Title

Stationary distributions for stochastic delay differential equations with non-negativity constraints

Author

Kinnally, Michael Sean

Publication Date

2009

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Stationary Distributions for Stochastic Delay Differential Equations with Non-Negativity Constraints

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Michael Sean Kinnally

Committee in charge:

Professor Ruth J. Williams, Chair
Professor Patrick J. Fitzsimmons
Professor Massimo Franceschetti
Professor Jun Liu
Professor Jason R. Schweinsberg

2009

The dissertation of Michael Sean Kinnally is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

2009

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I certainly owe a great deal of gratitude to Ruth Williams for her support, motivation, and guidance. This thesis would not have been possible without her analytical expertise and helpful suggestions. I'd also like to thank Patrick Fitzsimmons, Massimo Franceschetti, Jun Liu, and Jason Schweinsberg for taking the time to be on my advancement and defense committees. I am grateful to the UCSD Department of Mathematics for support as a graduate teaching assistant, and to the National Science Foundation for support as a graduate research assistant under NSF grant DMS 06-04537. I'll always be indebted to my mother Kathleen (a.k.a. Casey) and father Keith, in part because of their invaluable advise pertaining to both life and career ever since I was cognisant. Along with my brother Pete and sister Kate, they always believed in me and were there to help me in any way possible. Finally, I'd like to thank my friends, without whom my mind would surely have been lost over these long years in grad school.

VITA

2000	B. A. with honors in Mathematics, Columbia University, New York
2000-2002	Proprietary Trader, Worldco, L.L.C., New York
2003-2008	Graduate Teaching Assistant, Department of Mathematics, University of California, San Diego
2005	M. A. in Mathematics, University of California, San Diego
2006-2009	Graduate Research Assistant, University of California, San Diego
2009	Ph. D. in Mathematics, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Stationary Distributions for Stochastic Delay Differential Equations with Non-Negativity Constraints

by

Michael Sean Kinnally
Doctor of Philosophy in Mathematics

University of California San Diego, 2009

Professor Ruth J. Williams, Chair

Deterministic dynamic models with delayed feedback and state constraints arise in a variety of applications in science and engineering. Much of the analysis of such deterministic models has focussed on stability analysis of equilibrium points. There is interest in understanding what effect noise has on the behavior of such systems. Here we consider a multidimensional stochastic delay differential equation with normal reflection as a noisy analogue of a deterministic system with delayed feedback and non-negativity constraints. We obtain sufficient conditions for existence and uniqueness of stationary distributions. The results are applied to examples from Internet rate control and biochemical reaction systems.

Chapter 1

Introduction

1.1 Overview

Dynamical system models with delay are used in a variety of applications in science and engineering where the dynamics are subject to propagation delay. Examples of such application domains include packet level models of Internet rate control where the finiteness of transmission times leads to delay in receipt of congestion signals or prices [32, 44], neuronal models where the spatial distribution of neurons can result in delayed dynamics, epidemiological models where incubation periods result in delayed transmission of disease [7], and biochemical reactions in gene regulation where lengthy transcription and translation operations have been modeled with delayed dynamics [1, 6, 29]. There is an extensive literature, both theoretical and applied on ordinary delay differential equations. The book [17] by Hale and Lunel provides an introduction to this vast subject.

In some applications involving delayed dynamics, the quantities of interest are naturally non-negative. For instance, rates and prices in Internet models are non-negative, concentrations of ions or chemical species and proportions of a population that are infected are all naturally non-negative quantities. In deterministic differential equation models for the delayed dynamics of such quantities, the dynamics may naturally keep the quantities non-negative or they may need to be adapted to be so, sometimes leading to piecewise continuous delay differential dynamics (see e.g., [32, 33, 34, 35, 36]). There is some literature, especially applied, on the latter, although less than for unconstrained delay systems or naturally constrained ones.

Frequently in applications, noise is present in a system and it is desirable to understand its effect on the dynamics. For unconstrained systems, one can consider ordinary delay differential equations with an addition to the dynamics in the form of white noise or even a state dependent noise. There is a sizeable literature on such stochastic delay differential equations (SDDE) especially when the associated noiseless system has a globally attracting equilibrium [3, 11, 15, 20, 27, 28, 30, 31, 37, 41, 42, 43]. To obtain the analogue of such SDDE models with non-negativity constraints, it is not simply a matter of adding a noise term to the ordinary differential equation dynamics, as this will typically not lead to a solution respecting the state constraint (even if the deterministic model was naturally constrained).

As described above, there is natural motivation for considering stochastic differential equations where all three features, delay, non-negativity constraints and noise, are present. However, there has been little work on systematically studying such equations. One exception is the work of Kushner (see e.g., [25]), although this focuses on numerical methods for stochastic delay differential equations (including those with state constraints), especially those with bounded state space. We note that the behavior of constrained systems can be quite different from that of unconstrained analogues, e.g., in the deterministic delay equation case, the addition of a non-negativity constraint can turn an equation with unbounded oscillatory solutions into one with bounded periodic solutions, and in the stochastic delay equation case, transient behavior can be transformed into positive recurrence.

Here we seek conditions for existence and uniqueness of stationary distributions for stochastic delay differential equations with non-negativity constraints of the form:

$$X(t) = X(0) + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW(s) + Y(t), \quad t \geq 0, \tag{1.1}$$

where $X(t)$ takes values in the positive orthant of some Euclidean space, $\tau \in [0, \infty)$ is the length of the delay period, $X_s = \{X(s+u) : -\tau \leq u \leq 0\}$ tracks the history of the process over the delay period, W is a standard (multi-dimensional) Brownian motion noise source and the stochastic integral with respect to W is an Itô integral, Y is a vector-valued non-decreasing process which ensures that the non-negativity constraints on X are enforced, in particular, the i^{th} component of Y can increase only when the i^{th} component of X is zero. We refer to equations of the form (1.1) as stochastic delay differential equations with reflection, where the action of Y is termed reflection (at the boundary of the orthant).

This thesis is organized as follows. The rigorous definition of a solution of (1.1) and properties of solutions are given in Chapter 2. Stationary distributions are defined in Chapter 3, and a general condition guaranteeing their existence is described in Sections 3.2-3.4. This condition is in terms of uniform moment bounds. Conditions under which such moment bounds hold in terms of restrictions on b and σ are given in Chapter 4. While the results here are new, we do use some results and adapt some techniques developed by Itô and Nisio [20] and Mao [28]. The results of Chapters 3 and 4 are combined to give sufficient conditions for existence of a stationary distribution in Section 5.1. Conditions for uniqueness of such a stationary distribution are given in Section 5.2. Our proofs in that section are an adaptation of methods developed recently by Hairer, Mattingley, and Scheutzow [16] for proving uniqueness of stationary distributions for stochastic delay differential equations without constraints. An important aspect of the results in [16] is that they enable one to obtain uniqueness of stationary distributions for stochastic delay differential equations when the dispersion coefficient depends on the history of the process over the delay period. Previous results on uniqueness of stationary distributions were often restricted to cases where dispersion coefficients depended only on the current state of the process (cf. [11, 25, 37, 41, 43]), with an exception being [20]. This is in part due to potential reconstruction of the initial condition from the quadratic variation process ([37, 43]), which precludes ergodicity of the process and rules out use of Doob's theorem (see [11], Theorem 4.2.1). Some applications of our results to a few particular examples arising from biochemical reaction systems and Internet rate control are discussed in Chapter 6. Appendix A contains a list of the notation that appears throughout this work, Appendix B covers some inequalities that appear frequently throughout this work, and Appendix C discusses some conditions that imply that equation (1.1) is well-posed.

As an example of the applicability of these results, the one-dimensional equation (in differential form)

$$dX(t) \quad = \quad (A_1 - B_1 X(t-1))dt + (A_2 + B_2 X(t-1))dW(t) + dY(t), \quad t \geq 0, \tag{1.2}$$

where $A_1 \in \mathbb{R}$, $B_1 > 0$, $A_2 > 0$, $B_2 \in (0, \frac{1}{4})$, has a unique stationary distribution as long as $B_1 > \frac{B_2^2}{2(1-4B_2)}$. See Section 6.2 for justification of this result.

1.2 Notation

We shall use the following notation throughout this work.

For each positive integer d, let $\mathbb{R}^d$ denote d-dimensional Euclidean space, and let $\mathbb{R}^d_+$ denote the closed positive orthant in $\mathbb{R}^d$. When $d = 1$, we suppress the d and write $\mathbb{R}$ for $(-\infty, \infty)$ and $\mathbb{R}_+$ for $[0, \infty)$. For each $i = 1, \ldots, d$, the i^{th} component of a column vector $v \in \mathbb{R}^d$ will be denoted by v^i. For two vectors $u, v \in \mathbb{R}^d$, the statement $u \geq v$ will mean that $u^i \geq v^i$ for each $i = 1, \ldots, d$. For each $r \in \mathbb{R}$, define $r^+ = \max\{r, 0\}$ and $r^- = \max\{-r, 0\}$.

For any real numbers r, s, $\delta_{r,s}$ denotes the Kronecker delta, i.e., it is one if $r = s$ and zero otherwise.

Unless specified otherwise, we treat vectors $v \in \mathbb{R}^d$ as column vectors, i.e., $v = (v^1, \ldots, v^d)'$. For $u, v \in \mathbb{R}^d$, $u \cdot v = \sum_{i=1}^{d} u^i v^i$ denotes the dot product of u with v. Let $\mathbb{M}^{d\times m}$ denote the set of $d \times m$ matrices with real entries. For a given matrix $A \in \mathbb{M}^{d\times m}$, A^i_j denotes the entry of the i^{th} row and the j^{th} column, A^i denotes the i^{th} row, and A_j denotes the j^{th} column. The notation I_d will denote the $(d \times d)$-identity matrix.

We denote the maximum norm on $\mathbb{R}^d$ by

$$|v|_\infty = \max_{i=1,\ldots,d} |v^i|, \quad v = (v^1, \ldots, v^d)' \in \mathbb{R}^d.$$

For $p \in [1, \infty)$, we also have the corresponding p-norms:

$$|v|_p = \left(|v^1|^p + \cdots + |v^d|^p\right)^{\frac{1}{p}}, \quad v \in \mathbb{R}^d.$$

These norms can also be applied to row vectors, i.e., for $v = (v^1, \ldots, v^d)$, $|v|_p := |v'|_p$ for $p \in [1, \infty]$. We use some matrix norms as well. Given a matrix A, $\|A\|_\infty := \max_{i,j} |A^i_j|$ denotes the maximum norm of A, and $\|A\|_2 := \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{m} (A^i_j)^2}$ denotes the Frobenius norm, of A.

For any two metric spaces $\mathbb{E}_1, \mathbb{E}_2$, let $C(\mathbb{E}_1, \mathbb{E}_2)$ denote the space of continuous functions from $\mathbb{E}_1$ into $\mathbb{E}_2$. Here, $\mathbb{E}_1$ will often be a closed interval $F \subset (-\infty, \infty)$, and $\mathbb{E}_2$ will often be $\mathbb{R}^d$ or $\mathbb{R}^d_+$ for various dimensions d. For any metric space $\mathbb{E}$ with metric ρ, we use $B(x, r)$ (where $x \in \mathbb{E}$ and $r > 0$) to denote the open ball $\{y \in \mathbb{E} : \rho(x, y) < r\}$ of radius r around x, and we use $\mathcal{B}(\mathbb{E})$ to denote the associated collection of Borel sets of $\mathbb{E}$. The set of bounded Borel measurable real-valued functions on $\mathbb{E}$ will be denoted by $B_b(\mathbb{E})$, and $C_b(\mathbb{E})$ will denote the set of bounded continuous real-valued functions on $\mathbb{E}$.

For any integer d and closed interval I in $(-\infty, \infty)$, we endow $C(I, \mathbb{R}^d)$ and $C(I, \mathbb{R}^d_+)$ with the topologies of uniform convergence on compact intervals in I. These are Polish spaces. In the case of $C(I, \mathbb{R}^d_+)$, we use $\mathcal{M}_I$ to denote the associated Borel σ-algebra. We shall also use the abbreviations $\mathbb{C}_I = C(I, \mathbb{R}_+)$ and $\mathbb{C}^d_I = C(I, \mathbb{R}^d_+)$. For a given dimension m, let $C_0(\mathbb{R}_+, \mathbb{R}^m)$ denote the set of continuous functions $x : [0, \infty) \to \mathbb{R}^m$ such that $x(0) = 0$. For a given closed bounded interval I, metric space $\mathbb{E}$, and $t \in I$, we define the evaluation map $e_t : C(I, \mathbb{E}) \to \mathbb{E}$ by $e_t(f) = f(t)$.

Throughout this work, we fix $\tau \in (0, \infty)$, which will be referred to as the delay. Define $\mathbb{I} = [-\tau, 0]$ and $\mathbb{J} = [-\tau, \infty)$. As a subset of the vector space $C(\mathbb{I}, \mathbb{R}^d)$, $\mathbb{C}^d_{\mathbb{I}}$ has the equivalent norms

$$\|x\|_p := \sup_{t\in\mathbb{I}} |x(t)|_p, \quad x \in \mathbb{C}^d_{\mathbb{I}}, \quad p \in [1,\infty],$$

that induce its topology of uniform convergence on compact intervals. The associated Borel σ-algebra is $\mathcal{M}_{\mathbb{I}}$. For $x \in \mathbb{C}^d_{\mathbb{J}}$ and $t \geq 0$, define $x_t \in \mathbb{C}^d_{\mathbb{I}}$ by $x_t(s) = x(t+s)$ for all $s \in \mathbb{I}$. It should be emphasized that $x(t) \in \mathbb{R}^d_+$ is a point, while $x_t \in \mathbb{C}^d_{\mathbb{I}}$ is a continuous function on $\mathbb{I}$ taking values in $\mathbb{R}^d_+$. For each $t \in \mathbb{R}_+$, we define the projection $p_t : \mathbb{C}^d_{\mathbb{J}} \to \mathbb{C}^d_{\mathbb{I}}$ by $p_t(x) := x_t$ for each $x \in \mathbb{C}^d_{\mathbb{J}}$.

