 shows that $\{\mathcal{L}(X^{\epsilon_n}),\ n \in \mathbb{N}\}$ is also relatively compact. Finally, weak relative compactness of $\{\mathcal{L}(X^{\epsilon_n}, Z^{\epsilon_n}),\ n \in \mathbb{N}\}$ follows from Lemma (C.0.11) on page 205.

□

A.4 Supporting Results for the Proof of Proposition A.3.1 on page 98

Remark A.4.1. In this section we give the proofs of the various supporting results which were used to establish Proposition A.3.1 on page 98. To this end, for each $\epsilon \in (0,1]$ define

$$\Xi^\epsilon(\tau) \triangleq \epsilon^{-1/2}\int_0^\tau \mathcal{A}\Phi(X^\epsilon, Y^\epsilon)(s)ds, \qquad \forall \tau \in [0,1], \tag{A.4.116}$$

where $\Phi(x,y)$ is the C^2-function in Condition 3.2.8. Also, observe from Condition 3.2.18 that there is some constant $p \in (2,\infty)$ such that $p(1+q_2) \leq q_4$.

Proposition A.4.2. *Suppose that Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8 and 3.2.18 hold, and let $p \in (2,\infty)$ be the constant in Remark A.4.1. Then, for each $R \in (0,\infty)$, there is a constant $C(R) \in (0,\infty)$ such that*

$$E\left[\max_{\tau\in[\tau_1,\tau_2]} |Z^\epsilon(\tau\wedge T_R^\epsilon) - Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p\right] \leq C(R)\left[\epsilon^{(p-2)/2} + (\tau_2-\tau_1)^{p/2}\right], \tag{A.4.117}$$

for all $\epsilon \in (0,1]$ and $0 \leq \tau_1 < \tau_2 \leq 1$.

Proof: Fix $R \in (0,\infty)$. From (3.3.42), (3.3.44), (3.3.45) and (3.2.19), for each $\epsilon \in (0,1]$, we have

$$\begin{aligned}
Z^\epsilon(u \wedge T_R^\epsilon) &\triangleq \epsilon^{-1/2}\left[X^\epsilon(u \wedge T_R^\epsilon) - \overline{X}^\epsilon(u \wedge T_R^\epsilon)\right] \\
&= \epsilon^{-1/2}\left[\int_0^{u\wedge T_R^\epsilon} F(X^\epsilon(s),Y^\epsilon(s))ds + \int_0^{u\wedge T_R^\epsilon} G(X^\epsilon(s))dW^\epsilon(s)\right. \\
&\qquad \left. - \left(\int_0^{u\wedge T_R^\epsilon} \overline{F}(\overline{X}^\epsilon(s))ds + \int_0^{u\wedge T_R^\epsilon} G(\overline{X}^\epsilon(s))dW^\epsilon(s)\right)\right] \\
&= \epsilon^{-1/2}\left[\int_0^{u\wedge T_R^\epsilon} \left[\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right] ds\right. \\
&\qquad + \int_0^{u\wedge T_R^\epsilon} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s) \\
&\qquad \left. + \int_0^{u\wedge T_R^\epsilon} \left[F(X^\epsilon(s),Y^\epsilon(s)) - \overline{F}(X^\epsilon(s))\right] ds\right] \\
&= \epsilon^{-1/2}\left\{\int_0^{u\wedge T_R^\epsilon} \left[\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right] ds\right. \\
&\qquad + \int_0^{u\wedge T_R^\epsilon} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s) \\
&\qquad \left. - \int_0^{u\wedge T_R^\epsilon} \mathcal{A}\Phi(X^\epsilon(s),Y^\epsilon(s))\, ds\right\}, \qquad \forall u \in [0,1].
\end{aligned} \tag{A.4.118}$$

In view of (A.4.116) and (A.4.118):

$$\begin{aligned}
Z^\epsilon(u \wedge T_R^\epsilon) &= \epsilon^{-1/2}\int_0^{u\wedge T_R^\epsilon} \left[\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right] ds \\
&\quad + \epsilon^{-1/2}\int_0^{u\wedge T_R^\epsilon} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s) \\
&\quad - \Xi^\epsilon(u \wedge T_R^\epsilon),
\end{aligned} \tag{A.4.119}$$

$\forall u \in [0,1], \forall \epsilon \in (0,1]$. Hence

$$\begin{aligned} Z^\epsilon(u \wedge T_R^\epsilon) - Z^\epsilon(\tau_1 \wedge T_R^\epsilon) &= \epsilon^{-1/2} \int_{\tau_1 \wedge T_R^\epsilon}^{u \wedge T_R^\epsilon} \left[\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right] ds \\ &\quad + \epsilon^{-1/2} \int_{\tau_1 \wedge T_R^\epsilon}^{u \wedge T_R^\epsilon} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s) \\ &\quad - \left[\Xi^\epsilon(u \wedge T_R^\epsilon) - \Xi^\epsilon(\tau_1 \wedge T_R^\epsilon)\right], \qquad (A.4.120) \end{aligned}$$

$\forall \epsilon \in (0,1], \forall 0 \le \tau_1 \le u \le 1$. From (A.4.120) we can write

$$\begin{aligned} Z^\epsilon(u \wedge T_R^\epsilon) - Z^\epsilon(\tau_1 \wedge T_R^\epsilon) &= \epsilon^{-1/2} \int_{\tau_1}^{u} \left[\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right] I\{s \le T_R^\epsilon\} ds \\ &\quad + \epsilon^{-1/2} \int_{\tau_1}^{u} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] I\{s \le T_R^\epsilon\} dW^\epsilon(s) \\ &\quad - \left[\Xi^\epsilon(u \wedge T_R^\epsilon) - \Xi^\epsilon(\tau_1 \wedge T_R^\epsilon)\right], \qquad (A.4.121) \end{aligned}$$

$\forall \epsilon \in (0,1], \forall 0 \le \tau_1 \le u \le 1$. Thus, from (A.4.121), we get

$$\begin{aligned} &\max_{u \in [\tau_1, \tau]} |Z^\epsilon(u \wedge T_R^\epsilon) - Z^\epsilon(\tau_1 \wedge T_R^\epsilon)| \\ &\quad \le \epsilon^{-1/2} \int_{\tau_1}^{\tau} \left|\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right| I\{s \le T_R^\epsilon\} ds \\ &\quad + \epsilon^{-1/2} \max_{u \in [\tau_1, \tau]} \left| \int_{\tau_1}^{u} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] I\{s \le T_R^\epsilon\} dW^\epsilon(s) \right| \\ &\quad + \max_{u \in [\tau_1, \tau]} |\Xi^\epsilon(u \wedge T_R^\epsilon) - \Xi^\epsilon(\tau_1 \wedge T_R^\epsilon)|, \qquad (A.4.122) \end{aligned}$$

$\forall \epsilon \in (0,1], \forall 0 \le \tau_1 \le \tau \le 1$. Now, from (A.4.122), one has

$$\begin{aligned} &\max_{u \in [\tau_1, \tau]} |Z^\epsilon(u \wedge T_R^\epsilon) - Z^\epsilon(\tau_1 \wedge T_R^\epsilon)|^p \\ &\quad \le 3^p \left\{ \epsilon^{-p/2} \left[\int_{\tau_1}^{\tau} \left|\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right| I\{s \le T_R^\epsilon\} ds \right]^p \right. \\ &\quad + \epsilon^{-p/2} \max_{u \in [\tau_1, \tau]} \left| \int_{\tau_1}^{u} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] I\{s \le T_R^\epsilon\} dW^\epsilon(s) \right|^p \\ &\quad \left. + \max_{u \in [\tau_1, \tau]} |\Xi^\epsilon(u \wedge T_R^\epsilon) - \Xi^\epsilon(\tau_1 \wedge T_R^\epsilon)|^p \right\}, \qquad (A.4.123) \end{aligned}$$

$\forall \epsilon \in (0,1], \forall 0 \le \tau_1 \le \tau \le 1$. Taking expectation in (A.4.123), gives

$$E\left[\max_{u\in[\tau_1,\tau]} |Z^\epsilon(u\wedge T_R^\epsilon) - Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p\right]$$
$$\le 3^p \left\{ E\left[\epsilon^{-1/2}\left[\int_{\tau_1}^{\tau} \left|\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right| I\{s \le T_R^\epsilon\} ds\right]^p\right]\right.$$
$$+ E\left[\max_{u\in[\tau_1,\tau]} \left|\epsilon^{-1/2}\int_{\tau_1}^{u} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] I\{s \le T_R^\epsilon\} dW^\epsilon(s)\right|^p\right]$$
$$\left. + E\left[\max_{u\in[\tau_1,\tau]} |\Xi^\epsilon(u\wedge T_R^\epsilon) - \Xi^\epsilon(\tau_1\wedge T_R^\epsilon)|^p\right]\right\}, \qquad \text{(A.4.124)}$$

$\forall \epsilon \in (0,1], \forall 0 \le \tau_1 \le \tau \le 1$. Now fix $\tau_1 \in [0,1)$ and put

$$U_R^\epsilon(\tau) \triangleq E\left[\max_{u\in[\tau_1,\tau]} |Z^\epsilon(u\wedge T_R^\epsilon) - Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p\right], \qquad \forall \tau \in [\tau_1, 1]. \qquad \text{(A.4.125)}$$

Then, if L_1 denotes the global Lipschitz constant of $\overline{F}(.)$ (see Condition 3.2.3), one sees from Holder's inequality (with conjugate exponents $\frac{p-1}{p}$ and $\frac{1}{p}$) and (3.3.45) that

$$E\left[\epsilon^{-1/2}\int_{\tau_1}^{\tau} \left|\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right| I\{s \le T_R^\epsilon\} ds\right]^p$$
$$\le (\tau-\tau_1)^{p-1} E\left[\epsilon^{-p/2}\int_{\tau_1}^{\tau} \left|\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right|^p I\{s \le T_R^\epsilon\} ds\right]$$
$$\le (\tau-\tau_1)^{p-1}(L_1)^p E\left[\int_{\tau_1}^{\tau} \left|\epsilon^{-1/2}\left(X^\epsilon(s) - \overline{X}^\epsilon(s)\right)\right|^p I\{s \le T_R^\epsilon\} ds\right]$$
$$\le (\tau-\tau_1)^{p-1}(L_1)^p E\left[\int_{\tau_1}^{\tau} |Z^\epsilon(s)|^p I\{s \le T_R^\epsilon\} ds\right]$$
$$\le (\tau-\tau_1)^{p-1}(L_1)^p E\left[\int_{\tau_1}^{\tau} |Z^\epsilon(s\wedge T_R^\epsilon)|^p ds\right]$$

$$
\begin{aligned}
&\leq\ (\tau-\tau_1)^{p-1}(L_1)^p E\left[\int_{\tau_1}^{\tau} |Z^\epsilon(\tau_1\wedge T_R^\epsilon)+[Z^\epsilon(s\wedge T_R^\epsilon)-Z^\epsilon(\tau_1\wedge T_R^\epsilon)]|^p\,ds\right]\\
&\leq\ (\tau-\tau_1)^{p-1}(2L_1)^p E\left[\int_{\tau_1}^{\tau} [|Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p+|Z^\epsilon(s\wedge T_R^\epsilon)-Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p]\,ds\right]\\
&\leq\ (\tau-\tau_1)^{p-1}(2L_1)^p[(\tau-\tau_1)E\,|Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p\\
&\qquad\qquad +\int_{\tau_1}^{\tau} E[\max_{u\in[\tau_1,s]}|Z^\epsilon(u\wedge T_R^\epsilon)-Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p]ds]\\
&\leq\ (\tau-\tau_1)^{p-1}(2L_1)^p\left[(\tau-\tau_1)E\,|Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p+\int_{\tau_1}^{\tau} U_R^\epsilon(s)ds\right], \qquad \text{(A.4.126)}
\end{aligned}
$$

$\forall\epsilon\in(0,1],\tau\in[\tau_1,1]$. Considering the second term on the right side of (A.4.124), the Burkholder inequality gives a constant $C_p\in[0,\infty)$ such that

$$
\begin{aligned}
&E\left[\max_{u\in[\tau_1,\tau]}\left|\epsilon^{-1/2}\int_{\tau_1}^{\tau}\left[G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right]I\{s\leq T_R^\epsilon\}dW^\epsilon(s)\right|^p\right]\\
&\leq \epsilon^{-p/2}E\left[\max_{u\in[\tau_1,\tau]}\left|\int_{\tau_1}^{\tau}\left[G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right]I\{s\leq T_R^\epsilon\}dW^\epsilon(s)\right|^p\right]\\
&\leq \epsilon^{-p/2}C_pE\left[\left(\int_{\tau_1}^{\tau}\left[G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right]^2 I\{s\leq T_R^\epsilon\}ds\right)^{p/2}\right]\\
&\leq \epsilon^{-p/2}C_pE\left[(\tau-\tau_1)^{(p-2)/2}\left\{\int_{\tau_1}^{\tau}\left|G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right|^p I\{s\leq T_R^\epsilon\}ds\right\}\right] \qquad \text{(A.4.127)}
\end{aligned}
$$

$\forall\epsilon\in(0,1],\tau\in[\tau_1,1]$. Thus, if L_2 is the global Lipschitz constant of $G(.)$ (see Condition 3.2.1), then Holder inequality (with conjugate exponent $\frac{p}{p-2},\frac{p}{2}$) and (A.4.127) gives

$$
\begin{aligned}
&E\left[\max_{u\in[\tau_1,\tau]}\left|\epsilon^{-1/2}\int_{\tau_1}^{\tau}\left[G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right]I\{s\le T_R^\epsilon\}dW^\epsilon(s)\right|^p\right]\\
&\le \epsilon^{-p/2}C_pE\left[(\tau-\tau_1)^{(p-2)/2}\left\{\int_{\tau_1}^{\tau}\left|G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right|^p I\{s\le T_R^\epsilon\}ds\right\}\right]\\
&\le (\tau-\tau_1)^{(p-2)/2}C_pE\left[\int_{\tau_1}^{\tau}\epsilon^{-p/2}\left|G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right|^p I\{s\le T_R^\epsilon\}ds\right]\\
&\le (\tau-\tau_1)^{(p-2)/2}C_p(L_2)^pE\left[\int_{\tau_1}^{\tau}\left|\epsilon^{-1/2}\left(X^\epsilon(s)-\overline{X}^\epsilon(s)\right)\right|^p I\{s\le T_R^\epsilon\}ds\right]\\
&\le C_p(\tau-\tau_1)^{(p-2)/2}(L_2)^pE\left[\int_{\tau_1}^{\tau}|Z^\epsilon(s)|^p I\{s\le T_R^\epsilon\}ds\right]\\
&\le C_p(\tau-\tau_1)^{(p-2)/2}(L_2)^pE\left[\int_{\tau_1}^{\tau}|Z^\epsilon(s\wedge T_R^\epsilon)|^p\,ds\right]\\
&\le C_p(\tau-\tau_1)^{(p-2)/2}(L_2)^pE\left[\int_{\tau_1}^{\tau}|Z^\epsilon(\tau_1\wedge T_R^\epsilon)+(Z^\epsilon(s\wedge T_R^\epsilon)-Z^\epsilon(\tau_1\wedge T_R^\epsilon))|^p\,ds\right]\\
&\le C_p(\tau-\tau_1)^{(p-2)/2}(2L_2)^pE\left[\int_{\tau_1}^{\tau}|Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p+|Z^\epsilon(s\wedge T_R^\epsilon)\right.\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\left.-Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p\,ds\right]\\
&\le C_p(\tau-\tau_1)^{(p-2)/2}(2L_2)^p\left[(\tau-\tau_1)E\,|Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p\right.\\
&\qquad\qquad\qquad\left.+\int_{\tau_1}^{\tau}E\left[\max_{u\in[\tau_1,s]}|Z^\epsilon(u\wedge T_R^\epsilon)-Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p\right]ds\right]\\
&\le C_p(2L_2)^p(\tau-\tau_1)^{(p-2)/2}\left[(\tau-\tau_1)E\,|Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p+\int_{\tau_1}^{\tau}U_R^\epsilon(s)ds\right],
\end{aligned}
\tag{A.4.128}
$$

$\forall\epsilon\in(0,1],\forall\tau\in[\tau_1,1]$. Combining (A.4.124) with (A.4.126) and (A.4.128):

$$
\begin{aligned}
U_R^\epsilon(\tau) \;\le\; 3^p\Big\{&(\tau-\tau_1)^{p-1}(2L_1)^p\left[(\tau-\tau_1)E\,|Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p+\int_{\tau_1}^{\tau}U_R^\epsilon(s)ds\right]\\
&+C_p(\tau-\tau_1)^{(p-2)/2}(2L_2)^p\left[(\tau-\tau_1)E\,|Z^\epsilon(\tau_1\wedge T_R^\epsilon)|^p+\int_{\tau_1}^{\tau}U_R^\epsilon(s)ds\right]\\
&+E\left[\max_{u\in[\tau_1,\tau]}|\Xi^\epsilon(u\wedge T_R^\epsilon)-\Xi^\epsilon(u\wedge T_R^\epsilon)|^p\right]\Big\}
\end{aligned}
\tag{A.4.129}
$$

$\forall \epsilon \in (0,1], \forall \tau \in [\tau_1, 1]$. Then, from Proposition A.4.3 on page 107 there is a constant $C_1(R) \in [0,\infty)$ such that

$$\begin{aligned} U_R^\epsilon(\tau) \leq\ & C_1(R)\left[\int_{\tau_1}^{\tau} U_R^\epsilon(s)ds + (\tau-\tau_1)^{p/2}\right. \\ & \left. + E[\max_{u\in[\tau_1,\tau]} |\Xi^\epsilon(u\wedge T_R^\epsilon) - \Xi^\epsilon(u\wedge T_R^\epsilon)|^p]\right]. \end{aligned} \tag{A.4.130}$$

$\forall \epsilon \in (0,1], \forall \tau \in [\tau_1, 1]$. Thus, from (A.4.129) and Proposition A.4.6 on page 117 there are some constants $C_2(R), C_3(R) \in [0,\infty)$ such that

$$\begin{aligned} U_R^\epsilon(\tau) &\leq C_1(R)\left\{(\tau-\tau_1)^{p/2} + C_2(R)\left[\epsilon^{(p-2)/2} + (\tau-\tau_1)^{p/2}\right] + \int_{\tau_1}^{\tau} U_R^\epsilon(s)ds\right\} \\ &\leq C_3(R)\left\{(\tau_2-\tau_1)^{p/2} + \epsilon^{(p-2)/2} + \int_{\tau_1}^{\tau} U_R^\epsilon(s)ds\right\}. \end{aligned} \tag{A.4.131}$$

$\forall \epsilon \in (0,1], \forall 0 \leq \tau_1 \leq \tau_2 \leq 1$. Applying the Gronwall's inequality in (A.4.131), one has

$$U_R^\epsilon(\tau) \leq C(R)\left[\epsilon^{(p-2)/2} + (\tau_2-\tau_1)^{p/2}\right], \tag{A.4.132}$$

for some constant $C(R) \in [0,\infty)$, thus (A.4.117) follows.

□

The next result is used for the proofs of Proposition A.4.2 on page 101 and Proposition A.3.1 on page 98.

Proposition A.4.3. *Suppose that Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8 and 3.2.18 hold, and let $p \in (2,\infty)$ be the constant in Remark A.4.1. Then, for each $R \in (0,\infty)$, there is a constant $C(R) \in (0,\infty)$ such that*

$$E\left[\max_{\tau\in[0,1]} |Z^\epsilon(\tau\wedge T_R^\epsilon)|^p\right] \leq C(R), \qquad \forall \epsilon \in (0,1]. \tag{A.4.133}$$

Proof: Fix $R \in (0, \infty)$. Using $Z^\epsilon(u \wedge T_R^\epsilon)$ given by (A.4.119), we have

$$\begin{aligned} |Z^\epsilon(u \wedge T_R^\epsilon)| \leq & \max_{u_1 \in [0,\tau]} |\Xi^\epsilon(u_1 \wedge T_R^\epsilon)| \\ & + \epsilon^{-1/2} \int_0^{\tau \wedge T_R^\epsilon} |\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))| \, ds \\ & + \epsilon^{-1/2} \max_{u_2 \in [0,\tau]} \left| \int_0^{u_2 \wedge T_R^\epsilon} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s) \right|, \end{aligned} \tag{A.4.134}$$

$\forall 0 \leq u \leq \tau \leq 1, \forall \epsilon \in (0, 1]$. From (A.4.134):

$$\begin{aligned} \max_{u \in [0,\tau]} |Z^\epsilon(u \wedge T_R^\epsilon)| \leq & \max_{u \in [0,\tau]} |\Xi^\epsilon(u \wedge T_R^\epsilon)| \\ & + \epsilon^{-1/2} \int_0^{\tau \wedge T_R^\epsilon} |\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))| \, ds \\ & + \epsilon^{-1/2} \max_{u \in [0,\tau]} \left| \int_0^{u \wedge T_R^\epsilon} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s) \right|, \end{aligned} \tag{A.4.135}$$

$\forall \tau \in [0, 1], \forall \epsilon \in (0, 1]$. Hence

$$\begin{aligned} \max_{u \in [0,\tau]} |Z^\epsilon(u \wedge T_R^\epsilon)|^p \leq 3^p \Bigg\{ & \max_{u \in [0,\tau]} |\Xi^\epsilon(u \wedge T_R^\epsilon)|^p \\ & + \epsilon^{-p/2} \left[\int_0^{\tau \wedge T_R^\epsilon} |\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))| \, ds \right]^p \\ & + \epsilon^{-p/2} \max_{u \in [0,\tau]} \left| \int_0^{u \wedge T_R^\epsilon} \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s) \right|^p \Bigg\}, \end{aligned} \tag{A.4.136}$$

$\forall \tau \in [0,1], \forall \epsilon \in (0,1]$. Upon taking expectation in (A.4.136):

$$
\begin{aligned}
&E\left[\max_{u\in[0,\tau]} |Z^\epsilon(u \wedge T_R^\epsilon)|^p\right] \\
&\le 3^p \left\{ E\left[\max_{u_1\in[0,1]} |\Xi^\epsilon(u_1 \wedge T_R^\epsilon)|^p\right] \right. \\
&+ E\left[\epsilon^{-1/2}\int_0^{\tau\wedge T_R^\epsilon} \left|\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right| ds\right]^p \\
&\left. + E\left[\max_{u\in[0,\tau]}\left|\epsilon^{-1/2}\int_0^u \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] I\{s \le T_R^\epsilon\} dW^\epsilon(s)\right|^p\right]\right\},
\end{aligned} \tag{A.4.137}
$$

$\forall \tau \in [0,1], \forall \epsilon \in (0,1]$. Now, for each $\tau \in [0,1]$, put

$$
U_R^\epsilon(\tau) \triangleq E\left[\max_{u\in[0,\tau]} |Z^\epsilon(u \wedge T_R^\epsilon)|^p\right], \quad \forall \epsilon \in (0,1]. \tag{A.4.138}
$$

From (3.3.45), (A.4.138), the global Lipschitz continuity of $\overline{F}(.)$ (see Condition 3.2.3), with Lipschitz Constant $L_1 \in [0,\infty)$, and Holder inequality, we find an upper bound for the second expectation on the right side of (A.4.137):

$$
\begin{aligned}
&E\left[\epsilon^{-1/2}\int_0^{\tau\wedge T_R^\epsilon} \left|\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right| ds\right]^p \\
&\qquad \le (L_1)^p E\left[\epsilon^{-1/2}\int_0^{\tau\wedge T_R^\epsilon} \left|X^\epsilon(s) - \overline{X}^\epsilon(s)\right| ds\right]^p \\
&\qquad \le (L_1)^p E\left[\int_0^{\tau\wedge T_R^\epsilon} \left(\epsilon^{-1/2}\left|X^\epsilon(s) - \overline{X}^\epsilon(s)\right|\right)^p ds\right] \\
&\qquad \le (L_1)^p E\left[\int_0^{\tau\wedge T_R^\epsilon} |Z^\epsilon(s)|^p ds\right] \\
&\qquad \le (L_1)^p E\left[\int_0^{\tau} |Z^\epsilon(s)|^p I\{s \le T_R^\epsilon\} ds\right] \\
&\qquad \le (L_1)^p E\left[\int_0^{\tau} |Z^\epsilon(s \wedge T_R^\epsilon)|^p ds\right] \\
&\qquad \le (L_1)^p E\left[\int_0^{\tau} \max_{u\in[0,s]} |Z^\epsilon(u \wedge T_R^\epsilon)|^p ds\right] \\
&\qquad \le (L_1)^p E\int_0^{\tau} U_R^\epsilon(s) ds,
\end{aligned} \tag{A.4.139}
$$

$\forall\tau \in [0,1], \forall\epsilon \in (0,\infty]$.

Also, using the Burkholder's inequality, for the third expectation on the right side of (A.4.137), there is a Constant $C_p \in [0,\infty)$ such that

$$E\left[\max_{u\in[0,\tau]}\left|\epsilon^{-1/2}\int_0^u \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] I\{s \le T_R^\epsilon\} dW^\epsilon(s)\right|^p\right]$$
$$\le C_p E\left[\left(\epsilon^{-1}\int_0^\tau \left|G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right|^2 I\{s \le T_R^\epsilon\} ds\right)^{p/2}\right] \tag{A.4.140}$$

$\forall\tau \in [0,1], \forall\epsilon \in (0,1]$. If $L_2 \in [0,\infty)$ is a global Lipschitz constant for $G(.)$ (see Condition 3.2.1), then, applying Holder inequality (with conjugate exponents $\frac{p}{p-2}, \frac{p}{2}$), (3.3.45) and (A.4.138), one establishes an upper bound for (A.4.140) as follows

$$\begin{aligned}
E&\left[\max_{u\in[0,\tau]}\left|\epsilon^{-1/2}\int_0^u \left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] I\{s \le T_R^\epsilon\} dW^\epsilon(s)\right|^p\right] \\
&\le C_p E\left[\epsilon^{-p/2}\int_0^\tau \left|G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right|^p I\{s \le T_R^\epsilon\} ds\right] \\
&\le C_p (L_2)^p E\left[\int_0^\tau \left|\epsilon^{-1/2}(X^\epsilon(s) - \overline{X}^\epsilon(s)\right|^p I\{s \le T_R^\epsilon\} ds\right] \\
&\le C_p (L_2)^p E\left[\int_0^\tau |Z^\epsilon(s)|^p I\{s \le T_R^\epsilon\} ds\right] \\
&\le C_p (L_2)^p E\left[\int_0^\tau |Z^\epsilon(s \wedge T_R^\epsilon)|^p\right] \\
&\le C_p (L_2)^p E\left[\int_0^\tau \max_{u\in[0,s]} |Z^\epsilon(s \wedge T_R^\epsilon)|^p ds\right] \\
&\le C_p (L_2)^p E\int_0^\tau U_R^\epsilon(s) ds, \qquad \forall\tau \in [0,1].
\end{aligned} \tag{A.4.141}$$

$\forall\tau \in [0,1], \forall\epsilon \in (0,1]$. Hence, from (A.4.137), (A.4.139) and (A.4.141) we have

$$U_R^\epsilon(\tau) \le C_2\left\{\int_0^\tau U_R^\epsilon(s) ds + E\left[\max_{u\in[0,1]} |\Xi^\epsilon(u \wedge T_R^\epsilon)|^p\right]\right\}, \tag{A.4.142}$$

$\forall \tau \in [0,1], \forall \epsilon \in (0,1]$ and for some constant $C_2 \in [0,\infty)$. Now, by Proposition A.4.6 on page 117, there is a constant $C_3(R) \in [0,\infty)$ such that

$$E\left[\max_{u\in[0,1]} |\Xi^\epsilon(u \wedge T_R^\epsilon)|^p\right] \leq C_3(R), \quad \forall \epsilon \in (0,1]. \tag{A.4.143}$$

Hence, from (A.4.142) and (A.4.143):

$$U_R^\epsilon(\tau) \leq C_2 \left\{ \int_0^\tau U_R^\epsilon(s)ds + C_3(R) \right\}, \tag{A.4.144}$$

$\forall \tau \in [0,1], \forall \epsilon \in (0,1]$. Thus (A.4.133) follows from (A.4.138), (A.4.144) and Gronwall's inequality.

□

The next result is used for Proposition A.3.1 on page 98 and Theorem 3.3.3 on page 42.

Proposition A.4.4. *Suppose that Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8 and 3.2.18 hold. Then, for each $\eta \in (0,1]$, there is some $R(\eta) \in (0,\infty)$ and $\epsilon(\eta) \in (0,1]$ such that*

$$P\left[T_R^\epsilon = 1\right] \geq 1 - \eta, \quad \forall \epsilon \in (0,\epsilon(\eta)], \quad \forall R \in [R(\eta),\infty). \tag{A.4.145}$$

Proof: Fix arbitrary $\epsilon \in (0,1], R \in [0,\infty)$. From Remark A.0.26 one has $|X^\epsilon(T_R^\epsilon)| = R$ on $\{T_R^\epsilon < 1\}$. Now, clearly

$$|X^\epsilon(\tau \wedge T_R^\epsilon)| \leq |\overline{X}^\epsilon(\tau \wedge T_R^\epsilon)| + |X^\epsilon(\tau \wedge T_R^\epsilon) - \overline{X}^\epsilon(\tau \wedge T_R^\epsilon)|, \tag{A.4.146}$$

for each $\tau \in [0,1]$, and thus we have

$$\max_{\tau\in[0,1]} |X^\epsilon(\tau \wedge T_R^\epsilon)| \leq \max_{\tau\in[0,1]} |\overline{X}^\epsilon(\tau)| + \max_{\tau\in[0,1]} |X^\epsilon(\tau \wedge T_R^\epsilon) - \overline{X}^\epsilon(\tau \wedge T_R^\epsilon)| \tag{A.4.147}$$

From (A.4.147) one has

$$\begin{aligned} &\left\{\max_{\tau\in[0,1]} |X^\epsilon(\tau \wedge T_R^\epsilon)| \geq R\right\} \\ &\subset \left\{\max_{\tau\in[0,1]} |\overline{X}^\epsilon(\tau)| \geq R/2\right\} \cup \left\{\max_{\tau\in[0,1]} |X^\epsilon(\tau \wedge T_R^\epsilon) - \overline{X}^\epsilon(\tau \wedge T_R^\epsilon)| \geq R/2\right\} \end{aligned} \tag{A.4.148}$$

Hence

$$P\left\{\max_{\tau\in[0,1]}|X^\epsilon(\tau\wedge T_R^\epsilon)|\geq R\right\}$$
$$\leq P\left\{\max_{\tau\in[0,1]}|\overline{X}^\epsilon(\tau)|\geq R/2\right\}+P\left\{\max_{\tau\in[0,1]}\left|X^\epsilon(\tau\wedge T_R^\epsilon)-\overline{X}^\epsilon(\tau\wedge T_R^\epsilon)\right|\geq R/2\right\} \tag{A.4.149}$$

Since

$$\{T_R^\epsilon<1\}\subset\{|X^\epsilon(T_R^\epsilon)|=R\}=\left\{\max_{\tau\in[0,1]}|X^\epsilon(\tau\wedge T_R^\epsilon)|\geq R\right\}, \tag{A.4.150}$$

from (A.4.149) and (A.4.150), we see that

$$P\{T_R^\epsilon<1\}$$
$$\leq P\left\{\max_{\tau\in[0,1]}|\overline{X}^\epsilon(\tau)|\geq R/2\right\}+P\left\{\max_{\tau\in[0,1]}\left|X^\epsilon(\tau\wedge T_R^\epsilon)-\overline{X}^\epsilon(\tau\wedge T_R^\epsilon)\right|\geq R/2\right\} \tag{A.4.151}$$

Now fix some arbitrary $\eta\in(0,1]$. Since x_0 is nonrandom, it follows from Remark 3.3.1 and standard moment bounds for stochastic differential equations (see e.g. Lemma C.0.6 on page 203 that there is a constant $C_0\in[0,\infty)$ such that

$$E\left[\max_{\tau\in[0,1]}|\overline{X}^\epsilon(\tau)|^p\right]\leq C_0,\qquad\forall\epsilon\in(0,1], \tag{A.4.152}$$

where $p\in(2,\infty)$ is the constant in Remark A.4.1. Thus, using Chebyshev's inequality, there is $R(\eta)\in(0,\infty)$ such that

$$\begin{aligned}P\left[\max_{\tau\in[0,1]}|\overline{X}^\epsilon(\tau)|\geq R(\eta)/2\right] &< \frac{E\left[\max_{\tau\in[0,1]}|\overline{X}^\epsilon(\tau)|^p\right]}{(R(\eta)/2)^p}\\ &\leq \left(\frac{2}{R(\eta)}\right)^p C_0\cdot\epsilon^{p/2}\\ &\leq \eta/2,\qquad\forall\epsilon\in(0,1].\end{aligned} \tag{A.4.153}$$

Moreover, by Proposition A.4.5 on page 113 and Chebyshev's inequality, there is $\epsilon(\eta) \in (0,1]$ such that

$$P\left[\max_{\tau\in[0,1]} |X^\epsilon(\tau \wedge T^\epsilon_{R(\eta)}) - \overline{X}^\epsilon(\tau \wedge T^\epsilon_{R(\eta)})| \geq R(\eta)/2\right] \leq \eta/2, \quad \forall \epsilon \in (0, \epsilon(\eta)]. \tag{A.4.154}$$

Thus, one sees from (A.4.154), (A.4.153) and (A.4.151) that

$$P\left[T^\epsilon_{R(\eta)} < 1\right] < \eta, \quad \forall \epsilon \in (0, \epsilon(\eta)]. \tag{A.4.155}$$

Since T^ϵ_R increases with increasing R (see Remark A.0.26), thus the result follows; namely

$$P\left[T^\epsilon_R < 1\right] \leq P\left[T^\epsilon_{R(\eta)} < 1\right] \leq \eta, \quad \forall R \in (R(\eta), \infty), \ \forall \epsilon \in (0, \epsilon(\eta)]. \tag{A.4.156}$$

□

The next result is used for the proofs of the Propositions A.4.4 on page 111 and Theorem 3.3.3 on page 42.

Proposition A.4.5. *Suppose that Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8 and 3.2.18 hold, and let $p \in (2,\infty)$ be the constant in Remark A.4.1. Then, for each $R \in (0,\infty)$, there is a constant $C(R) \in (0,\infty)$ such that*

$$E\left[\max_{\tau\in[0,1]} \left|X^\epsilon(\tau \wedge T^\epsilon_R) - \overline{X}^\epsilon(\tau \wedge T^\epsilon_R)\right|^p\right] \leq C(R)\epsilon^{p/2}, \quad \forall \epsilon \in (0,1]. \tag{A.4.157}$$

Proof: Fix $\epsilon \in (0,1]$, $R \in (0,\infty)$. In view of (3.3.42), (3.3.44) and Condition 3.2.8, clearly

$$
\begin{aligned}
\left|X^\epsilon(u\wedge T_R^\epsilon) - \overline{X}^\epsilon(u\wedge T_R^\epsilon)\right| &\leq \left|\int_0^{\tau\wedge T_R^\epsilon}\left[F(X^\epsilon(s),Y^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right] ds\right| \\
&\quad + \left|\int_0^{u\wedge T_R^\epsilon}\left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s)\right| \\
&\leq \left|\int_0^{u\wedge T_R^\epsilon}\left[F(X^\epsilon(s),Y^\epsilon(s)) - \overline{F}(X^\epsilon(s))\right] ds\right| \\
&\quad + \left|\int_0^{u\wedge T_R^\epsilon}\left[\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right] ds\right| \\
&\quad + \left|\int_0^{u\wedge T_R^\epsilon}\left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s)\right| \\
&\leq \left|\int_0^{u\wedge T_R^\epsilon}\left[\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right] ds\right| \\
&\quad + \left|\int_0^{u\wedge T_R^\epsilon}\left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s)\right| \\
&\quad + \left|\int_0^{u\wedge T_R^\epsilon}\mathcal{A}\Phi(X^\epsilon(s),Y^\epsilon(s))ds\right|,
\end{aligned}
\tag{A.4.158}
$$

$\forall u \in [0,1]$. Hence, from (A.4.158):

$$
\begin{aligned}
&\max_{u\in[0,\tau]}\left|X^\epsilon(u\wedge T_R^\epsilon) - \overline{X}^\epsilon(u\wedge T_R^\epsilon)\right| \\
&\qquad \leq \max_{u\in[0,\tau]}\int_0^{u\wedge T_R^\epsilon}\left|\overline{F}(X^\epsilon(s)) - \overline{F}(\overline{X}^\epsilon(s))\right| ds \\
&\qquad\quad + \max_{u\in[0,\tau]}\left|\int_0^{u\wedge T_R^\epsilon}\left[G(X^\epsilon(s)) - G(\overline{X}^\epsilon(s))\right] dW^\epsilon(s)\right| \\
&\qquad\quad + \max_{u\in[0,\tau]}\left|\int_0^{u\wedge T_R^\epsilon}\mathcal{A}(X^\epsilon(s)),Y^\epsilon(s))ds\right|, \quad \forall \tau \in [0,1].
\end{aligned}
\tag{A.4.159}
$$

Moreover, from (A.4.159), one sees that

$$
\begin{aligned}
E&\left[\max_{u\in[0,\tau]}|X^\epsilon(u\wedge T_R^\epsilon)-\overline{X}^\epsilon(u\wedge T_R^\epsilon)|^p\right]\\
&\le 3^p\left\{E\left[\int_0^\tau\left|\overline{F}(X^\epsilon(s))-\overline{F}(\overline{X}^\epsilon))\right| I\left[s\le T_R^\epsilon\right]ds\right]^p\right.\\
&\quad+E\left[\max_{u\in[0,\tau]}\left|\int_0^u\left[G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right]I\left[s\le T_R^\epsilon\right]dW^\epsilon(s)\right|^p\right]\\
&\quad\left.+E\left[\max_{u\in[0,1]}\left|\int_0^{u\wedge T_R^\epsilon}\mathcal{A}\Phi(X^\epsilon(s),Y^\epsilon(s))ds\right|^p\right]\right\},\quad\forall\tau\in[0,1].
\end{aligned}
\tag{A.4.160}
$$

Now put

$$
U_R^\epsilon(\tau)\triangleq E\left[\max_{u\in[0,\tau]}|X^\epsilon(u\wedge T_R^\epsilon)-\overline{X}^\epsilon(u\wedge T_R^\epsilon)|^p\right],\quad\forall\tau\in[0,1].\tag{A.4.161}
$$

If $L_1, L_2 \in [0,\infty)$ are global Lipschitz constants for $\overline{F}(.)$ and $G(.)$ respectively (see Conditions 3.2.3 and 3.2.1), then it follows easily that the first and second expectations on the right side of (A.4.160) have upper bound given by

$$
\begin{aligned}
E&\left[\int_0^\tau\left|\overline{F}(X^\epsilon(s))-\overline{F}(\overline{X}^\epsilon))\right| I\left[s\le T_R^\epsilon\right]ds\right]^p\\
&\le E\left[\int_0^\tau\left|\overline{F}(X^\epsilon(s))-\overline{F}(\overline{X}^\epsilon(s))\right|^p I\left[s\le T_R^\epsilon\right]ds\right]\\
&\le E\left[\int_0^\tau\max_{u\in[0,s]}\left|\overline{F}(X^\epsilon(u))-\overline{F}(\overline{X}^\epsilon(u))\right|^p I\left[s\le T_R^\epsilon\right]ds\right]\\
&\le (L_1)^p\int_0^\tau E\left[\max_{u\in[0,s]}|X^\epsilon(u)-\overline{X}^\epsilon(u)|^p\right]I\left[s\le T_R^\epsilon\right]ds\\
&\le (L_1)^p\int_0^\tau E\left[\max_{u\in[0,s]}|X^\epsilon(u\wedge T_R^\epsilon)-\overline{X}^\epsilon(u\wedge T_R^\epsilon)|^p\right]ds\\
&\le (L_1)^p\int_0^\tau U_R^\epsilon(s)ds,\qquad\forall\tau\in[0,1],
\end{aligned}
\tag{A.4.162}
$$

and

$$E\left[\max_{u\in[0,\tau]}\left|\int_0^u \left[G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right] I\left[s\le T_R^\epsilon\right] dW^\epsilon(s)\right|^p\right]$$
$$\le C_pE\left[\int_0^\tau \left|G(X^\epsilon(s))-G(\overline{X}^\epsilon(s))\right|^2 I\left[s\le T_R^\epsilon\right] ds\right]^{p/2}$$
$$\le C_p(L_2)^pE\left[\int_0^\tau \left|X^\epsilon(s)-\overline{X}^\epsilon(s)\right|^2 I\left[s\le T_R^\epsilon\right] ds\right]^{p/2}$$
$$\le C_p(L_2)^pE\left[\int_0^\tau \left|X^\epsilon(s)-\overline{X}^\epsilon(s)\right|^p I\left[s\le T_R^\epsilon\right] ds\right]$$
$$\le C_p(L_2)^pE\left[\int_0^\tau \left|X^\epsilon(s\wedge T_R^\epsilon)-\overline{X}^\epsilon(s\wedge T_R^\epsilon)\right|^p ds\right]$$
$$\le C_p(L_2)^p\int_0^\tau E\left[\max_{u\in[0,s]}\left|X^\epsilon(u\wedge T_R^\epsilon)-\overline{X}^\epsilon(u\wedge T_R^\epsilon)\right|^p\right] ds$$
$$\le C_p(L_2)^p\int_0^\tau U_R^\epsilon(s)ds, \quad \forall\tau\in[0,1]. \tag{A.4.163}$$

Here we used Holder's inequality to get the first inequality in (A.4.162), the Burkholder inequality at the first inequality of (A.4.163), and Holder's inequality at the third inequality of (A.4.163) (where C_p is a constant resulting from Burkholder's inequality). Using (A.4.116) and Proposition A.4.6 on page 117, there is a constant $C(R)\in[0,\infty)$ such that the upper bound of the third expectation in (A.4.160) is given by

$$\begin{aligned} E\left[\max_{u\in[0,1]}\left|\int_0^{u\wedge T_R^\epsilon} \mathcal{A}\Phi(X^\epsilon(s),Y^\epsilon(s))ds\right|^p\right] &= E\left[\max_{u\in[0,1]}\left|\Xi^\epsilon(u\wedge T_R^\epsilon)\right|^p\right]\epsilon^{p/2} \\ &\le C(R)\epsilon^{p/2}, \quad \forall\epsilon\in(0,1]. \end{aligned} \tag{A.4.164}$$

Inserting the upper bounds of (A.4.164), (A.4.163) and (A.4.162) into (A.4.160), and using (A.4.161), we get

$$U_R^\epsilon(\tau)\le C_1\int_0^\tau U_R^\epsilon(s)ds+C(R)\epsilon^{p/2}, \quad \forall\tau\in[0,1],\ \forall\epsilon\in(0,1], \tag{A.4.165}$$

for some constant $C_1\in[0,\infty)$. Thus, from (A.4.165) and Gronwall's inequality one has

$$U_R^\epsilon(\tau)\le C(R)\epsilon^{p/2}\cdot e^{C_1\tau}, \quad \forall\tau\in[0,1],\ \forall\epsilon\in(0,1], \tag{A.4.166}$$

which establishes (A.4.157).