For a closed interval I in $(-\infty,\infty)$, $a_1 \leq a_2$ in I, and a path $x = (x^1,\ldots,x^d)' \in C(I,\mathbb{R}^d)$, we define the oscillation of x over $[a_1,a_2]$ by

$$\mathrm{Osc}(x,[a_1,a_2]) \quad := \sup_{s,t\in[a_1,a_2]} |x(t)-x(s)|_\infty \quad = \quad \max_{i=1}^{d} \sup_{s,t\in[a_1,a_2]} |x^i(t)-x^i(s)|, \tag{1.3}$$

the modulus of continuity of x over I by

$$w_I(x,\delta) := \max_{i=1}^{d} \sup_{\substack{s,t\in I\\ |s-t|<\delta}} |x^i(t)-x^i(s)|, \quad \delta > 0$$

and for each $p \in [1,\infty]$, the supremum p-norm of x over I by

$$\|x\|_{I,p} = \sup_{t\in I} |x(t)|_p.$$

When $I = \mathbb{I}$, the notation $\|\cdot\|_p$ described in the previous paragraph will be used as an abbreviation for $\|\cdot\|_{\mathbb{I},p}$. When $d = 1$, the maximum norm and all p-norms (for $p \in [1,\infty)$) are equal to the absolute value, so we abbreviate $\|\cdot\|_I := \|\cdot\|_{I,\infty} = \|\cdot\|_{I,p}$ in this case.

By a filtered probability space, we mean a quadruple $(\Omega,\mathcal{F},\{\mathcal{F}_t, t\geq 0\},P)$, where $\mathcal{F}$ is a σ-algebra on the set of possible outcomes Ω, P is a probability measure on the measurable space $(\Omega,\mathcal{F})$, and $\{\mathcal{F}_t, t \geq 0\}$ is a filtration of sub-σ-algebras of $\mathcal{F}$ where the *usual conditions* are satisfied, i.e., $(\Omega,\mathcal{F},P)$ is a complete probability space, and for each $t \geq 0$, $\mathcal{F}_t$ contains all P-null sets of $\mathcal{F}$ and $\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$. Given two σ-finite measures μ,ν on a measurable space $(\Omega,\mathcal{F})$, the notation $\mu \sim \nu$ will mean that μ and ν are mutually absolutely continuous, i.e., for any $\Lambda \in \mathcal{F}$, $\mu(\Lambda) = 0$ if and only if $\nu(\Lambda) = 0$.

Given a positive integer m, by a standard m-dimensional Brownian motion, we mean a continuous process $\{W(t) = (W^1(t),\ldots,W^m(t))', t \geq 0\}$ taking values in $\mathbb{R}^m$ such that

(i) $W(0) = 0$ a.s.,

(ii) the coordinate processes, $W^1,\ldots,W^m$, are independent,

(iii) for each $i = 1, \dots, m$, positive integer n and $0 \leq t_1 < t_2 < \dots t_n < \infty$, the increments

$$W^i(t_2) - W^i(t_1), \quad W^i(t_3) - W^i(t_2), \quad \dots\ , \quad W^i(t_n) - W^i(t_{n-1})$$

are independent, and

(iv) for each $i = 1, \dots, m$ and $0 \leq s < t < \infty$, $W^i(t) - W^i(s)$ is normally distributed with mean zero and variance $t - s$.

Given a function $f : \{1, 2, \dots\} \to \mathbb{R}$ and $a \in (-\infty, \infty]$, the notation $f(n) \nearrow a$ as $n \to \infty$ means that $\lim_{n\to\infty} f(n) = a$ and $f(n) \leq f(n+1)$ for each $n = 1, 2, \dots$.

A list of the preceding notation along with other notation that appears in this work can be found in the Appendix.

Chapter 2

Stochastic Delay Differential Equations with Reflection

In this chapter, we define a solution to equation (1.1) precisely, and we derive some useful properties of solutions.

2.1 Definition of a Solution

Recall from Section 1.2 that we are fixing a $\tau \in (0, \infty)$, which will be referred to as the delay, and we define $\mathbb{I} = [-\tau, 0]$, $\mathbb{J} = [-\tau, \infty)$, $\mathbb{C}^d_{\mathbb{I}} = C(\mathbb{I}, \mathbb{R}^d_+)$, and $\mathbb{C}^d_{\mathbb{J}} = C(\mathbb{J}, \mathbb{R}^d_+)$. Furthermore, we fix positive integers d and m, and functions $b : \mathbb{C}^d_{\mathbb{I}} \to \mathbb{R}^d$ and $\sigma : \mathbb{C}^d_{\mathbb{I}} \to \mathbb{M}^{d\times m}$ that satisfy the following continuity and linear growth assumptions.

Assumption 2.1.1. *The functions b and σ are continuous, and there exist non-negative constants C_1, C_2, C_3, and C_4 such that for each $x \in \mathbb{C}^d_{\mathbb{I}}$,*

$$|b(x)|_2 \le C_1 + C_2\|x\|_2, \quad \textit{and} \tag{2.1}$$

$$\|\sigma(x)\|_2^2 \le C_3 + C_4\|x\|_2^2. \tag{2.2}$$

Definition 2.1.1. A solution of the stochastic delay differential equation with reflection (SDDER) associated with (b, σ) is a d-dimensional continuous process $X = \{X(t), t \in \mathbb{J}\}$ defined on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \ge 0\}, P)$ that P-a.s. satisfies (1.1), where

(i) $X(t)$ is $\mathcal{F}_0$-measurable for each $t \in \mathbb{I}$, $X(t)$ is $\mathcal{F}_t$-measurable for each $t > 0$, and $X(t) \in \mathbb{R}^d_+$ for all $t \in \mathbb{J}$,

(ii) Y is a d-dimensional continuous and non-decreasing process such that $Y(0) = 0$ and $Y(t)$ is $\mathcal{F}_t$-measurable for each $t \geq 0$,

(iii) W is a standard m-dimensional Brownian motion such that $\{W(t), \mathcal{F}_t, t \geq 0\}$ is a martingale under P,

(iv) $\int_0^t X(s) \cdot dY(s) = 0$ for all $t \geq 0$, i.e., Y^i can increase only when X^i is at zero for $i = 1, \ldots, d$.

A natural initial condition for equation (1.1) is not an initial state such as $X(0) = v \in \mathbb{R}_+^d$ as the dynamics would be indeterminate for $t \in [0, \tau]$ in that case (assuming that b or σ depends on a delayed state). The natural initial condition is an initial segment $X_0 = x \in \mathbb{C}_{\mathbb{I}}^d$, or more generally, an initial distribution μ on $(\mathbb{C}_{\mathbb{I}}^d, \mathcal{M}_{\mathbb{I}})$, i.e., $P(X_0 \in \Lambda) = \mu(\Lambda)$ for each $\Lambda \in \mathcal{M}_{\mathbb{I}}$.

Remark. A solution of the SDDER (1.1) defines a stochastic process $\{X_t, t \geq 0\}$ with state space $\mathbb{C}_{\mathbb{I}}^d$. This process may be considered a more natural "state descriptor process" than the process $\{X(t), t \geq 0\}$.

Remark. As a consequence of condition (i) and the continuity of the paths of X, $\{X_t, t \geq 0\}$ is adapted to $\{\mathcal{F}_t, t \geq 0\}$, and $t \mapsto X_t(\omega)$ is continuous from $\mathbb{R}_+$ into $\mathbb{C}_{\mathbb{I}}^d$ for each $\omega \in \Omega$. It follows that the mapping $F : \mathbb{R}_+ \times \Omega \to \mathbb{C}_{\mathbb{I}}^d$, where $F(t, \omega) = X_t(\omega)$, is progressively measurable, being continuous in t and adapted (see Lemma II.73.10 of [39]). Therefore since $\sigma : \mathbb{C}_{\mathbb{I}}^d \to \mathbb{M}^{d \times m}$ is continuous, $(\sigma \circ F)(t, \omega) = \sigma(X_t(\omega))$ is progressively measurable. Since $\sigma(\cdot)$ is continuous and $X_\cdot(\omega)$ is continuous for each ω, $\sigma(X_t(\omega))$ is continuous in t, and thus bounded on compact time intervals, so that $P\left(\int_0^t \|\sigma(X_s)\|_2^2 ds < \infty\right) = 1$ for each $t \in \mathbb{R}_+$, so that $\left\{\int_0^t \sigma(X_s) dW(s), \mathcal{F}_t, t \geq 0\right\}$ is a continuous d-dimensional local martingale. Similarly, $b(X_t(\omega))$ is continuous in t for each $\omega \in \Omega$, so for each fixed $T \geq 0$, there is a constant $K_\omega \in (0, \infty)$ such that $|b(X_t(\omega))|_2 \leq K_\omega$ for each $t \in [0, T]$. Therefore, $\int_0^T |b(X_t(\omega))|_2 dt \leq TK_\omega < \infty$, so that $\{\int_0^t b(X_s) ds, t \geq 0\}$ is a continuous adapted process whose coordinates are locally of bounded variation. Since $\{Y(t), t \geq 0\}$ is continuous, has nondecreasing (and therefore locally of bounded variation) coordinates, and is adapted to the filtration $\{\mathcal{F}_t, t \geq 0\}$, $\{X(0) + \int_0^t b(X_s) ds + Y(t), \mathcal{F}_t, t \geq 0\}$ is a continuous adapted process that is locally of bounded variation. Therefore, $\{X(t), t \geq 0\}$ is a continuous semimartingale with respect to $\{\mathcal{F}_t, t \geq 0\}$.

Since this work is directed at proving existence and uniqueness of *stationary* distributions, we shall assume that the equation (1.1) is well-posed. That is, in addition to Assumption

2.1.1, we make the following minimal assumption on existence and uniqueness in law of solutions to (1.1).

Assumption 2.1.2. *For each deterministic initial condition* $X_0 = x \in \mathbb{C}^d_{\mathbb{I}}$*, on some filtered probability space* $(\Omega^x, \mathcal{F}^x, \{\mathcal{F}^x_t, t \geq 0\}, P^x)$*, there exist a Brownian motion martingale* $\{W^x(t), t \geq 0\}$ *and continuous processes* $X^x = \{X^x(t), t \in \mathbb{J}\}$ *and* $Y^x = \{Y^x(t), t \in \mathbb{R}_+\}$*, such that* X^x *is a solution to the SDDER (1.1) with* (X^x, Y^x, W^x) *in place of* (X, Y, W)*. Furthermore, the law of* X^x *is unique given* x*.*

Sufficient conditions for strong existence and pathwise uniqueness for solutions to the SDDER (1.1) are given in Appendix C. These conditions imply that Assumptions 2.1.1 and 2.1.2 hold.

2.2 Examples

Example 2.2.1. *Fix* $\alpha, \gamma, \varepsilon, C > 0$*. For* $x \in \mathbb{C}_{\mathbb{I}}$*, define*

$$b(x) \quad = \quad \frac{\alpha}{\left(1+\frac{x(-\tau)}{C}\right)^2} - \gamma, \quad \text{and} \quad \sigma(x) \quad = \quad \varepsilon\sqrt{\frac{\alpha}{\left(1+\frac{x(-\tau)}{C}\right)^2} + \gamma}\ .$$

The SDDER associated with this pair (b, σ) *is an example for* $d = 1$ *that arises in the study of biochemical reaction systems [29].*

Example 2.2.2. *Differential delay equations with linear or affine coefficients are used often in engineering. An example of an SDDER with affine coefficients is the following. For* $x \in \mathbb{C}_{\mathbb{I}}$*, let*

$$b(x) := b_0 - b_1 x(0) - \sum_{i=2}^{n} b_i x(-r_i) + \sum_{i=n+1}^{n'} b_i x(-r_i), \tag{2.3}$$

and

$$\sigma(x) := a_0 + \sum_{i=1}^{n''} a_i x(-s_i), \tag{2.4}$$

where $0 \leq r_i \leq \tau$ *and* $0 \leq s_i \leq \tau$ *for each* i*,* $n' > n \geq 2$*,* $n'' \geq 0$*, and* $b_0, \ldots, b_{n'}, a_0, \ldots, a_{n''} \geq 0$*.*

Example 2.2.3. *Paganini and Wang [33], Peet and Lall [36], and Papachristadolou, Doyle, and Low [34, 35], studied a multidimensional deterministic model of Internet rate control with* d *servers and* d' *sources. In this, the dynamics are given by*

$$dX(t) \quad = \quad \hat{b}(X_t)dt, \tag{2.5}$$

where the i^{th} *component of* $X(t)$ *represents the price at time* t *that server* i *charges for the transmission of a packet through it, assuming that the servers use Active Queue Management (AQM) and that the sources use the Transmission Control Protocol (TCP). The drift* $\hat{b}$ *is discontinuous; for each* $i = 1, \ldots, d$*, and* $x \in \mathbb{C}^d_{\mathbb{I}}$*,*

$$\hat{b}^i(x) = \begin{cases} -1 + \sum_{j=1}^{d'} A_{ij} \exp\left(-B_j \sum_{k=1}^{d} A_{kj} C_{kj} x^k(-r_{ijk})\right) & \text{if } x^i(0) > 0 \\ \left(-1 + \sum_{j=1}^{d'} A_{ij} \exp\left(-B_j \sum_{k=1}^{d} A_{kj} C_{kj} x^k(-r_{ijk})\right)\right)^+ & \text{if } x^i(0) = 0 \end{cases} \quad (2.6)$$

for some $B_1, \ldots, B_d > 0$*, and* $A_{ij} \geq 0$*,* $C_{kj} > 0$*, and* $r_{ijk} > 0$ *for all* $i, k \in \{1, \ldots, d\}$ *and* $j \in \{1, \ldots, d'\}$*. The constants depend on parameters such as the capacity of the queues at the servers, the maximal rate of transmission from each source, a routing matrix that determines which sources use which servers, and some other parameters. The solutions of (2.5) remain in the nonnegative orthant by the construction of* $\hat{b}$ *(for the meaning of a solution with such a discontinuous righthand side, see, e.g., [13]). It turns out that the solutions of the SDDER associated with* $\sigma \equiv 0$ *coincide with the solutions of (2.5) when the drift* b *is defined by*

$$b^i(x) := -1 + \sum_{j=1}^{d'} A_{ij} \exp\left(-B_j \sum_{k=1}^{d} A_{kj} C_{kj} x^k(-r_{ijk})\right), \quad i = 1, \ldots, d.$$

Allowing σ *to be non-zero yields a noisy version of the deterministic model.*

2.3 Reflection

To ensure that a solution of (1.1) remains non-negative, we employ Skorokhod's well-known mapping for constraining a continuous real-valued function to be non-negative by means of reflection at the origin. We apply this mapping to each component.

For each positive integer d, define $C_+(\mathbb{R}_+, \mathbb{R}^d) := \{x \in C(\mathbb{R}_+, \mathbb{R}^d) : x(0) \in \mathbb{R}^d_+\}$.

Definition 2.3.1. Given a path $x \in C_+(\mathbb{R}_+, \mathbb{R}^d)$, we say that a pair (z, y) of functions in $C_+(\mathbb{R}_+, \mathbb{R}^d)$ solves the Skorokhod problem for x with (normal) reflection if

(i) $z(t) = x(t) + y(t)$ for all $t \geq 0$ and $z(t) \in \mathbb{R}^d_+$ for each $t \geq 0$,

(ii) for each $i = 1, \ldots, d$, $y^i(0) = 0$ and y^i is nondecreasing,

(iii) for each $i = 1, \ldots, d$, $y^i(t) = \int_0^t 1_{\{0\}}(z^i(s))dy^i(s)$ for all $t \geq 0$, i.e., y^i can increase only when z^i is at zero.