□

The following result is used for the proof of Proposition A.4.5 on page 113, Proposition A.4.3 on page 107 and Proposition A.4.2 on page 101:

Proposition A.4.6. *Suppose that Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8 and 3.2.18 hold, and let* $p \in (2,\infty)$ *be the constant in Remark A.4.1. Then, for each* $R \in (0,\infty)$*, there is a constant* $C(R) \in (0,\infty)$ *such that*

$$E\left[\max_{\tau\in[\tau_1,\tau_2]} |\Xi^\epsilon(\tau \wedge T_R^\epsilon) - \Xi^\epsilon(\tau_1 \wedge T_R^\epsilon)|^p\right] \leq C(R)\left[\epsilon^{(p-2)/2} + (\tau_2 - \tau_1)^{p/2}\right], \tag{A.4.167}$$

for all $0 \leq \tau_1 < \tau_2 \leq 1$ *and* $\epsilon \in (0,1]$ *(recall* Ξ^ϵ *is defined in (A.4.116)).*

Proof: Take $d = D = M = N = 1$ in Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8 and 3.2.18; the proof for the general dimensions is unchanged and just involves more complicated notation. Now (3.2.19) reduces to

$$\mathcal{A}\Phi(x,y) = \overline{F}(x) - F(x,y), \quad \forall (x,y) \in \mathbb{R}\otimes\mathbb{R}, \tag{A.4.168}$$

where $\Phi : \mathbb{R}\otimes\mathbb{R} \to \mathbb{R}$ and

$$\mathcal{A}\Phi(x,y) \triangleq b(x,y)(\partial_y\Phi)(x,y) + \frac{1}{2}\sigma^2(x,y)(\partial_y\partial_y\Phi)(x,y), \quad \forall (x,y) \in \mathbb{R}\otimes\mathbb{R}. \tag{A.4.169}$$

Fix $R \in (0,1]$. Expanding $\Phi(X^\epsilon(\tau), Y^\epsilon(\tau))$ in (A.4.168) using Itô's formula, we have

$$\begin{aligned}
\Phi(X^\epsilon(\tau), Y^\epsilon(\tau)) &= \Phi(x_0, y_0) + \int_0^\tau (\partial_x \Phi)(X^\epsilon(s), Y^\epsilon(s)) dX^\epsilon(s) \\
&+ \int_0^\tau (\partial_y \Phi)(X^\epsilon(s), Y^\epsilon(s)) dY^\epsilon(s) \\
&+ \frac{1}{2} \int_0^\tau (\partial_x \partial_x \Phi)(X^\epsilon(s), Y^\epsilon(s)) d\left[X^\epsilon\right](s) \\
&+ \int_0^\tau (\partial_x \partial_y \Phi)(X^\epsilon(s), Y^\epsilon(s)) d\left[X^\epsilon, Y^\epsilon\right](s) \\
&+ \frac{1}{2} \int_0^\tau (\partial_y \partial_y \Phi)(X^\epsilon(s), Y^\epsilon(s)) d\left[Y^\epsilon\right](s) \qquad \text{(A.4.170)}
\end{aligned}$$

$\forall \epsilon \in (0,1], \forall \tau \in [0,1]$. Now, using (3.3.42), (3.3.43) and independence of $\{W^\epsilon(\tau),\ \tau \in [0,1]\}$ and $\{B^\epsilon(\tau),\ \tau \in [0,1]\}$ (as follows from (3.3.40) and Condition 3.2.2), the cross-variation processes in (A.4.170) are given by

$$\begin{aligned}
\left[X^\epsilon\right](\tau) &= \int_0^\tau (G(X^\epsilon(s)))^2 ds, \\
\left[X^\epsilon, Y^\epsilon\right](\tau) &= \int_0^\tau G(X^\epsilon(s)) \sigma(X^\epsilon(s), Y^\epsilon(s)) d\left[W^\epsilon, B^\epsilon\right](s) \\
&= 0, \\
\left[Y^\epsilon\right](\tau) &= \epsilon^{-1} \int_0^\tau (\sigma(X^\epsilon(s), Y^\epsilon(s)))^2 ds \qquad \text{(A.4.171)}
\end{aligned}$$

$\forall\epsilon \in (0,1], \forall\tau \in [0,1]$. From (A.4.170), (A.4.169), (A.4.171), (3.3.42) and (3.3.43), it follows that

$$\begin{aligned}\Phi(X^\epsilon(\tau), Y^\epsilon(\tau)) &= \Phi(x_0, y_0) + \int_0^\tau (\partial_x\Phi)(X^\epsilon(s), Y^\epsilon(s))F(X^\epsilon(s), Y^\epsilon(s))ds \\ &\quad + \int_0^\tau (\partial_x\Phi)(X^\epsilon(s), Y^\epsilon(s))G(X^\epsilon(s))dW^\epsilon(s) \\ &\quad + \epsilon^{-1}\int_0^\tau \mathcal{A}\Phi(X^\epsilon(s), Y^\epsilon(s))ds \\ &\quad + \frac{1}{2}\int_0^\tau (\partial_x\partial_x\Phi)(X^\epsilon(s), Y^\epsilon(s))G^2(X^\epsilon(s))ds \\ &\quad + \epsilon^{-\frac{1}{2}}\int_0^\tau (\partial_y\Phi)(X^\epsilon(s), Y^\epsilon(s))\sigma(X^\epsilon(s), Y^\epsilon(s))dB^\epsilon(s),\end{aligned} \tag{A.4.172}$$

$\forall\epsilon \in (0,1], \forall\tau \in [0,1]$. Hence, from (A.4.172) and (A.4.116):

$$\begin{aligned}&\epsilon^{1/2}\Phi(X^\epsilon(\tau\wedge T_R^\epsilon), Y^\epsilon(\tau\wedge T_R^\epsilon)) \\ &\quad = \epsilon^{1/2}\Phi(x_0, y_0) \\ &\quad + \epsilon^{1/2}\int_0^{\tau\wedge T_R^\epsilon} (\partial_x\Phi)(X^\epsilon(s), Y^\epsilon(s))F(X^\epsilon(s), Y^\epsilon(s))ds \\ &\quad + \epsilon^{1/2}\int_0^{\tau\wedge T_R^\epsilon} (\partial_x\Phi)(X^\epsilon(s), Y^\epsilon(s))G(X^\epsilon(s))dW^\epsilon(s) \\ &\quad + \Xi^\epsilon(\tau\wedge T_R^\epsilon) \\ &\quad + \frac{1}{2}\epsilon^{\frac{1}{2}}\int_0^{\tau\wedge T_R^\epsilon} (\partial_x\partial_x\Phi)(X^\epsilon(s), Y^\epsilon(s))G^2(X^\epsilon(s))ds \\ &\quad + \int_0^{\tau\wedge T_R^\epsilon} (\partial_y\Phi)(X^\epsilon(s), Y^\epsilon(s))\sigma(X^\epsilon(s), Y^\epsilon(s))dB^\epsilon(s),\end{aligned} \tag{A.4.173}$$

$\forall \epsilon \in (0,1], \forall \tau \in [0,1]$. Clearly, using (A.4.173) and rearranging yields

$$\begin{aligned}
&[\Xi^\epsilon(\tau \wedge T_R^\epsilon) - \Xi^\epsilon(\tau_1 \wedge T_R^\epsilon)] \\
&\quad = \epsilon^{1/2}[\Phi(X^\epsilon, Y^\epsilon)(\tau \wedge T_R^\epsilon) - \Phi(X^\epsilon, Y^\epsilon)(\tau_1 \wedge T_R^\epsilon)] \\
&\quad -\epsilon^{1/2}\int_{\tau_1}^{\tau}(\partial_x\Phi)(X^\epsilon, Y^\epsilon)(s)F(X^\epsilon, Y^\epsilon)(s)I\{s \le T_R^\epsilon\}ds \\
&\quad -\epsilon^{1/2}\int_{\tau_1}^{\tau}(\partial_x\Phi)(X^\epsilon, Y^\epsilon)(s)G(X^\epsilon(s))I\{s \le T_R^\epsilon\}dW^\epsilon(s) \\
&\quad -\frac{1}{2}\epsilon^{1/2}\int_{\tau_1}^{\tau}(\partial_x\partial_x\Phi)(X^\epsilon, Y^\epsilon)(s)G^2(X^\epsilon(s))I\{s \le T_R^\epsilon\}ds \\
&\quad -\int_{\tau_1}^{\tau}(\partial_y\Phi)(X^\epsilon, Y^\epsilon)(s)\sigma(X^\epsilon, Y^\epsilon)(s)I\{s \le T_R^\epsilon\}dB^\epsilon(s),
\end{aligned} \tag{A.4.174}$$

$\forall 0 \le \tau_1 \le \tau \le 1, \forall \epsilon \in (0,1]$. Also, taking expectation in (A.4.174), for the constant $p \in (2,\infty)$ in Remark A.4.1, one sees that

$$\begin{aligned}
&E\left[\max_{\tau\in[\tau_1,\tau_2]}|\Xi^\epsilon(\tau \wedge T_R^\epsilon) - \Xi^\epsilon(\tau_1 \wedge T_R^\epsilon)|^p\right] \\
&\le 6^p\left\{E\left[\max_{\tau\in[0,1]}\left|\epsilon^{1/2}\Phi(X^\epsilon, Y^\epsilon)(\tau \wedge T_R^\epsilon)\right|^p\right]\right. \\
&+ E\left|\epsilon^{1/2}\Phi(X^\epsilon, Y^\epsilon)(\tau_1 \wedge T_R^\epsilon)\right|^p \\
&+ \epsilon^{p/2}E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_x\Phi)(X^\epsilon, Y^\epsilon)(s)F(X^\epsilon, Y^\epsilon)(s)I\{s \le T_R^\epsilon\}ds\right|^p\right] \\
&+ \epsilon^{p/2}E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_x\Phi)(X^\epsilon, Y^\epsilon)(s)G(X^\epsilon(s))I\{s \le T_R^\epsilon\}dW^\epsilon(s)\right|^p\right] \\
&+ \frac{1}{2^p}\epsilon^{p/2}E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_x\partial_x\Phi)(X^\epsilon, Y^\epsilon)(s)G^2(X^\epsilon(s))I\{s \le T_R^\epsilon\}ds\right|^p\right] \\
&\left.+ E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_y\Phi)(X^\epsilon, Y^\epsilon)(s)\sigma(X^\epsilon, Y^\epsilon)(s)I\{s \le T_R^\epsilon\}dB^\epsilon(s)\right|^p\right]\right\},
\end{aligned} \tag{A.4.175}$$

$\forall 0 \le \tau_1 < \tau_2 \le 1, \forall \epsilon \in (0,1]$. First consider the sixth term on the right side of (A.4.175). By the inequality of Burkholder there is a constant $C_1 \in [0,\infty)$ such that

$$E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_y\Phi)(X^\epsilon,Y^\epsilon)(s)\sigma(X^\epsilon,Y^\epsilon)(s)I\{s\le T_R^\epsilon\}dB^\epsilon(s)\right|^p\right]$$
$$\le C_1 E\left[\left(\int_{\tau_1}^{\tau_2}|(\partial_y\Phi)(X^\epsilon,Y^\epsilon)(s)\sigma(X^\epsilon,Y^\epsilon)(s)|^2 I\{s\le T_R^\epsilon\}ds\right)^{p/2}\right] \tag{A.4.176}$$

$\forall 0\le\tau_1<\tau_2\le 1, \forall\epsilon\in(0,1]$. Moreover, using Holder inequality (with conjugate exponents $\frac{p}{p-2}$ and $\frac{p}{2}$) and (A.4.176), we get

$$E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_y\Phi)(X^\epsilon,Y^\epsilon)(s)\sigma(X^\epsilon,Y^\epsilon)(s)I\{s\le T_R^\epsilon\}dB^\epsilon(s)\right|^p\right]$$
$$\le C_1(\tau_2-\tau_1)^{\frac{p-2}{2}} E\int_{\tau_1}^{\tau_2}|(\partial_y\Phi)(X^\epsilon,Y^\epsilon)(s)\sigma(X^\epsilon,Y^\epsilon)(s)|^p I\{s\le T_R^\epsilon\}ds, \tag{A.4.177}$$

$\forall 0\le\tau_1<\tau_2\le 1, \forall\epsilon\in(0,1]$. Now $(x,y)\to\partial_y\Phi(x,y)\sigma(x,y)$ is polynomially bounded of order $(1+q_2)$ in y and locally in x (by Conditions 3.2.1, 3.2.8), thus there is a constant $C_2(R)\in[0,\infty)$ such that

$$|\partial_y\Phi(x,y)\sigma(x,y)|\le C_2(R)\left[1+|y|^{(1+q_2)}\right],\quad \forall x\in S_R^d,\quad \forall y\in I\!R^D, \tag{A.4.178}$$

hence (since $p(1+q_2)\le q_4$, by Remark A.4.1)

$$|(\partial_y\Phi)(X^\epsilon,Y^\epsilon)(s)\sigma(X^\epsilon,Y^\epsilon)(s)|^p I\{s\le T_R^\epsilon\}$$
$$\le C_2(R)\left[1+|Y^\epsilon(s)|^{p(1+q_2)}\right]I\{s\le T_R^\epsilon\},$$
$$\le C_2(R)\left[1+|Y^\epsilon(s)|^{q_4} I\{s\le T_R^\epsilon\}\right], \tag{A.4.179}$$

$\forall s\in[0,1]$. Then, by (A.4.177), (A.4.179) and Lemma A.4.7 on page 123, we have

$$E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_y\Phi)(X^\epsilon,Y^\epsilon)(s)\sigma(X^\epsilon,Y^\epsilon)(s)I\{s\le T_R^\epsilon\}dB^\epsilon(s)\right|^p\right]$$
$$\le C_1(\tau_2-\tau_1)^{\frac{p-2}{2}}\left[C_2(R)\int_{\tau_1}^{\tau_2}E\left[1+|Y^\epsilon(s)|^{q_4} I\{s\le T_R^\epsilon\}\right]ds\right]$$
$$\le C_3(R)(\tau_2-\tau_1)^{p/2}, \tag{A.4.180}$$

$\forall 0 \le \tau_1 < \tau_2 \le 1$, and for some constant $C_3(R) \in [0, \infty)$. In the same way, one finds a similar bound on the stochastic integral of the fourth term on the right side of (A.4.175), namely

$$E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_x\Phi)(X^\epsilon,Y^\epsilon)(s)G(X^\epsilon(s))I\{s \le T_R^\epsilon\}dW^\epsilon(s)\right|^p\right]$$
$$\le C_4(R)\,(\tau_2-\tau_1)^{p/2}, \tag{A.4.181}$$

$\forall 0 \le \tau_1 < \tau_2 \le 1$, and for some constant $C_4(R) \in [0, \infty)$. Next, consider the third term on the right side of (A.4.175). Clearly: by Conditions 3.2.1 and 3.2.8

$$|\partial_x(X^\epsilon,Y^\epsilon)(s)F(X^\epsilon,Y^\epsilon)(s)|^p\, I\{s \le T_R^\epsilon\} \le C_5(R)\left[1+|Y^\epsilon(s)|^{q_4}\, I\{s \le T_R^\epsilon\}\right], \tag{A.4.182}$$

$\forall s \in [0,1]$, and for some constant $C_5(R) \in [0, \infty)$. Hence, by Holder's inequality, Lemma A.4.7 on page 123 and using (A.4.182), we have

$$\begin{aligned}
&E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_x\Phi)(X^\epsilon,Y^\epsilon)(s)F(X^\epsilon,Y^\epsilon)(s)I\{s \le T_R^\epsilon\}ds\right|^p\right]\\
&\le E\left[\left(\int_{\tau_1}^{\tau_2}|(\partial_x\Phi)\,(X^\epsilon,Y^\epsilon)(s)F(X^\epsilon,Y^\epsilon)(s))|\, I\{s \le T_R^\epsilon\}ds\right)^p\right]\\
&\le (\tau_2-\tau_1)^{p-1}E\left[\int_{\tau_1}^{\tau_2}|(\partial_x\Phi)(X^\epsilon,Y^\epsilon)(s)F(X^\epsilon,Y^\epsilon)(s)|^p\, I\{s \le T_R^\epsilon\}ds\right]\\
&\le (\tau_2-\tau_1)^{p-1}C_5(R)\int_{\tau_1}^{\tau_2}E\left[1+|Y^\epsilon(s)|^{q_4}\, I\{s \le T_R^\epsilon\}\right]ds\\
&\le C_6(R)(\tau_2-\tau_1)^p,
\end{aligned} \tag{A.4.183}$$

$\forall 0 \le \tau_1 < \tau_2 \le 1$, and for some constant $C_6(R) \in [0, \infty)$. Again, by a similar argument, one gets the same upper-bound for the fifth term on the right of (A.4.175), as follows

$$E\left[\max_{\tau\in[\tau_1,\tau_2]}\left|\int_{\tau_1}^{\tau}(\partial_x\partial_x\Phi)(X^\epsilon,Y^\epsilon)(s)G^2(X^\epsilon(s))I\{s \le T_R^\epsilon\}ds\right|^p\right]$$
$$\le C_7(R)(\tau_2-\tau_1)^p, \tag{A.4.184}$$

$\forall 0 \le \tau_1 < \tau_2 \le 1$, and for some constant $C_7(R) \in [0, \infty)$. Finally, consider the first term on the right side of (A.4.175): Since $(x, y) \to \Phi(x, y)$ is polynomially bounded of order q_2 in y locally in x and $pq_2 < q_4$, there is a constant $C_8(R) \in [0, \infty)$ such that

$$\max_{\tau\in[0,1]} |\Phi(X^\epsilon, Y^\epsilon)(\tau \wedge T_R^\epsilon)|^p \leq C_8(R) \left[1 + \max_{\tau\in[0,1]} (|Y^\epsilon(\tau)|^{q_4} I\{\tau \leq T_R^\epsilon\})\right]. \tag{A.4.185}$$

Thus, by Corollary A.4.8,

$$E\left[\max_{\tau\in[0,1]} \left|\epsilon^{1/2}\Phi(X^\epsilon, Y^\epsilon)(\tau \wedge T_R^\epsilon)\right|^p\right] \leq C_9(R)\epsilon^{(p-2)/2}, \tag{A.4.186}$$

$\forall \epsilon \in (0, 1]$, and for some constant $C_9(R) \in [0, \infty)$. Combining (A.4.175), (A.4.180), (A.4.181), (A.4.183), (A.4.184) , and (A.4.186), one finds a constant $C(R) \in (0, \infty]$ such that

$$\begin{aligned} E&\left[\max_{\tau\in[\tau_1,\tau_2]} |\Xi^\epsilon(\tau \wedge T_R^\epsilon) - \Xi^\epsilon(\tau_1 \wedge T_R^\epsilon)|^p\right] \\ &\leq 6^p \left\{C_9(R)\epsilon^{(p-2)/2} + C_9(R)\epsilon^{(p-2)/2}\right. \\ &\quad + C_6(R)\epsilon^{p/2}(\tau_2 - \tau_1)^p + C_4(R)\epsilon^{p/2}(\tau_2 - \tau_1)^{p/2} \\ &\quad \left. + C_7(R)\epsilon^{p/2}(\tau_2 - \tau_1)^p + C_3(R)(\tau_2 - \tau_1)^{p/2}\right\} \\ &\leq C(R) \left[\epsilon^{(p-2)/2} + (\tau_2 - \tau_1)^{p/2}\right], \end{aligned} \tag{A.4.187}$$

$\forall 0 \leq \tau_1 < \tau_2 \leq 1, \forall \epsilon \in (0, 1]$. Thus (A.4.167) holds.

□

The following result controls the q_4-th order moments of $Y^\epsilon(\tau)$ (recall Condition 3.2.15) when $X^\epsilon(\tau)$ is bounded, and is used in several proofs in this section. This Lemma is suggested by, and extends, Exercise 5.4.35 of Karatzas and Shreve [19].

Lemma A.4.7. *Suppose that Conditions 3.2.1, 3.2.2, 3.2.3, 3.2.8 and 3.2.18 hold. Then, for each $R \in [0, \infty)$ there is a constant $C(R) \in [0, \infty)$ such that*

$$E\left[I\{\tau \leq T_R^\epsilon\} |Y^\epsilon(\tau)|^{q_4}\right] \leq C(R), \quad \forall \epsilon \in (0, 1], \quad \forall \tau \in [0, 1]. \tag{A.4.188}$$

Proof: Fix $\epsilon \in (0,1], R \in [0,\infty)$, and let $\varphi(y) \triangleq |y|^{q_4}$, $\forall y \in \mathbb{R}^d$, be the C^2-function defined in Condition 3.2.18. Now, put

$$M^\epsilon(t) \triangleq \int_0^t (\partial_y \varphi)(y^\epsilon(s))\sigma(x^\epsilon, y^\epsilon)(s)d\beta(s), \quad \forall t \in [0,\infty). \tag{A.4.189}$$

In view of Condition 3.2.1 we see that $(\partial_y\varphi)(y)\sigma(x,y)$ is a row vector of length N. By elementary calculus we have

$$|(\partial_y\varphi)(y)\sigma(x,y)|^2 = q_4 \left(\sum_{l=1}^{D} (y^l)^2\right)^{(q-2)} \sum_{j=1}^{N} \left\{\sum_{i=1}^{D} y^i \sigma^{i,j}(x,y)\right\}^2. \tag{A.4.190}$$

From Condition 3.2.1 there is a constant $C_1 \in [0,\infty)$ such that

$$\left|\sigma^{i,j}(x,y)\right| \leq C_1 [1 + |x| + |y|], \quad \forall (x,y) \in (\mathbb{R}^d \otimes \mathbb{R}^D). \tag{A.4.191}$$

Combining (A.4.190) and (A.4.191) we easily see that there is a constant $C_2 \in [0,\infty)$ such that

$$|(\partial_y\varphi)(y)\sigma(x,y)|^2 \leq C_2 \left[1 + |(x,y)|^{(2q_4)}\right], \quad \forall x \in \mathbb{R}^d,\ \forall y \in \mathbb{R}^D, \tag{A.4.192}$$

(where $|(x,y)|$ denotes the Euclidean length of the $(d+D)$-vector (x,y)).

Since the coefficients in (3.1.1) and (3.1.2) are linearly bounded (see Condition 3.2.1) and x_0, y_0, are non-random, it follows from well-known moment bounds for stochastic differential equations (see Lemma C.0.6 on page 203) that

$$\begin{aligned} E\int_0^t |(\partial_y\varphi)(y^\epsilon(s))\sigma(x^\epsilon, y^\epsilon)(s)|^2 ds \\ &= \int_0^t E\,|(\partial_y\varphi)(y^\epsilon(s))\sigma(x^\epsilon, y^\epsilon)(s)|^2 ds \\ &< \infty \qquad \forall t \in [0,\infty). \end{aligned} \tag{A.4.193}$$

Thus, $\{(M^\epsilon(t), \mathcal{F}_t),\ \forall t \in [0,\infty)\}$ is a martingale on $(\Omega, \mathcal{F}, P)$, hence by the optional sampling theorem $\{(M^\epsilon(t \wedge t_R^\epsilon),\ \mathcal{F}_t),\ \forall t \in [0,\infty)\}$ is a martingale (recall Remark

A.0.26), hence

$$E\left[M^\epsilon(t\wedge t_R^\epsilon)\right] = E\left[M^\epsilon(0)\right] = 0, \qquad \forall t\in[0,\infty). \tag{A.4.194}$$

Using Itô's formula and (3.1.2) to expand $\varphi(y^\epsilon(t))$, one has

$$\begin{aligned}\varphi(y^\epsilon(t\wedge t_R^\epsilon)) &= \varphi(y_0) + \int_0^{t\wedge t_R^\epsilon} \mathcal{A}\varphi(x^\epsilon, y^\epsilon)(s)ds + M^\epsilon(t\wedge t_R^\epsilon)\\ &= \varphi(y_0) + \int_0^t I\{s\le t_R^\epsilon\}\mathcal{A}\varphi(x^\epsilon, y^\epsilon)(s)ds + M^\epsilon(t\wedge t_R^\epsilon)\end{aligned} \tag{A.4.195}$$

$\forall t\in[0,\infty)$, where $\mathcal{A}\varphi(x,y)$ and $M^\epsilon(t)$ are given by (3.2.14) and (A.4.189) respectively. Clearly, from (A.4.195):

$$\begin{aligned}&\varphi(y^\epsilon(\tau_2\wedge t_R^\epsilon)) - \varphi(y^\epsilon(\tau_1\wedge t_R^\epsilon))\\ &\qquad = \int_{t_1}^{t_2} I\{s\le t_R^\epsilon\}\mathcal{A}\varphi(x^\epsilon, y^\epsilon)(s)ds + M^\epsilon(t_2\wedge t_R^\epsilon) - M^\epsilon(t_1\wedge t_R^\epsilon),\end{aligned} \tag{A.4.196}$$

$\forall 0\le t_1 < t_2 < \infty$. From Condition 3.2.18 and (A.0.1),

$$\begin{aligned}&I\{s\le t_R^\epsilon\}\left(\mathcal{A}\varphi\right)(x^\epsilon, y^\epsilon)(s)\\ &\qquad \le I\{s\le t_R^\epsilon\}\left[\alpha_R - \lambda_R\varphi\left(y^\epsilon\left(s\wedge t_R^\epsilon\right)\right)\right], \quad \forall s\in[0,\infty).\end{aligned} \tag{A.4.197}$$

Upon taking expectation in (A.4.196), by (A.4.197) and (A.4.194), we get

$$\begin{aligned}&E\left[\varphi(y^\epsilon(t_2\wedge t_R^\epsilon)) - \varphi(y^\epsilon(t_1\wedge t_R^\epsilon))\right]\\ &\qquad = \int_{t_1}^{t_2} E\left[I\{s\le t_R^\epsilon\}(\mathcal{A}\varphi)(x^\epsilon(s), y^\epsilon(s))\right]ds\\ &\qquad \le \alpha_R(t_2-t_1) - \lambda_R\int_{t_1}^{t_2} E\left[I\{s\le t_R^\epsilon\}\varphi(y^\epsilon(s\wedge t_R^\epsilon))\right]ds,\end{aligned} \tag{A.4.198}$$

$\forall 0 \leq t_1 < t_2 < \infty$. We now show that

$$I\left\{t_2 \leq t_R^\epsilon\right\} \varphi\left(y^\epsilon\left(t_2 \wedge t_R^\epsilon\right)\right) - I\left\{t_1 \leq t_R^\epsilon\right\} \varphi\left(y^\epsilon\left(t_1 \wedge t_R^\epsilon\right)\right) \leq \varphi\left(y^\epsilon(t_2 \wedge t_R^\epsilon)\right) - \varphi\left(y^\epsilon(t_1 \wedge t_R^\epsilon)\right), \tag{A.4.199}$$

$\forall 0 \leq t_1 < t_2 < \infty$. To see this consider the following cases:

(i) $0 \leq t_R^\epsilon < t_1 < t_2 < \infty$.

Here

$$\begin{aligned} \text{left hand side of (A.4.199)} &= 0, \\ \text{right hand side of (A.4.199)} &= \varphi\left(y^\epsilon(t_R^\epsilon)\right) - \varphi\left(y^\epsilon(t_R^\epsilon)\right) = 0, \end{aligned}$$

hence (A.4.199) holds.

(ii) $0 \leq t_1 \leq t_R^\epsilon < t_2 < \infty$.

Here

$$\begin{aligned} \text{left hand side of (A.4.199)} &= -\varphi\left(y^\epsilon(t_1)\right), \\ \text{right hand side of (A.4.199)} &= \varphi\left(y^\epsilon(t_R^\epsilon)\right) - \varphi\left(y^\epsilon(t_1)\right), \end{aligned}$$

hence, since $\varphi\left(y^\epsilon(t_R^\epsilon)\right) \geq 0$, we get (A.4.199).

(iii) $0 \leq t_1 < t_2 \leq t_R^\epsilon < \infty$.

Here

$$\begin{aligned} \text{left hand side of (A.4.199)} &= \varphi\left(y^\epsilon(t_2)\right) - \varphi\left(y^\epsilon(t_1)\right), \\ \text{right hand side of (A.4.199)} &= \varphi\left(y^\epsilon(t_2)\right) - \varphi\left(y^\epsilon(t_1)\right), \end{aligned}$$

hence (A.4.199) follows. Thus we have established (A.4.199). From (A.4.199) and (A.4.198), one has

$$\begin{aligned} E\left[I\{t_2 \le t^\epsilon_R\}\varphi(y^\epsilon(t_2 \wedge t^\epsilon_R)) - I\{t_1 \le t^\epsilon_R\}\varphi(y^\epsilon(t_1 \wedge t^\epsilon_R))\right] \\ \le E\left[\varphi(y^\epsilon(t_2 \wedge t^\epsilon_R)) - \varphi(y^\epsilon(t_1 \wedge t^\epsilon_R))\right] \\ \le \alpha_R(t_2 - t_1) - \lambda_R \int_{t_1}^{t_2} E\left[I\{s \le t^\epsilon_R\}\varphi(y^\epsilon(s \wedge t^\epsilon_R))\right] ds. \end{aligned} \tag{A.4.200}$$

$\forall 0 \le t_1 < t_2 < \infty$. Now put

$$\psi(t) \triangleq E\left[I\{t \le t^\epsilon_R\}\varphi(y^\epsilon(t \wedge t^\epsilon_R))\right], \quad \forall t \in [0, \infty), \tag{A.4.201}$$

where the dependence of $\psi(t)$ on ϵ, R, is omitted to simplify the notation. From (A.4.200) and (A.4.201), we can write

$$\psi(t_2) - \psi(t_1) \le \alpha_R \cdot (t_2 - t_1) - \lambda_R \int_{t_1}^{t_2} \psi(s)ds, \quad 0 \le t_1 < t_2 < \infty. \tag{A.4.202}$$

Now we check that $t \to \psi(t)$ is left-continuous on $[0, \infty)$. Since $t \to I\{t \le t^\epsilon_R\}$ is clearly left-continuous on $[0, \infty)$, and $t \to \varphi(y^\epsilon(t \wedge t^\epsilon_R))$ is continuous on $[0, \infty)$, we see that

$t \to I\{t \le t^\epsilon_R\}\varphi(y^\epsilon(t \wedge t^\epsilon_R))$ is left-continuous on $[0, \infty)$. Now fix $\{t_n\} \subset [0, \infty)$ such that $\lim_{n\to\infty} t_n = t < \infty$, with $t_n \le t$. Clearly

$$I\{t_n \le t^\epsilon_R\}\,\varphi\left(y^\epsilon(t_n \wedge t^\epsilon_R)\right) \le \max_{0\le s\le t} \varphi\left(y^\epsilon(s)\right) \tag{A.4.203}$$

for all $n \in \mathbb{N}$. By Lemma C.0.6 on page 203 and linear boundedness of the coefficients in (3.1.1) and (3.1.2) (recall Condition 3.2.1) we have

$$E\left[\max_{0\le s\le t} \varphi\left(y^\epsilon(s)\right)\right] < \infty. \tag{A.4.204}$$

From (A.4.203), (A.4.204), Lebesgue Dominated Convergence Theorem and left-continuity of $t \to I\{t \le t_R^\epsilon\}\varphi(y^\epsilon(t \wedge t_R^\epsilon))$ it follows that $t \to \varphi(t)$ is left-continuous on $[0, \infty)$. We must now use (A.4.202) to upper-bound $\varphi(t)$. Notice that we can not apply the Gronwall inequality directly since $-\lambda_R \le 0$. Now put

$$v(t) \triangleq \varphi(t) - u(t), \qquad \forall t \in [0, \infty), \tag{A.4.205}$$

where $u(t)$ solves the equation

$$\dot{u}(t) = -\lambda_R u(t) + \alpha_R, \qquad u(0) \triangleq |y_0|^{q_4}. \tag{A.4.206}$$

Then the solution of (A.4.206) is clearly given by

$$u(t) = e^{-\lambda_R t} \cdot |y_0|^{q_4} + \frac{\alpha_R}{\lambda_R}\left(1 - e^{-\lambda_R t}\right), \quad \forall t \in [0, \infty). \tag{A.4.207}$$

Also, by (A.4.206), we can write

$$u(t_2) - u(t_1) = \alpha_R(t_2 - t_1) - \lambda_R \int_{t_1}^{t_2} u(s)ds. \tag{A.4.208}$$

In view of (A.4.208) and (A.4.202), one has

$$\begin{aligned} [\varphi(t_2) - \varphi(t_1)] &- [u(t_2) - u(t_1)] \\ &\le \left[\alpha_R(t_2 - t_1) - \lambda_R \int_{t_1}^{t_2} \varphi(s)ds\right] \\ &\quad - \left[\alpha_R(t_2 - t_1) - \lambda_R \int_{t_1}^{t_2} u(s)ds\right], \end{aligned} \tag{A.4.209}$$

$\forall 0 \le t_1 < t_2 < \infty$. Hence, from (A.4.209) and (A.4.205):

$$v(t_2) - v(t_1) \le -\lambda_R \int_{t_1}^{t_2} v(s)ds, \quad \forall 0 \le t_1 < t_2 < \infty, \tag{A.4.210}$$

where $t \to v(t)$ is clearly left-continuous and $v(0) = 0$. Then it is easy to see that $v(t) \le 0, \forall t \in [0, \infty)$. Indeed, we argue to contrary: let

$$A \triangleq \{t \in [0, \infty),\ v(t) > 0\}. \tag{A.4.211}$$

We must show that $A = \emptyset$, thus suppose $A \neq \emptyset$, and let $t_0 \in A$, i.e. $v(t_0) > 0$. Put

$$t_* \triangleq \sup \{ s \in [0, t_0] : v(s) \leq 0 \} . \tag{A.4.212}$$

Since $v(.)$ is left-continuous there is $\delta > 0$ such that

$$v(t) > 0, \quad \forall t \in (t_0 - \delta, t_0] \tag{A.4.213}$$

In view of (A.4.213) clearly

$$t_* \leq t_0 - \delta < t_0. \tag{A.4.214}$$

Also, by definition of t_* and the fact that $v(.)$ is left-continuous, one sees that $v(t_*) \leq 0$ and $v(s) > 0, \forall s \in (t_*, t_0]$. Thus, from (A.4.210) and $\lambda_R \in (0, \infty)$(see Condition 3.2.18),

$$v(t) \leq v(t_*) - \lambda_R \int_{t_*}^{t} v(s) ds < v(t_*) \leq 0, \quad \forall t \in (t_*, t_0], \tag{A.4.215}$$

which contradicts the fact that $v(s) > 0, \forall s \in (t_*, t_0]$. Thus $A = \emptyset$, hence $v(t) \leq 0, \forall t \in [0, \infty)$, hence from (A.4.205) and (A.4.207)

$$\varphi(t) \leq u(t) = e^{-\lambda_R t} |y_0|^{q_4} + (\alpha_R / \lambda_R) \left[1 - e^{-\lambda_R t} \right], \quad \forall t \in [0, \infty), \quad \forall \epsilon \in (0, 1]. \tag{A.4.216}$$

Now from (A.4.216) and (A.4.201) we have

$$\begin{aligned} &E \left[I\{ t \leq t_R^\epsilon \} |y^\epsilon(t)|^{q_4} \right] \\ &\qquad \leq e^{-\lambda_R t} |y_0|^{q_4} + (\alpha_R / \lambda_R) \left[1 - e^{-\lambda_R t} \right], \quad \forall t \in [0, \infty), \quad \forall \epsilon \in (0, 1], \end{aligned} \tag{A.4.217}$$

thus there is a constant $C(R) \in [0, \infty)$ such that

$$E \left[I\{ t \leq t_R^\epsilon \} |y^\epsilon(t)|^{q_4} \right] \leq C(R), \quad \forall t \in [0, \infty), \quad \forall \epsilon \in (0, 1], \tag{A.4.218}$$

thus

$$E \left[I\{ \tau\epsilon^{-1} \leq t_R^\epsilon \} \left| y^\epsilon(\tau\epsilon^{-1}) \right|^{q_4} \right] \leq C(R), \quad \forall \tau \in [0, 1], \quad \forall \epsilon \in (0, 1], \tag{A.4.219}$$

Now (A.4.188) follows from (A.4.219), (A.0.3), and (3.3.40).

□

The following result is used for the proof of Proposition A.4.6 on page 117.

Corollary A.4.8. *Suppose that Conditions 3.2.1, 3.2.2 and 3.2.18 hold. Then, for each $R \in (0, \infty)$, there is a constant $C(R) \in [0, \infty)$ such that*

$$E\left[\max_{\tau \in [0,1]} I\left\{\tau \leq T_R^\epsilon\right\} |Y^\epsilon(\tau)|^{q_4}\right] \leq \frac{C(R)}{\epsilon}, \qquad \forall \epsilon \in (0, 1]. \tag{A.4.220}$$

Proof: Fix $R \in (0, \infty)$, and put

$$U_n^{\epsilon,R} \triangleq \max_{n \leq t \leq n+1} I\{t \leq t_R^\epsilon\} |y^\epsilon(t)|^{q_4}, \quad n = 0, 1, 2, \ldots, \quad \forall \epsilon \in (0, 1], \tag{A.4.221}$$

where $y^\epsilon(t)$ is given by (3.1.2) and t_R^ϵ is the stopping time in Remark A.0.26. Clearly, from (A.4.221) we have

$$\max_{0 \leq t \leq \frac{1}{\epsilon}} I\{t \leq t_R^\epsilon\} |y^\epsilon(t)|^{q_4} \leq \sum_{n=0}^{[\epsilon^{-1}]} U_n^{\epsilon,R}. \tag{A.4.222}$$

Hence

$$E\left[\max_{0 \leq t \leq \frac{1}{\epsilon}} I\{t \leq t_R^\epsilon\} |y^\epsilon(t)|^{q_4}\right] \leq \sum_{n=0}^{[\epsilon^{-1}]} E\left[U_n^{\epsilon,R}\right]. \tag{A.4.223}$$

Now, it is enough to find an upper bound for $U_n^{\epsilon,R}$ as follows: one can write

$$|y^\epsilon(t)| \leq |y^\epsilon(n)| + |y^\epsilon(t) - y^\epsilon(n)|, \quad \forall t \in [n, n+1], \ \forall n = 0, 1, 2, \ldots, \ \forall \epsilon \in (0, 1]. \tag{A.4.224}$$

Hence for each $\epsilon \in (0, 1]$ and $n = 0, 1, 2, \ldots$

$$|y^\epsilon(t)|^{q_4} \leq 2^{q_4}\left[|y^\epsilon(n)|^{q_4} + |y^\epsilon(t) - y^\epsilon(n)|^{q_4}\right], \quad \forall t \in [n, n+1], \tag{A.4.225}$$

thus from (A.4.225),

$$\begin{aligned} & I\{t \le t_R^\epsilon\} \, |y^\epsilon(t)|^{q_4} \\ & \le 2^{q_4} \left[I\{t \le t_R^\epsilon\} \, |y^\epsilon(n)|^{q_4} + I\{t \le t_R^\epsilon\} \, |y^\epsilon(t) - y^\epsilon(n)|^{q_4} \right] \\ & \le 2^{q_4} \left[I\{n \le t_R^\epsilon\} \, |y^\epsilon(n)|^{q_4} + I\{t \le t_R^\epsilon\} \, |y^\epsilon(t) - y^\epsilon(n)|^{q_4} \right], \qquad \forall t \in [n, n+1]. \end{aligned} \tag{A.4.226}$$

The second inequality in (A.4.226) follows because, for each $t \in [n, n+1], n = 0, 1, 2, \dots$, we have

$$I\{t \le t_R^\epsilon\} \le I\{n \le t_R^\epsilon\}, \quad \forall \epsilon \in (0, 1]. \tag{A.4.227}$$

From (A.4.226), we have

$$\begin{aligned} & \max_{t \in [n, n+1]} I\{t \le t_R^\epsilon\} \, |y^\epsilon(t)|^{q_4} \\ & \qquad \le 2^{q_4} \left[I\{n \le t_R^\epsilon\} \, |y^\epsilon(n)|^{q_4} + \max_{t \in [n, n+1]} I\{t \le t_R^\epsilon\} \, |y^\epsilon(t) - y^\epsilon(n)|^{q_4} \right], \end{aligned} \tag{A.4.228}$$

$\forall \epsilon \in (0, 1]$, $\forall n = 0, 1, 2, \dots$. Thus, from (A.4.221), (A.4.223), and (A.4.228), one observes that

$$\begin{aligned} & E\left[\max_{0 \le t \le \frac{1}{\epsilon}} I\{t \le t_R^\epsilon\} \, |y^\epsilon(t)|^{q_4} \right] \\ & \le \sum_{n=0}^{[\epsilon^{-1}]} E\left[\max_{t \in [n, n+1]} I\{t \le t_R^\epsilon\} \, |y^\epsilon(t)|^{q_4} \right] \\ & \le \sum_{n=0}^{[\epsilon^{-1}]} 2^{q_4} \left\{ E\left[I\{n \le t_R^\epsilon\} \, |y^\epsilon(n)|^{q_4} \right] + E\left[\max_{t \in [n, n+1]} |y^\epsilon(t) - y^\epsilon(n)|^{q_4} \, I\{t \le t_R^\epsilon\} \right] \right\}, \end{aligned} \tag{A.4.229}$$

$\forall \epsilon \in (0,1], \forall n = 0,1,2,\ldots$. By Lemma A.4.7 on page 123, for the first expectation of the right side of (A.4.229), we get

$$\begin{aligned} E\left[I\{n \le t_R^\epsilon\} |y^\epsilon(n)|^{q_4}\right] &= E\left[I\{\epsilon n \le \epsilon t_R^\epsilon\} \left|y^\epsilon(\frac{\epsilon n}{\epsilon})\right|^{q_4}\right] \\ &= E\left[I\{\epsilon n \le T_R^\epsilon\} |Y^\epsilon(\epsilon n)|^{q_4}\right] \\ &\le C_1(R), \end{aligned} \tag{A.4.230}$$

$\forall \epsilon \in (0,1], \forall n = 0,1,2,\ldots$. Also, for each $t \in [n, n+1]$, we have

$$\begin{aligned} &I\{t \le t_R^\epsilon\} |y^\epsilon(t) - y^\epsilon(n)|^{q_4} \\ &\quad = I\{t \le t_R^\epsilon\} |y^\epsilon(t \wedge t_R^\epsilon) - y^\epsilon(n \wedge t_R^\epsilon)|^{q_4} \\ &\quad \le |y^\epsilon(t \wedge t_R^\epsilon) - y^\epsilon(n \wedge t_R^\epsilon)|^{q_4}. \end{aligned} \tag{A.4.231}$$

Hence, using (3.1.2), for the second expectation of the right side of (A.4.229), one has

$$\begin{aligned} &E\left[\max_{t\in[n,n+1]} I\{t \le t_R^\epsilon\} |y^\epsilon(t) - y^\epsilon(n)|^{q_4}\right] \\ &\quad \le E\left[\max_{t\in[n,n+1]} |y^\epsilon(t \wedge t_R^\epsilon) - y^\epsilon(n \wedge t_R^\epsilon)|^{q_4}\right] \\ &\quad \le 2^{q_4} E\left\{\max_{t\in[n,n+1]} \left|\int_{n\wedge t_R^\epsilon}^{t\wedge t_R^\epsilon} b(x^\epsilon, y^\epsilon)(s)ds\right|^{q_4}\right. \\ &\qquad \left. + \max_{t\in[n,n+1]} \left|\int_{n\wedge t_R^\epsilon}^{t\wedge t_R^\epsilon} \sigma(x^\epsilon, y^\epsilon)(s)d\beta(s)\right|^{q_4}\right\}. \end{aligned} \tag{A.4.232}$$

Now when $t \in [n, n+1]$ we have

$$\begin{aligned} \int_{n\wedge t_R^\epsilon}^{t\wedge t_R^\epsilon} b(x^\epsilon, y^\epsilon)(s)ds &= 0, \quad \text{when } t_R^\epsilon \le n, \\ \int_{n\wedge t_R^\epsilon}^{t\wedge t_R^\epsilon} b(x^\epsilon, y^\epsilon)(s)ds &= \int_n^{t_R^\epsilon} b(x^\epsilon, y^\epsilon)(s)ds, \quad \text{when } n \le t_R^\epsilon \le t, \\ \int_{n\wedge t_R^\epsilon}^{t\wedge t_R^\epsilon} b(x^\epsilon, y^\epsilon)(s)dsds &= \int_n^t b(x^\epsilon, y^\epsilon)(s)ds, \quad \text{when } t < t_R^\epsilon \end{aligned} \tag{A.4.233}$$

From these three cases,

$$\int_{n\wedge t_R^\epsilon}^{t\wedge t_R^\epsilon} b(x^\epsilon, y^\epsilon)(s)ds = \int_n^t I\{s \le t_R^\epsilon\} b(x^\epsilon, y^\epsilon)(s)ds. \tag{A.4.234}$$

Similarly

$$\int_{n\wedge t_R^\epsilon}^{t\wedge t_R^\epsilon} \sigma(x^\epsilon, y^\epsilon)(s)ds = \int_n^t I\{s \le t_R^\epsilon\} \sigma(x^\epsilon, y^\epsilon)(s)d\beta(s). \tag{A.4.235}$$

Thus, from (A.4.232), (A.4.234) and (A.4.235)

$$\begin{aligned} E&\left[\max_{t\in[n+n+1]} I\{t \le t_R^\epsilon\} |y^\epsilon(t) - y^\epsilon(n)|^{q_4}\right] \\ &\le 2^{q_4} E\left\{\max_{t\in[n,n+1]} \left|\int_n^t I\{s \le t_R^\epsilon\} b(x^\epsilon, y^\epsilon)(s)ds\right|^{q_4} \right. \\ &\qquad \left. + \max_{t\in[n,n+1]} \left|\int_n^t I\{s \le t_R^\epsilon\} \sigma(x^\epsilon, y^\epsilon)(s)d\beta(s)\right|^{q_4}\right\} \\ &\le 2^{q_4}\left\{E\left[\int_n^{n+1} I\{s \le t_R^\epsilon\} |b(x^\epsilon, y^\epsilon)(s)|\, ds\right]^{q_4} \right. \\ &\qquad \left. + E\left[\max_{t\in[n,n+1]} \left|\int_n^t I\{s \le t_R^\epsilon\} \sigma(x^\epsilon, y^\epsilon)(s)d\beta(s)\right|^{q_4}\right]\right\} \end{aligned} \tag{A.4.236}$$

$\forall \epsilon \in (0, 1], \forall n = 0, 1, 2, \ldots$. By Condition 3.2.1 and Jensen's inequality, there are constants $C_1, C_2(R) \in [0, \infty)$ such that

$$\begin{aligned} E&\left(\int_n^{n+1} I\{s \le t_R^\epsilon\} |b(x^\epsilon, y^\epsilon)(s)|\, ds\right)^{q_4} \\ &\le E\left[\int_n^{n+1} I\{s \le t_R^\epsilon\} |b(x^\epsilon, y^\epsilon)(s)|^{q_4}\, ds\right] \\ &\le C_1 E\left[\int_n^{n+1} I\{s \le t_R^\epsilon\} (1 + |x^\epsilon(s)| + |y^\epsilon(s)|)^{q_4}\, ds\right] \\ &\le 3^{q_4} C_1 \int_n^{n+1} E\left[1 + I\{s \le t_R^\epsilon\} |x^\epsilon(s)|^{q_4} + I\{s \le t_R^\epsilon\} |y^\epsilon(s)|^{q_4}\right] ds \\ &\le 3^{q_4} C_1 \int_n^{n+1} \left[1 + |R|^{q_4} + EI\{s \le t_R^\epsilon\} |y^\epsilon(s)|^{q_4}\right] ds \\ &\le 3^{q_4} C_2(R) \int_n^{n+1} \left[1 + EI\{s \le t_R^\epsilon\} |y^\epsilon(s)|^{q_4}\right] ds \end{aligned} \tag{A.4.237}$$

$\forall \epsilon \in (0,1], \forall n = 0,1,2,\ldots$. Clearly, from (A.4.237), Lemma A.4.7 on page 123, (A.0.3) and (3.3.40), it follows that

$$\begin{aligned} E\left(\int_n^{n+1} I\{s \le t_R^\epsilon\}\, |b(x^\epsilon, y^\epsilon)(s)|\, ds\right)^{q_4} &\le 3^{q_4} C_2(R) \int_n^{n+1} \left[1 + E\left[I\{s \le \frac{T_R^\epsilon}{\epsilon}\} \left|y^\epsilon(\frac{\epsilon s}{\epsilon})\right|^{q_4}\right]\right] ds \\ &\le 3^{q_4} C_2(R) \int_n^{n+1} \left[1 + E\left[I\{\epsilon s \le T_R^\epsilon\}\, |Y^\epsilon(\epsilon s)|^{q_4}\right]\right] ds \\ &\le 3^{q_4} C_2(R)\, [1 + C_3(R)] \\ &\le C_4(R), \end{aligned} \tag{A.4.238}$$

$\forall \epsilon \in (0,1], \forall n = 0,1,2,\ldots$, and constants $C_3(R), C_4(R) \in [0,\infty)$. Also, by Jensen and Burkholder inequalities, there is a constant $C_5 \in [0,\infty)$ such that

$$\begin{aligned} E\left[\max_{t\in[n,n+1]} \left|\int_n^t I\{s \le t_R^\epsilon\} \sigma(x^\epsilon, y^\epsilon)(s) d\beta(s)\right|^{q_4}\right] &\le C_5 E\left[\left(\int_n^t I\{s \le t_R^\epsilon\} \operatorname{Trace}\left(\sigma\sigma^T(x^\epsilon, y^\epsilon)(s)\right) ds\right)^{q_4/2}\right] \\ &\le C_5 E\left[\int_n^t I\{s \le t_R^\epsilon\} \left(\operatorname{Trace}\left(\sigma\sigma^T(x^\epsilon, y^\epsilon)(s)\right)\right)^{q_4/2} ds\right] \end{aligned} \tag{A.4.239}$$

$\forall \epsilon \in [0,\infty), \forall n = 0,1,2,\ldots$. From Condition 3.2.1 we easily see that

$$\left(\operatorname{Trace}\left(\sigma\sigma^T(x,y)\right)\right)^{q_4/2} \le C\left[1 + |x|^{q_4} + |y|^{q_4}\right], \tag{A.4.240}$$

$\forall(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$, for a constant $C \in [0,\infty)$. Thus, combining (A.4.240), (A.4.239), (A.0.3), (3.3.40) and Lemma A.4.7 on page 123 we have

$$\begin{aligned} E\left[\max_{t\in[n,n+1]}\left|\int_n^t I\{s\le t_R^\epsilon\}\sigma(x^\epsilon,y^\epsilon)(s)d\beta(s)\right|^{q_4}\right] &\le C_5 E\left[\int_n^t I\{s\le t_R^\epsilon\}\left(1+|x^\epsilon(s)|+|y^\epsilon(s)|\right)^{q_4} ds\right] \\ &\le 3^{q_4}C_6(R)\int_n^t \left[1+E\left(I\{s\le t_R^\epsilon\}\,|y^\epsilon(s)|^{q_4}\right)\right]ds \\ &\le 3^{q_4}C_6(R)\int_n^t \left[1+E\left(I\{s\le \frac{T_R^\epsilon}{\epsilon}\}\left|y^\epsilon(\frac{\epsilon s}{\epsilon})\right|^{q_4}\right)\right]ds \\ &\le 3^{q_4}C_6(R)\int_n^t \left[1+E\left(I\{\epsilon s\le T_R^\epsilon\}\,|Y^\epsilon(s)|^{q_4}\right)\right]ds \\ &\le C_7(R), \end{aligned} \tag{A.4.241}$$

$\forall\epsilon \in (0,1], \forall n = 0,1,2,\ldots$, and constants $C_6(R), C_7(R) \in [0,\infty)$. Now, from (A.4.236), (A.4.238) and (A.4.241), there is a constant $C_8(R) \in [0,\infty)$ such that

$$E\left[\max_{t\in[n,n+1]} I\{t\le t_R^\epsilon\}\,|y^\epsilon(t)-y^\epsilon(n)|^{q_4}\right] \le C_8(R), \tag{A.4.242}$$

$\forall\epsilon \in (0,1], \forall n = 0,1,2,\ldots$. Thus, from (A.4.229), (A.4.230), (A.4.242), it follows that

$$E\left[\max_{0\le t\le\frac{1}{\epsilon}} I\{t\le t_R^\epsilon\}\,|y^\epsilon(t)|^{q_4}\right] \le \frac{C_9(R)}{\epsilon}, \tag{A.4.243}$$

$\forall\epsilon \in (0,1]$, and some constant $C_9(R) \in [0,\infty)$. In view of (A.4.243), (A.0.3), and (3.3.40) we get

$$\begin{aligned} E\left[\max_{0\le\tau\le 1} I\{\tau\le T_R^\epsilon\}\,|Y^\epsilon(\tau)|^{q_4}\right] &= E\left[\max_{0\le t\le 1/\epsilon} I\{t\le t_R^\epsilon\}\,|y^\epsilon(t)|^{q_4}\right] \\ &\le \frac{C_9(R)}{\epsilon}, \quad \forall\epsilon\in(0,1], \end{aligned} \tag{A.4.244}$$

as required.

□

Appendix B

Proofs for Section 3.4

Proof of Lemma 3.4.5 on page 47: The proof is just a tedious but elementary computation. Fix $R \in [0, \infty)$, $x \in S_R^d$. Using the Frobenius norm (see I of "Basic Notation and Terminology"), for all $\xi_1, \xi_2 \in \mathbb{R}^D$, we have

$$|\sigma(x, \xi_1) - \sigma(x, \xi_2)|^2 = \sum_{n=1}^{N} \sum_{k=1}^{D} \left|\sigma^{k,n}(x, \xi_1) - \sigma^{k,n}(x, \xi_2)\right|^2 . \qquad \text{(B.0.1)}$$

By the mean value theorem, we expand each term on the right side of (B.0.1):

$$\sigma^{k,n}(x, \xi_1) - \sigma^{k,n}(x, \xi_2) = \left[\int_0^1 (\partial_\xi \sigma^{k,n})(x, \xi_2 + \alpha(\xi_1 - \xi_2))\, d\alpha\right] (\xi_1 - \xi_2) . \qquad \text{(B.0.2)}$$

for $k = 1, 2, \ldots D, n = 1, 2, \ldots N$, and $\forall \xi_1, \xi_2 \in \mathbb{R}^D$. Thus, by Jensen inequality and (B.0.2), we get

$$\begin{aligned} &\left|\sigma^{k,n}(x, \xi_1) - \sigma^{k,n}(x, \xi_2)\right|^2 \\ &\qquad \le \int_0^1 \left|(\partial_\xi \sigma^{k,n})(x, \xi_2 + \alpha(\xi_1 - \xi_2))(\xi_1 - \xi_2)\right|^2 d\alpha \end{aligned} \qquad \text{(B.0.3)}$$

From (B.0.3) and using Cauchy-Schwarz inequality, it follows that

$$\begin{aligned}
&\left|\sigma^{k,n}(x,\xi_1) - \sigma^{k,n}(x,\xi_2)\right|^2 \\
&\quad \le \int_0^1 \left| \sum_{l=1}^{D} \left\{ \left(\partial_{\xi^l}\sigma^{k,n}\right)(x, \xi_2 + \alpha(\xi_1 - \xi_2)) \left({\xi_1}^l - {\xi_2}^l\right) \right\} \right|^2 d\alpha \\
&\quad \le \left[\int_0^1 \sum_{l=1}^{D} \left| \left(\partial_{\xi^l}\sigma^{k,n}\right)(x, \xi_2 + \alpha(\xi_1 - \xi_2)) \right|^2 d\alpha \right] |\xi_1 - \xi_2|^2
\end{aligned} \tag{B.0.4}$$

$\forall \xi_1, \xi_2 \in I\!R^D$. By (B.0.1) and (B.0.4), we can write

$$\begin{aligned}
&|\sigma(x,\xi_1) - \sigma(x,\xi_2)|^2 \\
&\quad \le \left[\int_0^1 \sum_{n=1}^{N} \sum_{k=1}^{D} \sum_{l=1}^{D} \left| \left(\partial_{\xi^l}\sigma^{k,n}\right)(x, \xi_2 + \alpha(\xi_1 - \xi_2)) \right|^2 d\alpha \right] |\xi_1 - \xi_2|^2 \\
&\quad \le \left[\int_0^1 \sum_{l=1}^{D} \left| \left(\partial_{\xi^l}\sigma\right)(x, \xi_2 + \alpha(\xi_1 - \xi_2)) \right|^2 d\alpha \right] |\xi_1 - \xi_2|^2
\end{aligned} \tag{B.0.5}$$

Now put

$$\Lambda_0(R) \triangleq \left[\sup_{(x,\xi) \in S_R^d \otimes I\!R^D} \left\{ \sum_{l=1}^{D} \left| (\partial_{\xi^l}\sigma)(x,\xi) \right|^2 \right\} \right]^{1/2}, \quad \forall R \in [0,\infty). \tag{B.0.6}$$

Thus, in view of (B.0.5) and (B.0.6), we have

$$|\sigma(x,\xi_1) - \sigma(x,\xi_2)|^2 \le \Lambda_0^2(R)\, |\xi_1 - \xi_2|^2, \quad \forall x \in S_R^d, \quad \forall \xi_1, \xi_2 \in I\!R^D, \tag{B.0.7}$$

as required for (3.4.63). Finally, from (B.0.6) and Condition 3.4.4, we see that $\Lambda_0(R) < \infty, \forall R \in [0,\infty)$.

□

Proof of Proposition 3.4.12 on page 50: The proof just uses easy calculus. Consider the second order differential operator $\mathcal{A}(x,y)$ defined by (3.2.14); recall

$$\mathcal{A}\varphi(x,y) \triangleq \sum_{i=1}^{D} b^i(x,y)(\partial_{y^i}\varphi)(y) + \frac{1}{2} \sum_{i=1}^{D} \sum_{j=1}^{D} \left[\sigma\sigma^T(x,y)\right]^{i,j} (\partial_{y^i}\partial_{y^j}\varphi)(y), \tag{B.0.8}$$

$\forall(x,y) \in I\!R^d \otimes I\!R^D$. Using the mean value theorem, for the first term on the right hand side of (B.0.8) we have

$$b^i(x,y) = b^i(x,0) + \int_0^1 (\partial_y b^i)(x,\alpha y)y \, d\alpha, \quad \forall i = 1,2,\ldots,D. \tag{B.0.9}$$

Put

$$J(x,y) \triangleq (\partial_y b)(x,y), \quad \forall(x,y) \in I\!R^d \otimes I\!R^D. \tag{B.0.10}$$

Thus, from (B.0.9) and (B.0.10):

$$b(x,y) = b(x,0) + \int_0^1 J(x,\alpha y)y \, d\alpha, \quad \forall(x,y) \in I\!R^d \otimes I\!R^D. \tag{B.0.11}$$

Hence

$$y^T b(x,y) = y^T b(x,0) + \int_0^1 y^T J(x,\alpha y)y \, d\alpha, \tag{B.0.12}$$

$\forall(x,y) \in I\!R^d \otimes I\!R^D$. Now, obviously, one has

$$\begin{aligned} y^T J(x,\alpha y)y &= y^T J^T(x,\alpha y)y \\ &= \frac{1}{2} y^T \left[J(x,\alpha y) + J^T(x,\alpha y)\right] y, \end{aligned} \tag{B.0.13}$$

$\forall(x,y) \in I\!R^d \otimes I\!R^D, \forall\alpha \in [0,1]$. Fix arbitrary $R \in [0,\infty)$. Then Remark 3.4.6, (3.4.66), and Lemma C.0.12 on page 205 for the symmetric matrix $\frac{1}{2}\left[J(x,\alpha y) + J^T(x,\alpha y)\right]$ in (B.0.13) ensure that

$$\begin{aligned} y^T J(x,\alpha y)y &\leq \text{maximum eigenvalue}\left\{1/2[J(x,\alpha y) + J^T(x,\alpha y)]\right\}|y|^2 \\ &\leq \Lambda_1(R)\,|y|^2, \quad \forall\alpha \in [0,1], \quad \forall(x,y) \in S_R^d \otimes I\!R^D, \end{aligned} \tag{B.0.14}$$

where $\Lambda_1(R)$ is defined by (3.4.66) in Remark 3.4.6. Thus, from (B.0.12) and (B.0.14), we get

$$\begin{aligned} y^T b(x,y) &\leq y^T b(x,0) + \Lambda_1(R)\,|y|^2 \\ &\leq |b(x,0)|\,|y| + \Lambda_1(R)\,|y|^2, \end{aligned} \tag{B.0.15}$$

$\forall(x,y) \in S_R^d \otimes I\!R^D$. For use here and later, we record the following elementary

Fact B.0.9. *For $\varphi : \mathbb{R}^D \to \mathbb{R}$ given by $\varphi(y) = |y|^q$, $q \in [2, \infty)$ a constant, we have $\varphi \in C^2(\mathbb{R}^D)$ with*

$$(\partial_{y^i}\varphi)(y) = qy^i |y|^{q-2}, \quad \forall y \in \mathbb{R}^D \tag{B.0.16}$$

$$(\partial_{y^i}\partial_{y^j}\varphi)(y) = q|y|^{q-2}\delta_{i,j} + q(q-2)y^i y^j |y|^{q-4}, \quad \forall y \in \mathbb{R}^D, \tag{B.0.17}$$

(where $\delta_{i,j}$ is the Kronecker δ).

Hence, from (B.0.15) and (B.0.16) one can write

$$\begin{aligned} \sum_{i=1}^{D} b^i(x,y)(\partial_{y^i}\varphi)(y) &= q|y|^{q-2}\sum_{i=1}^{D} y^i b^i(x,y) \\ &= q|y|^{q-2} y^T b(x,y) \\ &\le q|b(x,0)|\,|y|^{q-1} + q\Lambda_1(R)|y|^q, \end{aligned} \tag{B.0.18}$$

$\forall (x, y) \in S_R^d \otimes \mathbb{R}^D$. Now consider the second term on the right hand side of (B.0.8): From (B.0.17) it follows that

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^{D}\sum_{j=1}^{D}\left[\sigma(x,y)\sigma^T(x,y)\right]^{i,j}(\partial_{y^i}\partial_{y^j}\varphi)(y) \\ &= \frac{1}{2}\sum_{i=1}^{D}\sum_{j=1}^{D}\left[q|y|^{q-2}\delta_{i,j} + q(q-2)y^i y^j|y|^{q-4}\right]\left[\sigma\sigma^T(x,y)\right]^{i,j} \\ &= \frac{1}{2}\sum_{i=1}^{D} q|y|^{q-2}\left[\sigma\sigma^T(x,y)\right]^{i,i} \\ &\qquad + \frac{1}{2}\sum_{i=1}^{D}\sum_{j=1}^{D} q(q-2)y^i y^j |y|^{q-4}\left[\sigma\sigma^T(x,y)\right]^{i,j} \\ &= \frac{q}{2}|y|^{q-2}\,\mathrm{Trace}\left[\sigma\sigma^T(x,y)\right] \\ &\qquad + \frac{q}{2}(q-2)|y|^{q-4}\cdot y^T\left[\sigma\sigma^T(x,y)\right]y, \end{aligned} \tag{B.0.19}$$

$\forall (x,y) \in S_R^d \otimes \mathbb{R}^D$. By Lemma C.0.12 on page 205 we have

$$\begin{aligned} y^T[\sigma\sigma^T(x,y)]y &\leq \text{maximum eigenvalue}[\sigma\sigma^T(x,y)]|y|^2 \\ &\leq \text{Trace}[\sigma\sigma^T(x,y)]|y|^2, \quad \forall (x,y) \in S_R^d \otimes \mathbb{R}^D. \end{aligned} \tag{B.0.20}$$

In view of (B.0.20) and (B.0.19) one has

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^{D}\sum_{j=1}^{D}\left[\sigma(x,y)\sigma^T(x,y)\right]^{ij}(\partial_{y^i}\partial_{y^j}\varphi)(y) \\ &\qquad \leq \frac{q}{2}|y|^{q-2}\,\text{Trace}\left[\sigma\sigma^T(x,y)\right] \\ &\qquad\qquad + \frac{q}{2}(q-2)\,|y|^{q-2}\,\text{Trace}\left[\sigma\sigma^T(x,y)\right] \\ &\qquad \leq \frac{q}{2}(q-1)\,|y|^{q-2}\,\text{Trace}\left[\sigma\sigma^T(x,y)\right] \\ &\qquad \leq \frac{q}{2}(q-1)\,|y|^{q-2}\,|\sigma(x,y)|^2, \end{aligned} \tag{B.0.21}$$

$\forall (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Also, (3.4.63) together with triangle inequality for the Frobenius norm $|.|$ gives

$$\begin{aligned} |\sigma(x,y)| &\leq |\sigma(x,0)| + |\sigma(x,y) - \sigma(x,0)| \\ &\leq |\sigma(x,0)| + \Lambda_0(R)|y|, \quad \forall (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D. \end{aligned} \tag{B.0.22}$$

Thus, from (B.0.22) and (B.0.21), we have

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^{D}\sum_{j=1}^{D}\left[\sigma(x,y)\sigma^T(x,y)\right]^{ij}(\partial_{y^i}\partial_{y^j}\varphi)(y) \\ &\qquad \leq \frac{q}{2}(q-1)\left(\Lambda_0^2(R)|y|^q + 2\Lambda_0(R)\,|\sigma(x,0)|\,|y|^{q-1} + |\sigma(x,0)|^2\,|y|^{q-2}\right), \end{aligned} \tag{B.0.23}$$

$\forall (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Define

$$C_1(R) \triangleq \max_{x\in S_R^d} |b(x,0)|,$$
$$C_2(R) \triangleq \max_{x\in S_R^d} |\sigma(x,0)|. \tag{B.0.24}$$

Combining (B.0.8), (B.0.18), (B.0.23) and (B.0.24) we get

$$\begin{aligned} \mathcal{A}\varphi(x,y) \quad \leq \quad & [qC_1(R) + q(q-1)\Lambda_0(R)C_2(R)]\,|y|^{q-1} \\ & + \left[q\Lambda_1(R) + \frac{q}{2}(q-1)\Lambda_0^2(R)\right]|y|^q \\ & + \frac{q}{2}(q-1)C_2(R)^2|y|^{q-2}. \end{aligned} \tag{B.0.25}$$

We now require the following elementary Fact B.0.10.

Fact B.0.10. *Suppose* $0 < C_1 < C_2 < \infty$. *Then, for each* $\delta \in (0,\infty)$, *one has*

$$|z|^{C_1} \leq \delta^{C_2 - C_1}|z|^{C_2} + (1/\delta)^{C_1}, \quad \forall z \in \mathbb{R}^D. \tag{B.0.26}$$

Proof of Fact B.0.10: To see this, observe that

$$\zeta^{C_1} \leq 1 + \zeta^{C_2}, \quad \forall \zeta \in [0,\infty). \tag{B.0.27}$$

Fix $\delta \in (0,\infty)$ and put $\zeta \triangleq \delta\,|z|$, for $z \in \mathbb{R}^D$. Thus, by (B.0.26), we have

$$\delta^{C_1}|z|^{C_1} \leq 1 + \delta^{C_2}|z|^{C_2}, \quad \forall z \in \mathbb{R}^D, \tag{B.0.28}$$

hence

$$|z|^{C_1} \leq \delta^{C_2 - C_1}|z|^{C_2} + (1/\delta)^{C_1}, \quad \forall z \in \mathbb{R}^D. \tag{B.0.29}$$

□

Now fix $\delta \in (0,\infty)$. From Fact B.0.10 on page 141 we have

$$\begin{aligned} |y|^{q-1} & \leq \delta|y|^q + (1/\delta)^{q-1}, \ \forall y \in \mathbb{R}^D \\ |y|^{q-2} & \leq \delta^2|y|^q + (1/\delta)^{q-2}, \ \forall y \in \mathbb{R}^D \end{aligned} \tag{B.0.30}$$

Also, define

$$\begin{aligned} C_3(R,\delta) &\triangleq q\Lambda_1(R) + \frac{q}{2}(q-1)[\Lambda_0(R)]^2 + qC_1(R)\delta + \delta q(q-1)\Lambda_0(R)C_2(R) \\ &\qquad + \delta^2\frac{q}{2}(q-1)[C_2(R)]^2, \\ C_4(R,\delta) &\triangleq qC_1(R)(1/\delta)^{q-1} + q(q-1)\Lambda_0(R)C_2(R)(1/\delta)^{q-1} \\ &\qquad + \frac{q}{2}(q-1)[C_2(R)]^2(1/\delta)^{q-2} \end{aligned} \tag{B.0.31}$$

Thus, from (B.0.25) and (B.0.30), one gets

$$\mathcal{A}\varphi(x,y) \leq C_3(R,\delta)|y|^q + C_4(R,\delta), \quad \forall (x,y) \in S_R^d \otimes \mathbb{R}^D. \tag{B.0.32}$$

Now Condition 3.4.7 ensures that

$$q\Lambda_1(R) + \frac{q}{2}(q-1)[\Lambda_0(R)]^2 < 0. \tag{B.0.33}$$

Hence, for each $R \in [0,\infty)$, we can fix $\delta \triangleq \delta(R)$ small enough that $\lambda_R \triangleq -C_3(R,\delta(R)) > 0$. Now put $\alpha_R \triangleq C_4(R,\delta(R))$ in (B.0.32) to get (3.4.73).