The path z is called the reflection of x, and the path y is called the regulator of x.

Remark. Here, we consider only normal reflection as described in the above definition, but there is a substantial theory for oblique reflection. For a survey up through 1995, see [45], and for some applications, see [24]. We have some partial results still under development for oblique reflection. In the following, when we use the term reflection, we mean normal reflection.

We summarize some basic facts about the Skorokhod problem in the next proposition. With normal reflection, the problem can be solved component by component in an explicit way.

Proposition 2.3.1. *For each path $x \in C(\mathbb{R}_+, \mathbb{R}^d)$, there exists a unique solution (z, y) to the Skorokhod problem for x. Thus there exists a pair of functions $(\phi, \psi) : C_+(\mathbb{R}_+, \mathbb{R}^d) \to C_+(\mathbb{R}_+, \mathbb{R}^{2d})$ defined by $(\phi(x), \psi(x)) = (z, y)$. The pair (ϕ, ψ) satisfies the following:*

(i) $\mathrm{Osc}(\phi(x), [a, b]) \leq \mathrm{Osc}(x, [a, b])$.

(ii) There exists a constant $K_\ell > 0$ such that for each $x, y \in C_+(\mathbb{R}_+, \mathbb{R}^d)$, we have for each $t \geq 0$,

$$\|\psi(x) - \psi(y)\|_{[0,t],2} \leq K_\ell \|x - y\|_{[0,t],2}, \text{ and}$$

$$\|\phi(x) - \phi(y)\|_{[0,t],2} \leq K_\ell \|x - y\|_{[0,t],2}.$$

Proof. These properties follow from the well-known construction of y:

$$y^i(t) \quad = \quad \left(- \min_{0 \leq s \leq t} x^i(s) \right)^+, \quad i = 1, \ldots, d.$$

For more details, see [12, 18, 46]. We note that $K_\ell \leq 2$, but we keep the notation K_ℓ for convenience. □

Thus the Skorokhod problem with reflection is well-posed, the solution map (ϕ, ψ) is Lipschitz continuous, and oscillations of the reflection $\phi(x)$ are bounded by the oscillations of x.

For notational convenience, given a continuous adapted stochastic process $\{\xi(t), t \geq -\tau\}$ taking values in $\mathbb{R}^d_+$ and an m-dimensional Brownian motion W defined on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, , we define

$$\mathcal{I}(\xi)(t) \quad := \quad \xi(0) + \int_0^t b(\xi_s) ds + \int_0^t \sigma(\xi_s) dW(s), \quad t \geq 0. \tag{2.7}$$

For a solution X of the SDDER, $X(t) = \mathcal{I}(X)(t) + Y(t), t \geq 0$. In other words, $\{X(t), t \geq 0\}$ can be obtained by reflecting $\mathcal{I}(X)$, i.e., $X = \phi(\mathcal{I}(X))$, and $Y = \psi(\mathcal{I}(X))$, because of the uniqueness of solutions to the Skorokhod problem. Then, as a consequence of Proposition 2.3.1(i), for any $0 \le a < b < \infty$

$$\text{Osc}(X, [a, b]) \le \text{Osc}(\mathcal{I}(X), [a, b]). \tag{2.8}$$

2.4 Bounds for Moments of Supremum Norm over Bounded Intervals

We now state the Burkholder-Davis-Gundy inequalities. A proof can be found, e.g., in [22], Theorem 3.3.28.

Proposition 2.4.1. *For each $p > 0$, there exist constants $c_p, \tilde{c}_p > 0$ such that for any real-valued continuous local martingale $\{M(t), \mathcal{F}_t, t \ge 0\}$ with quadratic variation process $\{\langle M\rangle(t), t \ge 0\}$, and each stopping time η taking values in $\mathbb{R}_+$,*

$$c_p E\left[(\langle M\rangle(\eta))^{\frac{p}{2}}\right] \le E\left[\|M\|^p_{[0,\eta]}\right] \le \tilde{c}_p E\left[(\langle M\rangle(\eta))^{\frac{p}{2}}\right].$$

Under Assumption 2.1.1, any solution X to (1.1) satisfies the following supremum bound.

Lemma 2.4.1. *For each $p \in [2, \infty)$, there exists a continuous function $F_p : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ that is nondecreasing in each argument, such that*

$$E\left[\|X\|^p_{[-\tau,T],p}\right] \quad \le \quad F_p(E[\|X_0\|^p_p], T) \text{ for each } T > 0. \tag{2.9}$$

In fact,

$$F_p(r, s) \quad = \quad k_p(s) + \tilde{k}_p(s) r,$$

where the functions k_p and $\tilde{k}_p$ are non-decreasing on $(0, \infty)$, and they depend only on p, the dimensions d, m, and the linear growth constants C_1, C_2, C_3, C_4 from Assumption 2.1.1.

Proof. For any $T > 0$,

$$\begin{aligned}
\|X\|^p_{[0,T],p} &\leq \|X^1\|^p_{[0,T]} + \cdots + \|X^d\|^p_{[0,T]} \\
&\leq |X^1(0) + \text{Osc}(X^1,[0,T])|^p + \cdots + |X^d(0) + \text{Osc}(X^d,[0,T])|^p \\
&\leq 2^{p-1}\Big(|X^1(0)|^p + (\text{Osc}(X^1,[0,T]))^p + \cdots + |X^d(0)|^p + (\text{Osc}(X^d,[0,T]))^p\Big) \\
&\leq 2^{p-1}\left(|X(0)|^p_p + d(\text{Osc}(X,[0,T]))^p\right) \\
&\leq 2^{p-1}\left(|X(0)|^p_p + d(\text{Osc}(\mathcal{I}(X),[0,T]))^p\right) \\
&\leq 2^{p-1}|X(0)|^p_p \\
&\quad + 2^{p-1} d \max_i \left(\int_0^T |b^i(X_t)|dt + 2\sup_{s\in[0,T]}\left|\int_0^s \sigma^i(X_t)dW(t)\right|\right)^p \\
&\leq 2^{p-1}|X(0)|^p_p \\
&\quad + 2^{p-1} d \left(\int_0^T |b(X_t)|_2 dt + 2\sup_{s\in[0,T]}\left|\int_0^s \sigma(X_t)dW(t)\right|_2\right)^p \\
&\leq 2^{p-1}|X(0)|^p_p \\
&\quad + 2^{2p-2} d \left(\left(\int_0^T |b(X_t)|_2 dt\right)^p + 2^p \sup_{s\in[0,T]}\left|\int_0^s \sigma(X_t)dW(t)\right|^p_2\right).
\end{aligned} \tag{2.10}$$

Here, we have used Proposition B.0.1 for the third and eighth inequalities.

For each integer $n \geq 1$, define the stopping time $\eta_n := \inf\{t \geq 0 : \|X\|_{[-\tau,t],2} \geq n\}$, with the convention that $\inf \emptyset = \infty$. The Burkholder-Davis-Gundy inequalities imply that for each n, for any $T > 0$,

$$\begin{aligned}
&\sum_{i=1}^d \sum_{j=1}^m E\left[\sup_{s\in[0,T\wedge\eta_n]}\left|\int_0^s \sigma^i_j(X_t)dW^j(t)\right|^p\right] \\
&\qquad \leq \tilde{c}_p \sum_{i=1}^d \sum_{j=1}^m E\left[\left(\int_0^{T\wedge\eta_n} \left|\sigma^i_j(X_t)\right|^2_2 dt\right)^{\frac{p}{2}}\right] \\
&\qquad \leq \tilde{c}_p E\left[\left(\int_0^{T\wedge\eta_n} \|\sigma(X_t)\|^2_2\, dt\right)^{\frac{p}{2}}\right] \\
&\qquad \leq \tilde{c}_p E\left[\left(\int_0^{T\wedge\eta_n} \left(C_3 + C_4\|X_t\|^2_2\right) dt\right)^{\frac{p}{2}}\right] \\
&\qquad \leq \tilde{c}_p 2^{\frac{p}{2}-1}\left(C_3^{\frac{p}{2}} T^{\frac{p}{2}} + C_4^{\frac{p}{2}} E\left[\left(\int_0^T \|X\|^2_{[-\tau,t\wedge\eta_n],2} dt\right)^{\frac{p}{2}}\right]\right).
\end{aligned} \tag{2.11}$$

The third inequality follows from repeated application of inequality (B.2) and the linearity of integration, the third inequality follows from the linear growth condition (2.2), and the fourth inequality follows from Proposition B.0.1.

Then, on replacing T by $T\wedge\eta_n$ in (2.10), taking expectations, using the linear growth assumption (2.1) on b, inequality (2.11), and repeated application of Proposition B.0.1 and Hölder's inequality, we have for each n and $T>0$,

$$\begin{aligned}
&E\left[\|X\|^p_{[0,T\wedge\eta_n],p}\right]\\
\leq\ & 2^{p-1}E\left[\|X_0\|_p^p\right]+2^{2p-2}dE\left[\left(\int_0^{T\wedge\eta_n}|b(X_t)|_2dt\right)^p\right]\\
&+2^{3p-2}dE\left[\sup_{s\in[0,T\wedge\eta_n]}\left|\int_0^s\sigma(X_t)dW(t)\right|_2^p\right]\\
\leq\ & 2^{p-1}E\left[\|X_0\|_p^p\right]+2^{2p-2}dE\left[\left(\int_0^{T\wedge\eta_n}(C_1+C_2\|X_t\|_2)\,dt\right)^p\right]\\
&+2^{3p-2}dE\left[\sup_{s\in[0,T\wedge\eta_n]}\left(\sum_{i=1}^d\left|\sum_{j=1}^m\int_0^s\sigma_j^i(X_t)dW^j(t)\right|^2\right)^{\frac{p}{2}}\right]\\
\leq\ & 2^{p-1}E\left[\|X_0\|_p^p\right]+2^{3p-3}dC_1^pT^p+2^{3p-3}dC_2^pE\left[\left(\int_0^{T\wedge\eta_n}\|X_t\|_2dt\right)^p\right]\\
&+2^{3p-2}dm^{\frac{p}{2}}E\left[\sup_{s\in[0,T\wedge\eta_n]}\left(\sum_{i=1}^d\sum_{j=1}^m\left|\int_0^s\sigma_j^i(X_t)dW^j(t)\right|^2\right)^{\frac{p}{2}}\right]\\
\leq\ & 2^{p-1}E\left[\|X_0\|_p^p\right]+2^{3p-3}dC_1^pT^p+2^{3p-3}dC_2^pT^{\frac{p-1}{p}}E\left[\int_0^{T\wedge\eta_n}\|X_t\|_2^pdt\right]\\
&+2^{3p-2}dm^{\frac{p}{2}}(dm)^{\frac{p-2}{2}}\sum_{i=1}^d\sum_{j=1}^mE\left[\sup_{s\in[0,T\wedge\eta_n]}\left|\int_0^s\sigma_j^i(X_t)dW^j(t)\right|^p\right]\\
\leq\ & 2^{p-1}E\left[\|X_0\|_p^p\right]+2^{3p-3}dC_1^pT^p+2^{3p-3}dC_2^pT^{\frac{p-1}{p}}\int_0^TE\left[\|X\|^p_{[-\tau,t\wedge\eta_n],2}\right]dt\\
&+2^{3p-2}d^{\frac{p}{2}}m^{p-1}\tilde{c}_p2^{\frac{p}{2}-1}\left(C_3^{\frac{p}{2}}T^{\frac{p}{2}}+C_4^{\frac{p}{2}}T^{\frac{p-2}{p}}E\left[\int_0^T\|X\|^p_{[-\tau,t\wedge\eta_n],2}dt\right]\right)\\
\leq\ & K_1(T)+K_2(T)E\left[\|X_0\|_p^p\right]+K_3(T)\int_0^TE\left[\|X\|^p_{[0,t\wedge\eta_n],p}\right]dt,
\end{aligned}\tag{2.12}$$

where

$$\begin{aligned}
K_1(T) &= 2^{3p-3}dC_1^pT^p+2^{3p-2}d^{\frac{p}{2}}m^{p-1}\tilde{c}_p2^{\frac{p}{2}-1}C_3^{\frac{p}{2}}T^{\frac{p}{2}},\\
K_2(T) &= 2^{p-1}+2^{3p-3}dC_2^pT^{\frac{2p-1}{p}}d^{\frac{p}{2}-1}+2^{3p-2}d^{\frac{p}{2}}m^{p-1}\tilde{c}_p2^{\frac{p}{2}-1}C_4^{\frac{p}{2}}T^{\frac{2p-2}{p}}d^{\frac{p}{2}-1},\\
K_3(T) &= 2^{3p-3}dC_2^pT^{\frac{p-1}{p}}d^{\frac{p}{2}-1}+2^{3p-2}d^{\frac{p}{2}}m^{p-1}\tilde{c}_p2^{\frac{p}{2}-1}C_4^{\frac{p}{2}}T^{\frac{p-2}{p}}d^{\frac{p}{2}-1}.
\end{aligned}$$

For the last inequality, we used the fact that inequality (B.5) implies

$$\|X\|^p_{[-\tau,t],2}\ \leq\ \|X_0\|_2^p+\|X\|^p_{[0,t],2}\ \leq\ d^{\frac{p}{2}-1}\|X_0\|_p^p+d^{\frac{p}{2}-1}\|X\|^p_{[0,t],p}.\tag{2.13}$$

Note that K_1, K_2, K_3 are all increasing in T, so that for each $t \in [0,T]$, we have

$$E\left[\|X\|^p_{[0,t\wedge\eta_n],p}\right] \leq K_1(T) + K_2(T)E\left[\|X_0\|^p_p\right] + K_3(T)\int_0^t E\left[\|X\|^p_{[0,s\wedge\eta_n],p}\right] ds. \quad (2.14)$$

Gronwall's inequality (Proposition B.0.4) now yields for each $t \in [0,T]$,

$$E\left[\|X\|^p_{[0,t\wedge\eta_n],p}\right] \leq \left(K_1(T) + K_2(T)E\left[\|X_0\|^p_p\right]\right) e^{tK_3(T)}, \quad (2.15)$$

so that

$$\begin{aligned} E\left[\|X\|^p_{[-\tau,T\wedge\eta_n],p}\right] &\leq E\left[\|X_0\|^p_p\right] + E\left[\|X\|^p_{[0,T\wedge\eta_n],p}\right] \\ &\leq E\left[\|X_0\|^p_p\right] + \left(K_1(T) + K_2(T)E\left[\|X_0\|^p_p\right]\right) e^{TK_3(T)}. \quad (2.16) \end{aligned}$$

The monotone convergence theorem can now be invoked to obtain

$$E\left[\|X\|^p_{[-\tau,T],p}\right] \leq E\left[\|X_0\|^p_p\right] + \left(K_1(T) + K_2(T)E\left[\|X_0\|^p_p\right]\right) e^{TK_3(T)}. \quad (2.17)$$

Thus, the result holds with $F_p(r,s) = K_1(s)e^{sK_3(s)} + (1 + K_2(s)e^{sK_3(s)})r$ for $p \geq 2$.