□

Proof of Proposition 3.4.14 on page 50: Fix some $x \in \mathbb{R}^d$ and some $R \in [0,\infty)$ such that $x \in S_R^d$. From Condition 3.4.7 and Lemma 3.4.5 on page 47 we obviously have

$$\sup_{\xi \in \mathbb{R}^D} \lambda_{\max}(x,\xi) < -(1/2)\left[\Lambda_0(R)\right]^2. \tag{B.0.34}$$

and

$$|\sigma(x,\xi_1) - \sigma(x,\xi_2)| \leq \Lambda_0(R)\,|\xi_1 - \xi_2|\,, \qquad \forall \xi_1,\xi_2 \in \mathbb{R}^D.$$

Hence Theorem 3.4.1 on page 43 applied to the coefficients $b(x,.)$ and $\sigma(x,.)$ shows that there exists a unique invariant probability measure π_x on $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ for the

Markov process $\{\xi(t,x)\}$ defined by (3.4.74). We next establish (3.4.75). To this end, fix some $y \in \mathbb{R}^D$. From Proposition B.1.1 on page 175 we have

$$E|\xi(t,x,y)|^q \le |y|^q + C(q,R), \quad \forall t \in [0,\infty). \tag{B.0.35}$$

Now put

$$f^N(\xi) \triangleq |\xi|^q \wedge N, \quad \forall \xi \in \mathbb{R}^D, \;\; N = 1,2,3,\dots . \tag{B.0.36}$$

From (B.0.35) and (B.0.36),

$$Ef^N(\xi(t,x,y)) \le E|\xi(t,x,y)|^q \le |y|^q + C(q,R), \quad \forall t \in [0,\infty). \tag{B.0.37}$$

But, $f^N : \mathbb{R}^D \to \mathbb{R}$ is bounded and continuous, so that Theorem 3.4.1 on page 43 gives

$$\lim_{t\to\infty} Ef^N(\xi(t,x,y)) = \int_{\mathbb{R}^D} f^N(\xi)d\pi_x(\xi). \tag{B.0.38}$$

From (B.0.37) and (B.0.38),

$$\int_{\mathbb{R}^D} f^N(\xi)d\pi_x(\xi) \le |y|^q + C(q,R), \quad \forall N = 1,2,3,\dots . \tag{B.0.39}$$

Now take $N \to \infty$ in (B.0.39) and use the monotone convergence theorem to obtain (3.4.75). It remains to establish (3.4.76). Fix some $r_1 \in [0, q-1]$ and some $h \in Li(r_1)$. From Remark 3.4.9 we have

$$|h(\xi(t,x,y))| \le M_{r_1}(h)\left[1 + |\xi(t,x,y)|^{1+r_1}\right], \tag{B.0.40}$$

$\forall t \in [0,\infty), \forall y \in \mathbb{R}^D$. Clearly $r_2 \triangleq 1 + r_1 \le q$, hence from Proposition B.1.1 on page 175 there is a constant $C_1(R) \in [0,\infty)$ such that

$$E\,|\xi(t,x,y)|^{1+r_1} \le C_1(R)[1 + |y|^{1+r_1}] \tag{B.0.41}$$

$\forall t \in [0,\infty), \forall y \in \mathbb{R}^D$. Upon taking expectation in (B.0.40) and using (B.0.41) one has

$$E\,|h\,(\xi(t,x,y))| \le C_2(R)M_{r_1}(h)\left(1 + |y|^{1+r_1}\right), \tag{B.0.42}$$

$\forall t \in [0,\infty), \forall y \in I\!R^D$, where $C_2(R) \in [0,\infty)$ is a constant. Moreover, by Remark 3.4.9 we have

$$|h(\xi(t,x,y_1)) - h(\xi(t,x,y_2)| \\ \leq M_{r_1}(h)\,|\xi(t,x,y_1) - \xi(t,x,y_2)|\,[1 + |\xi(t,x,y_1)|^{r_1} + |\xi(t,x,y_2)|^{r_1}]\,. \tag{B.0.43}$$

$\forall t \in [0,\infty), \forall y_1, y_2 \in I\!R^D$. First suppose $r_1 = 0$, and take $r_2 \triangleq 1 + r_1$. Using (B.0.43), one has

$$\begin{aligned} &|Eh\,(\xi(t,x,y_1)) - Eh\,(\xi(t,x,y_2))| \\ &\qquad \leq E\,|h\,(\xi(t,x,y_1)) - h\,(\xi(t,x,y_2))| \\ &\qquad \leq 3M_0(h)E\,|\xi(t,x,y_1) - \xi(t,x,y_2)|\,, \end{aligned} \tag{B.0.44}$$

and hence, from Proposition B.1.1 on page 175 (see (B.1.202)), we get

$$E\,|h(\xi(t,x,y_1)) - h(\xi(t,x,y_2))| \leq 3M_0(h)\; e^{-\gamma_0(R)t}\,|y_1 - y_2|\,, \tag{B.0.45}$$

$\forall t \in [0,\infty), \forall y_1, y_2 \in I\!R^D$, for some constant $\gamma_0(R) \in (0,\infty)$. Now suppose $r_1 \in (0, q-1]$. By (B.0.43) and Holder's inequality, with conjugate exponents $\alpha \triangleq 1 + r_1$ and $\beta \triangleq (1+r_1)/r_1$, one has

$$\begin{aligned} &|Eh\,(\xi(t,x,y_1)) - Eh\,(\xi(t,x,y_2))| \\ &\qquad \leq E\,|h\,(\xi(t,x,y_1)) - h\,(\xi(t,x,y_2))| \\ &\qquad \leq M_{r_1}(h)\;E\,\{|\xi(t,x,y_1) - \xi(t,x,y_2)| \\ &\qquad\qquad [1 + |\xi(t,x,y_1)|^{r_1} + |\xi(t,x,y_2)|^{r_1}]\} \\ &\qquad \leq M_{r_1}(h)\,\{E\,|\xi(t,x,y_1) - \xi(t,x,y_2)|^{\alpha}\}^{1/\alpha} \\ &\qquad\qquad \left\{E\,[1 + |\xi(t,x,y_1)|^{r_1} + |\xi(t,x,y_2)|^{r_1}]^{\beta}\right\}^{1/\beta}. \end{aligned} \tag{B.0.46}$$

$\forall t \in [0,\infty), \forall y_1, y_2 \in I\!\!R^D$. We have $\alpha = 1 + r_1 \leq q$ and $\beta r_1 = (1 + r_1) \leq q$. Thus, from Proposition B.1.1 on page 175, (B.0.46), and the fact that $[|a| + |b| + |c|]^\beta \leq 3^\beta \left[|a|^\beta + |b|^\beta + |c|^\beta\right]$ we can write

$$
\begin{aligned}
&|Eh\left(\xi(t,x,y_1)\right) - Eh\left(\xi(t,x,y_2)\right)| \\
&\quad \leq M_{r_1}(h) \left\{E\left|\xi(t,x,y_1) - \xi(t,x,y_2)\right|^\alpha\right\}^{1/\alpha} \\
&\qquad\qquad \left\{3^\beta\left[1 + E\left|\xi(t,x,y_1)\right|^{\beta r_1} + E\left|\xi(t,x,y_2)\right|^{\beta r_1}\right]\right\}^{1/\beta} \\
&\quad \leq M_{r_1}(h)\left\{e^{-\gamma_1(R)t}\,|y_1 - y_2|^{r_1+1}\right\}^{1/(r_1+1)} \\
&\qquad\qquad \left\{C_5(R)\left[1 + |y_1|^{1+r_1} + |y_2|^{1+r_1}\right]\right\}^{r_1/(r_1+1)} \\
&\quad \leq C_6(R) M_{r_1}(h)\, e^{-\gamma_2(R)t}\, |y_1 - y_2|\left[1 + |y_1|^{r_1} + |y_2|^{r_1}\right],
\end{aligned}
\tag{B.0.47}
$$

$\forall t \in [0,\infty), \forall y_1, y_2 \in I\!\!R^D$, where $C_5(R), C_6(R) \in [0,\infty)$ and $\gamma_1(R), \gamma_2(R) \in (0,\infty)$ are constants. Now put

$$
\theta(t,x,y) \triangleq E\left[h\left(\xi(t,x,y)\right)\right], \tag{B.0.48}
$$

$\forall t \in [0,\infty), \forall y \in I\!\!R^D$. By the Markov property of $\{\xi(t,x,y)\}$(recall (3.4.74)) (see e.g. page 240, Theorem 10.11, of Chung and Williams [7]) we have

$$
E\left[h\left(\xi(s+t,x,y)\right)|\mathcal{F}_s^\beta\right] = \theta(t,x,\xi(s,x,y)), \quad \text{a.s.} \tag{B.0.49}
$$

$\forall s,t \in [0,\infty), \forall y \in I\!\!R^D$, where

$$
\mathcal{F}_s^\beta \triangleq \sigma\left\{\beta(u),\ u \in [0,s]\right\} \vee \left\{P - \text{null events in } \mathcal{F}\right\}. \tag{B.0.50}
$$

Thus, from (B.0.49),

$$
E\left[h\left(\xi(s+t,x,y)\right)\right] = E\left[\theta\left(t,x,\xi(s,x,y)\right)\right]. \tag{B.0.51}
$$

Hence

$$\begin{aligned}
&|Eh(\xi(s+t,x,y)) - Eh(\xi(t,x,y))| \\
&\quad = |E\theta(t,x,\xi(s,x,y)) - Eh(\xi(t,x,y))| \\
&\quad = |E\{\theta(t,x,\xi(s,x,y)) - Eh(\xi(t,x,y))\}| \\
&\quad \leq E|\theta(t,x,\xi(s,x,y)) - Eh(\xi(t,x,y))| \qquad \text{(B.0.52)}
\end{aligned}$$

$\forall s,t \in [0,\infty), \forall y \in \mathbb{R}^D$. Now, by (B.0.48) and (B.0.47), one has

$$\begin{aligned}
&|\theta(t,x,\tilde{y}) - Eh(\xi(t,x,y))| \\
&\quad = |Eh(\xi(t,x,\tilde{y})) - Eh(\xi(t,x,y))| \\
&\quad \leq M_{r_1}(h)C_6(R)\, e^{-\gamma_2(R)t}|y-\tilde{y}|\left[1+|y|^{r_1}+|\tilde{y}|^{r_1}\right] \qquad \text{(B.0.53)}
\end{aligned}$$

$\forall t \in [0,\infty), \forall y,\tilde{y} \in \mathbb{R}^D$. Now take $\tilde{y} \triangleq \xi(s,x,y)$ in (B.0.53) to get

$$\begin{aligned}
&|\theta(t,x,\xi(s,x,y)) - Eh(\xi(t,x,y))| \\
&\quad \leq M_{r_1}(h)C_6(R)\, e^{-\gamma_2(R)t}|y-\xi(s,x,y)|\left[1+|y|^{r_1}+|\xi(s,x,y)|^{r_1}\right], \qquad \text{(B.0.54)}
\end{aligned}$$

$\forall s,t \in [0,\infty), \forall y \in \mathbb{R}^D$. Combining (B.0.52) and (B.0.54), and also using Proposition B.1.1 on page 175, the fact that $r_2 \triangleq 1+r_1 \leq q$, and the Liapunov inequality, it follows that

$$\begin{aligned}
&|Eh(\xi(s+t,x,y)) - Eh(\xi(t,x,y))| \\
&\quad \leq M_{r_1}(h)C_6(R)e^{-\gamma_2(R)t}E\left\{|y|+|y|^{1+r_1}+|y||\xi(s,x,y)|^{r_1}\right. \\
&\qquad\quad \left. +|\xi(s,x,y)|+|y|^{r_1}|\xi(s,x,y)|+|\xi(s,x,y)|^{1+r_1}\right\} \\
&\quad \leq M_{r_1}(h)C_7(R)\, e^{-\gamma_2(R)t}\left[1+|y|^{1+r_1}\right], \qquad \text{(B.0.55)}
\end{aligned}$$

$\forall s,t \in [0,\infty), \forall y \in \mathbb{R}^D$, where $C_7(R) \in [0,\infty)$ is a constant. We next observe that

$$\int_{\mathbb{R}^D} |h(\xi)|\; d\pi_x(\xi) < \infty. \qquad \text{(B.0.56)}$$

Indeed, since $h \in Li(r_1)$ for some $r_1 \in [0, \infty)$, we have that

$$|h(\xi)| \leq C_8 [1 + |\xi|^q], \quad \forall \xi \in I\!R^D, \tag{B.0.57}$$

for some constant $C_8 \in [0, \infty)$ (see Remark 3.4.9), thus we get (B.0.56) from (3.4.75) (which we have just established). Now put

$$h_+^N(\xi) \stackrel{\Delta}{=} h_+(\xi) \wedge N, \quad \forall \xi \in I\!R^D, \ N = 1, 2, 3, \dots . \tag{B.0.58}$$

where $h_+(\xi)$ is the usual positive part of $h(\xi)$. Then one trivially checks that

$$h_+^N(\xi) \in Li(r_1), \quad M_{r_1}(h_+^N) \leq M_{r_1}(h). \tag{B.0.59}$$

Since (B.0.55) holds for arbitrary $h \in Li(r_1)$, in view of (B.0.59) we get

$$\left| Eh_+^N(\xi(s+t, x, y)) - Eh_+^N(\xi(t, x, y)) \right| \leq M_{r_1}(h) C_7(R) e^{-\gamma_2(R)t} \left[1 + |y|^{1+r_1} \right]. \tag{B.0.60}$$

Since $h_+^N : I\!R^D \to I\!R$ is bounded and continuous, from Theorem 3.4.1 on page 43 we have

$$\lim_{s \to \infty} Eh_+^N(\xi(s+t, x, y)) = \int_{I\!R^D} h_+^N(\xi) d\pi_x(\xi). \tag{B.0.61}$$

Hence, taking $s \to \infty$ in (B.0.60) gives

$$\left| Eh_+^N(\xi(t, x, y) - \int_{I\!R^D} h_+^N(\xi) d\pi_x(\xi) \right| \leq M_{r_1}(h) C_7(R) e^{-\gamma_2(R)t} \left[1 + |y|^{1+r_1} \right]. \tag{B.0.62}$$

Also, by (B.0.42), one has

$$\begin{aligned} E\left[h_+(\xi(t, x, y))\right] &\leq E\left|h(\xi(t, x, y))\right| \\ &\leq C_2(R) M_{r_1}(h) \left[1 + |y|^{1+r_1} \right] \\ &< \infty, \quad \forall t \in [0, \infty), \ \forall y \in I\!R^D. \end{aligned} \tag{B.0.63}$$

By (B.0.63), and the monotone convergence theorem,

$$\lim_{N\to\infty}\left|Eh_+^N(\xi(t,x,y)) - \int_{\mathbb{R}^D} h_+^N(\xi)d\pi_x(\xi)\right|$$
$$= \left|Eh_+(\xi(t,x,y)) - \int_{\mathbb{R}^D} h_+(\xi)d\pi_x(\xi)\right|. \qquad \text{(B.0.64)}$$

From (B.0.62) and (B.0.64),

$$\left|Eh_+(\xi(t,x,y)) - \int_{\mathbb{R}^D} h_+(\xi)d\pi_x(\xi)\right| \le M_{r_1}(h)C_7(R)e^{-\gamma_2(R)t}\left[1+|y|^{1+r_1}\right]. \qquad \text{(B.0.65)}$$

An identical form to (B.0.65), with h_+ replaced by h_-, also holds, so that

$$\left|Eh(\xi(t,x,y)) - \int_{\mathbb{R}^D} h(\xi)d\pi_x(\xi)\right| \le 2M_{r_1}(h)C_7(R)e^{-\gamma_2(R)t}\left[1+|y|^{1+r_1}\right]. \qquad \text{(B.0.66)}$$

Finally, identify $h(.)$ in (B.0.66) with $f(x,.)$, for $f \in Li(r_1)_{loc}, r_1 \in [0, q-1]$, to obtain (3.4.76).

□

Proof of Proposition 3.4.16 on page 51:

Step 1 : Here it is shown that $\Theta(.,.) \in Li(r)_{loc}$, that $y \to \Theta(x,y)$ is a C^2 - function for each $x \in \mathbb{R}^d$, and (3.4.80) holds. Since $r \in [0, q/2]$ and $q \in (8, \infty)$ (see Condition 3.4.7) we have that $r_2 \triangleq 1 + r < q$. Thus Remark 3.4.15 ensures that $\bar{f}(x)$ in (3.4.77) and $\Theta(x,y)$ in (3.4.78) exist for each $x \in \mathbb{R}^d, y \in \mathbb{R}^D$. Fix some $R \in [0, \infty)$, and fix arbitrary $x \in S_R^d$. Since $f \in Li(r)_{loc}$, we have $f(x,.) \in Li(r)$, hence we can repeat the simple calculation which gave (B.0.47) (but with $r, f(x,.)$ in place of $r_1, h(.)$) to get

$$|Ef(x,\xi(t,x,y_1)) - Ef(x,\xi(t,x,y_2))|$$
$$\le C_2(R)e^{-\gamma_1(R)t}|y_1 - y_2|\left[1 + |y_1|^r + |y_2|^r\right], \qquad \text{(B.0.67)}$$

$\forall t \in [0,\infty), \forall y_1, y_2 \in \mathbb{R}^D$, where $C_2(R) \in [0,\infty)$ and $\gamma_1(R) \in (0,\infty)$ are constants. From (3.4.78), and (B.0.67), there is a constant $C_3(R) \in [0,\infty)$ such that

$$\begin{aligned}
&|\Theta(x,y_1) - \Theta(x,y_2)| \\
&\qquad \leq \int_0^\infty |E[f(x,\xi(t,x,y_1)) - f(x,\xi(t,x,y_2))]| \; dt \\
&\qquad \leq \int_0^\infty C_2(R)e^{-\gamma_1(R)t} |y_1 - y_2| [1 + |y_1|^r + |y_2|^r] \; dt \\
&\qquad \leq C_3(R)|y_1 - y_2| [1 + |y_1|^r + |y_2|^r],
\end{aligned} \tag{B.0.68}$$

$\forall y_1, y_2 \in \mathbb{R}^D$. Thus, from (B.0.68) and Remark 3.4.9, one has

$$[\Theta(x,.)]_r \leq C_3(R) < \infty, \qquad \forall x \in S_R^d. \tag{B.0.69}$$

Moreover, by (3.4.76), (3.4.77), (3.4.78), and the fact that $r < q - 1$, there is a constant $C_4(R) \in [0,\infty)$ such that

$$|\Theta(x,y)| \leq C_4(R) \left[1 + |y|^{r+1}\right], \quad \forall (x,y) \in S_R^d \otimes \mathbb{R}^D. \tag{B.0.70}$$

Hence, by Remark 3.4.9:

$$\|\Theta(x,.)\|_{r+1} \leq C_4(R), \quad \forall x \in S_R^d. \tag{B.0.71}$$

Now from (B.0.69), (B.0.71), and Remark 3.4.9, we have

$$M_r(\Theta(x,.)) \leq \max\{C_3(R), C_4(R)\}, \quad \forall x \in S_R^d, \tag{B.0.72}$$

whence $\Theta \in Li(r)_{loc}$.

Next, consider smoothness of the mapping $y \to \Theta(x,y)$ for fixed $x \in \mathbb{R}^d$. Define

$$\theta(t,x,y) \triangleq E\left[\tilde{f}(x,\xi(t,x,y))\right], \quad \forall (t,x,y) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D, \tag{B.0.73}$$

where

$$\tilde{f}(x,\xi) \triangleq f(x,\xi) - \bar{f}(x), \quad \forall (x,\xi) \in \mathbb{R}^d \otimes \mathbb{R}^D. \tag{B.0.74}$$

By Corollary F.0.34, it follows that $y \to \theta(t,x,y)$ is a C^1 - function with

$$(\partial_{y^l}\theta)(t,x,y) = E\left[\sum_{k=1}^{D}(\partial_{\xi^k}\tilde{f})(x,\xi(t,x,y))\,(\partial_{y^l}\xi^k)(t,x,y)\right], \tag{B.0.75}$$

$\forall(t,x,y) \in [0,\infty)\otimes \mathbb{R}^d \otimes \mathbb{R}^D$, $l = 1,2,\ldots,D$ (here $(\partial_{y^l}\xi^k)(t,x,y)$ is the first L_2 - derivative of $\xi^k(t,x,y)$ with respect to y^l, as formulated in Remark F.0.32). Now fix some $R \in [0,\infty)$; from hypothesis (ii) of Proposition 3.4.16 on page 51 and (B.0.74), there is a constant $C_5(R) \in [0,\infty)$ such that

$$\begin{aligned}\left|(\partial_{\xi^k}\tilde{f})(x,\xi)\right| &= \left|(\partial_{\xi^k}f)(x,\xi)\right| \\ &\le C_5(R)\left[1+|\xi|^r\right], \quad \forall(x,\xi)\in S_R^d\otimes\mathbb{R}^D.\end{aligned} \tag{B.0.76}$$

In view of (B.0.76) together with (B.0.75), the Cauchy-Schwarz inequality, Proposition B.1.3 on page 184 (see (B.1.245)), and (B.1.201) of Proposition B.1.1 on page 175 (with $r_2 \triangleq 2r \le q$), we see that there are constants $C_6(R), C_7(R) \in [0,\infty)$, such that

$$\begin{aligned}\left|(\partial_{y^l}\theta)(t,x,y)\right| &\le \sum_{k=1}^{D} E\left|(\partial_{\xi^k}\tilde{f})(x,\xi(t,x,y))(\partial_{y^l}\xi^k)(t,x,y)\right| \\ &\le C_5(R)\sum_{k=1}^{D} E\left[(1+|\xi(t,x,y)|^r)\,\left|(\partial_{y^l}\xi^k)(t,x,y)\right|\right] \\ &\le C_6(R)\sum_{k=1}^{D} E^{1/2}\left[1+|\xi(t,x,y)|^{2r}\right]E^{1/2}\left[\left|(\partial_{y^l}\xi^k)(t,x,y)\right|^2\right] \\ &\le C_7(R)\left[1+|y|^r\right]e^{-\gamma_2(R)t},\end{aligned} \tag{B.0.77}$$

$\forall(t,x,y) \in [0,\infty)\otimes S_R^d\otimes\mathbb{R}^D$, where $\gamma_2(R) \in (0,\infty)$ is a constant. Since Corollary F.0.34 shows that $y \to \theta(t,x,y)$ is a C^1 - mapping, for each $\alpha \in \mathbb{R}$, we have

$$\theta(t,x,y+\alpha e_l) - \theta(t,x,y) = \int_0^\alpha (\partial_{y^l}\theta)(t,x,y+\eta e_l)\,d\eta, \tag{B.0.78}$$

$\forall(t,x,y) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$ (here e_l denotes the l-th basis vector in $\mathbb{R}^D$). Now from (B.0.73), (B.0.74), and (3.4.78), we have

$$\Theta(x,y) = \int_0^\infty \theta(t,x,y)\,dt, \quad \forall(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D. \tag{B.0.79}$$

Thus, it follows from (B.0.78) and (B.0.79) that

$$\begin{aligned} \Theta(x,y+\alpha e_l) - \Theta(x,y) &= \int_0^\infty [\theta(t,x,y+\alpha e_l) - \theta(t,x,y)]\,dt \\ &= \int_0^\infty \left\{ \int_0^\alpha (\partial_{y^l}\theta)(t,x,y+\eta e_l)\,d\eta \right\} dt, \end{aligned} \tag{B.0.80}$$

$\forall(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D, \forall\alpha \in \mathbb{R}$. From (B.0.77) clearly:

$$\begin{aligned} &\int_0^\infty \left\{ \int_0^{|\alpha|} \left|(\partial_{y^l}\theta)(t,x,y+\eta e_l)\right| d\eta \right\} dt \\ &\leq C_7(R) \int_0^\infty \left\{ \int_0^{|\alpha|} [1 + |y+\eta e_l|^r]\, e^{-\gamma_2(R)t} d\eta \right\} dt \\ &\leq C_7(R) \left\{ \int_0^\infty e^{-\gamma_2(R)t} dt \right\} \left\{ \int_0^{|\alpha|} [1 + |y+\eta e_l|^r]\, d\eta \right\} \\ &< \infty, \end{aligned} \tag{B.0.81}$$

$\forall(x,y) \in S_R^d \otimes \mathbb{R}^D, \forall\alpha \in \mathbb{R}$. Since $R \in [0,\infty)$ is arbitrary in (B.0.81) this inequality in fact holds for all $\alpha \in \mathbb{R}, (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Hence, from (B.0.80), (B.0.81), and Fubini's theorem, one has

$$\Theta(x,y+\alpha e_l) - \Theta(x,y) = \int_0^\alpha \left\{ \int_0^\infty (\partial_{y^l}\theta)(t,x,y+\eta e_l)\,dt \right\} d\eta, \tag{B.0.82}$$

$\forall\alpha \in \mathbb{R}, (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Now, by (B.0.77), the Dominated Convergence Theorem, and the fact that $y \to \theta(t,x,y)$ is C^1, one easily sees that the mapping

$$\eta \to \int_0^\infty (\partial_{y^l}\theta)(t,x,y+\eta e_l)\,dt, \tag{B.0.83}$$

is continuous (for fixed (t,x,y)). In view of this fact, (B.0.82), and the Fundamental theorem of calculus, one has

$$\lim_{\alpha\to 0}\frac{\Theta(x,y+\alpha e_l)-\Theta(x,y)}{\alpha}=\int_0^\infty(\partial_{y^l}\theta)(t,x,y)dt. \tag{B.0.84}$$

$\forall(x,y)\in I\!R^d\otimes I\!R^D$. Now an argument identical to that showing that the mapping in (B.0.83) is continuous, also shows that

$$y\to\int_0^\infty(\partial_{y^l}\theta)(t,x,y)dt : I\!R^D\to I\!R$$

is continuous; it follows from (B.0.84) that the $y\to\Theta(x,y)$ is a C^1 - mapping, and

$$(\partial_{y^l}\Theta)(x,y)=\int_0^\infty(\partial_{y^l}\theta)(t,x,y)\,dt,\quad \forall(x,y)\in I\!R^d\otimes I\!R^D. \tag{B.0.85}$$

From (B.0.77) and (B.0.85) it follows that $(\partial_{y^l}\Theta)(x,y)$ is polynomially y-bounded of order r locally in x, i.e. for each $R\in[0,\infty)$ there is a constant $C_8(R)\in[0,\infty)$ such that

$$\begin{aligned}\left|(\partial_{y^l}\Theta)(x,y)\right| &\le \int_0^\infty C_7(R)\left[1+|y|^r\right]e^{-\gamma_2(R)t}dt\\ &\le C_8(R)\left[1+|y|^r\right],\quad \forall(x,y)\in S_R^d\otimes I\!R^D.\end{aligned} \tag{B.0.86}$$

Next consider the second partial y-derivatives of $\Theta(x,y)$. Using Corollary F.0.34, one sees that $y\to\theta(t,x,y)$ is a C^2-mapping, with

$$\begin{aligned}&(\partial_{y^l}\partial_{y^k}\theta)(t,x,y)\\ &\quad= E\left[\sum_{i=1}^D(\partial_{\xi^i}\bar f)(x,\xi(t,x,y))(\partial_{y^l}\partial_{y^k}\xi^i)(t,x,y)\right],\\ &\qquad+E\left[\sum_{i=1}^D\sum_{j=1}^D(\partial_{y^l}\xi^i)(t,x,y)(\partial_{\xi^i}\partial_{\xi^j}\bar f)(x,\xi(t,x,y))(\partial_{y^k}\xi^j)(t,x,y)\right].\end{aligned} \tag{B.0.87}$$

$\forall(t,x,y)\in[0,\infty)\otimes I\!R^d\otimes I\!R^D$. By hypothesis and (B.0.74), we know that $(\partial_{\xi^l}\bar f)(x,\xi)$ and $(\partial_{\xi^l}\partial_{\xi^k}\bar f)(x,y)$ are polynomially ξ-bounded of order r locally in x. Thus from

(B.0.87) and Holder's inequality, for each $R \in [0, \infty)$ there is a constant $C_9(R) \in [0, \infty)$ such that

$$\begin{aligned}
&\left|(\partial_{y^l}\partial_{y^k}\theta)(x, y)\right| \\
&\leq C_9(R)E^{1/2}\left[1 + |\xi(t, x, y)|^{2r}\right] E^{1/2}\left[\left|(\partial_{y^l}\partial_{y^k}\xi)(t, x, y)\right|^2\right] \\
&+C_9(R)E^{1/2}\left[1 + |\xi(t, x, y)|^{2r}\right] E^{1/4}\left[\left|(\partial_{y^l}\xi)(t, x, y)\right|^4\right] E^{1/4}\left[\left|(\partial_{y^k}\xi)(t, x, y)\right|^4\right].
\end{aligned} \tag{B.0.88}$$

$\forall (t, x, y) \in [0, \infty) \otimes S_R^d \otimes \mathbb{R}^D$. By Proposition B.1.3 on page 184, (B.1.201) of Proposition B.1.1 on page 175 (with $r_2 \triangleq 2r \leq q$), and (B.0.88), for each $R \in [0, \infty)$ there are constants $C_{10}(R) \in [0, \infty), \gamma_3(R) \in (0, \infty)$, such that

$$\left|(\partial_{y^l}\partial_{y^k}\theta)(t, x, y)\right| \leq C_{10}(R)e^{-\gamma_3(R)t}\left[1 + |y|^r\right], \tag{B.0.89}$$

$\forall (t, x, y) \in [0, \infty) \otimes S_R^d \otimes \mathbb{R}^d$. Now the mapping $y \to (\partial_{y^l}\theta)(t, x, y)$ has been seen to be a C^1 - function (by Corollary F.0.34), thus for each $\alpha \in \mathbb{R}$ we have

$$(\partial_{y^l}\theta)(t, x, y + \alpha e_j) - (\partial_{y^l}\theta)(t, x, y) = \int_0^\alpha (\partial_{y^j}\partial_{y^l}\theta)(t, x, y + \eta e_j)\, d\eta, \tag{B.0.90}$$

$\forall (t, x, y) \in [0, \infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$. By (B.0.85) and (B.0.90) one has

$$\begin{aligned}
&(\partial_{y^l}\Theta)(x, y + \alpha e_j) - (\partial_{y^l}\Theta)(x, y) \\
&\quad = \int_0^\infty \left[(\partial_{y^l}\theta)(t, x, y + \alpha e_j) - (\partial_{y^l}\theta)(t, x, y)\right]\, dt \\
&\quad = \int_0^\infty \left\{\int_0^\alpha (\partial_{y^j}\partial_{y^l}\theta)(t, x, y + \eta e_j)\, d\eta\right\}\, dt
\end{aligned} \tag{B.0.91}$$

$\forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D, \forall \alpha \in \mathbb{R}$. Moreover, by (B.0.89) for each $R \in [0, \infty)$, we see that

$$\begin{aligned}
&\int_0^\infty \left\{ \int_0^{|\alpha|} \left|(\partial_{y^j}\partial_{y^l}\theta)(t, x, y + \eta e_j)\right| d\eta \right\} dt \\
&\quad \le C_{10}(R) \int_0^\infty \left\{ \int_0^{|\alpha|} [1 + |y + \eta e_l|^r] \, e^{-\gamma_3(R)t} d\eta \right\} dt \\
&\quad \le C_{10}(R) \left\{ \int_0^\infty e^{-\gamma_3(R)t} dt \right\} \left\{ \int_0^{|\alpha|} [1 + |y + \eta e_l|^r] \, d\eta \right\} \\
&\quad < \infty, \qquad \text{(B.0.92)}
\end{aligned}$$

$\forall (x, y) \in S_R^d \otimes \mathbb{R}^D$. Thus, from (B.0.92), one can use Fubini's theorem to write (B.0.91) as

$$\begin{aligned}
&(\partial_{y^l}\Theta)(x, y + \alpha e_j) - (\partial_{y^l}\Theta)(x, y) \\
&\qquad = \int_0^\alpha \left\{ \int_0^\infty (\partial_{y^j}\partial_{y^l}\theta)(t, x, y + \eta e_j) \, dt \right\} d\eta
\end{aligned} \qquad \text{(B.0.93)}$$

$\forall \alpha \in \mathbb{R}, \forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Now one easily observes from (B.0.89), the fact that $y \to \theta(t, x, y)$ is C^2, and the Lebesgue Dominated Convergence Theorem that the mapping

$$\eta \to \int_0^\infty (\partial_{y^j}\partial_{y^l}\theta)(t, x, y + \eta e_j) \, dt \qquad \text{(B.0.94)}$$

is continuous. Thus from (B.0.93) and the fundamental theorem of calculus one has

$$\begin{aligned}
&\lim_{\alpha \to 0} \frac{1}{\alpha} \left[(\partial_{y^l}\Theta)(x, y + \alpha e_j) - (\partial_{y^l}\Theta)(x, y) \right] \\
&\qquad = \int_0^\infty (\partial_{y^j}\partial_{y^l}\theta)(t, x, y) \, dt. \qquad \text{(B.0.95)}
\end{aligned}$$

$\forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. From (B.0.89), the fact that $y \to \theta(t, x, y)$ is C^2, and the Lebesgue Dominated Convergence Theorem, it follows that

$$y \to \int_0^\infty (\partial_{y^j}\partial_{y^l}\theta)(t, x, y) dt : \mathbb{R}^D \to \mathbb{R}$$

is continuous, so we see from (B.0.95) that $y \to \Theta(x,y)$ is a C^2 - function with

$$(\partial_{y^j}\partial_{y^i}\Theta)(x,y) = \int_0^\infty (\partial_{y^j}\partial_{y^i}\theta)(t,x,y)\, dt, \tag{B.0.96}$$

$\forall (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. From (B.0.89) and (B.0.96), for each $R \in [0,\infty)$ there is a constant $C_{11}(R) \in [0,\infty)$ such that

$$\left|(\partial_{y^j}\partial_{y^i}\Theta)(x,y)\right| \le C_{11}(R)\,[1 + |y|^r], \tag{B.0.97}$$

$\forall (x,y) \in S_R^d \otimes \mathbb{R}^D$. We have seen that $y \to \Theta(x,y)$ is a C^2-mapping, and also, from (B.0.86) and (B.0.97), the mappings $(\partial_{y^i}\Theta)(x,y)$ and $(\partial_{y^j}\partial_{y^i}\Theta)(x,y)$ are polynomially y-bounded of order r locally in x. This completes Step 1.

Step 2 : It remains to show that (3.4.81) holds. We have seen that $\Theta \in Li(r)_{loc}$, hence it follows that $\Theta(x,y)$ is polynomially y-bounded of order $r+1$ locally in x (see Remark 3.4.11), thus for each $R \in [0,\infty)$ there is a constant $C_{12}(R) \in [0,\infty)$ such that

$$|\Theta(x,y)| \le C_{12}(R)\left[1 + |y|^{r+1}\right], \quad \forall (x,y) \in S_R^d \otimes \mathbb{R}^D. \tag{B.0.98}$$

Moreover we have seen that $(\partial_{y^i}\Theta)(x,y)$ and $(\partial_{y^i}\partial_{y^k}\Theta)(x,y)$ are polynomially y-bounded of order r locally in x. In view of these facts one sees from (3.2.14), (3.4.74), and Theorem 3 on page 293 of Gihman and Skorohod [14],

$$\mathcal{A}\Theta(x,y) = \lim_{t_1 \downarrow 0} \frac{E\Theta(x, \xi(t_1,x,y)) - \Theta(x,y)}{t_1}, \quad \forall (x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D. \tag{B.0.99}$$

Now evaluate quantity on right side of (B.0.99). Fix $x \in \mathbb{R}^d$, fix $R \in [0,\infty)$ such that $x \in S_R^d$. From (B.0.74) and (3.4.77) we have

$$\int_{\mathbb{R}^D} \tilde{f}(x,\xi)\,d\pi_x(\xi) = 0, \quad \forall x \in \mathbb{R}^d. \tag{B.0.100}$$

Then, from (B.0.100), (B.0.73) and Proposition 3.4.14 on page 50, there are constants $C_{13}(R) \in [0, \infty)$ and $\gamma_4(R) \in (0, \infty)$ such that

$$\begin{aligned} |\theta(t, x, \tilde{y})| &= \left| E\tilde{f}(x, \xi(t, x, \tilde{y}))\right| \\ &\leq C_{13}(R)e^{-\gamma_4(R)t}\left[1 + |\tilde{y}|^{1+r}\right], \end{aligned} \qquad \text{(B.0.101)}$$

$\forall (t, \tilde{y}) \in [0, \infty) \otimes \mathbb{R}^D$. Fix $t_1 \in (0, \infty)$ and $y \in \mathbb{R}^D$, and take $\tilde{y} \triangleq \xi(t_1, x, y)$. Thus, from (B.0.101), we can write

$$|\theta(t, x, \xi(t_1, x, y)| \leq C_{13}(R)e^{-\gamma_4(R)t}\left[1 + |\xi(t_1, x, y)|^{1+r}\right], \qquad \text{(B.0.102)}$$

$\forall (t, y) \in [0, \infty) \otimes \mathbb{R}^D$. From Proposition B.1.1 on page 175 (with $r_2 \triangleq 1 + r < q$) and (B.0.102) we get

$$\begin{aligned} E\left|\theta(t, x, \xi(t_1, x, y))\right| &\leq C_{13}(R)e^{-\gamma_4(R)t}E\left[1 + |\xi(t_1, x, y)|^{1+r}\right], \\ &\leq C_{14}(R)e^{-\gamma_4(R)t}\left[1 + |y|^{1+r}\right], \end{aligned} \qquad \text{(B.0.103)}$$

$\forall (t, y) \in [0, \infty) \otimes \mathbb{R}^D$, for some constant $C_{14}(R) \in [0, \infty)$. Hence from (B.0.103),

$$\int_0^\infty E\left|\theta(t, x, \xi(t_1, x, y))\right| dt < \infty, \qquad \text{(B.0.104)}$$

$\forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Now from (B.0.79) we have

$$\Theta(x, \xi(t_1, x, y)) = \int_0^\infty \theta(t, x, \xi(t_1, x, y))dt, \qquad \text{(B.0.105)}$$

$\forall (t, x, y) \in [0, \infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$. By (B.0.104), (B.0.105), and Fubini theorem one gets

$$E\Theta\left(x, \xi(t_1, x, y)\right) = \int_0^\infty E\theta(t, x, \xi(t_1, x, y))\, dt, \qquad \text{(B.0.106)}$$

$\forall (t_1, x, y) \in (0, \infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$. From (B.0.73), we have

$$\begin{aligned} \theta(t + t_1, x, y) &= E\left[\tilde{f}(x, \xi(t + t_1, x, y))\right] \\ &= E\left[E\left[\tilde{f}(x, \xi(t + t_1, x, y)|\mathcal{F}_{t_1}\right]\right], \end{aligned} \qquad \text{(B.0.107)}$$

$\forall t, t_1 \in [0, \infty), \forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Also, from the Markov property of the solution of (3.4.74) (see Theorem 10.11 of Chung and Williams [7]) and (B.0.73), we have

$$E\left[\tilde{f}(x, \xi(t+t_1, x, y)) \mid \mathcal{F}_{t_1}\right] = E\left[\tilde{f}(x, \xi(t+t_1, x, y)) \mid \xi(t_1, x, y)\right] = \theta(t, x, \xi(t_1, x, y)) \quad \text{(B.0.108)}$$

$\forall t, t_1 \in [0, \infty), \forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Thus, from (B.0.107) and B.0.108),

$$\theta(t + t_1, x, y) = E\left[\theta(t, x, \xi(t_1, x, y))\right], \quad \text{(B.0.109)}$$

$\forall t, t_1 \in [0, \infty), \forall (x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. From (B.0.106) and (B.0.107), for each $(x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$, we have

$$E\left[\Theta(x, \xi(t_1, x, y))\right] = \int_0^\infty \theta(t + t_1, x, y)dt, \quad \forall t_1 \in [0, \infty). \quad \text{(B.0.110)}$$

Combining (B.0.79) and (B.0.110) shows that

$$\frac{E\Theta(x, \xi(t_1, x, y)) - \Theta(x, y)}{t_1} = \int_0^\infty \frac{\theta(t + t_1, x, y) - \theta(t, x, y)}{t_1} dt, \quad \text{(B.0.111)}$$

$(t_1, x, y) \in (0, \infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$. Now from Theorem F.0.37 on page 223 (with $h \triangleq \tilde{f}$), one sees that $t \to \theta(t, x, y)$ is a C^1-mapping, that $y \to \theta(t, x, y)$ is a C^2 - mapping, and the following relation holds

$$(\partial_t \theta)(t, x, y) = \mathcal{A}\theta(t, x, y), \quad \text{(B.0.112)}$$

$\forall (t, x, y) \in [0, \infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$, for $\mathcal{A}\theta$ given by (F.0.11). In the view of Condition 3.2.1, (F.0.11), (B.0.77), (B.0.89), and (B.0.112), there are constants $C_{16}(R) \in [0, \infty)$ and $\gamma_6(R) \in (0, \infty)$ such that.

$$\begin{aligned} |\partial_t \theta(t, x, y)| &\leq \sum_{i=1}^{D} \left|b^i(x, y)\right| \left|(\partial_{y^i}\theta)(t, x, y)\right| \\ &\quad + \frac{1}{2}\sum_{i=1}^{D}\sum_{j=1}^{D} \left|\left[\sigma(x, y)\sigma^T(x, y)\right]_{i,j}\right| \left|(\partial_{y^i}\partial_{y^j}\theta)(t, x, y)\right| \\ &\leq C_{16}(R)e^{-\gamma_6(R)t}\left[1 + |y|^{2+r}\right], \end{aligned} \quad \text{(B.0.113)}$$

$\forall (t, x, y) \in [0, \infty) \otimes S_R^d \otimes I\!R^D$. From the mean - value theorem, and the fact that $t \to \theta(t, x, y)$ is a C^1-mapping one has

$$\theta(t + t_1, x, y) - \theta(t, x, y) = t_1 (\partial_t \theta)(s_1, x, y), \quad \text{for some } s_1 \in [t, t + t_1]. \tag{B.0.114}$$

Using (B.0.113) and (B.0.114), we get

$$\begin{aligned} \left| \frac{\theta(t + t_1, x, y) - \theta(t, x, y)}{t_1} \right| &= |(\partial_t \theta)(s_1, x, y)| \\ &\le C_{16}(R) e^{-\gamma_6(R) s_1} \left[1 + |y|^{2+r} \right], \end{aligned}$$

$\forall t_1 \in (0, \infty)$, $\forall (t, x, y) \in [0, \infty) \otimes S_R^d \otimes I\!R^D$. Thus, since $s_1 \in [t, t + t_1]$, we find

$$\left| \frac{\theta(t + t_1, x, y) - \theta(t, x, y)}{t_1} \right| \le C_{16}(R) e^{-\gamma_6(R) t} \left[1 + |y|^{2+r} \right], \tag{B.0.115}$$

$\forall t_1 \in (0, \infty), \forall (t, x, y) \in [0, \infty) \otimes S_R^d \otimes I\!R^D$. From (B.0.115) and the Dominated Convergence Theorem one has

$$\begin{aligned} &\lim_{\substack{t_1 \to 0 \\ t_1 \in (0, \infty)}} \int_0^\infty \frac{\theta(t + t_1, x, y) - \theta(t, x, y)}{t_1} \, dt \\ &= \int_0^\infty \left[\lim_{\substack{t_1 \to 0 \\ t_1 \in (0, \infty)}} \frac{\theta(t + t_1, x, y) - \theta(t, x, y)}{t_1} \right] dt \\ &= \int_0^\infty (\partial_t \theta)(t, x, y) \, dt, \quad \forall (x, y) \in I\!R^d \otimes I\!R^D, \end{aligned} \tag{B.0.116}$$

where the third line follows since $t \to \theta(t, x, y)$ is a C^1 - function. Now

$$\begin{aligned} \int_0^\infty (\partial_t \theta)(t, x, y) dt &= \lim_{T \to \infty} \int_0^T (\partial_t \theta)(t, x, y) dt \\ &= \lim_{T \to \infty} \left[\theta(T, x, y) - \theta(0, x, y) \right], \end{aligned} \tag{B.0.117}$$

$\forall(x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. From (B.0.73) and (B.0.74) one has

$$\begin{aligned} \theta(0, x, y) &= E\left[\tilde{f}(x, \xi(0, x, y))\right] \\ &= \tilde{f}(x, y) \quad (\text{since } \xi(0, x, y) = y) \\ &= f(x, y) - \bar{f}(x), \quad \forall(x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D. \end{aligned} \tag{B.0.118}$$

Moreover one finds from (B.0.101) that

$$\lim_{T\to\infty} \theta(T, x, y) = 0, \quad \forall(x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D. \tag{B.0.119}$$

Combining (B.0.99), (B.0.111), (B.0.116), (B.0.117), (B.0.118), and (B.0.119), it follows that

$$\begin{aligned} \mathcal{A}\Theta(x, y) &= \lim_{t_1\to 0} \frac{E\Theta(x, \xi(t_1, x, y)) - \Theta(x, y)}{t_1} \\ &= \lim_{t_1\to 0} \int_0^\infty \frac{\theta(t + t_1, x, y) - \theta(t, x, y)}{t_1}\, dt \\ &= \int_0^\infty (\partial_{t_1}\theta)(t, x, y)\, dt \\ &= \lim_{T\to\infty} [\theta(T, x, y) - \theta(0, x, y)] \\ &= \bar{f}(x) - f(x, y), \quad \forall(x, y) \in \mathbb{R}^d \otimes \mathbb{R}^D, \end{aligned} \tag{B.0.120}$$

thus (3.4.81) follows.

Step 3 : Here we establish joint continuity in (x, y) of the mappings

$$(x, y) \to (\partial_{y^l}\Theta)(x, y), \quad (x, y) \to (\partial_{y^l}\partial_{y^k}\Theta)(x, y).$$

Fix some sequence $\{(x_n, y_n)\}$ in $\mathbb{R}^d \otimes \mathbb{R}^D$ such that

$$\lim_{n\to\infty} (x_n, y_n) = (x_0, y_0).$$

Using Theorem 5.2 on page 118 of Friedman [12] (on parametric dependence in the L_2 - sense of solutions of stochastic differential equations) it is easy, although tedious, to show that

$$\lim_{n\to\infty} E\left|(\partial_{\xi^k}\tilde{f})(x_n, \xi(t, x_n, y_n)) - (\partial_{\xi^k}\tilde{f})(x_0, \xi(t, x_0, y_0))\right|^2 = 0 \tag{B.0.121}$$

and

$$\lim_{n\to\infty} E\left|(\partial_{y^l}\xi^k)(t,x_n,y_n) - (\partial_{y^l}\xi^k)(t,x_0,y_0)\right|^2 = 0 \qquad \text{(B.0.122)}$$

for each $t \in [0,\infty)$. From (B.0.121), (B.0.122) and (B.0.75) it follows that

$$\lim_{n\to\infty}(\partial_{y^l}\theta)(t,x_n,y_n) \;=\; (\partial_{y^l}\theta)(t,x_0,y_0), \qquad \text{(B.0.123)}$$

for each $t \in [0,\infty)$. In (B.0.77) fix $R \in [0,\infty)$ large enough that $x_0, x_n \in S^d_R, n = 1,2,3,\ldots$. Then from (B.0.77), (B.0.85), (B.0.123), and the Lebesgue Dominated Convergence Theorem, we get

$$\lim_{n\to\infty}(\partial_{y^l}\Theta)(x_n,y_n) = (\partial_{y^l}\Theta)(x_0,y_0) \qquad \text{(B.0.124)}$$

as required for joint continuity of $(x,y) \to (\partial_{y^l}\Theta)(x,y)$. The case of joint continuity of the second derivative $(x,y) \to (\partial_{y^l}\partial_{y^k}\Theta)(x,y)$ follows similarly, but uses the bound (B.0.89) to justify use of the Lebesgue Dominated Convergence Theorem in (B.0.96).

□

Proof of Proposition 3.4.17 on page 52: Expanding $\tilde{\Theta}(x,\xi(t,x,y))$ by Itô's formula and (3.4.74), yields

$$\begin{aligned}\tilde{\Theta}(x,\xi(t,x,y)) \;&=\; \tilde{\Theta}(x,y) + \int_0^t \mathcal{A}\tilde{\Theta}(x,\xi(s,x,y))\,ds \\ &\quad + \sum_{k=1}^{D}\sum_{n=1}^{N}\int_0^t \sigma^{k,n}(x,\xi(s,x,y))(\partial_{\xi^k}\tilde{\Theta})(x,\xi(s,x,y))\,d\beta^n(s),\end{aligned} \qquad \text{(B.0.125)}$$

$\forall(t,x,y) \in [0,\infty)\otimes \mathbb{R}^d \otimes \mathbb{R}^D$, where $\mathcal{A}\Theta(.,.)$ is given by (3.2.14). Since by hypothesis the partial derivative functions $\partial_{y^k}\tilde{\Theta}(x,y)$ and $\partial_{y^l}\partial_{y^k}\tilde{\Theta}(x,y)$ are polynomially y-bounded locally in x, it is easily checked that all integrals in (B.0.125) exist, and we have

$$E\int_0^t \left|\sigma^{k,n}(x,\xi(s,x,y))(\partial_{\xi^k}\tilde{\Theta})(x,\xi(s,x,y))\right|^2 ds < \infty, \tag{B.0.126}$$

$\forall(t,x,y) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$. Thus, the stochastic integrals on right hand side of (B.0.125) are martingales null at the origin, hence

$$E\left\{\sum_{k=1}^{D}\sum_{n=1}^{N}\int_0^t \sigma^{k,n}(x,\xi(s,x,y))(\partial_{\xi^k}\tilde{\Theta})(x,\xi(s,x,y))\, d\beta^n(s)\right\} = 0, \tag{B.0.127}$$

$\forall(t,x,y) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$. Fix arbitrary $x \in \mathbb{R}^d$, choose $R \in [0,\infty)$ such that $x \in S^d_R$. From (3.2.14), Condition 3.2.1, and the hypotheses, there is a constant $C_1(R) \in [0,\infty)$ such that

$$\left|\mathcal{A}\tilde{\Theta}(x,y)\right| \le C_1(R)\left[1+|y|^{r+2}\right], \quad \forall y \in \mathbb{R}^D. \tag{B.0.128}$$

Thus one sees from (B.0.128) and (C.0.7) (see Appendix C), that

$$\int_0^t E\left|\mathcal{A}\tilde{\Theta}(x,\xi(s,x,y))\right| ds < \infty, \quad \forall(t,y) \in [0,\infty) \otimes \mathbb{R}^D. \tag{B.0.129}$$

Taking expectations in (B.0.125), using (3.4.82), (B.0.127), (B.0.129), and Fubini theorem, gives

$$E\tilde{\Theta}(x,\xi(t,x,y)) = \tilde{\Theta}(x,y) + \int_0^t E\left[\bar{f}(x) - f(x,\xi(s,x,y))\right] ds, \tag{B.0.130}$$

$\forall(t,y) \in [0,\infty) \otimes \mathbb{R}^D$. Since $\tilde{\Theta} \in Li(r)_{loc}$, for some $r \le q/2 < q-1$ (where q is given by Condition 3.4.7) from Proposition 3.4.14 on page 50 we get

$$\left|E\tilde{\Theta}(x,\xi(t,x,y)) - \int_{\mathbb{R}^D}\tilde{\Theta}(x,\xi)d\pi_x(\xi)\right| \le C(R)e^{-\gamma(R)t}\left[1+|y|^{1+r}\right] \tag{B.0.131}$$

$\forall(t,y) \in [0,\infty) \otimes \mathbb{R}^D$, where $C(R) \in [0,\infty)$ and $\gamma(R) \in (0,\infty)$ are constants which depend on our choice of $R \in [0,\infty)$ to ensure that $x \in S^d_R$. Now take $t \to \infty$ in

(B.0.131) and use (3.4.84) to get

$$\lim_{t\to\infty} E\left[\tilde{\Theta}(x,\xi(t,x,y))\right] = \int_{\mathbb{R}^D} \tilde{\Theta}(x,\xi)\, d\pi_x(\xi) \triangleq \bar{\Theta}(x). \tag{B.0.132}$$

Moreover, since $f \in Li(r)_{loc}$, for some $r \leq q/2 < q-1$, one sees from Remark 3.4.15 that

$$\begin{aligned} \lim_{t\to\infty} \int_0^t \left[\bar{f}(x) - f(x,\xi(s,x,y))\right] ds &= \int_0^\infty \left[\bar{f}(x) - f(x,\xi(s,x,y))\right] ds \\ &= -\Theta(x,y), \quad \forall y \in \mathbb{R}^D. \end{aligned} \tag{B.0.133}$$

Thus, using (B.0.132), (B.0.133), and taking $t \to \infty$ in (B.0.130), we have

$$\bar{\Theta}(x) = \tilde{\Theta}(x,y) - \Theta(x,y), \quad \forall y \in \mathbb{R}^D. \tag{B.0.134}$$

Now (3.4.83) follows from (B.0.134) and our arbitrary choice of $x \in \mathbb{R}^D$.

□

Proof of Proposition 3.4.18 on page 53: The proof of this proposition involves many steps that are very similar to arguments deployed elsewhere in this thesis. Accordingly, in this proof we shall depart from our usual custom of exhaustively presenting all details, and will merely summarize the main steps, indicating where arguments and calculations are similar to those elsewhere in the thesis.

Fix an arbitrary $l \in \{1, 2, \ldots, d\}$. For each $(t,x,y,z) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^D$ let $\{\eta_l(t,x,y,z)\}$ be the solution of the stochastic differential equation

$$\begin{aligned} \eta_l(t,x,y,z) &= z + \int_0^t (\partial_{x^l} b)(x,\xi(s,x,y))ds + \int_0^t (\partial_{x^l}\sigma)(x,\xi(s,x,y))d\beta(s) \\ &\quad + \int_0^t J(x,\xi(s,x,y))\eta_l(s,x,y,z)ds \\ &\quad + \int_0^t B\left(x,\xi(s,x,y),\eta_l(s,x,y,z)\right) d\beta(s), \end{aligned} \tag{B.0.135}$$

where $J(x, \xi)$ is defined by (3.4.65), and $B(x, \xi, \eta)$ is the $D \times N$ matrix whose (k, n) - element is given by

$$B^{k,n}(x, \xi, \eta) \stackrel{\Delta}{=} (\partial_\xi \sigma^{k,n})(x, \xi)\eta, \quad \forall k = 1, 2, \dots, D, \ \forall n = 1, 2, \dots, N. \tag{B.0.136}$$

From Condition 3.4.4 and Theorem 4 on page 55, 56 of Gihman and Skorohod [14] one sees that partial derivative $(\partial_{x^l}\xi)(t, x, y)$ of the solution $\xi(t, x, y)$ of (3.4.74) exists in the L_2 - sense and is given by

$$(\partial_{x^l}\xi)(t, x, y) = \eta_l(t, x, y, 0). \tag{B.0.137}$$

Indeed, the right hand side of (B.0.135), with $z \stackrel{\Delta}{=} 0$, is just the formal derivative of the right - hand side of (3.4.74) with respect to x^l, and Theorem 4 on page 55, 56 of Gihman and Skorohod [14] justifies this formal differentiation in the light of the smoothness hypotheses postulated in Condition 3.4.4.

STEP 1 : In this step we shall establish that $\bar{f}(x)$ given by (3.4.77) is a C^2 - mapping on $\mathbb{R}^d$. To simplify the notation, put

$$g(x, \xi, \eta) \stackrel{\Delta}{=} (\partial_\xi f)(x, \xi)\eta \tag{B.0.138}$$

and

$$h(x, \xi, \eta) \stackrel{\Delta}{=} (\partial_{x^l} f)(x, \xi) + g(x, \xi, \eta), \tag{B.0.139}$$

$\forall (x, \xi, \eta) \in \mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^D$. Also put

$$\psi(t, x, y) \stackrel{\Delta}{=} E\left[f(x, \xi(t, x, y))\right], \tag{B.0.140}$$

$\forall (t, x, y) \in [0, \infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$. From (B.0.137) and Corollary 1 on page 62 of Gihman and Skorohod [14] we see that $x \to \psi(t, x, y)$ is a C^1 - mapping, with

$$(\partial_{x^l}\psi)(t, x, y) = E\left[h(x, \xi(t, x, y), \eta_l(t, x, y, 0))\right]. \tag{B.0.141}$$

Hence from (B.0.141) and the fundamental theorem of calculus, one has, for fixed $(t, x_0, y) \in [0, \infty) \otimes I\!R^d \otimes I\!R^D$, that

$$\begin{aligned}\psi(t, x_0 + \alpha e_l, y) &- \psi(t, x_0, y) \\ &= \int_0^\alpha (\partial_{x^l}\psi)(t, x_0 + se_l, y)ds \\ &= \int_0^\alpha Eh(x_0 + se_l, \xi(t, x_0 + se_l, y), \eta_l(t, x_0 + se_l, y, 0))ds,\end{aligned} \tag{B.0.142}$$

for $\alpha \in I\!R$ (here $e_l = (0, \ldots, 0, 1, 0, \ldots, 0)$ denotes the usual l - th canonical basis vector of $I\!R^d$). In light of Proposition 3.4.14 on page 50 (see (3.4.76)) and (B.0.140) we have

$$\bar{f}(x) = \lim_{t \to \infty} \psi(t, x, y), \quad \forall (x, y) \in I\!R^d \otimes I\!R^D. \tag{B.0.143}$$

We will shortly establish existence of a mapping $\bar{h}^l : I\!R^d \to I\!R$ with the following property : for each $R \in [0, \infty)$ there are constants $C_1(R) \in [0, \infty)$ and $\gamma_1(R) \in (0, \infty)$ such that

$$\left|Eh(x, \xi(t, x, y), \eta_l(t, x, y, 0)) - \bar{h}^l(x)\right| \leq C_1(R)e^{-\gamma_1(R)t}\left[1 + |y|^{1+r}\right], \tag{B.0.144}$$

$\forall (t, x, y) \in [0, \infty) \otimes S_R^d \otimes I\!R^D$. Then, since the mapping in (B.0.141) is continuous in x (for arbitrary (t, y)) and the convergence in (B.0.144) is clearly uniform with respect to x in balls S_R^d, we see that $\bar{h}^l(.)$ is necessarily *continuous* on $I\!R^d$. Now fix some x_0, y, and α in (B.0.142), and make $R \in [0, \infty)$ large enough that $x_0 + se_l \in S_R^d$, for all $s \in [-|\alpha|, |\alpha|]$. Then, from (B.0.144), we have

$$\begin{aligned}|Eh(x_0 + se_l, \xi(t, x_0 + se_l, y), \eta_l(t, x_0 + se_l, 0))| \\ \leq \sup_{x \in S_R^d} |\bar{h}^l(x)| + C_1(R)\left[1 + |y|^{1+r}\right],\end{aligned} \tag{B.0.145}$$

for all $s \in [-|\alpha|, |\alpha|], t \in [0, \infty)$, and $y \in I\!R^D$. Since the quantity on the right of (B.0.145) is finite and uniform with respect to $t \in [0, \infty)$ and $|s| \leq |\alpha|$, we can take $t \to \infty$ in (B.0.142), and use (B.0.143), (B.0.144) and the Lebesgue bounded convergence theorem to get

$$\bar{f}(x_0 + \alpha e_l) - \bar{f}(x_0) = \int_0^\alpha \bar{h}^l(x_0 + s e_l) ds, \quad \forall \alpha \in I\!R. \tag{B.0.146}$$

In view of the observed continuity of $\bar{h}^l(.)$, we see from (B.0.146) that $\bar{f}(.)$ is a C^1 - function, with

$$(\partial_{x^l} \bar{f})(x) = \bar{h}^l(x), \quad \forall x \in I\!R^d. \tag{B.0.147}$$

Thus, it remains to show existence of $\bar{h}^l : I\!R^d \to I\!R$ such that (B.0.144) holds. An argument similar to that used for establishing (B.1.201) shows that, for each $R \in [0, \infty)$, there is a constant $C_2(R) \in [0, \infty)$ such that

$$E\,|\eta_l(t, x, y, z)|^q \le |z|^q + C_2(R) \tag{B.0.148}$$

$\forall (t, x, y, z) \in [0, \infty) \otimes S_R^d \otimes I\!R^D \otimes I\!R^D$. In view of (B.0.148), (B.1.201), and the Chebyshev inequality, it follows that $\{\xi(t, x, y), \eta_l(t, x, y, z), t \in [0, \infty)\}$ is a tight family of $I\!R^{2D}$ - valued random vectors for each $(x, y, z) \in I\!R^d \otimes I\!R^D \otimes I\!R^D$. Thus, for each $(x, y, z) \in I\!R^d \otimes I\!R^D \otimes I\!R^D$, there is a sequence $t_n \to \infty$, and a probability measure $\mu^l((x, y, z), .)$ on $I\!R^{2D}$, such that

$$\text{weak} - \lim_{n \to \infty} (\xi(t_n, x, y), \eta_l(t_n, x, y, z)) = \mu^l((x, y, z), .). \tag{B.0.149}$$

We are now going to see that the limit $\mu^l((x, y, z), .)$ in (B.0.149) does not in fact depend on (y, z) : An application of Condition 3.4.7 shows that, for each $R \in [0, \infty)$, there is a constant $\gamma_2(R) \in (0, \infty)$ such that

$$\begin{aligned} &E\,|\eta_l(t, x, y_1, z_1) - \eta_l(t, x, y_2, z_2)|^{q/2} \\ &\quad \le e^{-\gamma_2(R)t} \left\{|z_1 - z_2|^{q/2} + |y_1 - y_2| \left[1 + |y_1 - y_2|\right] \left[1 + |z_1|^{q/2} + |z_2|^{q/2}\right]\right\}, \end{aligned} \tag{B.0.150}$$

$\forall x \in S_R^d, \forall y_1, z_1, y_2, z_2 \in I\!R^D, \forall t \in [0, \infty)$ (the technical details for establishing (B.0.150) are fairly lengthy, but essentially parallel the steps by which we obtained

(B.1.202)). In view of (B.1.202) and (B.0.150) we see that

$$\lim_{n\to\infty} E\left\{\left[|\xi(t_n,x,y_1)-\xi(t_n,x,y)|+|\eta_l(t_n,x,y_1,z_1)-\eta_l(t_n,x,y,z)|\right]^{q/2}\right\}=0. \tag{B.0.151}$$

for arbitrary $y_1, z_1 \in I\!R^D$. In view of (B.0.151) and (B.0.149) we have

$$\text{weak} - \lim_{n\to\infty} (\xi(t_n,x,y_1), \eta_l(t_n,x,y_1,z_1)) = \mu^l((x,y,z),.), \tag{B.0.152}$$

for arbitrary $y_1, z_1 \in I\!R^D$. It follows that the limit $\mu^l((x,y,z),.)$ in (B.0.149) does not depend on (y,z), and will be denoted by $\mu^l(x,.)$. To summarize, for each $x \in I\!R^d$, there is a sequence $t_n \to \infty$ and a probability measure $\mu^l(x,.)$ on $I\!R^{2D}$ such that

$$\text{weak} - \lim_{n\to\infty} (\xi(t_n,x,y), \eta_l(t_n,x,y,z)) = \mu^l(x,.), \tag{B.0.153}$$

for all $(y,z) \in I\!R^{2D}$. Now, for each $x \in I\!R^d$, (3.4.74) and (B.0.135) is a pair of classical Itô stochastic differential equations, and hence defines a Markov diffusion (with state space $I\!R^{2D}$) whose transition operator $\{T^x_t\}$ is easily seen to have the Feller property (i.e. $(T^x_t\phi)(y,z)$ is a real - valued bounded continuous function in $(y,z) \in I\!R^{2D}$ when ϕ is a real - valued bounded continuous function on $I\!R^{2D}$). For bounded continuous $\phi : I\!R^{2D} \to I\!R$, we see from (B.0.153) that

$$\lim_{n\to\infty} (T^x_{t_n}\phi)(y,z) = \bar{\phi}(x) \triangleq \int_{I\!R^{2D}} \phi(y',z')\mu^l(x,d(y',z')), \tag{B.0.154}$$

for each $(x,y,z) \in I\!R^d \otimes I\!R^D \otimes I\!R^D$. Then, by Lebesgue Dominated Convergence Theorem and (B.0.154),

$$\begin{aligned}
&\lim_{n\to\infty} T^x_t(T^x_{t_n}\phi)(y,z) \\
&\quad = \lim_{n\to\infty} E\left[(T^x_{t_n}\phi)(\xi(t,x,y),\eta_l(t,x,y,z))\right] \\
&\quad = E\left[\lim_{n\to\infty}(T^x_{t_n}\phi)(\xi(t,x,y),\eta_l(t,x,y,z))\right] \\
&\quad = \bar{\phi}(x),
\end{aligned} \tag{B.0.155}$$

for arbitrary $(t,x,y,z) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^D$. Also, since $T_t^x\phi$ is bounded and continuous on $\mathbb{R}^{2D}$ (by the Feller property of $\{T_t^x\}$) we can apply (B.0.154), but with ϕ replaced by $T_t^x\phi$, to get

$$\lim_{n\to\infty} T_{t_n}^x(T_t^x\phi)(y,z) = \int_{\mathbb{R}^{2D}} (T_t^x\phi)(y',z')\mu^l(x,d(y',z')), \qquad \text{(B.0.156)}$$

for arbitrary $(t,x,y,z) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^D$. Now

$$T_t^x(T_{t_n}^x\phi) = T_{t_n}^x(T_t\phi)$$

(since $\{T_t^x\}$ is a semigroup), thus comparison of (B.0.155) and (B.0.156) gives

$$\int_{\mathbb{R}^{2D}} (T_t^x\phi)(y',z')\mu^l(x,d(y',z')) = \int_{\mathbb{R}^{2D}} \phi(y',z')\mu^l(x,d(y',z')), \qquad \text{(B.0.157)}$$

for each $(t,x) \in [0,\infty) \otimes \mathbb{R}^d$, each bounded continuous function $\phi : \mathbb{R}^{2D} \to \mathbb{R}$. It follows from (B.0.157) that, for each $x \in \mathbb{R}^d$, the measure $\mu^l(x,.)$ is an invariant probability for the Markov diffusion defined by (3.4.74) and (B.0.135). To see that $\mu^l(x,.)$ is the only such invariant probability, let $\nu(x,.)$ be some arbitrary invariant probability for the Markov diffusion defined by (3.4.74) and (B.0.135). Then for arbitrary bounded and continuous $\phi : \mathbb{R}^{2D} \to \mathbb{R}$, we have

$$\int_{\mathbb{R}^{2D}} (T_{t_n}^x)(y',z')\nu(x,d(y',z')) = \int_{\mathbb{R}^{2D}} \phi(y',z')\nu(x,d(y',z')), \qquad \text{(B.0.158)}$$

for all $n = 1,2,3,\ldots$. Then, from (B.0.154), (B.0.158) and Lebesgue Dominated Convergence Theorem,

$$\int_{\mathbb{R}^{2D}} \phi(y',z')\mu^l(x,d(y',z')) = \int_{\mathbb{R}^{2D}} \phi(y',z')\nu(x,d(y',z')) \qquad \text{(B.0.159)}$$

Since (B.0.159) holds for arbitrary bounded continuous ϕ, we see that

$$\nu(x,.) = \mu^l(x,.)$$

as required for uniqueness. To summarize, we see that, for each $x \in \mathbb{R}^d$, the Markov diffusion (3.4.74) and (B.0.135) has a unique invariant probability measure which is given by $\mu^l(x,.)$. We therefore conclude that

$$\text{weak} - \lim_{t\to\infty}(\xi(t,x,y),\eta_l(t,x,y,z)) = \mu^l(x,.), \tag{B.0.160}$$

for each $(x,y,z) \in \mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^D$. We next show that the integral

$$\bar{h}^l(x) \triangleq \int_{\mathbb{R}^{2D}} h(x,\xi,\eta)\mu^l(x,d(\xi,\eta)) \tag{B.0.161}$$

exists for each $x \in \mathbb{R}^d$, and that (B.0.144) holds for the mapping $\bar{h}^l(.)$ in (B.0.161). From (B.0.139) and (B.0.141) we see that

$$\begin{aligned} &E\left[h(x,\xi(t,x,y),\eta_l(t,x,y,0)\right] \\ &\qquad = E\left[(\partial_{x^l} f)(x,\xi(t,x,y))\right] + E\left[g(x,\xi(t,x,y),\eta_l(t,x,y,0)\right]. \end{aligned} \tag{B.0.162}$$

(that the expectations in (B.0.162) exist follows easily from (3.4.74), (B.0.135), the postulated polynomial boundedness in y of $(\partial_{x^l} f)(x,y)$ and $(\partial_y f)(x,y)$, and Lemma C.0.6) on page 203. Consider the second term on the right hand side of (B.0.162) : Put

$$\theta(t,x,y,z) \triangleq E\left[g(x,\xi(t,x,y),\eta_l(t,x,y,z))\right]. \tag{B.0.163}$$

By the Markov property of the diffusion $\{\xi(t,x,y),\eta_l(t,x,y,z)\}$, we see that

$$E\left[g(x,\xi(s+t,x,y),\eta_l(s+t,x,y,z))\right] = E\left[\theta(t,x,\xi(s,x,y),\eta_l(s,x,y,z)\right], \tag{B.0.164}$$

$\forall s,t \in [0,\infty), \forall (x,y,z) \in \mathbb{R}^d \otimes \mathbb{R}^D \otimes \mathbb{R}^D$. Now an application of Condition 3.4.7 shows that, for each $R \in [0,\infty)$, there are constants $C_3(R) \in [0,\infty)$ and $\gamma_3(R) \in (0,\infty)$ such that

$$\begin{aligned} &\left|E\left[g\left(x,\xi(s+t,x,y),\eta_l(s+t,x,y,0)\right] - E\left[g\left(x,\xi(t,x,y),\eta_l(t,x,y,0)\right)\right]\right| \\ &\qquad \le C_3(R)e^{-\gamma_3(R)t}\left[1+|y|^{1+r}\right], \end{aligned} \tag{B.0.165}$$

$\forall s, t \in [0, \infty), \forall (x, y) \in S_R^d \otimes \mathbb{R}^D$ (the technical details for establishing (B.0.165) are fairly lengthy, but parallel very closely the steps by which we obtained the inequality (B.0.55)). Likewise, considering the first term on the right hand side of (B.0.162), we can use Condition 3.4.7 to see that, for each $R \in [0, \infty)$ there exist some constants $C_4(R) \in [0, \infty)$ and $\gamma_4(R) \in (0, \infty)$ such that

$$\begin{aligned} &\left| E\left[(\partial_{x^l} f)(x, \xi(s+t, x, y)) \right] - E\left[(\partial_{x^l} f)(x, \xi(t, x, y)) \right] \right| \\ &\quad \leq C_4(R) e^{-\gamma_4(R) t} \left[1 + |y|^{1+r} \right], \end{aligned} \tag{B.0.166}$$

$\forall s, t \in [0, \infty), \forall (x, y) \in S_R^d \otimes \mathbb{R}^D$ (again, the technical details for establishing (B.0.166) are similar to those for the inequality (B.0.55)). From (B.0.165), (B.0.166), and (B.0.162), for each $R \in [0, \infty)$ there are constants $C_5(R) \in [0, \infty)$ and $\gamma_5(R) \in (0, \infty)$ such that

$$\begin{aligned} &\left| Eh\left(x, \xi(s+t, x, y), \eta_l(s+t, x, y, 0)\right) - Eh\left(x, \xi(t, x, y), \eta_l(t, x, y, 0)\right) \right| \\ &\quad \leq C_5(R) e^{-\gamma_5(R) t} \left[1 + |y|^{1+r} \right], \end{aligned} \tag{B.0.167}$$

$\forall s, t \in [0, \infty), \forall (x, y) \in S_R^d \otimes \mathbb{R}^D$. Now, in view of (B.0.138), (B.0.139), and the hypothesis that $(\partial_{x^l} f)(x, y)$ and $(\partial_{y^k} f)(x, y)$ are polynomially y - bounded of order r locally in x, we see that for each $R \in [0, \infty)$ there is a constant $C_6(R) \in [0, \infty)$ such that for $\beta \in (1, \infty)$

$$\begin{aligned} |h(x, \xi, \eta)|^\beta &\leq 2^\beta \left| (\partial_{x^l} f)(x, \xi) \right|^\beta + 2^\beta \left| (\partial_\xi f)(x, \xi) \right|^\beta |\eta|^\beta \\ &\leq C_6(R) \left[1 + |\xi|^{\beta r} \right] \left[1 + |\eta|^\beta \right], \end{aligned} \tag{B.0.168}$$

$\forall (x, \xi, \eta) \in S_R^d \otimes \mathbb{R}^D \otimes \mathbb{R}^D$. Thus from (B.0.168) and Holder's inequality,

$$\begin{aligned} &E\left| h\left(x, \xi(t, x, y), \eta_l(t, x, y, 0)\right) \right|^\beta \\ &\quad \leq C_7(R) \left\{ E\left[1 + |\xi(t, x, y)|^{2\beta r} \right] \right\}^{1/2} \left\{ E\left[1 + |\eta_l(t, x, y)|^{2\beta} \right] \right\}^{1/2}, \end{aligned} \tag{B.0.169}$$

$\forall(t,x,y) \in [0,\infty) \otimes S_R^d \otimes I\!R^D$ and some constant $C_7(R) \in [0,\infty)$. Now fix $\beta \in (1, q/2r) \cap (1, q/2)$ (recall that r is postulated in the range $[0, q/2)$, so that $q/2r > 1$). Since $2\beta < q$, we get from (B.0.148) that

$$E\,|\eta_l(t,x,y,z)|^{2\beta} \leq |z|^q + C_8(R), \tag{B.0.170}$$

$\forall(t,x,y,z) \in [0,\infty) \otimes S_R^d \otimes I\!R^D \otimes I\!R^D$, where $C_8(R) \in [0,\infty)$ is some constant. Since $2\beta r < q$, we see from (B.1.201) that

$$E\,|\xi(t,x,y)|^{2\beta r} \leq |y|^q + C_9(R), \tag{B.0.171}$$

$\forall(t,x,y) \in [0,\infty) \otimes S_R^d \otimes I\!R^D$, where $C_9(R) \in [0,\infty)$ is a constant. Combine (B.0.169), (B.0.170}), (B.0.171) to get

$$E\,|h(x,\xi(t,x,y),\eta_l(t,x,y,0))|^{\beta} \leq C_{10}(R)\left[1+|y|^{q/2}\right], \tag{B.0.172}$$

$\forall(t,x,y) \in [0,\infty) \otimes S_R^d \otimes I\!R^D$, where $C_{10}(R) \in [0,\infty)$ is a constant. Since $\beta > 1$, it follows from (B.0.172) that $\{h(x,\xi(t,x,y),\eta_l(t,x,y,0)),\ t \in [0,\infty)\}$ is uniformly integrable for each $(x,y) \in I\!R^d \otimes I\!R^D$. Thus, from (B.0.160) we see that the integral in (B.0.161) is defined and

$$\lim_{t\to\infty} Eh(x,\xi(t,x,y),\eta_l(t,x,y,0)) = \int_{I\!R^{2D}} h(x,\xi,\eta)\ \mu^l(x;d(\xi,\eta)) \tag{B.0.173}$$

for each $x \in I\!R^d$ (by Theorem 5.4 on page 32 of Billingsley [4]). Taking $s \to \infty$ in (B.0.167), together with (B.0.173) and (B.0.161), gives constants $C_1(R) \in [0,\infty)$ and $\gamma_1(R) \in (0,\infty)$ such that (B.0.144) holds. We have therefore established that $\bar{f}(.)$ defined by (3.4.77) is a C^1 - mapping on $I\!R^d$ (with derivative $(\partial_{x^l}\bar{f})(x)$ given by $\bar{h}^l(x)$ defined by (B.0.161)). To show that $\bar{f}(.)$ is a C^2 - mapping one procedes in much the same way as before, only now we must take a formal derivative of $\eta_l(t,x,y,z)$ with respect to (say) x^k, to get the L_2 - derivative $(\partial_{x^l}\partial_{x^k}\xi)(t,x,y)$, and then follow the preceding arguments.

STEP 2 : We show that $(\partial_{x^l}\Theta)(x,y)$ exists, is continuous in (x,y), and is polynomailly y - bounded of order $(r+1)$ locally in x. Put

$$\tilde{f}(x,\xi) \triangleq f(x,\xi) - \bar{f}(x) \tag{B.0.174}$$

and

$$\phi(t,x,y) \triangleq E\tilde{f}(x,\xi(t,x,y)). \tag{B.0.175}$$

Then, from (3.4.78),

$$\Theta(x,y) = \int_0^\infty \phi(t,x,y)dt \tag{B.0.176}$$

$\forall(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. Now fix some $x_0 \in \mathbb{R}^d$. From (B.0.175) and (B.0.176), for each $\alpha \in \mathbb{R}$, we can write

$$\Theta(x_0+\alpha e_l,y) - \Theta(x_0,y) = \int_0^\infty [\phi(t,x_0+\alpha e_l,y) - \phi(t,x_0,y)]\,dt. \tag{B.0.177}$$

Now we have seen in Step 1 that, for each $t \in [0,\infty), y \in \mathbb{R}^D$, the mappings $x \to Ef(x,\xi(t,x,y))$ and $x \to \bar{f}(x)$ are C^1 - functions on $\mathbb{R}^d$. Thus from (B.0.174) and (B.0.175) the mapping $x \to \phi(t,x,y)$ is a C^1 - function. Therefore:

$$\phi(t,x_0+\alpha e_l,y) - \phi(t,x_0,y) = \int_0^\alpha (\partial_{x^l}\phi)(t,x_0+se_l,y)ds, \tag{B.0.178}$$

whence from (B.0.177) we have

$$\Theta(x_0+\alpha e_l,y) - \Theta(x_0,y) = \int_0^\infty \left\{\int_0^\alpha (\partial_{x^l}\phi)(t,x_0+se_l,y)ds\right\}dt. \tag{B.0.179}$$

From (B.0.140), (B.0.141), (B.0.147), (B.0.174), (B.0.175),

$$\begin{aligned}(\partial_{x^l}\phi)(t,x,y) &= (\partial_{x^l}\psi)(t,x,y) - (\partial_{x^l}\bar{f})(x) \\ &= Eh(x,\xi(t,x,y),\eta_l(t,x,y,0)) - \bar{h}^l(x).\end{aligned} \tag{B.0.180}$$

Thus by (B.0.180) with (B.0.144) one obtains

$$|(\partial_{x^l}\phi)(t,x,y)| \le C_1(R)e^{-\gamma_1(R)t}\left[1+|y|^{1+r}\right], \tag{B.0.181}$$

$\forall x \in S^d_R, \forall y \in \mathbb{R}^D, \forall t \in [0, \infty)$. From (B.0.181) one sees that

$$\int_0^\infty \int_0^{|\alpha|} |(\partial_{x^l}\phi)(t, x_0 + se_l, y)ds|\, dt < \infty, \tag{B.0.182}$$

$\forall \alpha \in \mathbb{R}, \forall y \in \mathbb{R}^D$. Thus from Fubini's theorem with (B.0.179):

$$\Theta(x_0 + \alpha e_l, y) - \Theta(x_0, y) = \int_0^\alpha \left\{ \int_0^\infty (\partial_{x^l}\phi)(t, x_0 + se_l, y)dt \right\} ds, \tag{B.0.183}$$

$\forall \alpha \in \mathbb{R}, \forall y \in \mathbb{R}^D$. Now an easy but tedious computation shows that the mapping

$$(x, y) \to Eh(x, \xi(t, x, y), \eta_l(t, x, y, 0)) \tag{B.0.184}$$

is continuous on $\mathbb{R}^{d+D}$ for each $t \in [0, \infty)$, and since $x \to \bar{h}^l(x)$ was shown to be continuous on $\mathbb{R}^d$, we see from (B.0.180) that $(x, y) \to (\partial_{x^l}\phi)(t, x, y)$ is continuous on $\mathbb{R}^{d+D}$ for each $t \in [0, \infty)$. Thus, by (B.0.181) and Dominated Convergence Theorem one finds that

$$(x, y) \to \int_0^\infty (\partial_{x^l}\phi)(t, x, y)dt : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R} \tag{B.0.185}$$

is continuous. In particular, for each $y \in \mathbb{R}^D$, the mapping

$$s \to \int_0^\infty (\partial_{x^l}\phi)(t, x_0 + se_l, y)dt \tag{B.0.186}$$

is continuous. Thus from (B.0.183) and the fundamental theorem of calculus, we have

$$(\partial_{x^l}\Theta)(x_0, y) = \int_0^\infty (\partial_{x^l}\phi(t, x_0, y)dt, \tag{B.0.187}$$

and therefore (B.0.185) shows that

$$(x, y) \to (\partial_{x^l}\Theta)(x, y) : \mathbb{R}^d \otimes \mathbb{R}^D \to \mathbb{R}$$

is continuous. Moreover, from (B.0.187) with (B.0.181), for each $R \in [0, \infty)$, there is a constant $C_{11}(R) \in [0, \infty)$ such that

$$\begin{aligned} |(\partial_{x^l}\Theta)(x, y)| &\leq \int_0^\infty |(\partial_{x^l}\phi)(t, x, y)|\, dt \\ &\leq C_{11}(R)\left[1 + |y|^{1+r}\right], \end{aligned} \tag{B.0.188}$$

$\forall(x,y) \in S_R^d \otimes \mathbb{R}^D$, so that $(\partial_{x^l}\Theta)(.)$ is polynomially y - bounded of order $(r+1)$ locally in x. To conclude, we have shown that $(\partial_{x^l}\Theta)(x,y)$ exists for each $(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$, is continuous in (x,y), and is polynomially y - bounded of order $(r+1)$ locally in x. In the same way one establishes that the second derivative $(\partial_{x^l}\partial_{x^k}\Theta)(x,y)$ exists for each $(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$, is continuous in (x,y), and is polynomially y - bounded of order $(r+1)$ locally in x.