□

2.5 Feller Property

This section is devoted to proving a type of regularity in the initial condition of the solutions to (1.1) (referred to as Feller continuity). This will be used in later sections. Indeed, in the next section it will be shown that under Assumptions 2.1.1 and 2.1.2, the SDDER (1.1) generates a Feller continuous family of transition functions $\{P_t(x,\Lambda), t \geq 0, x \in \mathbb{C}^d_{\mathbb{I}}, \Lambda \in \mathcal{M}_{\mathbb{I}}\}$, where $P_t(x,\Lambda) = P(X^x_t \in \Lambda)$. The proof of this relies on a standard argument: prove tightness of solutions with initial conditions converging to $x \in \mathbb{C}^d_{\mathbb{I}}$ and show any limit point has law of the solution starting from x.

Recall the notation for the modulus of continuity of a function introduced in Section 1.2. We will use the following well-known criterion for tightness on $C(I,\mathbb{R}^d)$, where I is a closed interval in $\mathbb{R}$ and has left endpoint $t_0 \in \mathbb{R}$. A proof can be found in Theorem 2.4.10 of [22], or in the case of a bounded interval I, Theorem 7.3 of [5].

Proposition 2.5.1. *For any closed interval I of $\mathbb{R}$ with left endpoint $t_0 \in \mathbb{R}$, a sequence $\{P_n\}_{n=1}^\infty$ of probability measures on the path space $(C(I,\mathbb{R}^d), \mathcal{B}(C(I,\mathbb{R}^d)))$ is tight if and only if*

(i) $\lim_{a\to\infty}\sup_{n\geq 1} P_n\left(x \in C(I,\mathbb{R}^d) : |x(t_0)|_\infty > a\right) = 0,$ *and*

(ii) for each fixed $T > t_0$ and $\lambda > 0$, we have

$$\lim_{\delta\to 0}\sup_{n\geq 1} P_n\left(x \in C(I,\mathbb{R}^d) : w_{[t_0,T]\cap I}(x,\delta) \geq \lambda\right) = 0.$$

Remark. If the probability laws of a sequence $\{X^n\}_{n=1}^{\infty}$ of continuous stochastic processes are tight then we say that the sequence $\{X^n\}_{n=1}^{\infty}$ is tight.

The following two technical lemmas have a general form that will allow us to use them again in Section 3.3.

Lemma 2.5.1. *Assume that $-\tau \leq t_1 < t_2$, and that X is a solution to (1.1) on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ with associated Brownian motion W. Then for any $\delta, \lambda > 0$, we have*

$$\begin{aligned} &P\left(w_{[t_1,t_2]}(X,\delta) \geq \lambda\right) \\ &\leq \;\; P\left(w_{[t_1\wedge 0,t_2\wedge 0]}(X,\delta) + \delta(C_1 + C_2\|X\|_{[(t_1\vee 0)-\tau,t_2\vee 0],2}) \geq \frac{\lambda}{2}\right) \\ &\quad +P\left(\sup_{\substack{t_1\vee 0\leq s<t\leq t_2\vee 0\\ |s-t|<\delta}}\left|\int_s^t \sigma(X_r)dW(r)\right|_{\infty} \geq \frac{\lambda}{2}\right). \end{aligned} \tag{2.18}$$

Proof. The conclusion is obvious if $t_2 \leq 0$, so we assume that $t_2 > 0$. Fix $\delta > 0$ and $\lambda > 0$. Then,

$$\begin{aligned} \sup_{\substack{s,t\in[t_1,t_2]\\ |s-t|<\delta}} |X(s)-X(t)|_{\infty} &\leq w_{[t_1\wedge 0,t_2\wedge 0]}(X,\delta) + \sup_{\substack{s,t\in[t_1\vee 0,t_2]\\ |s-t|<\delta}} |X(s)-X(t)|_{\infty} \\ &\leq w_{[t_1\wedge 0,t_2\wedge 0]}(X,\delta) + \sup_{\substack{t_1\vee 0\leq s<t\leq t_2\\ |s-t|<\delta}} \mathrm{Osc}(X,[s,t]) \\ &\leq w_{[t_1\wedge 0,t_2\wedge 0]}(X,\delta) + \sup_{\substack{t_1\vee 0\leq s<t\leq t_2\\ |s-t|<\delta}} \mathrm{Osc}(\mathcal{I}(X),[s,t]) \\ &\leq w_{[t_1\wedge 0,t_2\wedge 0]}(X,\delta) + \sup_{\substack{t_1\vee 0\leq s<t\leq t_2\\ |s-t|<\delta}} \left|\int_s^t b(X_r)dr\right|_{\infty} \\ &\quad + \sup_{\substack{t_1\vee 0\leq s<t\leq t_2\\ |s-t|<\delta}} \left|\int_s^t \sigma(X_r)dW(r)\right|_{\infty} \\ &\leq w_{[t_1\wedge 0,t_2\wedge 0]}(X,\delta) + \delta\left(C_1 + C_2\|X\|_{[(t_1\vee 0)-\tau,t_2],2}\right) \\ &\quad + \sup_{\substack{t_1\vee 0\leq s<t\leq t_2\\ |s-t|<\delta}} \left|\int_s^t \sigma(X_r)dW(r)\right|_{\infty}. \end{aligned}$$

The third inequality follows from Proposition 2.3.1(i), while the fourth inequality follows from the form of $\mathcal{I}(X)$ as in (2.7). The linear growth condition (2.1) and the fact that $|v|_\infty \leq |v|_2$ for all $v \in \mathbb{R}^d$ were used for the fifth inequality. Therefore, for any $\delta > 0$, $\lambda > 0$, we have

$$\begin{aligned} P\left(w_{[t_1,t_2]}(X,\delta) \geq \lambda\right) &= P\left(\sup_{\substack{s,t\in[t_1,t_2] \\ |s-t|<\delta}} |X(s)-X(t)|_\infty \geq \lambda\right) \\ &\leq P\left(w_{[t_1\wedge 0,t_2\wedge 0]}(X,\delta) + \delta\left(C_1 + C_2\|X\|_{[(t_1\vee 0)-\tau,t_2\vee 0],2}\right) \geq \frac{\lambda}{2}\right) \\ &\quad + P\left(\sup_{\substack{t_1\vee 0\leq s<t\leq t_2 \\ |s-t|<\delta}} \left|\int_s^t \sigma(X_r)dW(r)\right|_\infty \geq \frac{\lambda}{2}\right). \end{aligned} \tag{2.19}$$

□

Lemma 2.5.2. *Assume that for some index set N and positive real number T, we have a collection of closed subintervals $\{[s_\nu, t_\nu], \nu \in N\}$ of $\mathbb{R}_+$ such that $t_\nu - s_\nu = T$ for all $\nu \in N$, and we have a collection $\{X^\nu\}_{\nu\in N}$ where for each $nu \in N$, X^ν defined on $(\Omega^\nu, \mathcal{F}^\nu, \{\mathcal{F}_t^\nu\}, P^\nu)$ is a solutions to (1.1) with associated Brownian motion W^ν, and $\{\|X^\nu\|_{[s_\nu-\tau,t_\nu],2}\}_{\nu\in N}$ is stochastically bounded, i.e.,*

$$\lim_{a\to\infty}\sup_{\nu\in N} P^\nu\left(\|X^\nu\|_{[s_\nu-\tau,t_\nu],2} > a\right) = 0.$$

Then for each $\varepsilon, \lambda > 0$, there is a $\delta_{\varepsilon,\lambda} > 0$ such that $\delta \in (0, \delta_{\varepsilon,\lambda}]$ implies that for all $\nu \in N$,

$$P^\nu\left(\sup_{\substack{s_\nu\leq s<t\leq t_\nu \\ |s-t|<\delta}} \left|\int_s^t \sigma(X_r^\nu)dW^\nu(r)\right|_\infty \geq \lambda\right) < \varepsilon. \tag{2.20}$$

Proof. We shall prove this by using a time change to transform the local martingales

$$\left\{\int_0^t \sigma^i(X_r^\nu)dW^\nu(r), t \geq 0\right\}$$

to Brownian motions (possibly run up to random times).

For each $\nu \in N$, define $M^\nu(t) = (M^{\nu,1}, \ldots, M^{\nu,d})'(t) := \int_0^t \sigma(X_s^\nu)dW^\nu(s)$ for all $t \geq 0$ so that

$$\langle M^{\nu,i}\rangle(t) = \int_0^t |\sigma^i(X_s^\nu)|_2^2 ds. \tag{2.21}$$

For $s \geq 0$, let $T_{\nu,i}(s) = \inf\{t \geq 0 : \langle M^{\nu,i}\rangle(t) > s\}$ and $\mathcal{G}_s^{\nu,i} := \mathcal{F}_{T_{\nu,i}(s)}^\nu$.

For each $\nu \in N$ and $i = 1, \ldots, d$, let

$$\left\{\{W^{(\nu,i)}(t), \mathcal{F}_t^{(\nu,i)}, t \geq 0\}, \left(\Omega^{(\nu,i)}, \mathcal{F}^{(\nu,i)}, P^{(\nu,i)}\right)\right\}$$

be a Brownian motion martingale on a filtered space satisfying the usual conditions. Define $\tilde{\Omega}^{\nu,i} := \Omega^{\nu} \times \Omega^{(\nu,i)}$, $\tilde{\mathcal{G}}^{\nu,i} := \mathcal{F}^{\nu} \otimes \mathcal{F}^{(\nu,i)}$, and $\tilde{P}^{\nu,i} := P^{\nu} \times P^{(\nu,i)}$, and let $\mathcal{N}^{\nu,i}$ be the set of all subsets of null sets of $\tilde{P}^{\nu,i}$ from $\tilde{\mathcal{G}}^{\nu,i}$. Set $\tilde{\mathcal{F}}^{\nu,i} := \sigma\left(\tilde{\mathcal{G}}^{\nu,i} \cup \mathcal{N}^{\nu,i}\right)$, and for each $t \geq 0$, $\tilde{\mathcal{G}}_t^{\nu,i} := \bigcap_{s>t} \sigma\left(\left(\mathcal{F}^{\nu}_{T_{\nu,i}(s)} \otimes \mathcal{F}_s^{(\nu,i)}\right) \cup \mathcal{N}^{\nu,i}\right)$, and $\tilde{\mathcal{F}}_t^{\nu,i} := \bigcap_{s>t} \sigma\left(\left(\mathcal{F}_s^{\nu} \otimes \mathcal{F}_0^{(\nu,i)}\right) \cup \mathcal{N}^{\nu,i}\right)$.

Consider the extended probability space $(\tilde{\Omega}^{\nu,i}, \tilde{\mathcal{F}}^{\nu,i}, \tilde{P}^{\nu,i})$ and on it, the processes $\tilde{M}^{\nu,i}(\omega^{\nu}, \omega^{(\nu,i)}) := M^{\nu,i}(\omega^{\nu})$ and $\tilde{W}^{(\nu,i)}(\omega^{\nu}, \omega^{(\nu,i)}) := W^{(\nu,i)}(\omega^{(\nu,i)})$, which are adapted to $\{\tilde{\mathcal{F}}_t^{\nu,i}, t \geq 0\}$ and $\{\tilde{\mathcal{G}}_t^{\nu,i}, t \geq 0\}$, respectively, and have the same distributions under $\tilde{P}^{\nu,i}$ as $M^{\nu,i}$ under P^{ν} and $W^{(\nu,i)}$ under $P^{(\nu,i)}$, respectively. Also define $\tilde{T}_{\nu,i}(s) := \inf\{t \geq 0 : \langle \tilde{M}^{\nu,i}\rangle(t) > s\}$ and $\tilde{S}_{\nu,i} := \lim_{t\to\infty} \langle \tilde{M}^{\nu,i}\rangle(t)$, so that, for instance, $\tilde{T}_{\nu,i}(s)(\omega^{\nu}, \omega^{(\nu,i)}) = T_{\nu,i}(s)(\omega^{\nu})$, $\{\tilde{T}_{\nu,i} \leq t\} = \{T_{\nu,i} \leq t\} \times \Omega^{(\nu,i)}$, and $\tilde{\mathcal{F}}^{\nu,i}_{\tilde{T}_{\nu,i}(t)} \subset \tilde{\mathcal{G}}_t^{\nu,i}$.

For each $t \geq 0$, $\langle \tilde{M}^{\nu,i}\rangle(t)$ is a $\{\tilde{\mathcal{G}}_s^{\nu,i}, s \geq 0\}$-stopping time because

$$\{\langle \tilde{M}^{\nu,i}\rangle(t) \leq s\} = \{\tilde{T}_{\nu,i}(s) \geq t\} \in \tilde{\mathcal{F}}_{\tilde{T}_{\nu,i}(s)} \subset \tilde{\mathcal{G}}_s^{\nu,i},$$

and thus $\tilde{S}_{\nu,i} = \lim_{n\to\infty} \langle \tilde{M}^{\nu,i}\rangle(n)$ is also a $\{\tilde{\mathcal{G}}_s^{\nu,i}\}$-stopping time.

The time-change theorem (see, e.g., Theorem 3.4.6 and Problem 3.4.7 in [22], or Theorem V.1.7 in [38]) implies that

$$\left\{\beta^{\nu,i}(t) := \tilde{W}^{(\nu,i)}(t) - \tilde{W}^{(\nu,i)}(t \wedge \tilde{S}_{\nu,i}) + \tilde{M}^{\nu,i}(\tilde{T}_{\nu,i}(t)), \tilde{\mathcal{G}}_t^{\nu,i}, t \geq 0\right\}$$

is a Brownian motion martingale, and that $\beta^{\nu,i}\left(\langle \tilde{M}^{\nu,i}\rangle(t)\right) = \tilde{M}^{\nu,i}(t)$.