□

Proof of Proposition 3.4.25 on page 64: Condition 3.2.1 and (3.2.27) ensure that $\tilde{F}^i \in L_i(0)_{loc}$, thus for each $R \in [0,\infty)$ we have

$$\left|\tilde{F}^i(x,y)\right| \leq C_1(R)\,[1+|y|]\,, \quad \forall(x,y) \in S_R^d \otimes \mathbb{R}^D,\; i = 1,2,\ldots,d \tag{B.0.189}$$

for some constant $C_1(R) \in [0,\infty)$. Then Proposition 3.4.14 on page 50 shows that certainly (see (3.4.75))

$$\int_{\mathbb{R}^D} |y|^2 \, d\pi_x(y) < \infty, \quad \forall x \in \mathbb{R}^d, \tag{B.0.190}$$

and there are constants $C_2(R) \in [0,\infty)$, $\gamma_1(R) \in (0,\infty)$, such that (see (3.2.30) and (3.4.76) with $r_1 \triangleq 0$ and $f \triangleq \tilde{F}^j$)

$$\left|E\tilde{F}^j(x,\xi(t,x,y))\right| \leq C_2(R)e^{-\gamma_1(R)t}\,[1+|y|]\,, \tag{B.0.191}$$

$\forall(t,x,y) \in [0,\infty) \otimes S_R^d \otimes \mathbb{R}^D$. From (B.0.189), (B.0.190) and (B.0.191), we see that

$$\begin{aligned} &\int_{\mathbb{R}^D} \left\{ \int_0^\infty \left|\tilde{F}^i(x,y)\right| \left|E\tilde{F}^j(x,\xi(s,x,y))\right| \, ds \right\} d\pi_x(y) \\ &\leq \int_{\mathbb{R}^D} C_1(R)\,[1+|y|] \left\{ \int_0^\infty C_2(R)e^{-\gamma(R)s}\,[1+|y|]\, ds \right\} d\pi_x(y) \\ &< \infty, \end{aligned} \tag{B.0.192}$$

$\forall(x,y) \in S_R^d \otimes \mathbb{R}^D$. Since $R \in [0,\infty)$ is arbitrary, this bound holds for all $(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$. From Proposition 3.4.17 on page 52 we know that the solution $\Phi^j(x,y)$

in (3.2.28) is a unique modulo function of x only. Thus, by Remark 3.2.12, without loss of generality we can take the $\Phi^j(x,y)$ in (3.2.28) to be given by

$$\Phi^j(x,y) \triangleq \int_0^\infty E\tilde{F}^j(x,\xi(s,x,y))\,ds, \quad \forall (x,y) \in I\!R^d \otimes I\!R^D. \qquad \text{(B.0.193)}$$

Thus, from (B.0.192), (B.0.193) and Fubini's theorem we have

$$\int_{I\!R^D} \tilde{F}^i(x,y)\Phi^j(x,y)\,d\pi_x(y) = \int_0^\infty \left[\int_{I\!R^D} \tilde{F}^i(x,y)E\tilde{F}^j(x,\xi(s,x,y))\,d\pi_x(y)\right] ds, \qquad \text{(B.0.194)}$$

$\forall (x,y) \in I\!R^d \otimes I\!R^D$. Now put

$$\theta(t,x,y) \triangleq E\left[\tilde{F}^j(x,\xi(t,x,y))\right], \quad \forall (t,x,y) \in [0,\infty) \otimes I\!R^d \otimes I\!R^D. \qquad \text{(B.0.195)}$$

By the Markov property of $\{\bar{\xi}(t;x)\}$ (see e.g. Lemma 10.10 of Chung and Williams [7]) we have

$$\bar{E}\left[\tilde{F}^j(x,\bar{\xi}(s;x)) \mid \bar{\xi}(0,x)\right] = \theta(s,x,\bar{\xi}(0,x)), \quad \text{a.s.}, \qquad \text{(B.0.196)}$$

$\forall (s,x) \in [0,\infty) \otimes I\!R^d$. Hence, by stationarity of $\{\bar{\xi}(t,x)\}$,

$$\begin{aligned}
\bar{E}\left[\tilde{F}^i(x,\bar{\xi}(0,x))\tilde{F}^j(x,\bar{\xi}(s,x))\right] &= \bar{E}\left[\bar{E}\left[\tilde{F}^i(x,\bar{\xi}(0,x))\tilde{F}^j(x,\bar{\xi}(s,x)) \mid \bar{\xi}(0,x)\right]\right] \\
&= \bar{E}\left[\tilde{F}^i(x,\bar{\xi}(0,x))\bar{E}\left[\tilde{F}^j(x,\bar{\xi}(s,x) \mid \bar{\xi}(0,x))\right]\right] \\
&= \bar{E}\left[\tilde{F}^i(x,\bar{\xi}(0,x))\theta(s,x,\bar{\xi}(0,x))\right] \\
&= \int_{I\!R^D} \tilde{F}^i(x,y)\theta(s,x,y)\,d\pi_x(y),
\end{aligned} \qquad \text{(B.0.197)}$$

$\forall (s,x) \in [0,\infty) \otimes I\!R^d$. Now combine (B.0.195) and (B.0.197) to get

$$\bar{E}\left[\tilde{F}^i(x,\bar{\xi}(0,x))\tilde{F}^j(x,\bar{\xi}(s,x))\right] = \int_{I\!R^D} \tilde{F}^i(x,y)E\tilde{F}^j(x,\xi(s,x,y))d\pi_x(y), \qquad \text{(B.0.198)}$$

$\forall (s,x) \in [0,\infty) \otimes \mathbb{R}^d$. By (B.0.194) and (B.0.198), one has

$$\int_{\mathbb{R}^D} \tilde{F}^i(x,y)\Phi^j(x,y)d\pi_x(y) = \int_0^\infty \bar{E}\left[\tilde{F}^i(x,\bar{\xi}(0,x))\tilde{F}^j(x,\bar{\xi}(s,x))\right] ds \tag{B.0.199}$$

$\forall x \in \mathbb{R}^d$. Now using (3.2.28), (3.2.29), and (B.0.199), we get

$$\begin{aligned}\bar{a}^{i,j}(x) \triangleq \int_0^\infty \Big\{ &\bar{E}\left[\tilde{F}^i(x,\bar{\xi}(0,x))\tilde{F}^j(x,\bar{\xi}(s,x))\right] \\ &+ \bar{E}\left[\tilde{F}^j(x,\bar{\xi}(0,x))\tilde{F}^i(x,\bar{\xi}(s,x))\right]\Big\} ds,\end{aligned} \tag{B.0.200}$$

$\forall x \in \mathbb{R}^d, i,j = 1,2,\ldots,d$, and thus (3.4.114) follows.

□

B.1 Some Useful Results

In this section, we present some results which are needed for proofs of Proposition 3.4.14 on page 50 and Proposition 3.4.16 on page 51.

Proposition B.1.1. *Suppose Conditions 3.2.1, 3.4.4, and 3.4.7 hold. Then, for each $R \in [0,\infty)$ and $r_2 \in [0,q]$ (q defined in Condition 3.4.7), there are constants $C(r_2,R) \in [0,\infty)$ and $\gamma(r_2,R) \in (0,\infty)$ such that*

$$E\left|\xi(t,x,y)\right|^{r_2} \le |y|^{r_2} + C(r_2,R), \tag{B.1.201}$$

$$E\left|\xi(t,x,y_1) - \xi(t,x,y_2)\right|^{r_2} \le e^{-\gamma(r_2,R)t}\left|y_1 - y_2\right|^{r_2}, \tag{B.1.202}$$

$\forall x \in S^d_R, \forall y, y_1, y_2 \in \mathbb{R}^D, \forall t \in [0,\infty)$ *(recall Remark 3.4.13 for the definition of $\xi(t,x,y)$).*

Proof: Fix some $R \in [0,\infty)$, and fix $x \in S_R^d$ and $y, y_1, y_2 \in I\!R^D$. Put

$$\varphi(\xi) \triangleq |\xi|^q, \quad \forall \xi \in I\!R^D, \tag{B.1.203}$$

(where q is given by Condition 3.4.7). From Remark 3.4.13, (B.1.203), and Itô formula, one can write

$$\begin{aligned}\varphi(\xi(t,x,y)) &= |y|^q + \int_0^t \mathcal{A}\varphi(x,\xi(t,x,y))ds \\ &\quad \sum_{l=1}^{D}\sum_{n=1}^{N}\int_0^t (\partial_{\xi^l}\varphi)\xi(t,x,y)\, \sigma^{l,n}(x,\xi(t,x,y))\, d\beta^n(s).\end{aligned} \tag{B.1.204}$$

$\forall t \in [0,\infty)$, where $\mathcal{A}\varphi(.,.)$ is defined by (3.2.14). In view of Condition 3.2.1, Lemma C.0.6 on page 203 (see (C.0.7)), and (B.0.16), we see that

$$\begin{aligned}\int_0^t E\left|(\partial_{\xi^l}\varphi)\xi(s,x,y)\sigma^{l,n}(x,\xi(s,x,y))\right|^2 ds \\ \leq C_1(R)\int_0^t E|\xi(s,x,y)|^{2q}ds \\ < \infty, \quad \forall t \in [0,\infty),\end{aligned} \tag{B.1.205}$$

for some constant $C_1(R) \in [0,\infty)$. Thus the sum of stochastic integrals on the right hand side of (B.1.204) is a continuous martingale which is null at the origin, hence has zero expectation. Moreover, from Proposition 3.4.12 on page 50 and (B.1.203), we have

$$|\mathcal{A}\varphi(x,\xi(s,x,y))| \leq \alpha(R) + \gamma(R)|\xi(s,x,y)|^q, \quad \forall s \in [0,1], \tag{B.1.206}$$

where $\alpha(R) \in [0,\infty)$ and $\gamma(R) \in (0,\infty)$ are constants. Thus, from Lemma C.0.6 on page 203 and (B.1.206):

$$\int_0^t E\left|\mathcal{A}\varphi(x,\xi(s,x,y))\right| ds < \infty, \quad \forall t \in [0,\infty). \tag{B.1.207}$$

In view of (B.1.204), (B.1.207), and using Fubini's theorem we get

$$E\varphi(\xi(t,x,y)) = |y|^q + \int_0^t E\left[\mathcal{A}\varphi(x,\xi(s,x,y))\right] ds, \quad \forall t \in [0,\infty). \tag{B.1.208}$$

Now, put

$$\theta(t) \triangleq E\varphi(\xi(t,x,y)), \quad \forall t \in [0,\infty). \tag{B.1.209}$$

In view of (B.1.206), Lemma C.0.6 on page 203 (see (C.0.7)), and the Dominated Convergence Theorem, we easily see that $s \to E\left[\mathcal{A}\varphi(x,\xi(s,x,y))\right] : [0,\infty) \to I\!R$ is continuous. Thus from (B.1.208), (B.1.209), and Proposition 3.4.12 on page 50, it follows that

$$\begin{aligned} \frac{d\theta(t)}{dt} &= E\left[\mathcal{A}\varphi(x,\xi(t,x,y))\right] \\ &\leq -\gamma(R)\theta(t) + \alpha(R), \quad \forall t \in [0,\infty). \end{aligned} \tag{B.1.210}$$

Thus from (B.1.210), one easily sees that

$$\theta(t) \leq e^{-\gamma(R)t}\theta(0) + \frac{\alpha(R)}{\gamma(R)}\left[1 - e^{-\gamma(R)t}\right], \quad \forall t \in [0,\infty). \tag{B.1.211}$$

Hence, by (B.1.211), (B.1.209) and (B.1.203),

$$E\left|\xi(t,x,y)\right|^q \leq e^{-\gamma(R)t}\left|y\right|^q + \frac{\alpha(R)}{\gamma(R)}, \quad \forall t \in [0,\infty). \tag{B.1.212}$$

Without loss of generality we take $r_2 \in (0,q]$. By the Liapunov inequality and (B.1.212), we find

$$\begin{aligned} \left[E\left|\xi(t,x,y)\right|^{r_2}\right]^{1/r_2} &\leq \left[E\left|\xi(t,x,y)\right|^q\right]^{1/q} \\ &\leq \left[e^{-\gamma(R)t}|y|^q + \frac{\alpha(R)}{\gamma(R)}\right]^{1/q} \\ &\leq |y| + C_2(R), \quad \forall t \in [0,\infty), \end{aligned} \tag{B.1.213}$$

where $C_2(R) \triangleq \left[\alpha(R)/\gamma(R)\right]^{1/q}$. Then, from (B.1.213) and (B.1.212), we see that, for each $r_2 \in [0,q]$,

$$E\left|\xi(t,x,y)\right|^{r_2} \leq |y|^{r_2} + C_3(r_2,R), \quad \forall t \in [0,\infty), \tag{B.1.214}$$

where $C_3(r_2,R) \in [0,\infty)$ is a constant. Thus (B.1.201) is proved.

Next we prove the inequality (B.1.202). To establish this result fix some $R \in [0, \infty)$, and fix $x \in S_R^d$, and $y, y_1, y_2 \in I\!R^D$, and also put

$$\Delta\xi(t) \triangleq \xi(t, x, y_1) - \xi(t, x, y_2), \tag{B.1.215}$$

$$\theta(t) \triangleq E\,|\Delta\xi(t)|^q, \tag{B.1.216}$$

$$\Delta b(t) \triangleq b(x, \xi(t, x, y_1)) - b(x, \xi(t, x, y_2)), \tag{B.1.217}$$

$$\Delta\sigma(t) \triangleq \sigma(x, \xi(t, x, y_1)) - \sigma(x, \xi(t, x, y_2)). \tag{B.1.218}$$

From (3.4.74), (B.1.215), (B.1.217) and (B.1.218) we have

$$\Delta\xi(t) \triangleq (y_1 - y_2) + \int_0^t \Delta b(s)ds + \int_0^t \Delta\sigma(s)\, d\beta(s), \quad \forall t \in [0, \infty). \tag{B.1.219}$$

Put

$$\varphi(\xi) \triangleq |\xi|^q, \quad \forall \xi \in I\!R^D. \tag{B.1.220}$$

Then, by (B.1.219), (B.1.220), and using Itô formula, one sees that

$$\begin{aligned} \varphi(\Delta(\xi(t))) &= |y_1 - y_2|^q + \int_0^t \sum_{j=1}^D (\partial_{\xi^j}\varphi)(\Delta\xi(s))\, d(\Delta\xi^j(s)) \\ &\quad + \frac{1}{2}\int_0^t \sum_{j=1}^D \sum_{k=1}^D (\partial_{\xi^j}\partial_{\xi^k}\varphi)(\Delta\xi(s))\, d\left[\Delta\xi^j, \Delta\xi^k\right](s), \quad \forall t \in [0, \infty). \end{aligned} \tag{B.1.221}$$

From Fact B.0.9 on page 139, we have

$$(\partial_{\xi^j}\varphi)(\xi) = q\,\xi^j\,|\xi|^{q-2}, \quad \forall \xi \in I\!R^D, \tag{B.1.222}$$

$$(\partial_{\xi^j}\partial_{\xi^k}\varphi)(\xi) = q\,\delta_{j,k}|\xi|^{q-2} + q(q-2)\xi^j\,\xi^k\,|\xi|^{q-4}, \quad \forall \xi \in I\!R^D. \tag{B.1.223}$$

Also from (B.1.219) we have

$$\begin{aligned}\left[\Delta\xi^j, \Delta\xi^k\right](t) &= \left[\sum_{m=1}^{N}\int_0^{\cdot} \Delta\sigma^{j,m}(s)d\beta^m(s), \sum_{n=1}^{N}\int_0^{\cdot} \Delta\sigma^{k,n}(s)d\beta^n(s)\right](t),\\ &= \sum_{n=1}^{N}\sum_{m=1}^{N}\int_0^t \Delta\sigma^{j,m}(s)\ \Delta\sigma^{k,n}(s)\ d[\beta^m,\beta^n](s)\\ &= \sum_{m=1}^{N}\int_0^t \Delta\sigma^{j,m}(s)\ \Delta\sigma^{k,m}(s)\ ds, \quad \forall t\in[0,\infty)\end{aligned}$$

(B.1.224)

Using (B.1.222) and (B.1.219), the second term in (B.1.221) can be written:

$$\begin{aligned}&\sum_{j=1}^{D}\int_0^t (\partial_{\xi^j}\varphi)(\Delta\xi(s))\ d(\Delta\xi^j(s))\\ &\qquad = q\sum_{j=1}^{D}\int_0^t |\Delta\xi(s)|^{q-2}\Delta\xi^j(s)\ d(\Delta\xi^j(s))\\ &\qquad = q\int_0^t |\Delta\xi(s)|^{q-2}(\Delta\xi(s))^T d(\Delta\xi(s))\\ &\qquad = q\int_0^t |\Delta\xi(s)|^{q-2}\ (\Delta\xi(s))^T\ \Delta b(s)\ ds\\ &\qquad\qquad +q\int_0^t |\Delta\xi(s)|^{q-2}\ (\Delta\xi(s))^T\ \Delta\sigma(s)\ d\beta(s),\end{aligned}$$

(B.1.225)

$\forall t\in[0,\infty)$. Moreover, by (B.1.223) and (B.1.224), for the third term in (B.1.221) we have

$$\begin{aligned}
&\frac{1}{2}\sum_{j=1}^{D}\sum_{k=1}^{D}\int_0^t (\partial_{\xi^j}\partial_{\xi^k}\varphi)(\Delta\xi(s))\, d\left[\Delta\xi^j,\Delta\xi^k\right](s) \\
&\quad = \frac{q}{2}\sum_{j=1}^{D}\int_0^t |\Delta\xi(s)|^{q-2} d[\Delta\xi^j](s) \\
&\qquad + \frac{q}{2}(q-2)\sum_{j=1}^{D}\sum_{k=1}^{D}\int_0^t \Delta\xi^j(s)\Delta\xi^k(s)|\Delta\xi(s)|^{q-4}\, d[\Delta\xi^j,\Delta\xi^k](s) \\
&\quad = \frac{q}{2}\sum_{j=1}^{D}\sum_{m=1}^{N}\int_0^t |\Delta\xi(s)|^{q-2}\, |\Delta\sigma^{j,m}(s)|^2\, ds \\
&\qquad + \frac{q}{2}(q-2)\sum_{j=1}^{D}\sum_{k=1}^{D}\sum_{m=1}^{N}\int_0^t |\Delta\xi(s)|^{q-4}\Delta\xi^j(s)\Delta\xi^k(s)\Delta\sigma^{j,m}(s)\Delta\sigma^{k,m}(s)\, ds \\
&\quad = \frac{q}{2}\int_0^t |\Delta\xi(s)|^{q-2}\, |\Delta\sigma(s)|^2\, ds \\
&\qquad + \frac{q}{2}(q-2)\int_0^t |\Delta\xi(s)|^{q-4}\, (\Delta\xi(s))^T\left[\Delta\sigma(s)\,(\Delta\sigma(s))^T\right]\,\Delta\xi(s)\, ds,
\end{aligned} \tag{B.1.226}$$

$\forall t\in[0,\infty)$.

Combining (B.1.221), (B.1.225), and (B.1.226) one has

$$\begin{aligned}
|\Delta\xi(t)|^q \;=\;& |y_1-y_2|^q + q\int_0^t |\Delta\xi(s)|^{q-2}\,(\Delta\xi(s))^T\,\Delta b(s)\, ds \\
&+ q\int_0^t |\Delta\xi(s)|^{q-2}\,(\Delta\xi(s))^T\,\Delta\sigma(s)\, d\beta(s) \\
&+ \frac{1}{2}q\int_0^t |\Delta\xi(s)|^{q-2}\,|\Delta\sigma(s)|^2\, ds \\
&+ \frac{q}{2}(q-2)\int_0^t |\Delta\xi(s)|^{q-4}\,(\Delta\xi(s))^T\left[\Delta\sigma(s)\,(\Delta\sigma(s))^T\right]\Delta\xi(s)\, ds,
\end{aligned} \tag{B.1.227}$$

$\forall t \in [0, \infty)$. Now by (B.1.217), (B.1.215), and the mean value theorem we get

$$\begin{aligned}\Delta b(s) &= \left[\int_0^1 (\partial_\xi b)\,(x, \xi(s, x, y_2) + \theta(\Delta\xi(s)))\,d\theta\right]\;\Delta\xi(s)\\ &= \int_0^1 J\,(x, \xi(s, x, y_2) + \theta\;\Delta\xi(s))\;\;\Delta\xi(s)\;d\theta. \qquad \text{(B.1.228)}\end{aligned}$$

(recall (3.4.65) where $J(x, \xi) \triangleq \partial_\xi b(x, \xi)$). Thus from (B.1.228),

$$(\Delta\xi(s))^T \Delta b(s) = \int_0^1 (\Delta\xi(s))^T\; J\,(x, \xi(s, x, y_2) + \theta\;\Delta\xi(s))\;\;\Delta\xi(s)\;d\theta. \qquad \text{(B.1.229)}$$

Now, by Lemma C.0.12 on page 205, one has for the integrand of (B.1.229),

$$\begin{aligned}&(\Delta\xi(s))^T\; J\,(x, \xi(s, x, y_2) + \theta\;\Delta\xi(s))\;\;\Delta\xi(s)\\ &= (\Delta\xi(s))^T\; J^T\,(x, \xi(s, x, y_2) + \theta\;\Delta\xi(s))\;\;\Delta\xi(s)\\ &= \frac{1}{2}(\Delta\xi(s))^T\;\left[J\,(x, \xi(s, x, y_2) + \theta\;\Delta\xi(s)) + J^T\,(x, \xi(s, x, y_2) + \theta\;\Delta\xi(s))\right]\Delta\xi(s)\\ &\le \Lambda_1(R)\;|\Delta\xi(s)|^2. \qquad \text{(B.1.230)}\end{aligned}$$

where $\Lambda_1(R)$ is defined by (3.4.66). Thus, from (B.1.230) and (B.1.229), we have

$$|\Delta\xi(s)|^{q-2}\;(\Delta\xi(s))^T\;\Delta b(s) \le \Lambda_1(R)\;|\Delta\xi(s)|^q. \qquad \text{(B.1.231)}$$

Moreover, from (B.1.215), (B.1.218) and Lemma 3.4.5 on page 47

$$|\Delta\xi(s)|^{q-2}\;|\Delta\sigma(s)|^2 \le \Lambda_0^2(R)\;|\Delta\xi(s)|^q. \qquad \text{(B.1.232)}$$

Also

$$\begin{aligned}&(\Delta\xi(s))^T\;\left[\Delta\sigma(s)\;(\Delta\sigma(s))^T\right]\;\Delta\xi(s)\\ &\quad\le \text{Trace}\left[\Delta\sigma(s)\;(\Delta\sigma(s))^T\right]\;|\Delta\xi(s)|^2\\ &\quad\le |\Delta\sigma(s)|^2\;|\Delta\xi(s)|^2. \qquad \text{(B.1.233)}\end{aligned}$$

Thus one has from (B.1.232) and (B.1.233)

$$\begin{aligned}|\Delta\xi(s)|^{q-4}\,(\Delta\xi(s))^T\left[\Delta\sigma(s)(\Delta\sigma(s))^T\right]\Delta\xi(s) \\ \le |\Delta\xi(s)|^{q-2}|\Delta\sigma(s)|^2 \\ \le \Lambda_0^2(R)|\Delta\xi(s)|^q. \end{aligned} \tag{B.1.234}$$

By Lemma C.0.6 on page 203 (see (C.0.7)) one easily sees that

$$\int_0^t E|\Delta\xi(s)|^{2q}\,ds < \infty, \quad \forall t \in [0,\infty). \tag{B.1.235}$$

Thus the stochastic integral on the right hand side of (B.1.227) is a continuous martingale null at the origin. Now, put

$$\theta(t) \stackrel{\Delta}{=} E|\Delta\xi(t)|^q. \tag{B.1.236}$$

Therefore, taking expectations in (B.1.227) and using Fubini's theorem, we get

$$\begin{aligned}\theta(t) &= |y_1-y_2|^q + q\int_0^t E\left[|\Delta\xi(s)|^{q-2}(\Delta\xi(s))^T\Delta b(s)\right]\,ds \\ &+\frac{q}{2}\int_0^t E\left[|\Delta\xi(s)|^{q-2}|\Delta\sigma(s)|^2\right]\,ds \\ &+\frac{q}{2}(q-2)\int_0^t E\left[|\Delta\xi(s)|^{q-4}(\Delta\xi(s))^T\left[\Delta\sigma(s)(\Delta\sigma(s))^T\right]\Delta\xi(s)\right]\,ds,\end{aligned} \tag{B.1.237}$$

$\forall t \in [0,\infty)$. Hence (B.1.237), (B.1.231), (B.1.232), and (B.1.234) implies

$$\theta(0) = |y_1-y_2|^q \tag{B.1.238}$$

and

$$\begin{aligned}\frac{d\theta(t)}{dt} &= qE\left[|\Delta\xi(t)|^{q-2}(\Delta\xi(t))^T\Delta b(t)\right] \\ &+\frac{q}{2}E\left[|\Delta\xi(t)|^{q-2}|\Delta\sigma(t)|^2\right] \\ &+\frac{q}{2}(q-2)E\left[|\Delta\xi(t)|^{q-4}(\Delta\xi(t))^T\left[\Delta\sigma(t)(\Delta\sigma(t))^T\right]\Delta\xi(t)\right] \\ &\le q\Lambda_1(R)\theta(t) + \frac{q}{2}\Lambda_0^2(R)\theta(t) + \frac{q}{2}(q-2)\Lambda_0^2(R)\theta(t), \quad \forall t\in[0,\infty)\end{aligned} \tag{B.1.239}$$

Thus, from (B.1.239), we have

$$\frac{d\theta(t)}{dt} \leq -\gamma_1(R)\,\theta(t), \quad \forall t \in [0,\infty), \tag{B.1.240}$$

where

$$\gamma_1(R) \triangleq -q\Lambda_1(R) - \frac{q}{2}(q-1)\,[\Lambda_0(R)]^2 . \tag{B.1.241}$$

By Condition 3.4.7, we observe that $\gamma_1(R) \in (0,\infty)$. Combining (B.1.238), (B.1.240), and (B.1.216), gives

$$E|\Delta\xi(t)|^q \leq e^{-\gamma_1(R)t}|y_1 - y_2|^q, \quad \forall t \in [0,\infty). \tag{B.1.242}$$

Without loss of generality take $r_2 \in (0,q]$. Now, by Liapunov's L_p-inequality and (B.1.242), for each $r_2 \in (0,q]$, one has

$$E|\Delta\xi(t)|^{r_2} \leq e^{-\gamma(r_2,R)t}|y_1 - y_2|^{r_2}, \quad \forall t \in [0,\infty). \tag{B.1.243}$$

where $\gamma(r_2,R) \in (0,\infty)$ is a constant. Thus from (B.1.243), (B.1.242), and (B.1.215), for each $r_2 \in [0,q]$, we have

$$E|\xi(t,x,y_1) - \xi(t,x,y_2)|^{r_2} \leq e^{-\gamma(r_2,R)t}|y_1 - y_2|^{r_2}, \quad \forall t \in [0,\infty), \tag{B.1.244}$$

as required.

□

Remark B.1.2. For the next proposition we will need the notion of the partial derivatives $(\partial_{y^i}\xi)(t,x,y)$ and $(\partial_{y^i}\partial_{y^k}\xi)(t,x,y)$ of the solution of (3.4.74) in a certain L_2 - sense which is made explicit in Appendix F. The bounds developed in Proposition B.1.3 on page 184 are essential for the proof of Proposition 3.4.16 on page 51.

Proposition B.1.3. *Suppose that Conditions 3.4.4 and 3.4.7 hold. Then, with reference to Remark B.1.2 and Appendix F, for each $R \in [0, \infty)$ there are constants $C(R) \in [0, \infty)$ and $\gamma_1(R), \gamma_2(R) \in (0, \infty)$, such that*

$$E\left|\partial_{y^l}\xi(t,x,y)\right|^4 \leq e^{-\gamma_1(R)t}, \tag{B.1.245}$$

$$E\left|\partial_{y^l}\partial_{y^k}\xi(t,x,y)\right|^2 \leq C(R)e^{-\gamma_2(R)t}, \tag{B.1.246}$$

$\forall(t,x,y) \in [0,\infty) \otimes S^d_R \otimes \mathbb{R}^D, \forall l, k = 1, 2, \ldots, D.$

Proof: For $i = 1, 2, \ldots, D$, define the $\mathbb{R}^D$-valued random vector

$$\vartheta_i(t,x,y) \triangleq (\partial_{y^i}\xi)(t,x,y), \quad \forall(t,x,y) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D. \tag{B.1.247}$$

Also, for $x \in \mathbb{R}^d$ and $\xi, \vartheta \in \mathbb{R}^D$, let $B(x,\xi,\vartheta)$ be the D by N matrix whose (k,n) - element is defined by

$$B^{k,n}(x,\xi,\vartheta) \triangleq (\partial_\xi \sigma^{k,n})(x,\xi)\vartheta, \tag{B.1.248}$$

$\forall k = 1, 2, \ldots, D, \forall n = 1, 2, 3, \ldots, N$ (recall that $(\partial_\xi \sigma^{k,n})(x,\xi)$ is a row vector with D entries given by $\partial_{\xi^l}\sigma^{k,n}(x,\xi), l = 1, 2, \ldots, D$). Moreover, write $\xi(t)$ for $\xi(t,x,y)$ and $\vartheta_i(t)$ for $\vartheta_i(t,x,y)$, when there is no risk of confusion. By Remark F.0.32 and (3.4.74), one has

$$\begin{aligned}(\partial_{y^i}\xi^k)(t,x,y) &= \delta_{i,k} + \sum_{l=1}^{D}\int_0^t (\partial_{\xi^l}b^k)(x,\xi(s,x,y))(\partial_{y^i}\xi^l)(s,x,y)ds \\ &\quad + \sum_{n=1}^{N}\sum_{l=1}^{D}\int_0^t (\partial_{\xi^l}\sigma^{k,n})(x,\xi(s,x,y))(\partial_{y^i}\xi^l)(s,x,y)d\beta^n(s),\end{aligned} \tag{B.1.249}$$

$\forall(t,x,y) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$, $\forall i, k = 1, 2, 3, \ldots, D$, (where $\delta_{i,k}$ is the Kronecker delta). Using (B.1.247), (B.1.248), and the $D \times D$ Jacobian matrix $J(.,.)$ defined by (3.4.65), the system of equations in (B.1.249) can be written in vector form as follows

$$\vartheta_i(t) = e_i + \int_0^t J(x,\xi(s))\vartheta_i(s)\, ds + \int_0^t B\left(x,\xi(s),\vartheta_i(s)\right)\, d\beta(s), \tag{B.1.250}$$

$\forall i = 1, 2, \ldots, D$, (where $e_i = (0, \ldots, 1, 0, \ldots, 0)^T$ is the usual i-th canonical basis vector in $\mathbb{R}^D$). Now put

$$\varphi(\vartheta) \triangleq |\vartheta|^4, \quad \forall \vartheta \in \mathbb{R}^D. \tag{B.1.251}$$

By (B.1.250), (B.1.251), and using Itô's formula, one has

$$\begin{aligned} \varphi(\vartheta_i(t)) \;\triangleq\; & |e_i|^4 + \sum_{j=1}^{D} \int_0^t \left(\partial_{\vartheta^j}\varphi\right)(\vartheta_i(s))\, d\vartheta_i^j(s) \\ & + \frac{1}{2}\sum_{j=1}^{D}\sum_{k=1}^{D} \int_0^t \left(\partial_{\vartheta^j}\partial_{\vartheta^k}\varphi\right)(\vartheta_i(s))\, d\left[\vartheta_i^j, \vartheta_i^k\right](s), \end{aligned} \tag{B.1.252}$$

$\forall t \in [0, \infty), \forall i = 1, 2, \ldots, D$. From Fact B.0.9 on page 139 and (B.1.251) we have

$$(\partial_{\vartheta^j}\varphi)(\vartheta) = 4\vartheta^j |\vartheta|^2, \tag{B.1.253}$$

$$(\partial_{\vartheta^j}\partial_{\vartheta^k}\varphi)(\vartheta) = 4|\vartheta|^2\delta_{j,k} + 8\vartheta^j\vartheta^k, \tag{B.1.254}$$

$\forall j, k = 1, 2, \ldots, D$. Thus, from (B.1.250) and (B.1.253), for the second term on right hand side of (B.1.252) we have

$$\begin{aligned} & \sum_{j=1}^{D} \int_0^t \left(\partial_{\vartheta^j}\varphi\right)(\vartheta_i(s))\, d\vartheta_i^j(s) \\ & = 4\int_0^t |\vartheta_i(s)|^2 (\vartheta_i(s))^T\, d\vartheta_i(s) \\ & = 4\int_0^t |\vartheta_i(s)|^2 (\vartheta_i(s))^T J(x,\xi(s))\vartheta_i(s)\, ds \\ & \quad + 4\int_0^t |\vartheta_i(s)|^2 (\vartheta_i(s))^T B(x,\xi(s),\vartheta_i(s))\, d\beta(s), \end{aligned} \tag{B.1.255}$$

From (B.1.250), the co-quadratic variation in (B.1.252) is given by

$$
\begin{aligned}
&\left[\vartheta_i^j, \vartheta_i^k\right](t) \\
&= \Bigg[\int_0^{\cdot} \sum_{n=1}^{N} B^{j,n}(x, \xi(s), \vartheta_i(s))\, d\beta^n(s), \\
&\qquad\qquad \int_0^{\cdot} \sum_{m=1}^{N} B^{k,m}(x, \xi(s), \vartheta_i(s))\, d\beta^m(s)\Bigg](t) \\
&= \sum_{m=1}^{N} \sum_{n=1}^{N} \int_0^t B^{j,n}(x, \xi(s), \vartheta_i(s)) B^{k,m}(x, \xi(s), \vartheta_i(s))\, d\left[\beta^n, \beta^m\right](s) \\
&= \sum_{n=1}^{N} \int_0^t B^{j,n}(x, \xi(s), \vartheta_i(s)) B^{k,n}(x, \xi(s), \vartheta_i(s))\, ds \\
&= \int_0^t \left[B(x, \xi(s), \vartheta_i(s)) B^T(x, \xi(s), \vartheta_i(s))\right]^{j,k} ds \\
&\triangleq \int_0^t \left[BB^T(x, \xi(s), \vartheta_i(s))\right]^{j,k} ds
\end{aligned}
\tag{B.1.256}
$$

$\forall t \in [0, \infty)$. Thus, from (B.1.256) and (B.1.254), for the third term in (B.1.252), one gets

$$\frac{1}{2}\sum_{j=1}^{D}\sum_{k=1}^{D}\int_0^t (\partial_{\vartheta^j}\partial_{\vartheta^k}\varphi)(\vartheta_i(s))\, d\left[\vartheta_i^j, \vartheta_i^k\right](s)$$
$$= \frac{1}{2}\sum_{j=1}^{D}\sum_{k=1}^{D}\int_0^t (\partial_{\vartheta^j}\partial_{\vartheta^k}\varphi)(\vartheta_i(s))\left[BB^T(x,\xi(s),\vartheta_i(s))\right]^{j,k} ds$$
$$= 2\sum_{j=1}^{D}\int_0^t \left[BB^T(x,\xi(s),\vartheta_i(s))\right]^{j,j}|\vartheta_i(s)|^2\, ds$$
$$+4\sum_{j=1}^{D}\sum_{k=1}^{D}\int_0^t \left[BB^T(x,\xi(s),\vartheta_i(s))\right]^{j,k}\vartheta_i^j(s)\vartheta_i^k(s)\, ds$$
$$= 2\int_0^t \text{Trace}\left[BB^T(x,\xi(s),\vartheta_i(s))\right]|\vartheta_i(s)|^2\, ds$$
$$+4\int_0^t (\vartheta_i(s))^T\left[BB^T(x,\xi(s),\vartheta_i(s))\right]\vartheta_i(s)\, ds \tag{B.1.257}$$

Now, by (B.1.251), (B.1.252), (B.1.255) and (B.1.257), it follows that

$$\begin{aligned}|\vartheta_i(t)|^4 \;=\; & |e_i|^4 + 4\int_0^t |\vartheta_i(s)|^2(\vartheta_i(s))^T J(x,\xi(s))\vartheta_i(s)\, ds \\ & +2\int_0^t \text{Trace}\left[BB^T(x,\xi(s),\vartheta_i(s))\right]|\vartheta_i(s)|^2\, ds \\ & +4\int_0^t (\vartheta_i(s))^T\left[BB^T(x,\xi(s),\vartheta_i(s))\right]\vartheta_i(s)\, ds \\ & +4\int_0^t |\vartheta_i(s)|^2(\vartheta_i(s))^T B(x,\xi(s),\vartheta_i(s))\, d\beta(s).\end{aligned} \tag{B.1.258}$$

$\forall t \in [0,\infty)$. One easily sees from Lemma C.0.6 on page 203 that the stochastic integral on right hand side of (B.1.258) is a continuous martingale, hence

$$E\left\{\int_0^t |\vartheta_i(s)|^2(\vartheta_i(s))^T B(x,\xi(s),\vartheta_i(s))\, d\beta(s)\right\} = 0, \tag{B.1.259}$$

$\forall (t, x) \in [0, \infty) \otimes I\!R^d$. Now put

$$\theta_1(t, x, y) \triangleq E\left[|\vartheta_i(t, x, y)|^4\right], \quad \forall (t, x, y) \in [0, \infty) \otimes I\!R^d \otimes I\!R^D. \quad \text{(B.1.260)}$$

Thus, from (B.1.258), (B.1.259) and (B.1.260), we get

$$\begin{aligned} \theta_1(t, x, y) &= 1 + 4 \int_0^t E\left[|\vartheta_i(s)|^2 (\vartheta_i(s))^T J(x, \xi(s))\vartheta_i(s)\right] ds \\ &\quad + 2 \int_0^t E\left[\text{Trace}\left(BB^T(x, \xi(s), \vartheta_i(s))\right) |\vartheta_i(s)|^2\right] ds \\ &\quad + 4 \int_0^t E\left[(\vartheta_i(s))^T \left(BB^T(x, \xi(s), \vartheta_i(s))\right) \vartheta_i(s)\right] ds. \end{aligned} \quad \text{(B.1.261)}$$

Thus

$$\begin{aligned} (\partial_t \theta_1)(t, x, y) &= 4E\left[|\vartheta_i(t)|^2 (\vartheta_i(t))^T J(x, \xi(t))\vartheta_i(t)\right] \\ &\quad + 2E\left[\text{Trace}\left(BB^T(x, \xi(t), \vartheta_i(t))\right) |\vartheta_i(t)|^2\right] \\ &\quad + 4E\left[(\vartheta_i(t))^T \left(BB^T(x, \xi(t), \vartheta_i(t))\right) \vartheta_i(t)\right], \end{aligned} \quad \text{(B.1.262)}$$

$\forall t \in [0, \infty)$, where $\theta_1(0, x, y) = 1$. Now fix $R \in [0, \infty)$. From the definition of $\Lambda_1(R)$ (see (3.4.66)) and Lemma C.0.12 on page 205:

$$\begin{aligned} \vartheta^T J(x, \xi)\vartheta &= \vartheta^T \left[\frac{J(x, \xi) + J^T(x, \xi)}{2}\right] \vartheta \\ &\leq \Lambda_1(R)|\vartheta|^2, \quad \forall x \in S_R^d, \ \forall \xi, \vartheta \in I\!R^D. \end{aligned} \quad \text{(B.1.263)}$$

Moreover, since $BB^T(x, \xi, \varphi)$ is positive semi-definite, one easily sees that

$$\vartheta^T \left[BB^T(x, \xi, \vartheta)\right] \vartheta \leq \text{Trace}\left[BB^T(x, \xi, \vartheta)\right] |\vartheta|^2, \quad \forall x \in I\!R^d, \ \forall \xi, \vartheta \in I\!R^D. \quad \text{(B.1.264)}$$

Combining (B.1.262), (B.1.263), (B.1.264) and (B.1.260) one has

$$(\partial_t \theta_1)(t, x, y) \leq 4\Lambda_1(R)\theta_1(t, x, y) + 6E\left[\text{Trace}\left[BB^T(x, \xi(t), \vartheta_i(t))\right] |\vartheta_i(t)|^2\right], \quad \text{(B.1.265)}$$

$\forall(t,x,y) \in [0,\infty) \otimes S_R^d \otimes I\!R^D$. From (B.1.248), (3.4.64), and Cauchy-Schwarz, one observes that

$$\begin{aligned}
\text{Trace}\left[BB^T(x,\xi,\vartheta)\right] &= \sum_{n=1}^{N}\sum_{k=1}^{D}\left[B^{k,n}(x,\xi,\vartheta)\right]^2 \\
&= \sum_{n=1}^{N}\sum_{k=1}^{D}\left[\left(\partial_\xi \sigma^{k,n}\right)(x,\xi)\vartheta\right]^2 \\
&\leq \sum_{n=1}^{N}\sum_{k=1}^{D}\left\{\sum_{l=1}^{D}\left[\left(\partial_{\xi^l}\sigma^{k,n}(x,\xi)\right)\right]^2\right\}|\vartheta|^2 \\
&\leq \left[\sum_{l=1}^{D}\left|\left(\partial_{\xi^l}\sigma\right)(x,\xi)\right|^2\right]|\vartheta|^2 \\
&\leq \Lambda_0^2(R)\,|\vartheta|^2, \quad \forall x \in S_R^d, \forall \xi,\vartheta \in I\!R^D.
\end{aligned} \tag{B.1.266}$$

By (B.1.265), (B.1.266) and (B.1.260) we have

$$\begin{aligned}
(\partial_t\theta_1)(t,x,y) &\leq 4\Lambda_1(R)\theta_1(t,x,y) + 6\Lambda_0^2(R)\theta_1(t) \\
&\leq -\gamma_1(R)\theta_1(t),
\end{aligned} \tag{B.1.267}$$

for

$$\gamma_1(R) \triangleq -4\Lambda_1(R) - 6\Lambda_0^2(R). \tag{B.1.268}$$

From Condition 3.4.7 we have

$$4\Lambda_1(R) < 2(1-q)\Lambda_0^2(R)$$

for a constant $q \in (8,\infty)$, thus

$$0 < [2(q-1)-6]\Lambda_0^2(R) < -4\Lambda_1(R) - 6\Lambda_0^2(R)$$

so that $\gamma_1(R) \in (0,\infty)$. Now, from (B.1.267), (B.1.268), and the fact that $\theta_1(0,x,y) = 1$, one sees that

$$\begin{aligned}
\theta_1(t,x,y) &\leq \theta_1(0,x,y)e^{-\gamma_1(R)t} \\
&= e^{-\gamma_1(R)t}, \quad \forall(t,x,y) \in [0,\infty) \otimes S_R^d \otimes I\!R^d.
\end{aligned} \tag{B.1.269}$$

Thus, the first inequality of (B.1.245) follows from (B.1.260) and (B.1.269), i.e.

$$E\left|\vartheta_i(t,x,y)\right|^4 \le e^{-\gamma_1(R)t}, \quad \forall (t,x,y) \in [0,\infty)\otimes S_R^d \otimes \mathbb{R}^D. \tag{B.1.270}$$

Next, we establish the second inequality of (B.1.246). put

$$\chi_{i,j}(t,x,y) \stackrel{\Delta}{=} \partial_{y^i}\partial_{y^j}\xi(t,x,y), \tag{B.1.271}$$

for $i,j = 1,2,\dots,D$. From Remark F.0.32 and (B.1.249) the second derivative $\left(\partial_{y^j}\partial_{y^i}\xi\right)(t,x,y)$ can be computed formally as the solution of the following equation

$$\begin{aligned}
\left(\partial_{y^j}\partial_{y^i}\xi^k\right)(t,x,y) &= \sum_{l=1}^{D}\int_0^t (\partial_{\xi^l}b^k)(x,\xi(s,x,y))\ \left(\partial_{y^j}\partial_{y^i}\xi^l\right)(s,x,y)\,ds \\
&+\sum_{l=1}^{D}\sum_{p=1}^{D}\int_0^t \left[\left(\partial_{\xi^p}\partial_{\xi^l}b^k\right)(x,\xi(s,x,y))\times\right. \\
&\qquad \left.\left(\partial_{y^j}\xi^p\right)(s,x,y)\left(\partial_{y^i}\xi^l\right)(s,x,y)\right]\,ds \\
&+\sum_{n=1}^{N}\sum_{l=1}^{D}\int_0^t \left(\partial_{\xi^l}\sigma^{k,n}\right)(x,\xi(s,x,y))\left(\partial_{y^j}\partial_{y^i}\xi^l\right)(s,x,y)\,d\beta^n(s) \\
&+\sum_{n=1}^{N}\sum_{l=1}^{D}\sum_{p=1}^{D}\int_0^t \left[\left(\partial_{\xi^p}\partial_{\xi^l}\sigma^{k,n}\right)(x,\xi(s,x,y))\times\right. \\
&\qquad \left.(\partial_{y^j}\xi^p)(s,x,y)\,(\partial_{y^i}\xi^l)(s,x,y)\right]\,d\beta^n(s),
\end{aligned} \tag{B.1.272}$$

$\forall (t,x,y) \in [0,\infty)\otimes \mathbb{R}^d \otimes \mathbb{R}^D, \forall i,j,k = 1,2,\dots,D$. Rearranging (B.1.272), for arbitrary $i, j, k \in \{1, 2, \ldots, D\}$, we get

$$\begin{aligned}
\left(\partial_{y^j}\partial_{y^i}\xi^k\right)(t,x,y) \;=\; & \int_0^t (\partial_\xi b^k)(x,\xi(s,x,y)) \; \left(\partial_{y^j}\partial_{y^i}\xi\right)(s,x,y)\, ds \\
& + \int_0^t \left[(\partial_{y^j}\xi)(s,x,y)\right]^T \left[(\partial_\xi\partial_\xi b^k)(x,\xi(s,x,y))\right] \times \\
& \qquad \left[(\partial_{y^i}\xi)(s,x,y)\right] ds \\
& + \sum_{n=1}^{N} \int_0^t \left[(\partial_\xi \sigma^{k,n})(x,\xi(s,x,y)\right] \left[(\partial_{y^j}\partial_{y^i}\xi)(s,x,y)\right] d\beta^n(s) \\
& + \sum_{n=1}^{N} \int_0^t \left[(\partial_{y^j}\xi)(s,x,y)\right]^T \left[(\partial_\xi\partial_\xi\sigma^{k,n})(x,\xi(s,x,y)\right] \times \\
& \qquad \left[(\partial_{y^i}\xi)(s,x,y)\right] d\beta^n(s)
\end{aligned}$$

(B.1.273)

Here we use $(\partial_\xi\partial_\xi b^k)(x,\xi)$ to denote the $D \times D$ symmetric matrix whose (p,l) element is given by $(\partial_{\xi^p}\partial_{\xi^l} b^k)(x,\xi), p, l = 1, 2, \ldots, D$. The notation $(\partial_\xi\partial_\xi\sigma^{k,n})(x,\xi)$ likewise indicates the $D \times D$ symmetric matrix whose (p,l) element is given by $(\partial_{\xi^p}\partial_{\xi^l}\sigma^{k,n})(x,\xi)$. Let $A(s,x,y)$ be the $D \otimes 1$ vector whose $k-th$ scalar entry is $A^k(s,x,y)$ defined by

$$A^k(s,x,y) \triangleq \left[(\partial_{y^j}\xi)(s,x,y)\right]^T \left[(\partial_\xi\partial_\xi b^k)(x,\xi(s,x,y)\right] \left[(\partial_{y^i}\xi)(s,x,y)\right],$$

(B.1.274)

$\forall k = 1, 2, \ldots, D$, and also let $C(s,x,y)$ be the $D \otimes N$ matrix whose (k,n)-entry is $C^{k,n}(s,x,y)$ given by

$$C^{k,n}(s,x,y) \triangleq \left[(\partial_{y^j}\xi)(s,x,y)\right]^T \left[(\partial_\xi\partial_\xi\sigma^{k,n})(x,\xi(s,x,y))\right] \left[(\partial_{y^i}\xi)(s,x,y)\right],$$

(B.1.275)

$\forall k = 1, 2, \ldots, D, n = 1, 2, \ldots, N$. Now using (B.1.271), (B.1.248), and (3.4.65), one can write (B.1.273) more compactly in vector form as follows

$$\begin{aligned}\chi_{i,j}(t,x,y) &= \int_0^t J(x,\xi(s,x,y))\chi_{i,j}(s,x,y)\,ds \\ &\quad+\int_0^t B(x,\xi(s,x,y),\chi_{i,j}(s,x,y))\,d\beta(s) \\ &\quad+\int_0^t A(s,x,y)\,ds+\int_0^t C(s,x,y)\,d\beta(s),\end{aligned} \tag{B.1.276}$$

$\forall(t,x,y)\in[0,\infty)\otimes\mathbb{R}^d\otimes\mathbb{R}^D, \forall i,j=1,2,\ldots,D$. Again, we write $\chi_{i,j}(t)$ for $\chi_{i,j}(t,x,y)$, $A(s)$ for $A(s,x,y)$, and $C(s)$ for $C(s,x,y)$ when there is no risk of confusion. Now, put

$$\varphi(\chi)\triangleq|\chi|^2,\quad\forall\chi\in\mathbb{R}^D. \tag{B.1.277}$$

Thus, by (B.1.276) and Itô formula, one gets

$$\begin{aligned}\varphi(\chi_{i,j}(t)) &= \varphi(\chi_{i,j}(0))+\sum_{k=1}^D\int_0^t\left(\partial_{\chi^k}\varphi\right)(\chi_{i,j}(s))\,d\chi_{i,j}^k(s) \\ &\quad+\frac{1}{2}\sum_{l=1}^D\sum_{k=1}^D\int_0^t\left(\partial_{\chi^l}\partial_{\chi^k}\varphi\right)(\chi_{i,j}(s))\,d\left[\chi_{i,j}^l,\chi_{i,j}^k\right](s).\end{aligned} \tag{B.1.278}$$

$\forall t\in[0,\infty)$. Now evaluate terms on right hand side of (B.1.278): From Fact B.0.9 on page 139, for each $\chi\in\mathbb{R}^D$, we have

$$\begin{aligned}\left(\partial_{\chi^k}\varphi\right)(\chi) &= 2\chi^k \\ \left(\partial_{\chi^l}\partial_{\chi^k}\varphi\right)(\chi) &= 2\delta_{l,k}.\end{aligned} \tag{B.1.279}$$

Moreover, from (B.1.276), one sees that

$$\chi_{i,j}(0)=0. \tag{B.1.280}$$

Combine (B.1.278) (B.1.279) and (B.1.280) to get

$$|\chi_{i,j}(t)|^2=2\int_0^t(\chi_{i,j}(s))^T\,d\chi_{i,j}(s)+\sum_{k=1}^D\left[\chi_{i,j}^k\right](t), \tag{B.1.281}$$

$\forall t \in [0,\infty), \forall i,j = 1,2,\dots,D$. In view of (B.1.276), the first term on the right hand side of (B.1.281) can be written:

$$\begin{aligned}\int_0^t (\chi_{i,j}(s))^T\, d\chi_{i,j}(s) &= \int_0^t (\chi_{i,j}(s))^T J(x,\xi(s))\chi_{i,j}(s)\, ds \\ &+ \int_0^t (\chi_{i,j}(s))^T A(s)\, ds + \int_0^t (\chi_{i,j}(s))^T B(x,\xi(s),\chi_{i,j}(s))\, d\beta(s) \\ &+ \int_0^t (\chi_{i,j}(s))^T C(s)\, d\beta(s). \end{aligned} \tag{B.1.282}$$

From (B.1.276) one has

$$\begin{aligned}\left[\chi_{i,j}^k\right](t) &= \left[\sum_{n=1}^N \int_0^\cdot \left[B^{k,n}(x,\xi(s),\chi_{i,j}(s)) + C^{k,n}(s)\right]\ d\beta^n(s)\right](t) \\ &= \sum_{n=1}^N \int_0^t \left[B^{k,n}(x,\xi(s),\chi_{i,j}(s))) + C^{k,n}(s)\right]^2\, ds \end{aligned} \tag{B.1.283}$$

$\forall k = 1,2,\dots,D$. Thus, for the second term on the right hand side of (B.1.281) we have

$$\begin{aligned}\sum_{k=1}^D \left[\chi_{i,j}^k\right](t) &= \sum_{k=1}^D \sum_{n=1}^N \int_0^t \left[B^{k,n}(x,\xi(s),\chi_{i,j}(s)) + C^{k,n}(s)\right]^2\, ds \\ &= \int_0^t \text{Trace}\left\{\left[B(s)+C(s)\right]\left[B(s)+C(s)\right]^T\right\}\, ds, \end{aligned} \tag{B.1.284}$$

where

$$B(s) \triangleq B(x,\xi(s),\chi_{i,j}(s)). \tag{B.1.285}$$

Now it is easily seen that the stochastic integrals in (B.1.282) are martingales, null at the origin, hence have expectation equal to zero. Combining (B.1.281), (B.1.282), (B.1.284) and taking expectations, it follows that

$$\begin{aligned} E\left|\chi_{i,j}(t)\right|^2 &= 2\int_0^t E\left[(\chi_{i,j}(s))^T J(x,\xi(s))\chi_{i,j}(s)\right] ds \\ &+2\int_0^t E\left[(\chi_{i,j}(s))^T A(s)\right]\ ds \\ &+\int_0^t E\left\{\text{Trace}\left([B(s)+C(s)]\,[B(s)+C(s)]^T\right)\right\}\ ds. \end{aligned} \tag{B.1.286}$$

Now put

$$\theta_2(t,x,y) \triangleq E\left|\chi_{i,j}(t,x,y)\right|^2, \quad \forall (t,x,y)\in[0,\infty)\otimes I\!R^d\otimes I\!R^D. \tag{B.1.287}$$

Thus, using (B.1.286) and (B.1.287), we get

$$\theta_2(0,x,y)=0, \quad \forall (x,y)\in I\!R^d\otimes I\!R^D, \tag{B.1.288}$$

and

$$\begin{aligned} (\partial_t\theta_2)(t,x,y) &= 2E\left[(\chi_{i,j}(t,x,y))^T J(x,\xi(t,x,y))\chi_{i,j}(t,x,y)\right] \\ &+2E\left[(\chi_{i,j}(t,x,y))^T A(t,x,y)\right] \\ &+E\left\{\text{Trace}\left([B(t,x,y)+C(t,x,y)]\,[B(t,x,y)+C(t,x,y)]^T\right)\right\}, \end{aligned} \tag{B.1.289}$$

$\forall (t,x,y)\in[0,\infty)\otimes I\!R^d\otimes I\!R^D$. Now fix $R\in[0,\infty)$. Using the definition of $\Lambda_1(R)$ (see (3.4.66)), (B.1.287), and Lemma C.0.12 on page 205, for the first term on right hand side of (B.1.289), one has

$$\begin{aligned} E\left[(\chi_{i,j}(t,x,y))^T J(x,\xi(t,x,y))\chi_{i,j}(t,x,y)\right] &\le E\left[\Lambda_1(R)|\chi_{i,j}(t,x,y)|^2\right] \\ &\le \Lambda_1(R)\theta_2(t,x,y), \end{aligned} \tag{B.1.290}$$

$\forall (t,x,y)\in[0,\infty)\otimes S_R^d\otimes I\!R^D$. For second term on right hand side of (B.1.289): clearly

$$\left|(\chi_{i,j}(t,x,y))^T A(t,x,y)\right| \le \left|\chi_{i,j}(t,x,y)\right|\left|A(t,x,y)\right|. \tag{B.1.291}$$

$\forall(t,x,y) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$. Now, from (B.1.291), (B.1.274), (B.1.247), and Condition 3.4.4, there is a constant $C_1(R) \in [0,\infty)$ such that

$$\begin{aligned} \left|(\chi_{i,j}(t,x,y))^T A(t,x,y)\right| &\leq C_1(R)\,|\chi_{i,j}(t,x,y)|\,\left|(\partial_{y^j}\xi)(t,x,y)\right|\,\left|(\partial_{y^i}\xi)(t,x,y)\right| \\ &= C_1(R)\,|\chi_{i,j}(t,x,y)|\,|\vartheta_j(t,x,y)|\,|\vartheta_i(t,x,y)|\,, \end{aligned} \tag{B.1.292}$$

$\forall(t,x,y) \in [0,\infty) \otimes S_R^d \otimes \mathbb{R}^D$. By (B.1.292), the Cauchy-Schwarz inequality, and (B.1.245) one finds that

$$\begin{aligned} & E\left|(\chi_{i,j}(t,x,y))^T A(t,x,y)\right| \\ &\quad \leq C_1(R) E\left[|\chi_{i,j}(t,x,y)|\,|\vartheta_j(t,x,y)||\vartheta_i(t,x,y)|\right] \\ &\quad \leq C_1(R) E^{1/2}\left[|\chi_{i,j}(t,x,y)|^2\right] E^{1/4}\left[|\vartheta_j(t,x,y)|^4\right] E^{1/4}\left[|\vartheta_i(t,x,y)|^4\right] \\ &\quad \leq C_1(R) e^{\frac{-\gamma_1(R)}{2}t}\,(\theta_2(t,x,y))^{1/2}, \end{aligned} \tag{B.1.293}$$

$\forall(t,x,y) \in [0,\infty) \otimes S_R^d \otimes \mathbb{R}^D$. As for the third term on right hand side of (B.1.289) we have

$$\begin{aligned} & \text{Trace}\left[(B(t)+C(t))\,(B(t)+C(t))^T\right] \\ &\quad = \text{Trace}\left[B(t)(B(t))^T + B(t)(C(t))^T + C(t)(B(t))^T + C(t)(C(t))^T\right] \\ &\quad = \text{Trace}\left[B(t)(B(t))^T\right] + 2\text{Trace}\left[B(t)(C(t))^T\right] + \text{Trace}\left[C(t)(C(t))^T\right] \end{aligned} \tag{B.1.294}$$

Now evaluate the terms on right hand side of (B.1.294): Exactly as in (B.1.266), for the first term on right hand side of (B.1.294) we get

$$\begin{aligned} \text{Trace}\left[B(t)(B(t))^T\right] &= \text{Trace}\left[B\left(x,\xi(t),\chi_{i,j}(t)\right)\left(B\left(x,\xi(t),\chi_{i,j}(t)\right)\right)^T\right] \\ &\leq \Lambda_0^2(R)\,|\chi_{i,j}(t)|^2. \end{aligned} \tag{B.1.295}$$

As for second term on right hand side of (B.1.294) we can write:

$$\text{Trace}\left[B(t)(C(t))^T\right] = \sum_{k=1}^{D}\left\{\sum_{n=1}^{N} B^{k,n}(t)C^{k,n}(t)\right\}. \tag{B.1.296}$$

Now, by (B.1.275) and Condition 3.4.4, there is a Constant $C_2(R) \in [0, \infty)$ such that

$$\left|C^{k,n}(t)\right| \leq C_2(R)\left|(\partial_{y_j}\xi)(t)\right|\left|(\partial_{y_i}\xi)(t)\right|, \tag{B.1.297}$$

Moreover, from (B.1.248), (B.1.285), and using Condition 3.4.4 one has

$$\left|B^{k,n}(t)\right| \leq C_3(R)\left|\chi_{i,j}(t)\right|, \tag{B.1.298}$$

for some constant $C_3(R) \in [0, \infty)$. Thus, by (B.1.298), (B.1.297), (B.1.296), and (B.1.247) there is a constant $C_4(R) \in [0, \infty)$ such that

$$\text{Trace}\left[B(t)(C(t))^T\right] \leq C_4(R)\left|\chi_{i,j}(t)\right|\left|\vartheta_j(t)\right|\left|\vartheta_i(t)\right| \tag{B.1.299}$$

Upon taking expectation in (B.1.299), using the Cauchy-Schwarz inequality and (B.1.245) we have

$$\begin{aligned} E\text{Trace}\left[B(t)(C(t))^T\right] &\leq C_4(R)E\left\{\left|\chi_{i,j}(t)\right|\left|\vartheta_j(t)\right|\left|\vartheta_i(t)\right|\right\} \\ &\leq C_4(R)E^{1/2}\left[\left|\chi_{i,j}(t)\right|^2\right]E^{1/4}\left[\left|\vartheta_j(t)\right|^4\right]E^{1/4}\left[\left|\vartheta_i(t)\right|^4\right] \\ &\leq C_4(R)e^{-\gamma_3(R)t}(\theta_2(t,x,y))^{1/2}, \end{aligned} \tag{B.1.300}$$

where $\gamma_3(R) \triangleq \frac{\gamma_1(R)}{2}$. From (B.1.297) and (B.1.247) one sees that

$$\begin{aligned} \text{Trace}\left[C(t)(C(t))^T\right] &= \sum_{n=1}^{N}\sum_{k=1}^{D}\left|C^{k,n}(t)\right|^2 \\ &\leq C_5(R)\left|\vartheta_j(t)\right|^2\left|\vartheta_i(t)\right|^2 \end{aligned} \tag{B.1.301}$$

Thus, from (B.1.301), (B.1.245) and the Cauchy - Schwarz inequality one has

$$\begin{aligned} E\text{Trace}\left[C(t)(C(t))^T\right] &\le C_5(R)E^{1/2}\left[|\vartheta_j(t)|^4\right]E^{1/2}\left[|\vartheta_i(t)|^4\right] \\ &\le C_5(R)e^{-\gamma_1(R)t}, \end{aligned} \tag{B.1.302}$$

for constant $C_5(R) \in [0, \infty)$. Combining (B.1.287), (B.1.294), (B.1.295), (B.1.300) and (B.1.302) we get

$$\begin{aligned} &E\left\{\text{Trace}\left[(B(t)+C(t))(B(t)+C(t))^T\right]\right\} \\ &\qquad \le \Lambda_0^2(R)\theta_2(t,x,y) + 2C_4(R)e^{-\gamma_3(R)t}\left(\theta_2(t,x,y)\right)^{1/2} + C_5(R)e^{-\gamma_1(R)t} \end{aligned} \tag{B.1.303}$$

Thus, from (B.1.287), (B.1.289), (B.1.290), (B.1.293) and (B.1.303) it follows that

$$\begin{aligned} &(\partial_t\theta_2)(t,x,y) \\ &\quad \le 2\Lambda_1(R)\theta_2(t,x,y) + 2C_1(R)e^{-\gamma_3(R)t}\left(\theta_2(t,x,y)\right)^{1/2} \\ &\quad +\Lambda_0^2(R)\theta_2(t,x,y) + 2C_4(R)e^{-\gamma_3(R)t}\left(\theta_2(t,x,y)\right)^{1/2} + C_5(R)e^{-\gamma_1(R)t}, \end{aligned} \tag{B.1.304}$$

$\forall (t,x,y) \in [0,\infty) \otimes S_R^d \otimes \mathbb{R}^D$. Simplifying (B.1.304) and using the fact that $\theta^{1/2} \le 1+\theta$, for $\theta \in [0,\infty)$, it is easily seen that

$$\begin{aligned} &(\partial_t\theta_2)(t,x,y) \\ &\le \left[2\Lambda_1(R)+\Lambda_0^2(R)\right]\theta_2(t,x,y) + C_6(R)e^{-\gamma_3(R)t}\left(\theta_2(t,x,y)\right)^{1/2} + C_5(R)e^{-\gamma_1(R)t} \\ &\le \left[2\Lambda_1(R)+\Lambda_0^2(R)\right]\theta_2(t,x,y) + C_6(R)e^{-\gamma_3(R)t}\left[1+\theta_2(t,x,y)\right] + C_5(R)e^{-\gamma_1(R)t} \\ &\le \left[2\Lambda_1(R)+\Lambda_0^2(R)+C_6(R)e^{-\gamma_3(R)t}\right]\theta_2(t,x,y) + C_7(R)e^{-\gamma_3(R)t}, \end{aligned} \tag{B.1.305}$$

$\forall (t,x,y) \in [0,\infty) \otimes S_R^d \otimes \mathbb{R}^D$, for some constants $C_6(R), C_7(R) \in [0,\infty)$. Clearly Condition 3.4.7 implies that

$$0 < [q-2]\Lambda_0^2(R) < -\left[2\Lambda_1(R)+\Lambda_0^2(R)\right]. \tag{B.1.306}$$

Now fix some $t_0(R) \in [0, \infty)$ such that

$$\left[2\Lambda_1(R) + \Lambda_0^2(R) + C_6(R)e^{-\gamma_3(R)t}\right] < \frac{2\Lambda_1(R) + \Lambda_0^2(R)}{2}, \quad \forall t \in [t_0(R), \infty) \tag{B.1.307}$$

From (B.1.306)

$$\gamma_5(R) \triangleq -\frac{2\Lambda_1(R) + \Lambda_0^2(R)}{2} \in (0, \infty). \tag{B.1.308}$$

By (B.1.305), (B.1.307) and (B.1.308) we have

$$(\partial_t\theta_2)(t, x, y) \leq -\gamma_5(R)\theta_2(t, x, y) + C_7(R)e^{-\gamma_3(R)t}, \quad \forall t \in [t_0(R), \infty). \tag{B.1.309}$$

$\forall (t, x, y) \in [0, \infty) \otimes S_R^d \otimes \mathbb{R}^D$. Now fix $\gamma_6(R)$ such that

$$0 < \gamma_6(R) < \min\{\gamma_5(R), \gamma_3(R)\}. \tag{B.1.310}$$

Then $0 < \gamma_6(R) < \gamma_3(R)$ hence from (B.1.309), we have

$$(\partial_t\theta_2)(t, x, y) \leq -\gamma_5(R)\theta_2(t, x, y) + C_7(R)e^{-\gamma_6(R)t}. \tag{B.1.311}$$

Then, by (B.1.311) one gets

$$\begin{aligned}
&\frac{d}{dt}\left[e^{\gamma_5(R)t}\theta_2(t, x, y)\right] \\
&= \gamma_5(R)e^{\gamma_5(R)t}\theta_2(t, x, y) + e^{\gamma_5(R)t}(\partial_t\theta_2)(t, x, y) \\
&\leq \gamma_5(R)e^{\gamma_5(R)t}\theta_2(t, x, y) + \left[-\gamma_5(R)e^{\gamma_5(R)t}\theta_2(t, x, y) + C_7(R)e^{-[\gamma_6(R)-\gamma_5(R)]t}\right] \\
&\leq C_7(R)e^{-[\gamma_6(R)-\gamma_5(R)]t},
\end{aligned} \tag{B.1.312}$$

$\forall(t,x,y) \in [t_0(R),\infty)\otimes S_R^d \otimes I\!R^D$. Now by (B.1.312), for all $t \in [t_0(R),\infty)$, we have

$$\begin{aligned} & e^{\gamma_5(R)t}\theta_2(t,x,y) - e^{\gamma_5(R)t_0(R)}\theta_2(t_0(R),x,y) \\ & \quad = \int_{t_0(R)}^{t} \frac{d}{ds}\left[e^{\gamma_5(R)s}\theta_2(s,x,y)\right]\, ds \\ & \quad \le \frac{C_7(R)\left[e^{(\gamma_5(R)-\gamma_6(R))t} - e^{(\gamma_5(R)-\gamma_6(R))t_0(R)}\right]}{\gamma_5(R)-\gamma_6(R)}. \end{aligned} \tag{B.1.313}$$

$\forall(t,x,y) \in [t_0(R),\infty)\otimes S_R^d \otimes I\!R^D$. Hence, since $\gamma_5(R) > \gamma_6(R)$, we get

$$\begin{aligned} \theta_2(t,x,y) \ \le\ & e^{\gamma_5(R)t_0(R)}\theta_2(t_0(R),x,y)e^{-\gamma_5(R)t} \\ & + \frac{C_7(R)\left[e^{-\gamma_6(R)t}\right]}{\gamma_5(R)-\gamma_6(R)} + \frac{C_7(R)e^{(\gamma_5(R)-\gamma_6(R))t_0(R)}}{\gamma_5(R)-\gamma_6(R)}e^{-\gamma_5(R)t}, \end{aligned} \tag{B.1.314}$$

$\forall(t,x,y) \in [t_0(R),\infty)\otimes S_R^d \otimes I\!R^D$. Thus there are constants $C_8(R), C_9(R), C_{10}(R) \in [0,\infty)$ such that

$$\theta_2(t,x,y) \le C_8(R)\theta_2(t_0(R),x,y)e^{-\gamma_5(R)t} + C_9(R)e^{-\gamma_6(R)t} + C_{10}(R)e^{-\gamma_5(R)t}, \tag{B.1.315}$$

$\forall(t,x,y) \in [t_0(R),\infty)\otimes S_R^d \otimes I\!R^D$. We next show that

$$\sup_{(t,x,y)\in[0,t_0(R)]\otimes S_R^d\otimes I\!R^D} \theta_2(t,x,y) < \infty. \tag{B.1.316}$$

Let $\Psi(t)$ solve the ordinary differential equation

$$(\partial_t\Psi)(t) = \alpha(t,R)\Psi(t) + C_7(R)e^{-\gamma_3(R)t} \tag{B.1.317}$$

$$\Psi(0) = 0 \tag{B.1.318}$$

for

$$\alpha(t,R) \triangleq 2\Lambda_1(R) + \Lambda_0^2(R) + C_6(R)e^{-\gamma_3(R)t} \tag{B.1.319}$$

and fix arbitrary $(x, y) \in S_R^d \otimes \mathbb{R}^D$. Put

$$\xi(t) \triangleq \Psi(t) - \theta_2(t, x, y). \tag{B.1.320}$$

Then, from (B.1.320), (B.1.319), (B.1.317), and (B.1.305) we have

$$(\partial_t \xi)(t) \geq \alpha(t, R)\xi(t) \tag{B.1.321}$$

with

$$\xi(0) = \Psi(0) - \theta_2(0, x, y) = 0 \tag{B.1.322}$$

(see (B.1.288)). Thus, from (B.1.321)

$$\begin{aligned}
&\frac{d}{dt}\left[\exp\left\{\int_0^t -\alpha(\tau, R)d\tau\right\}\xi(t)\right] \\
&= -\alpha(t, R)\exp\left\{\int_0^t -\alpha(\tau, R)d\tau\right\}\xi(t) + \exp\left\{\int_0^t -\alpha(\tau, R)d\tau\right\}(\partial_t\xi)(t) \\
&\geq -\alpha(t, R)\exp\left\{\int_0^t -\alpha(\tau, R)d\tau\right\}\xi(t) \\
&\qquad\qquad +\alpha(t, R)\exp\left\{\int_0^t -\alpha(\tau, R)d\tau\right\}\xi(t) \\
&\geq 0.
\end{aligned} \tag{B.1.323}$$

From (B.1.323) and (B.1.322) we get

$$\exp\left[\int_0^t -\alpha(\tau, R)d\tau\right]\xi(t) \geq 0, \qquad \forall t \in [0, \infty), \tag{B.1.324}$$

hence $\xi(t) \geq 0, \forall t \in [0, \infty)$, hence by (B.1.320)

$$\Psi(t) \geq \theta_2(t, x, y), \qquad \forall (t, x, y) \in [0, \infty) \otimes S_R^d \otimes \mathbb{R}^D. \tag{B.1.325}$$

Thus

$$\sup_{(t,x,y)\in[0,t_0(R)]\otimes S_R^d\otimes\mathbb{R}^D} \theta_2(t, x, y) \leq \sup_{t\in[0,t_0(R)]} \Psi(t). \tag{B.1.326}$$

But (B.1.317) is a linear ordinary differential equation thus $\Psi(.)$ exists over all $t \in [0, \infty)$ hence the right hand side of (B.1.326) is finite, as required for (B.1.316). Now it follows from (B.1.315) and (B.1.316) that there are constants $C_{11}(R) \in [0, \infty)$ and $\gamma(R) \in (0, \infty)$ such that

$$\theta_2(t, x, y) \le C_{11}(R) e^{-\gamma(R)t}, \quad \forall (t, x, y) \in [t_0(R), \infty) \otimes S_R^d \otimes \mathbb{R}^D. \tag{B.1.327}$$

Also, again by (B.1.326), there is a constant $C_{12}(R) \in [0, \infty)$ such that

$$\theta_2(t, x, y) \le C_{12}(R) \le C_{12}(R) e^{\gamma(R)t_0(R)} e^{-\gamma(R)t} \tag{B.1.328}$$

for all $(t, x, y) \in [0, t_0(R)] \otimes S_R^d \otimes \mathbb{R}^D$. Now (B.1.327) and (B.1.328) gives a constant $C(R) \in [0, \infty)$ such that

$$\theta_2(t, x, y) \le C(R) e^{-\gamma(R)t}, \quad \forall (t, x, y) \in [0, \infty) \otimes S_R^d \otimes \mathbb{R}^D, \tag{B.1.329}$$

and this gives (B.1.246).

□

Appendix C

Miscellaneous Technical Results

Here we list for easy reference a miscellany of simple technical results that are needed for the thesis.

Fact C.0.4. *If (A.1.16) holds for $\Gamma(X, Z)$ given by (A.1.15) when the $h_i : \mathbb{R}^d \otimes \mathbb{R}^d \to \mathbb{R}$ are uniformly bounded and continuous, then*

$$E^{P^*}\left[(\Lambda_g(X, Z)(\tau_2) - \Lambda_g(X, Z)(\tau_1))\, ;\; B\right] = 0, \qquad \forall B \in \mathcal{B}(\tau_1). \tag{C.0.1}$$

Proof: To show (C.0.1) it is enough to show that (A.1.16) holds for $\Gamma(X, Z)$ given by (A.1.15) when the h_i have the form

$$h_i = I_{F_i}, \tag{C.0.2}$$

for closed sets $F_i \subset \mathbb{R}^d \otimes \mathbb{R}^d$, since $\mathcal{B}(\mathbb{R}^d \otimes \mathbb{R}^d)$ is the minimal σ-algebra which includes all of the closed subsets of $\mathbb{R}^d \otimes \mathbb{R}^d$. But for closed set $F_i \subset \mathbb{R}^d \otimes \mathbb{R}^d$, there is a sequence of bounded and continuous functions $g_k^i : \mathbb{R}^d \otimes \mathbb{R}^d \to \mathbb{R}$ such that $0 \le g_k^i \le 1$, $\forall k \in \mathbb{N}$, and

$$\lim_{k \to \infty} g_k^i(x) = I_{F_i}(x), \qquad \forall x \in \mathbb{R}^d \otimes \mathbb{R}^d. \tag{C.0.3}$$

Hence, by the dominated convergence theorem, we see that (A.1.16) holds for h_i given by I_{F_i}, F_i closed.

□

Fact C.0.5. *Suppose $\{x_n\}$ is a sequence in a metric space S. If each subsequence $\{x_{n(k)}\}_k$ of $\{x_n\}$ contains a further subsequence $\{x_{n_{(k(r))}}\}_r$ such that*

$$\lim_{r\to\infty} x_{n_{(k(r))}} = x^*, \tag{C.0.4}$$

then $\lim_{n\to\infty} x_n = x^*$.

Lemma C.0.6. (Problem 3.15 on page 306 of Karatzas and Shreve [19]). *Suppose that the following holds:*

(i) *The mappings $b^i(\xi)$ and $\sigma^{ij}(\xi)$, $1 \le i \le D$, $1 \le j \le N$, are Borel measurable functions from $\mathbb{R}^D$ into $\mathbb{R}$ satisfying*

$$|b(\xi)|^2 + |\sigma(\xi)|^2 \le K\left(1 + |\xi(s)|^2\right), \qquad \forall \xi \in \mathbb{R}^D, \tag{C.0.5}$$

where $K \in [0, \infty)$ is a constant.

(ii) *$\{(\beta(t), \mathcal{F}_t),\ t \in [0, \infty)\}$ is an $\mathbb{R}^N$-valued standard Wiener process on $(\Omega, \mathcal{F}, P)$ and*

$\{(\xi(t, y), \mathcal{F}_t),\ t \in [0, \infty)\}$ is an $\mathbb{R}^D$-valued adapted process on $(\Omega, \mathcal{F}, P)$ such that

$$\xi(t, y) = y + \int_0^t b(\xi(s, y))ds + \int_0^t \sigma(\xi(s, y))d\beta(s), \quad 0 \le t < \infty, \tag{C.0.6}$$

for some(non-random) $y \in \mathbb{R}^D$. Then, for each $T \in (0, \infty)$ and $m = 1, 2, 3\ldots$ we have

$$E\left(\max_{0\le s\le t} |\xi(s, y)|^{2m}\right) \le C\left(1 + |y|^{2m}\right)e^{Ct}; \quad 0 \le t \le T, \tag{C.0.7}$$

and

$$E|\xi(t, y) - \xi(s, y)|^{2m} \le C\left(1 + |y|^{2m}\right)(t - s)^m; \quad 0 \le s < t \le T, \tag{C.0.8}$$

where C is a positive constant depending only on m, T, K, and D.

Remark C.0.7. Let $\{X_n,\ n \in \mathbb{N}\}$ be a sequence of processes from $(\Omega, \mathcal{F}, P)$ into $(C[0,1], \mathcal{B}(C[0,1]))$. The sequence $\{X_n,\ n \in \mathbb{N}\}$ is by definition tight when the sequence of corresponding distribution is tight. According to Theorem 8.2 on page 55 and Theorem 8.3 on 56 of Billingsley [4], we establish the following results.

Theorem C.0.8. *The sequence $\mathcal{L}(X_n)$ is tight if and only if the following two conditions hold:*

(i) *For each $\eta \in (0,1]$, there is a $\zeta \in (0,\infty)$ such that*

$$P\{|X_n(0)| > \zeta\} \le \eta, \quad n = 1,2,3,... \tag{C.0.9}$$

(ii) *For each $\epsilon \in (0,\infty)$ and $\eta \in (0,1]$, there exists a $\gamma \in (0,1]$ and a positive integer n_0 such that*

$$P\left\{\sup_{|s-t|<\gamma} |X_n(s) - X_n(t)| \ge \epsilon\right\} \le \eta, \quad n \ge n_0. \tag{C.0.10}$$

Remark C.0.9. Condition (i) stipulates that $\{\mathcal{L}(X_n(0)),\ n \in \mathbb{N}\}$ be tight. Condition (ii) says that the processes $\{X_n(t),\ t \in [0,\infty)\}$ do not vary too rapidly. Next theorem can be stated similar to Theorem 8.2, assuming a stronger assumption in Condition (ii).

Theorem C.0.10. *The sequence $\mathcal{L}(X_n)$ is tight if these conditions are satisfied:*

(i) *For each $\eta \in (0,1]$, there is a $\zeta \in (0,\infty)$ such that*

$$P\{|X_n(0)| > \zeta\} \le \eta, \quad n = 1,2,3,... \tag{C.0.11}$$

(ii) *For each $\epsilon \in (0,\infty)$ and $\eta \in (0,1]$, there exists a $\gamma \in (0,1)$ and a positive integer n_0 such that*

$$\frac{1}{\gamma} P\left\{\sup_{t \le s \le t+\gamma} |X_n(s) - X_n(t)| \ge \epsilon\right\} \le \eta, \quad n \ge n_0. \tag{C.0.12}$$

for each $t \in [0,1]$.

Lemma C.0.11. (Problem 6 on page 41 of Billingsley [4]) *Let S_1 and S_2 be metric spaces, and $\{P_\alpha\}$ be the collection of probability measures on $S_1 \otimes S_2$. Let P_α^1 be the marginal of P_α on S_1 (i.e. $P_\alpha^1(A_1) \triangleq P_\alpha(A_1 \otimes S_2)$, $\forall A_1 \in \mathcal{B}(S_1)$) and let P_α^2 be the marginal of P_α on S_2. Then, the family $\{P_\alpha\}$ of probability measures on $S_1 \otimes S_2$ is tight if and only if the family $\{P_\alpha^1\}$ of probability measures on S_1 is tight and the family $\{P_\alpha^1\}$ of probability measures on S_2 is tight.*

In this thesis we repeatedly use the following result which is a special case of "Rayleigh's Principle" in linear algebra:

Lemma C.0.12. (see Theorem 10.25 of Noble and Daniel [27])
Let M be a $D \times D$ symmetric matrix with eigenvalues (necessarily real)

$$\lambda_{\min} = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_D = \lambda_{\max}.$$

Then we have

$$\lambda_{\min} \|x\|^2 \leq x^T M x \leq \lambda_{\max} \|x\|^2, \quad \forall x \in \mathbb{R}^D. \tag{C.0.13}$$

Appendix D

Ergodicity and Mixing

In this appendix we summarize, for the sake of completeness, some of the most basic definitions and results on ergodicity and mixing in stochastic processes. Let $(X, \mathcal{B}, m)$ be a probability space, and let $T : X \to X$ be a transformation. T is measurable when

$$T^{-1}(A) \in \mathcal{B}, \quad \forall A \in \mathcal{B}. \tag{D.0.1}$$

Definition D.0.13. A measurable transformation $T : X \to X$ is said to be *measure preserving* when

$$P(T^{-1}A) = P(A), \quad \forall A \in \mathcal{B}. \tag{D.0.2}$$

Definition D.0.14. A set $B \in \mathcal{B}$ is said to be an *invariant set* when

$$T^{-1}(B) = B. \tag{D.0.3}$$

Remark D.0.15. let $\mathcal{B}_I$ be the family of invariant sets in $\mathcal{B}$, namely

$$\mathcal{B}_I \triangleq \left\{A \in \mathcal{B} : T^{-1}(A) = A\right\}. \tag{D.0.4}$$

We see easily that $\mathcal{B}_I$ is a σ-algebra.