Equation (2.21) and the linear growth bound (2.2) imply that for each $a > 0$, on the event $A_{a,\nu,i} := \{\|X^{\nu}\|_{[s_{\nu}-\tau,t_{\nu}],\infty} \leq a\} \times \Omega^{(\nu,i)} \in \tilde{\mathcal{F}}^{\nu,i}$, we have for any $s_{\nu} \leq s < t \leq t_{\nu}$,

$$(C_3 + C_4 a^2)(t - s) \geq \langle \tilde{M}^{\nu,i}\rangle(t) - \langle \tilde{M}^{\nu,i}\rangle(s).$$

Since $\tilde{M}^{\nu,i}(t) = \beta^{\nu,i}\left(\langle \tilde{M}^{\nu,i}\rangle(t)\right)$ for each $t \geq 0$, we have on each $A_{a,\nu,i}$ that

$$\begin{aligned}
w_{[s_{\nu},t_{\nu}]}(\tilde{M}^{\nu,i}, \delta) &= w_{[s_{\nu},t_{\nu}]}(\beta^{\nu,i}(\langle \tilde{M}^{\nu,i}\rangle(\cdot)), \delta) \\
&\leq w_{[\langle \tilde{M}^{\nu,i}\rangle(s_{\nu}), \langle \tilde{M}^{\nu,i}\rangle(t_{\nu})]}(\beta^{\nu,i}, \delta(C_3 + C_4 a^2)) \\
&\leq w_{[\langle \tilde{M}^{\nu,i}\rangle(s_{\nu}), \langle \tilde{M}^{\nu,i}\rangle(s_{\nu}) + (C_3 + C_4 a^2)(t_{\nu} - s_{\nu})]}(\beta^{\nu,i}, \delta(C_3 + C_4 a^2)).
\end{aligned}$$

Therefore for every $a > 0$,

$$
\begin{aligned}
P^{\nu}\left(\sup_{\substack{s_\nu \le s < t \le t_\nu \\ |s-t|<\delta}} \left|\int_s^t \sigma(X_r^{\nu}) dW^{\nu}(r)\right|_{\infty} \ge \lambda\right) &= P^{\nu}\left(w_{[s_\nu,t_\nu]}(M^{\nu},\delta) \ge \lambda\right) \\
&= P^{\nu}\left(\max_i w_{[s_\nu,t_\nu]}(M^{\nu,i},\delta) \ge \lambda\right) \\
&\le \sum_{i=1}^d P^{\nu}\left(w_{[s_\nu,t_\nu]}(M^{\nu,i},\delta) \ge \lambda\right) \\
&= \sum_{i=1}^d \tilde{P}^{\nu,i}\left(w_{[s_\nu,t_\nu]}(\tilde{M}^{\nu,i},\delta) \ge \lambda\right) \\
&\le \sum_{i=1}^d \tilde{P}^{\nu,i}\left(w_{[\langle\tilde{M}^{\nu,i}\rangle(s_\nu),\langle\tilde{M}^{\nu,i}\rangle(s_\nu)+(C_3+C_4a^2)T]}(\beta^{\nu,i},\delta(C_3+C_4a^2)) \ge \lambda, A_{a,\nu,i}\right) \\
&\quad + \sum_{i=1}^d \tilde{P}^{\nu,i}(A^c_{a,\nu,i}) \\
&\le \sum_{i=1}^d \tilde{P}^{\nu,i}\left(w_{[0,(C_3+C_4a^2)T]}(\beta^{\nu,i},\delta(C_3+C_4a^2)) \ge \lambda\right) + \sum_{i=1}^d \tilde{P}^{\nu,i}(A^c_{a,\nu,i}). \qquad (2.22)
\end{aligned}
$$

The third inequality follows from the fact that $\langle\tilde{M}^{\nu,i}\rangle(s_\nu)$ is a $\{\tilde{\mathcal{G}}_t^{\nu,i}\}$-stopping time, and Brownian motion restarted at a stopping time is a another Brownian motion. Since the set of random variables $\{\|X^{\nu}\|_{[s_\nu-\tau,t_\nu],2}\}_{\nu\in N}$ is stochastically bounded, there is an a_ε big enough that

$$\sup_{\nu\in N} \tilde{P}^{\nu,i}(A^c_{a_\varepsilon,\nu,i}) = \sup_{\nu\in N} P^{\nu}(\|X^{\nu}\|_{[s_\nu-\tau,t_\tau],2} > a_\varepsilon) < \frac{\varepsilon}{2d} \text{ for each } i = 1,\dots,d.$$

Then since a single measure is tight, and $\{\beta^{\nu,i}\}$ all have the same distribution under their respective probability measures $\tilde{P}^{\nu,i}$, there is a $\delta_{\varepsilon,\lambda} > 0$ such that $\delta \in (0,\delta_{\varepsilon,\lambda}]$ implies that

$$\sup_{\nu\in N} \tilde{P}^{\nu,i}\left(w_{[0,(C_3+C_4a_\varepsilon^2)T]}(\beta^{\nu,i},\delta(C_3+C_4a_\varepsilon^2)) \ge \lambda\right) < \frac{\varepsilon}{2d} \text{ for each } i = 1,\dots,d.$$

The result follows.

$\square$

Lemma 2.5.3. *Assume* $\{x_n\}_{n=1}^{\infty} \subset \mathbb{C}_{\mathbb{I}}^d$ *such that* $\lim_{n\to\infty} x_n = x \in \mathbb{C}_{\mathbb{I}}^d$*, and for each* $n \ge 1$*, let* P^n *be the distribution on the space* $(\mathbb{C}_{\mathbb{J}}^d \times C_0(\mathbb{R}_+,\mathbb{R}^m), \mathcal{M}_{\mathbb{J}} \otimes \mathcal{B}(C_0(\mathbb{R}_+,\mathbb{R}^m)))$ *of the pair* (X^{x_n}, W^{x_n}) *associated with a solution of (1.1) that has initial condition* x_n*. Then* $\{P^n\}_{n=1}^{\infty}$ *is tight.*

Proof. Fix $\varepsilon > 0$. Since $x_n \to x$, the set $\{\|x_n\|_2 : n \ge 1\}$ is bounded,so that for each fixed $T \ge 0$, Lemma 2.4.1 implies that $\sup_{n\ge1} E[\|X^{x_n}\|^2_{[-\tau,T],2}] < \infty$,so that the collection

$\{\|X^{x_n}\|_{[-\tau,T],2}\}_{n=1}^{\infty}$ is stochastically bounded. Because of the boundedness of $\{\|x_n\|_\infty : n \geq 1\}$, which follows from $x_n \to x$ as $n \to \infty$, we have

$$\begin{aligned}\lim_{a\to\infty}\sup_{n\geq 1} P(|X^{x_n}(-\tau)|_\infty \geq a) &= \lim_{a\to\infty}\sup_{n\geq 1} P(|x_n(-\tau)|_\infty \geq a) \\ &= 0. \end{aligned} \tag{2.23}$$

Lemma 2.5.1 implies that for any $\delta, \lambda > 0$ and $n \geq 1$, we have

$$\begin{aligned} P\left(w_{[-\tau,T]}(X^{x_n},\delta) \geq \lambda\right) &\leq P\left(w_{\mathbb{I}}(x_n,\delta) + \delta\left(C_1 + C_2\|X^{x_n}\|_{[-\tau,T],2}\right) \geq \frac{\lambda}{2}\right) \\ &\quad + P\left(\sup_{\substack{0\leq s<t\leq T \\ |s-t|<\delta}} \left|\int_s^t \sigma(X_r^{x_n})dW^{x_n}(r)\right|_\infty \geq \frac{\lambda}{2}\right). \end{aligned} \tag{2.24}$$

Since $x_n \to x$ in $\mathbb{C}_{\mathbb{I}}^d$, the set $\{x_n : n \geq 1\}$ is precompact in $\mathbb{C}_{\mathbb{I}}^d$, so that the Arzelà-Ascoli theorem implies that there is a $\delta_{\varepsilon,\lambda}^{(1)} > 0$ such that $w_{\mathbb{I}}(x_n,\delta) < \frac{\lambda}{4}$ for every $n \geq 1$ and $\delta \in (0, \delta_{\varepsilon,\lambda}^{(1)})$. Then for each $\delta \in (0, \frac{\lambda}{4C_1} \wedge \delta_{\varepsilon,\lambda}^{(1)})$, we have

$$\begin{aligned} &P\left(w_{\mathbb{I}}(x_n,\delta) + \delta\left(C_1 + C_2\|X^{x_n}\|_{[-\tau,T],2}\right) \geq \frac{\lambda}{2}\right) \\ &\qquad \leq P\left(\delta\left(C_1 + C_2\|X^{x_n}\|_{[-\tau,T],2}\right) \geq \frac{\lambda}{4}\right) \\ &\qquad = P\left(\|X^{x_n}\|_{[-\tau,T],2} \geq \frac{\frac{\lambda}{4\delta} - C_1}{C_2}\right). \end{aligned} \tag{2.25}$$

Since $\frac{\frac{\lambda}{4\delta}-C_1}{C_2} \to \infty$ as $\delta \to 0$, the stochastic boundedness of $\{\|X^{x_n}\|_{[-\tau,T],2}\}$ implies that there is a $\delta_{\varepsilon,\lambda}^{(2)} \in (0, \frac{\lambda}{4C_1} \wedge \delta_{\varepsilon,\lambda}^{(1)})$ such that for all $\delta \in (0, \delta_{\varepsilon,\lambda}^{(2)})$,

$$P\left(w_{\mathbb{I}}(x_n,\delta) + \delta\left(C_1 + C_2\|X^{x_n}\|_{[-\tau,T],2}\right) \geq \frac{\lambda}{2}\right) < \frac{\varepsilon}{2} \text{ for all } n \geq 1. \tag{2.26}$$

Lemma 2.5.2 implies that there is a $\delta_{\varepsilon,\lambda}^{(3)} > 0$ such that $\delta \in (0, \delta_{\varepsilon,\lambda}^{(3)}]$ implies that

$$P\left(\sup_{\substack{0\leq s<t\leq T \\ |s-t|<\delta}} \left|\int_s^t \sigma(X_r^{x_n})dW^{x_n}(r)\right|_\infty \geq \frac{\lambda}{2}\right) < \frac{\varepsilon}{2} \text{ for all } n \geq 1. \tag{2.27}$$

Therefore, whenever $\delta \in (0, \delta_{\varepsilon,\lambda}^{(2)} \wedge \delta_{\varepsilon,\lambda}^{(3)}]$, from (2.24) we have for all $n \geq 1$,

$$P\left(w_{[-\tau,T]}(X^{x_n},\delta) \geq \lambda\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{2.28}$$

Therefore, by Proposition 2.5.1, the sequence of stochastic processes $\{X^{x_n}\}$ is tight. Since each element of the sequence $\{W^{x_n}\}$ has the same distribution, this sequence is tight as well. Therefore the sequence $\{(X^{x_n}, W^{x_n})\}$ is also tight (see [20], Lemma 3.1), which shows that $\{P^n\}$ is tight.

□

Given a continuous function $f : \mathbb{R}_+ \to \mathbb{R}$ that is locally of bounded variation, and $t \geq 0$, denote the total variation of f up to time t by $\mathcal{TV}_t(f)$. We will use the following proposition proved in [23] (Theorem 2.2 and Remark 2.3).

Proposition 2.5.2. *Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and for each integer $n \geq 1$, let $\{\mathcal{F}_t^n, t \geq 0\}$ be a filtration satisfying the usual conditions, and let X^n, S^n be $\{\mathcal{F}_t^n\}$-adapted processes with sample paths in $C(\mathbb{R}_+, \mathbb{R}^d)$ and $C(\mathbb{R}_+, \mathbb{R}^m)$, respectively, such that S^n is an $\{\mathcal{F}_t^n\}$-semimartingale with decomposition $S^n = M^n + A^n$, where M^n is an $\{\mathcal{F}_t^n\}$-local martingale and A^n is an $\{\mathcal{F}_t^n\}$-adapted process locally of bounded variation such that for each $t \geq 0$, the sequence of random variables $\{\mathcal{TV}_t(A^n)\}_{n=1}^\infty$ is stochastically bounded. If $(X^n, S^n) \to (X, S)$ in probability in the Skorokhod topology as $n \to \infty$, then S is a semimartingale with respect to a filtration to which X and S are adapted, and $(X^n, S^n, \int_0^\cdot X^n(s)dS^n(s)) \to (X, S, \int_0^\cdot X(s)dS(s))$ in probability in the Skorokhod topology as $n \to \infty$.*

Remark. Kurtz and Protter actually prove a more general theorem, and the above proposition is a simplification tailored to our needs.

Lemma 2.5.4. *Assume that $\{x_n\}_{n=1}^\infty \subset \mathbb{C}_{\mathbb{I}}^d$ is such that $x_n \to x \in \mathbb{C}_{\mathbb{I}}^d$ as $n \to \infty$. Let P^n be the law of the pair (X^{x_n}, W^{x_n}) associated with a solution to (1.1) having initial condition $X_0^{x_n} = x_n$. Let Q be any weak limit point of the sequence $\{P^n\}_{n=1}^\infty$. Then, Q is the law of a pair (X^*, W^*) associated with a solution to (1.1) having initial condition $X_0^* = x$.*

Proof. To simplify notation, we assume (by passing to a subsequence) that $P^n \to Q$ weakly as $n \to \infty$. The by the Skorokhod representation theorem, there is a complete probability space $(\Omega^*, \mathcal{F}^*, P^*)$ and $\mathbb{C}_{\mathbb{J}}^d \times C_0(\mathbb{R}_+, \mathbb{R}^m)$-valued random elements $\{(X^n, W^n)\}_{n=1}^\infty$, and (X^*, W^*) on that probability space such that $(X^n, W^n) \to (X^*, W^*)$ P^*-a.s. in the topology of uniform convergence on compact time intervals, and such that P^n is the law of (X^n, W^n) under P^* for each $n \geq 1$, and Q is the law of (X^*, W^*) under P^*. In particular,

$$P^*(X_0^* = x) = P^*\left(\lim_{n\to\infty} X_0^n = x\right) \geq 1 - \sum_{n=1}^{\infty} P^*(X_0^n \neq x_n) = 1.$$

For each $t \geq 0$, define $\mathcal{F}_t^n = \bigcap_{s>t} \mathcal{F}_s^{n,o}$ where $\mathcal{F}_s^{n,o}$ is the sub-σ-algebra of $\mathcal{F}^*$ generated by $\{(X^n(u), W^n(u)) : u \leq s\}$ and the null sets of $\mathcal{F}^*$.