Definition D.0.16. A measure preserving transformation $T : X \to X$ is said to be *ergodic* when

$$P(B) = 0 \text{ or } P(B) = 1, \quad \text{for each } B \in \mathcal{B}_I. \tag{D.0.5}$$

D.0.1 Ergodicity of Strictly Stationary Processes

We now formulate the notion of *ergodicity* in the context of a strictly stationary $\mathbb{R}^D$-valued process.

Let Ξ denote the set of all functions from $[0, \infty)$ to $\mathbb{R}^D$, namely

$$\Xi \triangleq \bigotimes_{t \in [0,\infty)} \mathbb{R}^D. \tag{D.0.6}$$

Also, let $\mathcal{G}$ be the minimal σ-algebra generated by the cylindrical subsets of Ξ. Let $\{\xi(t),\ t \in [0, \infty)\}$ be a strictly stationary $\mathbb{R}^D$-valued process on the probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ (whose paths are therefore members of Ξ). It is elementary to show that the mapping

$$\hat{\omega} \to \xi(., \hat{\omega}) : (\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}) \to \Xi \tag{D.0.7}$$

is $\mathcal{F}/\mathcal{G}$-measurable, hence we can define a probability measure $\hat{P}\xi^{-1}$ on $(\Xi, \mathcal{G})$ by

$$\left(\hat{P}\xi^{-1}\right)(\Gamma) \triangleq \hat{P}\{\xi \in \Gamma\}, \quad \forall \Gamma \in \mathcal{G}. \tag{D.0.8}$$

For each $t \in [0, \infty)$ define a shift operator $T_t : \Xi \to \Xi$ by

$$\left(T_t \psi\right)(s) \triangleq \psi(t + s), \quad \forall s \in [0, \infty), \quad \psi \in \Xi. \tag{D.0.9}$$

One trivially sees that T_t is $\mathcal{G}/\mathcal{G}$ measurable and the family $\{T_t,\ t \in [0, \infty)\}$ is a semi-group of operators, meaning that

$$T_{t_1 + t_2} = T_{t_1} T_{t_2} = T_{t_2} T_{t_1} \tag{D.0.10}$$

for all $t_1, t_2 \in [0, \infty)$. Also, since $\{\xi(t),\ t \in [0,\infty)\}$ is strictly stationary, it follows that each T_t is *measure preserving*, in the sense that, for each $t \in [0, \infty)$, we have

$$\left(\hat{P}\xi^{-1}\right)(T_t^{-1}(\Gamma)) = \left(\hat{P}\xi^{-1}\right)(\Gamma), \quad \forall \Gamma \in \mathcal{G}. \tag{D.0.11}$$

An additional property which the strictly stationary process $\{\xi(t),\ t \in [0,\infty)\}$ may possibly have, is that of *ergodicity*. To formulate this property, define the so-called *invariant σ-algebra* (of the sets in $\mathcal{G}$) by

$$\mathcal{I} \triangleq \left\{\Gamma \in \mathcal{G} :\ T_t^{-1}(\Gamma) = \Gamma,\ \forall t \in [0,\infty)\right\}, \tag{D.0.12}$$

(it is trivially verified that $\mathcal{I}$ is indeed a σ-algebra).

If, for each $\Gamma \in \mathcal{I}$, we have

$$\left(\hat{P}\xi^{-1}\right)(\Gamma) = 0, \quad \text{or} \quad \left(\hat{P}\xi^{-1}\right)(\Gamma) = 1, \tag{D.0.13}$$

then the stationary process $\{\xi(t),\ t \in [0,\infty)\}$ is an *ergodic process*, or *ergodic*. To understand what it means for the stationary process to be ergodic, suppose that the property asserted by ergodicity fails to hold. Then there is some set $\Gamma \in \mathcal{I}$ such that

$$\left(\hat{P}\xi^{-1}\right)(\Gamma) = \alpha, \quad \text{for some } \alpha \in (0,1). \tag{D.0.14}$$

Since $T_t^{-1}(\Gamma) = \Gamma, \forall t \in [0,\infty)$ we see that if $\psi \in \Gamma$, then $T_t\psi \in \Gamma$, $\forall t \in [0,\infty)$. Likewise if $\psi \notin \Gamma$, then since $\Gamma^c \in \mathcal{I}$, we have $T_t\psi \notin \Gamma, \forall t \in [0,\infty)$. Thus the space of paths Ξ can be partitioned into two sets Γ and Γ^c, each of *strictly positive* probability, but isolated from each other in that one can never transfer a path ψ inside Γ into a path $T_t\psi$ outside Γ by some shift operator T_t. Ergodicity eliminates this type of "non-mixing" behavior, and asserts the property that, for each $\Gamma \in \mathcal{G}$ with $\left(\hat{P}\xi^{-1}\right)(\Gamma) > 0$, the sets $T_t^{-1}(\Gamma), t \in [0,\infty)$ effectively "cover" the whole space Ξ. (More precisely it can be shown that ergodicity of $\{\xi(t),\ t \in [0,\infty)\}$ actually implies the following: if $\left(\hat{P}\xi^{-1}\right)(\Gamma) > 0$ then

$$\left(\hat{P}\xi^{-1}\right)\left[\bigcup_n T_{t_n}^{-1}(\Gamma)\right] = 1, \tag{D.0.15}$$

for any sequence $0 \le t_1 < t_2 < t_3 < \dots$ with $t_n \to \infty$).

It is generally difficult to check when a given strictly stationary process $\{\xi(t)\}$ is ergodic. However, two exceptions are when $\{\xi(t)\}$ is a strictly stationary Markov process and when $\{\xi(t)\}$ is a strictly stationary mixing process. We consider each of these cases next.

D.0.2 The Markov case

Let $\{\xi(t),\ t \in [0,\infty)\}$ be an $\mathbb{R}^D$-valued Markov process on probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ with homogeneous transition probability function $P_t(x, A)$ for $t \in [0,\infty)$, $x \in \mathbb{R}^D$, $A \in \mathcal{B}(\mathbb{R}^D)$, that is

$$\hat{P}\left[\xi(t) \in A \mid \xi(0) = x\right] = P_t(x, A), \quad \text{for } \hat{P}\xi(0)^{-1} - \text{almost all } x \in \mathbb{R}^D. \tag{D.0.16}$$

Let π be an invariant probability measure on $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ for the transition probability function $P_t(x, A)$, in the sense that, for each $t \in [0,\infty)$, we have

$$\pi(A) = \int_{\mathbb{R}^D} P_t(x, A)\ \pi(dx), \quad \forall A \in \mathcal{B}(\mathbb{R}^D). \tag{D.0.17}$$

We would like some characterization of the invariant probability measures π with the property that the corresponding strictly stationary Markov process $\{\xi(t),\ t \in [0,\infty)\}$ is ergodic. This is given by the next result. We write $\mathcal{P}(\mathbb{R}^D)$ for the set of probability measures on $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$. We also need the following terminology:

Definition D.0.17. *Given a homogeneous transition probability function* $P_t(x, A), t \in [0,\infty), x \in \mathbb{R}^D, A \in \mathcal{B}(\mathbb{R}^D)$, *a measure* $\pi \in \mathcal{P}(\mathbb{R}^D)$ *is called* ergodic *when*

(i) *π is an invariant probability measure for the transition probability function $P_t(x, A)$.*

(ii) *The $\mathbb{R}^D$-valued strictly stationary Markov process $\{\xi(t),\ t \in [0, \infty)\}$ with initial distribution π and transition probability $P_t(x, A)$ is ergodic.*

Theorem D.0.18. *(see Theorem 7.4.8 in Stroock [34]) Suppose that $P_t(x, A), t \in [0, \infty), x \in \mathbb{R}^D, A \in \mathcal{B}(\mathbb{R}^D)$ is a homogeneous transition probability function, and let $\mathcal{M}_1$ be the collection of all $\pi \in \mathcal{P}(\mathbb{R}^D)$ such that (D.0.17) holds (i.e. $\mathcal{M}_1$ is the collection of all invariant probability measures). If $\mathcal{M}_1$ is not empty, then $\mathcal{M}_1$ is a convex subset of $\mathcal{P}(\mathbb{R}^D)$. Moreover, $\pi \in \mathcal{M}_1$ is ergodic if and only if π is an extreme point of $\mathcal{M}_1$. Finally, if π_1 and π_2 are ergodic, then either $\pi_1 = \pi_2$ or π_1 and π_2 are mutually singular.*

The next result is immediate from the preceding:

Corollary D.0.19. *Suppose that the homogeneous transition probability function $P_t(x, A), t \in [0, \infty), x \in \mathbb{R}^D, A \in \mathcal{B}(\mathbb{R}^D)$, has a* unique *invariant probability measure π. Then π is ergodic.*

Remark D.0.20. Corollary D.0.19 is also established in 4.11.43 on page 271 of Ethier and Kurtz [10].

Remark D.0.21. In order to use Corollary D.0.19 it is necessary to have conditions which ensure existence and uniqueness of an invariant probability measure. This is typically ensured by conditions such as the *Doeblin condition* which, in the context of an $\mathbb{R}^D$-valued Markov process with transition probability $P_t(x, A)$, goes as follows: there exists a finite measure φ on $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$, a (small) number $\epsilon \in (0, 1)$, and some $t_0 \in (0, \infty)$ such that for each $A \in \mathcal{B}(\mathbb{R}^D)$ with $\varphi(A) < \epsilon$ we have

$$P_{t_0}(x, A) < 1 - \epsilon, \quad \forall x \in \mathbb{R}^D. \tag{D.0.18}$$

If $P_t(x, A)$ is non-degenerate in the sense that there exists a jointly continuous mapping

$$p : (0, \infty) \otimes \mathbb{R}^D \otimes \mathbb{R}^D \to (0, \infty) \tag{D.0.19}$$

such that

$$P_t(x, A) = \int_A p(t, x, z)\, dz, \quad \forall t \in [0, \infty), \ \forall x \in \mathbb{R}^D, \ \forall A \in \mathcal{B}(\mathbb{R}^D), \tag{D.0.20}$$

and the preceding Doeblin condition holds, then it can be shown that there exists a unique invariant probability measure π, together with numbers $a \in (0, \infty)$ and $\gamma \in (0, \infty)$ such that

$$\sup_x \|P_t(x, .) - \pi(.)\|_{TV} < ae^{-\gamma t}, \tag{D.0.21}$$

where $\|\mu\|_{TV}$ denotes total variation of a signed measure μ (this is a "continuous parameter" version of Theorem 16.0.2 of Meyn and Tweedie [26]). There are several objections to using this condition when the $\mathbb{R}^D$-valued Markov process $\{\xi(t)\}$ is a solution of an Itô stochastic differential equation of the form

$$d\xi(t) = b(\xi(t))dt + \sigma(\xi(t))d\beta(t). \tag{D.0.22}$$

First, we do not usually know the transition probability function for the the Markov process given by (D.0.22). Second, we can never use Lebesgue measure on $\mathbb{R}^D$ for φ in achieving the Doeblin's condition, (since it is not a finite measure), and there appear to be no guidelines indicating a reasonable choice for the finite measure $\varphi(.)$. Finally, we really have no need for the very strong convergence in (D.0.21). For these reasons we prefer to verify the existence of a unique invariant probability measure by following the approach of Bhattacharya and Waymire [3] to Markov process arising as solutions of Itô stochastic differential equations(see Theorem 3.4.1 on page 43).

D.0.3 The Mixing case

In this thesis we are concerned with an intrinsically Markovian situation and do not require any ideas from the classical theory of mixing processes. Nevertheless, for background information only, in this subsection we briefly indicate the links that exist between ergodicity and the classical notions of mixing. Let $\{\xi(t)\}$ be an $I\!R^D$-valued process on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ and define

$$\begin{aligned} \mathcal{F}_{t,\infty} &\stackrel{\Delta}{=} \sigma\{\xi(u),\ u \in [t,\infty)\}, \quad \forall t \in [0,\infty) \\ \mathcal{F}_{s,t} &\stackrel{\Delta}{=} \sigma\{\xi(u),\ u \in [s,t]\}, \quad 0 \leq s < t < \infty. \end{aligned} \tag{D.0.23}$$

The mapping $\alpha : [0,\infty) \to [0,\infty)$ defined by

$$\alpha(u) \stackrel{\Delta}{=} \sup_{t\in[0,\infty)} \sup_{\substack{A\in\mathcal{F}_{0,t} \\ B\in\mathcal{F}_{t+u,\infty}}} |P(AB) - P(A)P(B)|, \tag{D.0.24}$$

is called the *strong mixing function* of the $I\!R^D$-valued process $\{\xi(t)\}$. Then the $I\!R^D$-valued process $\{\xi(t)\}$ is said to be a *strong mixing* process when

$$\lim_{u\to\infty} \alpha(u) = 0. \tag{D.0.25}$$

Remark D.0.22. Observe that we have not assumed that $\{\xi(t)\}$ is strictly stationary, hence the notion of strong mixing of $\{\xi(t)\}$ makes sense even when $\{\xi(t)\}$ is not strictly stationary. Compare this with ergodicity, which requires strict stationarity as a prior condition for ergodicity to make sense.

Next, define the *tail σ-algebra* of $\{\xi(t)\}$ as

$$\mathcal{F}_{\infty,\infty} \stackrel{\Delta}{=} \bigcap_{t\in[0,\infty)} \mathcal{F}_{t,\infty}. \tag{D.0.26}$$

The $I\!R^D$-valued process $\{\xi(t)\}$ is called *regular* when, for each $A \in \mathcal{F}_{\infty,\infty}$, we have

$$\hat{P}(A) = 0 \quad \text{or} \quad \hat{P}(A) = 1. \tag{D.0.27}$$

By adapting the arguments for Theorem 17.1.1 on page 302 of Ibraginov and Linnik [15], we obtain:

Theorem D.0.23. *The stationary $\mathbb{R}^D$-valued process $\{\xi(t),\ t \in [0,\infty)\}$ is regular if and only if*

$$\lim_{t\to\infty} \sup_{A\in\mathcal{F}_{t,\infty}} \left|\hat{P}(AB) - \hat{P}(A)\hat{P}(B)\right| = 0, \quad \textit{for each } B \in \mathcal{F}_{0,\infty}. \tag{D.0.28}$$

Thus Theorem D.0.23 on page 213 gives an alternative characterization of $\{\xi(t),\ t \in [0,\infty)\}$ being regular. The relation between ergodicity and regularity of a process is given by Corollary 17.1.1 on page 302 of Ibraginov and Linnik [15], which we state as follows:

Theorem D.0.24. *If the $\mathbb{R}^D$-valued process $\{\xi(t),\ t \in [0,\infty)\}$ is regular and strictly stationary, then it is ergodic.*

Finally, the relation between strong mixing and regularity is given by the following simple result:

Theorem D.0.25. *If the $\mathbb{R}^D$-valued process $\{\xi(t),\ t \in [0,\infty)\}$ is strictly stationary and strong mixing, then (D.0.28) holds.*

Proof: Define

$$\mathcal{D} = \bigcup_{t\in[0,\infty)} \mathcal{F}_{0,t} \tag{D.0.29}$$

Then $\mathcal{D}$ is a π-class and we easily see that

$$\mathcal{F}_{0,\infty} = \sigma\{\mathcal{D}\}. \tag{D.0.30}$$

Put

$$\mathcal{C} \triangleq \left\{ A \in \mathcal{F}_{0,\infty} : \lim_{u\to\infty} \sup_{B\in\mathcal{F}_{u,\infty}} \left|\hat{P}(AB) - \hat{P}(A)\hat{P}(B)\right| = 0 \right\} \tag{D.0.31}$$

Fix $A \in \mathcal{D}$. From (D.0.29) we have

$$A \in \mathcal{F}_{0,t_1}, \qquad \text{for some } t_1 \in [0, \infty). \tag{D.0.32}$$

Since $\{\xi(t)\}$ is strong mixing we know that

$$\lim_{u \to \infty} \sup_{B \in \mathcal{F}_{t_1+u,\infty}} \left| \hat{P}(AB) - \hat{P}(A)\hat{P}(B) \right| = 0, \tag{D.0.33}$$

hence $A \in \mathcal{C}$. It follows that

$$\mathcal{D} \subset \mathcal{C}. \tag{D.0.34}$$

It remains to see that $\mathcal{C}$ is a λ-class: From (D.0.31) we have $\hat{\Omega} \in \mathcal{C}$. Next, fix $A_1, A_2 \in \mathcal{C}, A_2 \subset A_1$. One easily sees that

$$\lim_{u \to \infty} \sup_{B \in \mathcal{F}_{u,\infty}} \left| \hat{P}\left((A_1 - A_2)B\right) - \hat{P}(A_1 - A_2)\hat{P}(B) \right| = 0, \tag{D.0.35}$$

hence $A_1 - A_2 \in \mathcal{C}$. Finally, fix $A_n \in \mathcal{C}$, with $A_n \subset A_{n+1}$ and put $A = \bigcup_n A_n$. One easily shows that

$$\lim_{u \to \infty} \sup_{B \in \mathcal{F}_{u,\infty}} \left| \hat{P}(AB) - \hat{P}(A)\hat{P}(B) \right| = 0, \tag{D.0.36}$$

hence $A \in \mathcal{C}$. It follows that $\mathcal{C}$ is a λ-class, hence (D.0.34), (D.0.30) and Dynkin's λ-π theorem gives

$$\mathcal{C} = \mathcal{F}_{0,\infty}, \tag{D.0.37}$$

which establishes the result.

□

Appendix E

Solvability of Poisson-type Equations

The Case of Discrete-time Markov Chains

In this appendix we give an adaptation due to Benveniste, Métivier and Priouret [1] of a result originally due to Sunyach [35] on stability properties of discrete-parameter Markov chains, since this result is an important motivation for the approach that we take in Section 3.4. Basically, the result says the following: If Π is the transition probability function of a discrete-parameter $\mathbb{R}^D$-valued Markov chain, and its r-th iterate Π^r is a *strict contraction* on a properly defined collection of locally Lipschitz functions on $\mathbb{R}^D$ for some positive integer r, then the Markov chain has a unique invariant probability measure and the Poisson equation defined by the operator $1 - \Pi$ is uniquely solvable to within constants. Before giving this result we formulate the " properly defined collection of Lipschitz continuous functions" mentioned previously.

Remark E.0.26. Fix some $p \in [0, \infty)$, and let $h : \mathbb{R}^D \to \mathbb{R}$ be a Borel-measurable function. Define

$$\begin{aligned} \|h\|_{p+1} &\triangleq \sup_x \frac{|h(x)|}{1+|x|^{p+1}} \\ [h]_p &\triangleq \sup_{x_1 \neq x_2} \frac{|h(x_1)-h(x_2)|}{|x_1-x_2|\left[1+|x_1|^p+|x_2|^p\right]} \\ Li(p) &\triangleq \left\{h : \mathbb{R}^D \to \mathbb{R} \mid \quad [h]_p < \infty\right\}. \end{aligned} \tag{E.0.1}$$

One easily sees that

$$\|h\|_{p+1} \leq |h(0)| + 2\,[h]_p < \infty, \quad \text{for } h \in Li(p). \tag{E.0.2}$$

Thus, if $[h]_p < \infty$, then we have $\|h\|_{p+1} < \infty$. Hence, for each $h \in Li(p)$, we see that

$$M_p(h) \triangleq \max\left\{\|h\|_{p+1},\ [h]_p\right\}, \tag{E.0.3}$$

is finite, and

$$\begin{aligned} |h(x)| &\leq M_p(h)\left[1+|x|^{p+1}\right], \ \forall x \in \mathbb{R}^D, \\ |h(x_1)-h(x_2)| &\leq M_p(h)|x_1-x_2|\left[1+|x_1|^p+|x_2|^p\right], \ \forall x_1, x_2 \in \mathbb{R}^D. \end{aligned} \tag{E.0.4}$$

For every Borel-measurable $h : \mathbb{R}^D \to \mathbb{R}$ such that

$$\int_{\mathbb{R}^D} |h(y)|\,\Pi(x,dy) < \infty, \quad \forall x \in \mathbb{R}^D, \tag{E.0.5}$$

let Πh be defined by

$$\Pi h(x) \triangleq \int_{\mathbb{R}^D} h(y)\Pi(x,dy), \quad \forall x \in \mathbb{R}^D. \tag{E.0.6}$$

where Π denotes the transition probability function of a discrete - parameter Markov chain on $\mathbb{R}^D$.

Proposition E.0.27. *(see Proposition 3 on page 255 of Benveniste, Métivier and Priouret [1]) Suppose $\Pi(x,A)$ is one-step transition probability function of a discrete-parameter Markov chain on $\mathbb{R}^D$, and $p \in [0,\infty)$ is a constant, such that the following hold:*

(i) *We have*

$$\int_{\mathbb{R}^D} |\xi|^{p+1}\,\Pi(x,d\xi) < \infty, \quad \forall x \in \mathbb{R}^D. \tag{E.0.7}$$

(ii) *For each $h \in Li(p)$ we have $\Pi^n h \in Li(p), \forall n = 1,2,3,\ldots$.*

(iii) *There exists some integer $r \geq 1$, with some constant $\rho \in (0,1)$ such that*

$$[\Pi^r h]_p \leq \rho\,[h]_p, \quad \forall h \in Li(p). \tag{E.0.8}$$

Then:

(a) Π *has a unique invariant probability measure m on $\mathbb{R}^D$ such that*

$$\int_{\mathbb{R}^D} |\xi|^{p+1}\,dm(\xi) < \infty. \tag{E.0.9}$$

(b) *There exist constants $K_1 > 0$, $0 < \rho_1 < 1$ such that*

$$\left|\Pi^n h(x) - \int_{\mathbb{R}^D} h dm\right| \leq K_1 \rho_1^n\,[h]_p\left[1 + |x|^{p+1}\right], \quad \forall n = 1,2,3,\ldots, \tag{E.0.10}$$

for each $h \in Li(p)$ and $x \in \mathbb{R}^D$.

(c) *For each $h \in Li(p)$, the function*

$$u(x) \triangleq \sum_{n\geq 0}\left[\Pi^n h(x) - \int h\,dm\right], \quad \forall x \in \mathbb{R}^D. \tag{E.0.11}$$

is a member of $Li(p)$, and solves the operator equation

$$([1-\Pi]\,u)(x) = h(x) - \int h\,dm, \quad \forall x \in \mathbb{R}^D. \tag{E.0.12}$$

Furthermore, if u_1 is another member of $Li(p)$ which solves this operator equation, i.e.

$$([1-\Pi]\,u_1)(x) = h(x) - \int h\,dm, \quad \forall x \in \mathbb{R}^D, \tag{E.0.13}$$

then

$$u(x) - u_1(x) = \text{ a constant }, \quad \forall x \in \mathbb{R}^D. \tag{E.0.14}$$

Remark E.0.28. The crucial condition in Proposition E.0.27 on page 216 is (iii) which says that, for some integer $r \geq 1$, the r-th iterate Π^r is a *strict contraction* on $Li(p)$, where we measure the "size" of $h \in Li(p)$ by the quantity $[h]_p$. This condition is strong enough to ensure existence of a *unique* invariant *probability* measure for Π, together with solvability of the Poisson-type operator equation (E.0.12) uniquely within constants. The nice feature of this result is that it applies even to Markov chains which fail to be recurrent in the Harris sense (examples are given in the paper of Sunyach [35]). In Section 3.4 of the thesis we will establish a related result for $\mathbb{R}^D$-valued Markov processes $\{\xi(t,x)\}$ defined by the Itô stochastic differential equation (3.1.3), which will give us existence of a unique invariant probability measure as in Condition 3.2.3 of Section 3.2, together with solvability of the Poisson-type equations in Conditions 3.2.8 and 3.2.15 of Section 3.2.

Remark E.0.29. Proposition E.0.27 on page 216 is formulated in terms of functions in $Li(p)$, which are locally Lipschitz continuous and have "polynomial growth" of order $1 + p$. The reason for introducing this space is that, subject to Conditions (i), (ii) and (iii) of Proposition E.0.27 on page 216, we are guaranteed solvability to within constants of the Poisson equation (E.0.12) for each and every h that belongs to $Li(p)$.

Appendix F

L_2-Derivatives and the Backward Equation

In this appendix we discuss the issue of smoothness of the solutions of stochastic differential equations with respect to a parameter in the L_2-sense and then introduce the Kolmogorov Backward Equation for diffusions defined by classical Itô stochastic differential equations. These ideas will be needed in the course of establishing Proposition 3.4.16 on page 51, and we summarize them here for easy reference. First we need the following general definition:

Definition F.0.30. Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space, and let $g, f : \mathbb{R}^D \otimes \Omega \to \mathbb{R}$ be $\mathcal{B}(\mathbb{R}^D) \otimes \mathcal{F}$ measurable mappings. If

$$\lim_{h \to 0} E\left\{\left|\frac{g(y + e_i h) - g(y)}{h} - f(y)\right|^2\right\} = 0, \tag{F.0.1}$$

(where $e_i \triangleq (0, \ldots, 0, 1, 0, \ldots, 0)$ is i-th basis vector for $\mathbb{R}^D$), then $g(.)$ is said to have a *partial derivative with respect to its i-th argument in the L_2-sense* at y, and this partial derivative is given by $f(y)$.

Remark F.0.31. Notice that this partial derivative is itself a random variable on $(\Omega, \mathcal{F}, P)$. We will use classical calculus notation to indicate partial derivatives in the L_2-sense, so we put

$$(\partial_{y^i} g)(y) \triangleq f(y), \quad \forall y \in I\!R^D. \tag{F.0.2}$$

By replacing g with $(\partial_{y^i} g)$ in the preceding definition we can obviously also define the notion of the *double* partial derivative $(\partial_{y^j} \partial_{y^i} g)$ in the L_2-sense.

Remark F.0.32. The notion of L_2 - derivatives set forth previously is due to Gihman and Skorohod [14], as are the results on differentiability in the L_2-sense of solutions of stochastic differential equations that we summarize next. Recall the stochastic differential equation

$$\xi(t, x, y) = y + \int_0^t b(x, \xi(s, x, y))ds + \int_0^t \sigma(x, \xi(s, x, y))d\beta(s), \tag{F.0.3}$$

which defines the process $\{\xi(t, x, y)\}$ (see (3.4.74)). Then, it follows from Condition 3.4.4 and Theorem 1 on page 61 of Gihman and Skorohod [14], that the derivative $(\partial_{y^i} \xi^k)(t, x, y)$ exists in the L_2 - sense of Definition F.0.30 and satisfies the relation

$$\begin{aligned}(\partial_{y^i} \xi^k)(t, x, y) &= \delta_{i,k} + \sum_{l=1}^{D} \int_0^t (\partial_{\xi^l} b^k)(x, \xi(s, x, y))(\partial_{y^i} \xi^l)(s, x, y)ds \\ &\quad + \sum_{n=1}^{N} \sum_{l=1}^{D} \int_0^t (\partial_{\xi^l} \sigma^{k,n})(x, \xi(s, x, y))(\partial_{y^i} \xi^l)(s, x, y)d\beta^n(s),\end{aligned} \tag{F.0.4}$$

(see (3) on page 59 of Gihman and Skorohod [14]). Also, the double derivative $(\partial_{y^j} \partial_{y^i} \xi^k)(t, x, y)$ exists in the L_2 - sense of Definition F.0.30 and satisfies the relation

$$\begin{aligned}
\left(\partial_{y^j}\partial_{y^i}\xi^k\right)(t,x,y) &= \sum_{l=1}^{D}\int_0^t (\partial_{\xi^l}b^k)(x,\xi(s,x,y))\ \left(\partial_{y^j}\partial_{y^i}\xi^l\right)(s,x,y)\ ds \\
&+\sum_{l=1}^{D}\sum_{p=1}^{D}\int_0^t \left[\left(\partial_{\xi^p}\partial_{\xi^l}b^k\right)(x,\xi(s,x,y))\times\right. \\
&\qquad \left.\left(\partial_{y^j}\xi^p\right)(s,x,y)\left(\partial_{y^i}\xi^l\right)(s,x,y)\right]\ ds \\
&+\sum_{n=1}^{N}\sum_{l=1}^{D}\int_0^t \left(\partial_{\xi^l}\sigma^{k,n}\right)(x,\xi(s,x,y))\left(\partial_{y^j}\partial_{y^i}\xi^l\right)(s,x,y)\ d\beta^n(s) \\
&+\sum_{n=1}^{N}\sum_{l=1}^{D}\sum_{p=1}^{D}\int_0^t \left[\left(\partial_{\xi^p}\partial_{\xi^l}\sigma^{k,n}\right)(x,\xi(s,x,y))\times\right. \\
&\qquad \left.(\partial_{y^j}\xi^p)(s,x,y)\ (\partial_{y^i}\xi^l)(s,x,y)\right]\ d\beta^n(s),
\end{aligned} \tag{F.0.5}$$

(see (4) on page 60 of Gihman and Skorohod [14]).

Remark F.0.33. The significance of the preceding results of Gihman and Skorohod [14] is that the first and second derivatives $(\partial_{y^i}\xi^k)(t,x,y)$ and $(\partial_{y^j}(\partial_{y^i}\xi^k)(t,x,y)$ exist in the L_2 -sense and can be calculated by formally taking the first and second derivatives with respect to y of (F.0.3). The L_2 - derivatives are useful for the following reason. Suppose we have a smooth mapping $f : \mathbb{R}^D \to \mathbb{R}$, which is used to define another (non-random) mapping $\phi(t,x,y) \triangleq E\left[f(\xi(t,x,y))\right]$. It is frequently necessary to calculate the partial derivatives $(\partial_{y^j}\phi)(t,x,y)$ and $(\partial_{y^j}\partial_{y^i}\phi)(t,x,y)$. Indeed, such derivatives must be computed several times in the course of establishing Proposition 3.4.16 on page 51. The next result is a type of "chain rule" which clearly illustrates the key role played by the L_2 - derivatives $(\partial_{y^i}\xi)(t,x,y)$ and $(\partial_{y^j}\partial_{y^i}\xi)(t,x,y)$ in the computation of the first and second partial derivatives of $y \to \phi(t,x,y)$.

Corollary F.0.34. (Corollary 1 on page 62 of Gihman and Skorohod [14]) *Suppose condition 3.4.4 holds, and let $f : \mathbb{R}^D \to \mathbb{R}$ be a C^2-mapping such that*

$$|f(\xi)| + |(\partial_{\xi^i} f)(\xi)| + |(\partial_{\xi^j}\partial_{\xi^i} f)(\xi)| \leq C\,(1 + |\xi|^r), \quad \forall \xi \in \mathbb{R}^D.$$

for constants $C, r \in [0, \infty)$. Then the mapping

$$(t, x, y) \to \phi(t, x, y) \triangleq E\,[f(\xi(t, x, y))]$$

exists for each $(t, x, y) \in [0, \infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$, is twice - continuously differentiable in y for each $(t, x) \in [0, \infty) \otimes \mathbb{R}^d$, and has first and second y - derivatives given by

$$\partial_{y^i}\phi(t, x, y) = E\left[\sum_{k=1}^{D} (\partial_{\xi^k} f)(\xi(t, x, y))(\partial_{y^i}\xi^k)(t, x, y)\right] \tag{F.0.6}$$

and

$$\begin{aligned}(\partial_{y^j}\partial_{y^i}\phi)(t, x, y) \\ = E\Bigg[&\sum_{k=1}^{D} (\partial_{\xi^k} f)(\xi(t, x, y))(\partial_{y^j}\partial_{y^i}\xi^k)(t, x, y) \\ &+ \sum_{l=1}^{D}\sum_{k=1}^{D} (\partial_{\xi^l}\partial_{\xi^k} f)(\xi(t, x, y))(\partial_{y^j}\xi^l)(t, x, y)(\partial_{y^i}\xi^k)(t, x, y)\Bigg]\end{aligned} \tag{F.0.7}$$

for all $(t, x, y) \in [0, \infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$.

Remark F.0.35. Observe that (F.0.6) and (F.0.7) are really a "stochastic chain rule", and are reasonably consistent with what one would expect on the basis of the ordinary chain rule of calculus, except that $(\partial_{y^i}\xi^k)(t, x, y)$ and $(\partial_{y^j}\partial_{y^i}\xi^k)(t, x, y)$ are derivatives in the L_2 - sense.

Remark F.0.36. Corollary F.0.34 is also useful for establishing the following result, which is a trivial variant of Theorem 5 on page 297 of Gihman and Skorohod [14]:

Theorem F.0.37. *Suppose Condition 3.4.4 holds, and let $h : \mathbb{R}^D \to \mathbb{R}$ be a C^2 - mapping such that*

$$|h(\xi)| + |(\partial_{\xi^i} h)(\xi)| + |(\partial_{\xi^j}\partial_{\xi^i} h)(\xi)| \le C\left[1 + |\xi|^r\right], \tag{F.0.8}$$

for some constants $C, r \in [0, \infty)$. Put

$$\theta(t,x,y) \triangleq E\left[h(\xi(t,x,y))\right], \quad \forall (t,x,y) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D. \tag{F.0.9}$$

Then $y \to \theta(t,x,y) : \mathbb{R}^D \to \mathbb{R}$ is a C^2 - mapping for each $(t,x) \in [0,\infty) \otimes \mathbb{R}^d$, while $t \to \theta(t,x,y) : [0,\infty) \to \mathbb{R}$ is a C^1 - mapping for each $(x,y) \in \mathbb{R}^d \otimes \mathbb{R}^D$, and

$$(\partial_t \theta)(t,x,y) = \mathcal{A}\theta(t,x,y), \tag{F.0.10}$$

$\forall (t,x,y) \in [0,\infty) \otimes \mathbb{R}^d \otimes \mathbb{R}^D$, where (c.f. (3.2.14))

$$\mathcal{A}\theta(t,x,y) \triangleq \sum_{i=1}^{D} b^i(x,y)(\partial_{y^i}\theta)(t,x,y) + \frac{1}{2}\sum_{i=1}^{D}\sum_{j=1}^{D}\left[\sigma\sigma^T(x,y)\right]^{ij}(\partial_{y^i}\partial_{y^j}\theta)(t,x,y). \tag{F.0.11}$$

Remark F.0.38. The relation (F.0.10) is of course the well - known Kolmogorov Backward Equation for the diffusion $\{\xi(t,x,y)\}$ given by (3.4.74).

Bibliography

[1] A. Benveniste, M. Métivier, and P. Priouret. *Adaptive Algorithms and Stochastic Approximations.* Springer-Verlag, 1987.

[2] J. G. Besjes. On the Asymptotic Methods for Non-linear Differential Equations. *J. Mécanique*, 8(3):357–373, 1969.

[3] R. N. Bhattacharya and E. C. Waymire. *Stochastic Processes with Applications.* John Wiley and Sons, 1990.

[4] P. Billingsley. *Convergence of Probability Measures.* John Wiley & Sons, 1968.

[5] G. Blankenship and G. C. Papanicolaou. Stability and Control of Stochastic Systems with Wide-band Noise Disturbances. *I. SIAM J. Appl. Math.*, 34:437–476, 1978.

[6] N. Bogoliubov and Yu. Mitropol'skii. *Asymptotic Methods in the Theory of Non-linear Oscillations.* Hindustan Publ., 1961.

[7] K. L. Chung and R. J. Williams. *Introduction to Stochastic Integration, 2nd Edition.* Birkhauser, 1990.

[8] W. Doeblin. Sur les Propriétés Aymptotiques de Mouvement Régis par Certains Types de Chaines Simples. *Bull. Math. Soc. Roum. Sci.*, 39, no. 1, 57 - 115 ; no. 2, 3 - 61, 1937.

[9] J. L. Doob. *Stochastic Processes.* John Wiley and Sons, 1990.

[10] S.N. Ethier and T.G. Kurtz. *Markov Processes - Characterization and Convergence.* Wiley, 1986.

[11] M.I. Freidlin and A.D. Wentzell. *Random Perturbations of Dynamical Systems.* Springer-Verlag, 1984.

[12] D. Friedman. *Stochastic Differential Equations and Applications, Volume 1.* Academic Press, New York, 1975.

[13] I. I. Gihman. On a Theorem of N. N. Bogoliubov. *Ukrain. Math. Zh.*, 4:215–219, 1952 (Russian).

[14] I.I. Gihman and A.V. Skorokhod. *Stochastic Differential Equations.* Springer-Verlag, 1972.

[15] I. A. Ibragimov and Yu. V. Linnik. *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, 1971.

[16] J. Jacod and A. N. Shiryaev. *Limit Theorem for Stochastic Processes.* Springer-Verlag, New York, 1987.

[17] J. A. Jeans. *Astronomy and Cosmogony.* Cambridge Univ. Press Cambridge, 1928.

[18] G. Kallianpur. *Stochastic Filtering Theory.* Springer-Verlag, 1980.

[19] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, 1991.

[20] R. Z. Khas'minskii. On the Averaging principle for Ito Stochastic Differential Equations. *Kybernetika (Prague)*, 4:260–279, 1968 (Russian).

[21] H. Kunita. *Stochastic Flows and Stochastic Differential Equations.* Cambridge University Press, 1990.

[22] T.G. Kurtz. A Limit Theorem for Perturbed Operator Semigroups with Applications to Random Evolutions. *Journal of Functional Analysis*, 12:55–67, 1973.

[23] H. J. Kushner. *Approximation and Weak Convergence Methods for Random Processes.* MIT Press, 1984.

[24] R. S. Liptser and A. N. Shiryayev. *Theory of Martingales.* Kluwer Academic Publishers, 1989.

[25] R. S. Liptser and J. Stoyanov. Stochastic Version of the Averaging Principle for Diffusion Type Processes. *Stochastics and Stochastics Reports*, 32:145–163, 1990.

[26] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, 1993.

[27] B. Noble and J. W. Daniel. *Applied Linear Algebra.* Prentice-Hall, 1988.

[28] G. C. Papanicolaou. Some probabilistic problems and methods in singular perturbations. *Rocky Mountain J. of Mathematics*, 6:653–674, 1976.

[29] G.C. Papanicolaou, D.W. Stroock, and S.R.S. Varadhan. Martingale Approach to Some Limit Theorems. *Conference on Statistical Mechanics, Dynamical Systems, and Turbulence, Duke Univ. Math. Series III*, pages 1–116, 1977.

[30] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales, Volume 1, Second Edition.* John Wiley and Sons, 1994.

[31] J. A. Sanders and F. Verhulst. *Averaging Methods in Nonlinear Dynamical*

Systems, Applied Mathematical Sciences Series, Volume 59. Springer-Verlag, New York, 1985.

[32] S. S. Sastry. The Effects of Small Noise on Implicitly Defined Nonlinear Dynamical Systems. *IEEE Trans. Circuits Syst.*, 30(9):651–663, Sep. 1983.

[33] A. V. Skorokhod. *Asymptotic Methods in the Theory of Stochastic Differential Equations.* American Mathematical Society, Translations of Mathematical Monographs, Volume 78, 1989.

[34] D. W. Stroock. *Probability Theory.* Cambridge University Press, 1993.

[35] C. Sunyach. Une Classe de Chaînes de Markov Récurrentes Sur un Espace Métrique Complet. *Ann. Inst. Henri Poincaré*, XI(4):325–343, 1975.