Fix $\varepsilon, \lambda > 0$. Define $I(t) = x(0) + \int_0^t b(X_s^*)ds + \int_0^t \sigma(X_s^*)dW^*(s), t \geq 0$, and for each $n \geq 1$, $I^n(t) = x_n(0) + \int_0^t b(X_s^n)ds + \int_0^t \sigma(X_s^n)dW^n(s), t \geq 0$. By Section 2.3, $X^n(\cdot) = \phi(I^n)(\cdot)$.Then for each $t \geq 0$, $\lambda > 0$,

$$\begin{aligned} P^*(|X^*(t) - \phi(I)(t)|_2 > (2 + K_\ell)\lambda) \\ &\leq P^*(|X^*(t) - X^n(t)|_2 > \lambda) + P^*(|X^n(t) - \phi(I^n)(t)|_2 > \lambda) \\ &\quad + P^*(|\phi(I^n)(t) - \phi(I)(t)|_2 > K_\ell \lambda) \\ &\leq P^*(|X^*(t) - X^n(t)|_2 > \lambda) + P^*(\|I^n - I\|_{[0,t],2} > \lambda). \end{aligned} \tag{2.29}$$

The second inequality uses Proposition 2.3.1.

Since $X^n \to X^*$ P^*-a.s., we obtain $X^n \to X^*$ in probability. Therefore for each $t \geq 0$, since the evaluation map e_t is continuous, we have $X^n(t) \to X^*(t)$ in probability as $n \to \infty$. Thus there is a $N_{\varepsilon,\lambda}^{(1)}(t) > 0$ such that

$$P^*(|X^*(t) - X^n(t)|_2 > \lambda) \quad < \quad \frac{\varepsilon}{2} \text{ whenever } n \geq N_{\varepsilon,\lambda}^{(1)}(t). \tag{2.30}$$

For each $t \geq 0$, we have that

$$\begin{aligned} P^*(\|I^n - I\|_{[0,t],2} > \lambda) \\ &\leq P^*\left(|x_n(0) - x(0)|_2 + \int_0^t |b(X_s^n) - b(X_s^*)|_1 ds > \frac{\lambda}{2}\right) \\ &\quad + P^*\left(\sup_{s\in[0,t]} \left|\int_0^s \sigma(X_r^n)dW^n(r) - \int_0^s \sigma(X_r^*)dW^*(r)\right|_2 > \frac{\lambda}{2}\right). \end{aligned} \tag{2.31}$$

Since $x_n \to x$, we have $x_n(0) \to x(0)$, so that there is a $N_{\varepsilon,\lambda}^{(2)} > 0$ such that $|x_n(0) - x(0)|_2 < \frac{\lambda}{4}$ whenever $n \geq N_{\varepsilon,\lambda}^{(2)}$. Therefore, for $n \geq N_{\varepsilon,\lambda}^{(2)}$

$$\begin{aligned} P^*\left(|x_n(0) - x(0)|_2 + \int_0^t |b(X_s^n) - b(X_s^*)|_1 ds > \frac{\lambda}{2}\right) \\ \leq P^*\left(\int_0^t |b(X_s^n) - b(X_s^*)|_1 ds > \frac{\lambda}{4}\right). \end{aligned} \tag{2.32}$$

For each $t \geq 0$, there is a set $\Omega_t^* \in \mathcal{F}$ of P^*-measure 1 such that for all $\omega \in \Omega_t^*$, $\|X^n(\omega) - X^*(\omega)\|_{[-\tau,t],2} \to 0$ as $n \to \infty$. We then have from the linear growth bound on b that for each $n \geq 1$, $s \in [0, t]$, $\omega \in \Omega_t^*$,

$$\begin{aligned} |b(X_s^n(\omega)) - b(X_s^*(\omega))|_1 &\leq d^{\frac{1}{2}} |b(X_s^n(\omega)) - b(X_s^*(\omega))|_2 \\ &\leq 2d^{\frac{1}{2}} C_1 + d^{\frac{1}{2}} C_2 \left(\|X^n(\omega)\|_{[-\tau,t],2} + \|X^*(\omega)\|_{[-\tau,t],2}\right) \\ &\leq 2d^{\frac{1}{2}} C_1 + d^{\frac{1}{2}} C_2 \left(\sup_{n\geq 1} \|X^n(\omega)\|_{[-\tau,t],2} + \|X^*(\omega)\|_{[-\tau,t],2}\right), \end{aligned}$$

which is finite since $\|X^n(\omega) - X^*(\omega)\|_{[-\tau,t],2} \to 0$.

Since for each $s \in [0,t]$ and $\omega \in \Omega_t^*$, $X_s^n(\omega) \to X_s^*(\omega)$ in $\mathbb{C}_{\mathbb{I}}^d$, the continuity of b implies that $|b(X_s^n(\omega)) - b(X_s^*(\omega))|_1 \to 0$, so that the dominated convergence theorem implies that as $n \to \infty$, $\int_0^t |b(X_s^n) - b(X_s^*)|_1 ds \to 0$ on Ω_t^* (i.e., P^*-a.s.), and thus in probability. Therefore there exists a $N_{\varepsilon,\lambda}^{(3)}(t) \geq N_{\varepsilon,\lambda}^{(2)}$ such that $n \geq N_{\varepsilon,\lambda}^{(3)}(t)$ implies that

$$P^*\left(|x_n(0) - x(0)|_2 + \int_0^t |b(X_s^n) - b(X_s^*)|_1 ds > \frac{\lambda}{2}\right) \quad < \quad \frac{\varepsilon}{4}. \tag{2.33}$$

Set $\overline{\Omega} := \bigcap_{n=1}^{\infty} \Omega_n^*$, which also has P^*-measure one. For each $t \geq 0$, the function $h_t : [0,t] \times \mathbb{C}_{\mathbb{J}}^d \to \mathbb{C}_{\mathbb{I}}^d$, defined by $h_t(s,f) := f_s$, is continuous (see, for instance, the proof of Lemma 4.2 in [20]). Since for each $\omega \in \overline{\Omega}$, $\{X^n(\omega) : n \geq 1\} \cup \{X^*(\omega)\}$ is a compact set in $\mathbb{C}_{\mathbb{J}}^d$, $[0,t] \times (\{X^n(\omega) : n \geq 1\} \cup \{X^*(\omega)\})$ is also compact. Therefore its image $\mathcal{H}_t(\omega) := \{X_s^n(\omega) : n \geq 1, s \in [0,t]\} \cup \{X_s^*(\omega) : s \in [0,t]\}$ under h_t is also compact. Therefore the restriction of σ to $\mathcal{H}_t(\omega)$ is uniformly continuous (see, e.g., [10], Theorem II.5.15).

Since the function σ and the paths of X^n are continuous, the processes $\{X_s^n, s \geq 0\}$ and $\{\sigma(X_s^n), s \geq 0\}$ are continuous. As a consequence of the uniform continuity of σ when restricted to each $\mathcal{H}_t(\omega)$, $\sigma(X_\cdot^n) \to \sigma(X_\cdot^*)$, P^*-a.s. (and thus in probability), in the topology of uniform convergence on compact sets of $\mathbb{R}_+$ (the Skorokhod topology). To see this, let $t \geq 0$, $\omega \in \overline{\Omega}$, $\eta > 0$, $\delta(\eta,\omega,t) > 0$ such that $y, z \in \mathcal{H}_t(\omega)$ and $\|y - z\|_2 < \delta(\eta,\omega,t)$ imply $\|\sigma(y) - \sigma(z)\|_2 < \eta$, and for each $\delta > 0, \omega \in \overline{\Omega}$, let $N(\delta,\omega,t)$ be big enough so that $\sup_{s\in[-\tau,t]} |X^n(\omega)(s) - X^*(\omega)(s)|_2 < \delta$ for all $n \geq N(\delta,\omega,t)$. Then $\sup_{s\in[0,t]} \|\sigma(X_s^n(\omega)) - \sigma(X_s^*(\omega))\|_2 < \eta$ for all $n \geq N(\delta(\eta,\omega,t),\omega,t)$.

Since for each $\omega \in \Omega^*$, $\sigma(X_s^n(\omega))$ is a continuous function of s, $(W^n(s), \mathcal{F}_s^n, s \geq 0)$ is an m-dimensional martingale, and $(\sigma(X_\cdot^n), W^n) \to (\sigma(X_\cdot^*), W^*)$ in probability in the Skorokhod topology, Proposition 2.5.2 implies that there is a filtration $\{\mathcal{F}_t, t \geq 0\}$ to which X^* and W^* are adapted, and with respect to which $\{W^*(t), \mathcal{F}_t, t \geq 0\}$ is a semimartingale, and that

$$\left\|\int_0^\cdot \sigma(X_s^n) dW^n(s) - \int_0^\cdot \sigma(X_s^*) dW^*(s)\right\|_{[0,t],2} \to 0 \text{ in probability, as } n \to \infty.$$

Therefore there is an $N_{\varepsilon,\lambda}^{(4)}(t) > 0$ such that $n \geq N_{\varepsilon,\lambda}^{(4)}(t)$ implies that

$$P^*\left(\sup_{s\in[0,t]}\left|\int_0^s \sigma(X_r^n) dW^n(r) - \int_0^s \sigma(X_r^*) dW^*(r)\right|_2 > \frac{\lambda}{2}\right) \quad < \quad \frac{\varepsilon}{4}. \tag{2.34}$$

In fact, W^* is a martingale with respect to the filtration generated by (X^*, W^*) since this property holds for each W^n with respect to the filtration generated by (X^n, W^n), and it is preserved in the limit by uniform integrability conferred by the fact that W^n is a standard m-dimensional Brownian motion for each n. Thus, whenever $n \geq N^{(1)}_{\varepsilon,\lambda}(t) \vee N^{(3)}_{\varepsilon,\lambda}(t) \vee N^{(4)}_{\varepsilon,\lambda}(t)$, it follows from (2.29), (2.30), (2.31), (2.33), and (2.34) that

$$\begin{aligned} P^*(|X^*(t) - \phi(I)(t)| > (2+K_\ell)\lambda) & \\ \leq \quad & P^*(|X^*(t) - X^n(t)|_2 > \lambda) + P^*(\|I^n - I\|_{[0,t],2} > \lambda) \\ < \quad & \varepsilon. \end{aligned} \tag{2.35}$$

Since $\varepsilon, \lambda > 0$ were arbitrary, we have that $X^*(t) = \phi(I)(t)$, P^*-a.s. for each $t \geq 0$. By considering $t \in \mathbb{Q} \cap \mathbb{R}_+$, the continuity of the paths of X^* and $\phi(I)(t)$ imply that $X^*(t) = \phi(I)(t)$ for all $t \geq 0$, P^*-a.s., and thus X^* is a solution of the SDDER (1.1).

□

Corollary 2.5.1. *Assume that the sequence $\{x_n\} \subset \mathbb{C}^d_{\mathbb{I}}$ is such that $x_n \to x \in \mathbb{C}^d_{\mathbb{I}}$ as $n \to \infty$. Then if P^{x_n} is the (unique) law of a solution X^{x_n} to (1.1) with initial condition $X^{x_n}_0 = x_n$, $\{P^{x_n}\}$ converges weakly to P^x, the unique law of a solution X^x to (1.1) with initial condition $X^x_0 = x$.*

Proof. This follows by a standard argument. Tightness of $\{P^{x_n}\}$ follows from Lemma 2.5.3 since the sequence of marginal distributions of a tight sequence is also a tight sequence. Since each subsequence has a further subsequence that converges weakly to the same limit law, which is P^x by Lemma 2.5.4, it follows that the original sequence converges. □

Theorem 2.5.1. *Under Assumptions 2.1.1 and 2.1.2, for each continuous and bounded function $f : \mathbb{C}^d_{\mathbb{J}} \to \mathbb{R}$, the function $x \mapsto E[f(X^x)] = \int_{y \in \mathbb{C}^d_{\mathbb{J}}} f(y) P(X^x \in dy), x \in \mathbb{C}^d_{\mathbb{I}}$, is a continuous function on $\mathbb{C}^d_{\mathbb{I}}$.*

Proof. Let $\{x_n\}_{n=1}^\infty \subset \mathbb{C}^d_{\mathbb{I}}$ such that $\lim_{n\to\infty} x_n = x \in \mathbb{C}^d_{\mathbb{I}}$. Then it follows immediately from the Corollary that $E[f(X^{x_n})] \to E[f(X^x)]$ as $n \to \infty$.

□

Corollary 2.5.2. *Under Assumptions 2.1.1 and 2.1.2, for each $t \geq 0$ and continuous and bounded function $f : \mathbb{C}^d_{\mathbb{I}} \to \mathbb{R}$, the function $x \mapsto E[f(X^x_t)] = \int_{\mathbb{C}^d_{\mathbb{I}}} f(y) P(X^x_t \in dy)$, is a continuous function on $\mathbb{C}^d_{\mathbb{I}}$.*

Proof. This follows from the fact that the function $g := f \circ p_t$ is continuous and bounded on $\mathbb{C}^d_{\mathbb{J}}$ if f is continuous and bounded on $\mathbb{C}^d_{\mathbb{I}}$. □

2.6 Markov Property and Associated Semigroup

We devote this section to proving that under Assumptions 2.1.1 and 2.1.2, equation (1.1) generates a family of Markovian transition functions, generally defined as follows (see [11]).

Definition 2.6.1. Let $(\mathbb{E}, \mathcal{E})$ be a Polish space with Borel σ-algebra generated by a metric ρ, and recall that we denote by $C_b(\mathbb{E})$ (resp., $B_b(\mathbb{E})$) the bounded and continuous (resp. bounded and Borel-measurable) real-valued functions on $\mathbb{E}$, with norm $\|f\|_{\mathbb{E}} = \sup_{x\in\mathbb{E}}|f(x)|$, for $f \in B_b(\mathbb{E})$. A *family of Markovian transition functions on* $(\mathbb{E}, \mathcal{E})$ is a family $\{P_t(\cdot,\cdot), t \geq 0\}$ of functions $P_t : \mathbb{E} \times \mathcal{E} \to [0, 1], t \geq 0$, such that

(i) For each $t \geq 0, \Lambda \in \mathcal{E}$, the function $x \mapsto P_t(x, \Lambda)$ is measurable on $(\mathbb{E}, \mathcal{E})$,

(ii) For each $t \geq 0, x \in \mathbb{E}$, the function $\Lambda \mapsto P_t(x, \Lambda)$ is a probability measure on $\mathcal{E}$,

(iii) For each $s, t \geq 0, x \in \mathbb{E}, \Lambda \in \mathcal{E}$,

$$P_{s+t}(x, \Lambda) = \int_{\mathbb{E}} P_s(y, \Lambda) P_t(x, dy), \tag{2.36}$$

(iv) For each $x \in \mathbb{E}, \Lambda \in \mathcal{E}$, $P_0(x, \Lambda) = 1_\Lambda(x)$.

For each $f \in B_b(\mathbb{E}), t \geq 0$, we define

$$(P_t f)(x) = \int_{\mathbb{E}} f(y) P_t(x, dy), \text{ for } x \in \mathbb{E}. \tag{2.37}$$

We call $\{P_t, t \geq 0\}$ defined on $B_b(\mathbb{E})$ a *Markovian semigroup* of linear operators if $\{P_t(\cdot,\cdot), t \geq 0\}$ is a family of Markovian transition functions.

Definition 2.6.2. A Markovian semigroup $\{P_t, t \geq 0\}$ on $B_b(\mathbb{E})$ is called *stochastically continuous* if

$$\lim_{t\to 0} P_t(x, B(x, \varepsilon)) = 1$$

for each $x \in \mathbb{E}$ and $\varepsilon > 0$.

The following proposition is proved in [11] (Proposition 2.1.1).

Proposition 2.6.1. *A Markovian semigroup* $\{P_t, t \geq 0\}$ *is stochastically continuous if and only if for each* $f \in C_b(\mathbb{E})$ *and* $x \in \mathbb{E}$, $\lim_{t\to 0}(P_t f)(x) = f(x)$.

Lemma 2.6.1. *If a Markovian semigroup* $\{P_t, t \geq 0\}$ *is stochastically continuous, then for each* $x \in \mathbb{E}$ *and* $\Lambda \in \mathcal{E}$*, the function* $t \mapsto P_t(x, \Lambda)$ *is Borel measurable on* $[0, \infty)$.

Proof. Since $\{P_t\}$ is Markovian, Proposition 2.6.1 implies that the function $t \mapsto P_t f(x)$ on $[0,\infty)$ is right continuous for each $x \in \mathbb{E}$ and $f \in C_b(\mathbb{E})$. For an open set $\Lambda \in \mathcal{E}$, the sequence of continuous functions $\{f_n\}_{n=1}^{\infty}$, where $f_n(x) := 1 \wedge n\rho(x, \Lambda^c)$, $x \in \mathbb{E}$, increases pointwise to 1_Λ. Therefore by the monotone convergence theorem, as $n \to \infty$,

$$(P_t f_n)(x) = \int_{\mathbb{E}} f_n(y) P_t(x, dy) \nearrow \int_{\mathbb{E}} 1_\Lambda(y) P_t(x, dy) = P_t(x, \Lambda),$$

for each $x \in \mathbb{E}$, $t \geq 0$. For each $n \geq 1$, $x \in \mathbb{E}$, the function $t \mapsto (P_t f_n)(x)$ is right continuous and therefore measurable. Therefore the function $t \mapsto P_t(x, \Lambda) = \sup_n (P_t f_n)(x)$ is measurable. Since $\mathcal{E}$ is generated by the open sets in $\mathbb{E}$, the result follows by a standard invocation of the monotone class theorem.

□

We now explicitly define the family of Markovian transition functions (or equivalently, the associated Markovian semigroup) induced by the SDDER (1.1) that we will work with henceforth. For each $(x, \Lambda) \in \mathbb{C}_{\mathbb{I}}^d \times \mathcal{M}_{\mathbb{I}}$, define

$$P_t(x, \Lambda) \quad = \quad P^x\left(X_t^x \in \Lambda\right), \quad t \geq 0, \tag{2.38}$$

where $\{(X^x, Y^x, W^x), (\Omega^x, \mathcal{F}^x, \{\mathcal{F}_t^x\}, P^x)\}$ yields a solution to (1.1) with initial condition $X_0^x = x$. Uniqueness in law implies that P_t is well-defined. Then conditions (ii) and (iv) of Definition 2.6.1 are clearly satisfied.

The remainder of this section is devoted to proving that $\{P_t\}$ is a Feller continuous and stochastically continuous semigroup of linear operators on $B_b(\mathbb{C}_{\mathbb{I}}^d)$.

Lemma 2.6.2. *For each $\Gamma \in \mathcal{M}_{\mathbb{J}}$, the function $x \mapsto P^x(X^x \in \Gamma)$ is measurable.*

Proof. Let ρ be a metric on $\mathbb{C}_{\mathbb{J}}^d$ inducing the same topology as that of uniform convergence on compact sets. As in the proof of Lemma 2.6.1, we first assume that Γ is open, and we define the continuous and bounded functions $f_n(w) := 1 \wedge n\rho(w, \Gamma^c)$, $w \in \mathbb{C}_{\mathbb{J}}^d$, for $n = 1, 2, \ldots$. By Theorem 2.5.1, for each n, $E^{\cdot}[f_n(X^{\cdot})]$ is continuous and therefore measurable. The sequence of functions $f_n \nearrow 1_\Gamma$ pointwise, so by the monotone convergence theorem, $E^{\cdot}[f_n(X^{\cdot})] \nearrow E^{\cdot}[1_\Gamma(X^{\cdot})] = P^{\cdot}(X^{\cdot} \in \Gamma)$, so that $P^x(X^x \in \Gamma) = \sup_n E^x[f_n(X^x)]$ is measurable in x. Applying a monotone class theorem completes the proof.

□

Corollary 2.6.1. *For each $t \geq 0, \Lambda \in \mathcal{M}_{\mathbb{I}}$, the function $x \mapsto P_t(x, \Lambda)$ is measurable.*

Proof. This follows from the fact that $\{w \in \mathbb{C}^d_{\mathbb{J}} : w_t \in \Lambda\} = p_t^{-1}(\Lambda) \in \mathcal{M}_{\mathbb{J}}$. $\square$

We have therefore shown that condition (i) of Definition 2.6.1 holds.

Lemma 2.6.3 (Chapman-Kolmogorov Equation)**.** *For each* $s, t \geq 0, \Lambda \in \mathcal{M}_{\mathbb{I}}, x \in \mathbb{C}^d_{\mathbb{I}}$,

$$P_{s+t}(x, \Lambda) = \int_{\mathbb{C}^d_{\mathbb{I}}} P_s(y, \Lambda) P_t(x, dy).$$

Proof. Define the canonical space $(\Omega := \mathbb{C}^d_{\mathbb{J}} \times C_0(\mathbb{R}_+, \mathbb{R}^m), \mathcal{F} := \mathcal{B}(\Omega))$. For each $x \in \mathbb{C}^d_{\mathbb{I}}$, we have a well-defined probability measure, $\bar{P}^x := P^x \circ (X^x, W^x)^{-1}$, on $(\Omega, \mathcal{F})$. Define the coordinate mapping process (X, W) on $(\Omega, \mathcal{F})$ by $X(\omega^{(1)}, \omega^{(2)}) = \omega^{(1)} \in \mathbb{C}^d_{\mathbb{J}}$ and $W(\omega^{(1)}, \omega^{(2)}) = \omega^{(2)} \in C_0(\mathbb{R}_+, \mathbb{R}^m)$, and define $\bar{\mathcal{F}}^x := \sigma(\mathcal{F} \cup \mathcal{N}^x)$ and $\bar{\mathcal{F}}^x_t := \hat{\mathcal{F}}^x_{t+}$, where $\mathcal{N}^x$ is the collection of all subsets of $\bar{P}^x$-null sets of Ω, and $\hat{\mathcal{F}}^x_t := \sigma\left(\mathcal{N}^x \cup \sigma(X(s), W(s), s \leq t)\right)$.

We claim that the process $\{X(t), t \geq -\tau\}$ on $(\Omega, \bar{\mathcal{F}}^x, \{W(t), \bar{\mathcal{F}}^x_t, t \geq 0\}, \bar{P}^x)$ is a solution to (1.1) with initial condition $X_0 = x$, $\bar{P}^x$-a.s..

The usual conditions are satisfied by construction, as is the initial condition, and $\{W(t), \bar{\mathcal{F}}^x_t, t \geq 0\}$ is a Brownian motion martingale under $\bar{P}^x$ by the following.

For any $0 \leq s \leq t$ and $\Gamma_s \in \hat{\mathcal{F}}^x_s$, there is a $\bar{\Gamma}_s \in \sigma(X(r), W(r), r \leq s)$ such that $\bar{P}^x(\Gamma_s \Delta \bar{\Gamma}_s) = 0$, where Δ denotes the symmetric difference. Then we have

$$\begin{aligned} E^{\bar{P}^x}[W(t)1_{\Gamma_s}] &= E^{\bar{P}^x}[W(t)1_{\bar{\Gamma}_s}] = E^{\bar{P}^x}[W(t)1_{\{(X,W)\in\bar{\Gamma}_s\}}] \\ &= E^{P^x}[W^x(t)1_{\{(X^x,W^x)\in\bar{\Gamma}_s\}}] = E^{P^x}[W^x(s)1_{\{(X^x,W^x)\in\bar{\Gamma}_s\}}], \end{aligned} \tag{2.39}$$

since $\{W^x(t), \mathcal{F}^x_t, t \geq 0\}$ is a martingale under P^x and $1_{\{(X^x,W^x)\in\bar{\Gamma}_s\}} \in \mathcal{F}^x_s$. Proceeding in reverse, we obtain $E^{\bar{P}^x}[W(t)1_{\Gamma_s}] = E^{\bar{P}^x}[W(s)1_{\Gamma_s}]$, so that $\{W(t), \hat{\mathcal{F}}^x_t, t \geq 0\}$ is a martingale under $\bar{P}^x$. Then Theorem II.2.8 in [38] implies that $\{W(t), \bar{\mathcal{F}}^x_t, t \geq 0\}$ is a martingale under $\bar{P}^x$.

Since for each $t \geq 0$, $E^{P^x}\left[\int_0^t \|\sigma(X_s)\|_2^2 ds\right] < \infty$ by (2.2) and Lemma 2.4.1, we can take a suitable sequence of partitions of $\mathbb{R}_+$ such that approximations to the stochastic integral process $\left\{\int_0^t \sigma(X^x_s) dW^x(s), t \geq 0\right\}$ converge uniformly on compact time intervals P^x-a.s.. Thus, for each $x \in \mathbb{C}^d_{\mathbb{I}}$, there exists a subsequence $\{U^{x,n_{xi}}\}_{i=1}^\infty$ of the sequence of processes

$$\left\{U^{x,n}(t) := \sum_{i=1}^{n^2} \sigma\left(X^x_{\frac{i}{n}\wedge t}\right)\left(W^x\left(\frac{i+1}{n}\wedge t\right) - W^x\left(\frac{i}{n}\wedge t\right)\right), t \geq 0\right\}, \quad n \geq 1,$$

such that for each $T \geq 0$,

$$\lim_{i\to\infty} \sup_{0\leq s\leq T} \left| U^{x,n_{xi}}(s) - \int_0^s \sigma(X_r^x) dW^x(r) \right|_2 = 0, \quad P^x - a.s..$$

If we define the continuous processes

$$\left\{ U^n(t) := \sum_{i=1}^{n^2} \sigma\left(X_{\frac{i}{n}\wedge t}\right) \left(W\left(\frac{i+1}{n}\wedge t\right) - W\left(\frac{i}{n}\wedge t\right)\right), t \geq 0 \right\}, \quad n \geq 1, \tag{2.40}$$

then the triple $(X, W, \{U^n\}_{n=1}^\infty)$ has the same distribution under $\bar{P}^x$ as $(X^x, W^x, \{U^{x,n}\}_{n=1}^\infty)$ has under P^x. Therefore, for each $x \in \mathbb{C}_{\mathbb{I}}^d$, we can define, on the probability space $(\Omega, \bar{\mathcal{F}}^x, \bar{P}^x)$, the Itô integral process

$$\int_0^t \sigma(X_s) dW(s) \quad := \quad 1_{A_x} \lim_{i\to\infty} U^{n_{xi}}(t), \quad t \geq 0, \tag{2.41}$$

where

$$A_x \quad := \quad \left\{ \lim_{i\to\infty} \sup_{j\geq i} \sup_{0\leq s\leq T} |U^{n_{xi}}(s) - U^{n_{xj}}(s)|_2 = 0 \text{ for all } T \geq 0 \right\}. \tag{2.42}$$

Since $\{\bar{\mathcal{F}}_t^x, t \geq 0\}$ satisfies the usual conditions and $\bar{P}^x(A_x) = 1$, $\{\int_0^t \sigma(X_s) dW(s), t \geq 0\}$ is adapted to $\{\bar{\mathcal{F}}_t^x, t \geq 0\}$. Then for each $x \in \mathbb{C}_{\mathbb{I}}^d$,

$$\begin{aligned} & \bar{P}^x\left(X(t) = X(0) + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW(s) + Y(t), \text{ for all } t \geq 0\right) \\ = \quad & \bar{P}^x\left(X(t) = X(0) + \int_0^t b(X_s)ds + \lim_{i\to\infty} U^{n_{xi}}(t) + Y(t), \text{ for all } t \geq 0\right) \\ = \quad & P^x\left(X^x(t) = X^x(0) + \int_0^t b(X_s^x)ds + \lim_{i\to\infty} U^{x,n_{xi}}(t) + Y^x(t), \text{ for all } t \geq 0\right) \\ = \quad & P^x\left(X^x(t) = X^x(0) + \int_0^t b(X_s^x)ds + \int_0^t \sigma(X_s^x)dW^x(s) + Y^x(t), \text{ for all } t \geq 0\right) \\ = \quad & 1. \end{aligned} \tag{2.43}$$

Thus, (1.1) also holds for the process X on $\left(\Omega, \bar{\mathcal{F}}^x, \{\bar{\mathcal{F}}_t^x\}, \bar{P}^x\right)$ with Brownian motion martingale W, and the claim is proved.

Fix $s, t \geq 0$ and $x \in \mathbb{C}_{\mathbb{I}}^d$. Define the Brownian motion $\{W^t(r) := W(t+r) - W(t), r \geq 0\}$, which under $\bar{P}^x$ is a martingale with respect to the filtration $\{\hat{\mathcal{F}}_r^{x,t} := \hat{\mathcal{F}}_{t+r}^x, r \geq 0\}$ (and thus also with respect to $\{\bar{\mathcal{F}}_r^{x,t} := \bar{\mathcal{F}}_{t+r}^x, r \geq 0\}$), and define $X^t(r) = X(t+r)$ for $r \in \mathbb{J}$, so that X^t is a weak solution to (1.1) on the probability space $(\Omega, \bar{\mathcal{F}}^x, \{\bar{\mathcal{F}}_r^{x,t}, r \geq 0\}, \bar{P}^x)$ with Brownian motion martingale W^t and the (random) initial condition $X_0^t = X_t$.

Let $\{\bar{P}^x_\omega(\Gamma), \omega \in \Omega, \Gamma \in \mathcal{F}\}$ be a regular conditional probability distribution for $\bar{P}^x$ on $(\Omega, \mathcal{F})$ given the σ-algebra generated by X_t, $\mathcal{G}_t := \sigma(X(s) : s \in [t-\tau, t])$. Since $(\Omega, \mathcal{F})$ and $(\mathbb{C}^d_{\mathbb{I}}, \mathcal{M}_{\mathbb{I}})$ are countably determined standard spaces (they are Polish (and therefore Lusin) spaces with their Borel σ-algebras), we have that $\bar{P}^x_\omega\left(\left\{\omega' \in \Omega : X^t_0(\omega') = X_t(\omega)\right\}\right) = 1$ for $\bar{P}^x$-a.a $\omega \in \Omega$ (see [21], Theorem 1.3.2, or [39], Theorem II.89.1). Define $\mathcal{N}^{x,\omega}$ to be the set of all subsets of $\bar{P}^x_\omega$-null sets of $\mathcal{F}$, and define $\hat{\mathcal{F}}^{x,t,\omega}_r := \sigma\left(\mathcal{N}^{x,\omega} \cup \sigma(X(s), W(s), s \leq t+r)\right)$ for each $r \geq 0$. We now show that W^t is a Brownian motion martingale under $\bar{P}^x_\omega$ for $\bar{P}^x$-a.a. ω.

For $\bar{P}^x$-a.a. $\omega \in \Omega$, $\left\{W^t(r), \hat{\mathcal{F}}^{x,t,\omega}_r, r \geq 0\right\}$ is a martingale under $\bar{P}^x_\omega$. Indeed, fix $r_2 \geq r_1 \geq 0$. For any $\Gamma_1 \in \hat{\mathcal{F}}^{x,t,\omega}_{r_1}$ and $\Gamma_2 \in \mathcal{G}_t$,

$$
\begin{aligned}
\int_{\Gamma_2} E^{\bar{P}^x_\omega}&\left[\left(E^{\bar{P}^x_\omega}\left[W^t(r_2)\middle|\hat{\mathcal{F}}^{x,t,\omega}_{r_1}\right] - W^t(r_1)\right)1_{\Gamma_1}\right]\bar{P}^x(d\omega) \\
&= \int_{\Gamma_2}\int_{\Gamma_1}\left(E^{\bar{P}^x_\omega}\left[W^t(r_2)\middle|\hat{\mathcal{F}}^{x,t,\omega}_{r_1}\right](\omega') - W^t(r_1)(\omega')\right)\bar{P}^x_\omega(d\omega')\bar{P}^x(d\omega) \\
&= \int_{\Gamma_2}\int_{\Gamma_1}\left(W^t(r_2)(\omega') - W^t(r_1)(\omega')\right)\bar{P}^x_\omega(d\omega')\bar{P}^x(d\omega) \\
&= \int_{\Gamma_2} E^{\bar{P}^x_\omega}\left[1_{\Gamma_1}\left(W^t(r_2) - W^t(r_1)\right)\right]\bar{P}^x(d\omega) \\
&= \int_{\Gamma_2} E^{\bar{P}^x}\left[1_{\Gamma_1}\left(W^t(r_2) - W^t(r_1)\right)\middle|\mathcal{G}_t\right](\omega)\bar{P}^x(d\omega) \\
&= \int_{\Gamma_2} 1_{\Gamma_1}(\omega)\left(W^t(r_2)(\omega) - W^t(r_1)(\omega)\right)\bar{P}^x(d\omega) \\
&= \int_{\Omega} 1_{\Gamma_2\cap\Gamma_1}(\omega)\left(W^t(r_2)(\omega) - W^t(r_1)(\omega)\right)\bar{P}^x(d\omega) \\
&= 0,
\end{aligned}
\tag{2.44}
$$

since $\Gamma_1 \cap \Gamma_2 \in \hat{\mathcal{F}}^{x,t,\omega}_{r_1}$, and $\left\{W^t(r), \hat{\mathcal{F}}^{x,t,\omega}_r, r \geq 0\right\}$ is a martingale with respect to $\bar{P}^x$. By definition, the function $\omega \mapsto \bar{P}^x_\omega(\Gamma)$ is $\mathcal{G}_t$-measurable for each $\Gamma \in \mathcal{F}$. The integral comparison theorem then implies that for $\bar{P}^x$-a.a. $\omega \in \Omega$, $E^{\bar{P}^x_\omega}[W^t(r_2)|\hat{\mathcal{F}}^{x,t,\omega}_{r_1}] = W^t(r_1)$, $\bar{P}^x_\omega$-a.s.; i.e., $\left\{W^t(r), \hat{\mathcal{F}}^{x,t,\omega}_r, r \geq 0\right\}$ is a martingale with respect to $\bar{P}^x_\omega$ for $\bar{P}^x$-a.a. $\omega \in \Omega$. Again, Theorem II.2.8 in [38] implies that $\left\{W^t(r), \bar{\mathcal{F}}^{x,t,\omega}_r := \hat{\mathcal{F}}^{x,t,\omega}_{r+}, r \geq 0\right\}$ is also a martingale under $\bar{P}^x_\omega$ for $\bar{P}^x$-a.a. $\omega \in \Omega$.

Let $\Gamma_0 \in \mathcal{G}_t$, and for each $i = 1, \ldots, m$, let $\Gamma_i \in \sigma\left(\left(W^t(s)\right)^i : s \geq 0\right)$, the σ-algebra generated by $(W^t)^i$. Then,

$$\begin{aligned}
\int_{\Gamma_0} & \left(\bar{P}^x_\omega \left(\bigcap_{i=1}^m \Gamma_i \right) - \prod_{i=1}^m \bar{P}^x_\omega(\Gamma_i) \right) \bar{P}^x(d\omega) \\
&= \int_{\Gamma_0} \left(\bar{P}^x \left(\bigcap_{i=1}^m \Gamma_i \middle| \mathcal{G}_t \right)(\omega) - \prod_{i=1}^m \bar{P}^x(\Gamma_i|\mathcal{G}_t)(\omega) \right) \bar{P}^x(d\omega) \\
&= \int_{\Gamma_0} \left(\bar{P}^x \left(\bigcap_{i=1}^m \Gamma_i \right) - \prod_{i=1}^m \bar{P}^x(\Gamma_i) \right) \bar{P}^x(d\omega) \\
&= \bar{P}^x(\Gamma_0) \left(\bar{P}^x \left(\bigcap_{i=1}^m \Gamma_i \right) - \prod_{i=1}^m \bar{P}^x(\Gamma_i) \right) \\
&= 0, \qquad (2.45)
\end{aligned}$$

since W^t is a Brownian motion under $\bar{P}^x$. We used the fact that under $\bar{P}^x$, Γ_i is independent of $\hat{\mathcal{F}}_0^{x,t,\omega}$ for each i since W^t is (and $X_t \in \hat{\mathcal{F}}_0^{x,t,\omega}$). Thus $\{W^t(s), s \geq 0\}$ has independent coordinates under $\bar{P}^x_\omega$ for $\bar{P}^x$-a.a. ω.

Fix $i \in \{1, \ldots, m\}$, and let $0 \leq r_0 < r_1 < \cdots < r_n < \infty$ and $v \in \mathbb{R}^n$, and set $u^2 := \sum_{j=1}^n v_j^2 (r_j - r_{j-1})$. Then for any $\Gamma \in \mathcal{G}_t$,

$$\begin{aligned}
\int_\Gamma & \left(E^{\bar{P}^x_\omega} \left[\exp\left(\sqrt{-1} \sum_{j=1}^n v^j \left((W^t)^i(r_j) - (W^t)^i(r_{j-1}) \right) \right) \right] - \exp\left(\frac{-u^2}{2} \right) \right) \bar{P}^x(d\omega) \\
&= \int_\Gamma E^{\bar{P}^x} \left[\exp\left(\sqrt{-1} \sum_{j=1}^n v^j \left((W^t)^i(r_j) - (W^t)^i(r_{j-1}) \right) \right) \middle| \mathcal{G}_t \right] (\omega) \bar{P}^x(d\omega) \\
&\quad - \bar{P}^x(\Gamma) \exp\left(\frac{-u^2}{2} \right) \\
&= E^{\bar{P}^x} \left[1_\Gamma \exp\left(\sqrt{-1} \sum_{j=1}^n v^j \left((W^t)^i(r_j) - (W^t)^i(r_{j-1}) \right) \right) \right] \\
&\quad - \bar{P}^x(\Gamma) \exp\left(\frac{-u^2}{2} \right) \\
&= \bar{P}^x(\Gamma) \left(E^{\bar{P}^x} \left[\exp\left(\sqrt{-1} \sum_{j=1}^n v^j \left((W^t)^i(r_j) - (W^t)^i(r_{j-1}) \right) \right) \right] - \exp\left(\frac{-u^2}{2} \right) \right) \\
&= 0, \qquad (2.46)
\end{aligned}$$

and thus for $\bar{P}^x$-a.a. $\omega \in \Omega$, under $\bar{P}^x_\omega$, $((W^t)^i(r_1) - (W^t)^i(r_0), \ldots, (W^t)^i(r_n) - (W^t)^i(r_{n-1}))$ has a multivariate normal distribution with mean zero and covariances

$$E^{\bar{P}^x_\omega} \left[\left((W^t)^i(r_j) - (W^t)^i(r_{j-1}) \right) \left((W^t)^i(r_k) - (W^t)^i(r_{k-1}) \right) \right] = \delta_{jk}(r_j - r_{j-1}).$$

In conclusion, there is a set $\Omega^{x,t} \in \mathcal{F}$ such that $\bar{P}^x(\Omega^{x,t}) = 1$, and for each $\omega \in \Omega^{x,t}$, $(W^t(r), \bar{\mathcal{F}}_r^{x,t,\omega}, r \geq 0)$ is a Brownian motion martingale on the probability space $(\Omega, \bar{\mathcal{F}}^{x,\omega}, \bar{P}_\omega^x)$, where $\bar{\mathcal{F}}^{x,\omega}$ is the $\bar{P}_\omega^x$-completion of $\mathcal{F}$.

For each $x \in \mathbb{C}_{\mathbb{I}}^d$, define the process

$$\int_0^r \sigma(X_u^t) dW^t(u) \quad := \quad 1_{A_x} \lim_{i\to\infty} \left(U^{n_{xi}}(t+r) - U^{n_{xi}}(t)\right), \quad r \geq 0, \tag{2.47}$$

where A_x is defined as in line (2.42), and U^n is defined as in line (2.40). This process can be defined on the probability space $(\Omega, \bar{\mathcal{F}}^x, \bar{P}^x)$, or on $(\Omega, \bar{\mathcal{F}}^{x,\omega}, \bar{P}_\omega^x)$ for any $\omega \in \Omega^{x,t}$. Then as above, for $\bar{P}^x$-a.a. $\omega \in \Omega^{x,t}$, $\bar{P}_\omega^x(A_x) = 1$ and $\{\int_0^r \sigma(X_u^t) dW^t(u), r \geq 0\}$ is adapted to $\{\bar{\mathcal{F}}_r^{x,t,\omega}, r \geq 0\}$.

If we define $Y^t(r) := Y(t+r) - Y(t)$, $r \geq 0$, then for $\bar{P}^x$-a.a. $\omega \in \Omega$, (X^t, W^t, Y^t) on $(\Omega, \bar{\mathcal{F}}^{x,\omega}, \{\bar{\mathcal{F}}_r^{x,t,\omega}, r \geq 0\}, \bar{P}_\omega^x)$ satisfies (i)-(iv) of Definition 2.1.1. For each $\Gamma \in \mathcal{G}_t$,

$$\begin{aligned}
&\int_\Gamma \bar{P}_\omega^x \left(X^t(r) = X^t(0) + \int_0^r b(X_u^t) du + \int_0^r \sigma(X_u^t) dW^t(u) + Y^t(r), \text{ for all } r \geq 0 \right) \bar{P}^x(d\omega) \\
&= \int_\Gamma \bar{P}^x \left(X^t(r) = X^t(0) + \int_0^r b(X_u^t) du + \int_0^r \sigma(X_u^t) dW^t(u) + Y^t(r), r \geq 0 \,\middle|\, \mathcal{G}_t \right)(\omega) \bar{P}^x(d\omega) \\
&= \bar{P}^x \left(\Gamma \cap \left\{ X^t(r) = X^t(0) + \int_0^r b(X_u^t) du + \int_0^r \sigma(X_u^t) dW^t(u) + Y^t(r), r \geq 0 \right\} \right) \\
&= \bar{P}^x \left(\Gamma \cap \left\{ X(t+r) = X(t) + \int_t^{t+r} b(X_u) du + \int_t^{t+r} \sigma(X_u) dW(u) + Y(t+r) - Y(t), r \geq 0 \right\} \right) \\
&= \bar{P}^x(\Gamma).
\end{aligned} \tag{2.48}$$

The last equality follows from equality (2.43). Therefore, for $\bar{P}^x$-a.a. $\omega \in \Omega$, X^t solves (1.1) with initial condition $X_0^t = X_t(\omega)$ on the probability space $(\Omega, \bar{\mathcal{F}}^{x,\omega}, \{\bar{\mathcal{F}}_r^{x,t,\omega}, r \geq 0\}, \bar{P}_\omega^x)$ with driving Brownian motion martingale $\{W^t(r), \bar{\mathcal{F}}_r^{x,t,\omega}, r \geq 0\}$.

By uniqueness in law, for $\bar{P}^x$-a.a. $\omega \in \Omega$, $\bar{P}_\omega^x(X^t \in \Gamma) = \bar{P}^{X_t(\omega)}(X \in \Gamma)$ for all $\Gamma \in \mathcal{M}_{\mathbb{J}}$. Corollary 2.6.1 and the measurability of $\omega \mapsto X_t(\omega)$ imply that the map $\omega \mapsto$

$P_s(X_t(\omega),\Lambda) = \bar{P}^{X_t(\omega)}(X_s \in \Lambda)$ is $\mathcal{G}_t$-measurable for each $\Lambda \in \mathcal{M}_{\mathbb{I}}$. Thus,

$$\begin{aligned} P_{s+t}(x,\Lambda) &= \bar{P}^x(X_{s+t} \in \Lambda) \\ &= \bar{P}^x(X_s^t \in \Lambda) \\ &= \int_\Omega \bar{P}^x_\omega(X_s^t \in \Lambda)\bar{P}^x(d\omega) \\ &= \int_\Omega \bar{P}^{X_t(\omega)}(X_s \in \Lambda)\bar{P}^x(d\omega) \\ &= \int_{\mathbb{C}^d_{\mathbb{I}}} \bar{P}^y(X_s \in \Lambda)P_t(x,dy) \\ &= \int_{\mathbb{C}^d_{\mathbb{I}}} P_s(y,\Lambda)P_t(x,dy). \end{aligned} \tag{2.49}$$

□

We have therefore shown that each condition (i)-(iv) of Definition 2.6.1 holds, and therefore $\{P_t(x,\Lambda), t \geq 0, x \in \mathbb{C}^d_{\mathbb{I}}, \Lambda \in \mathcal{M}_{\mathbb{I}}\}$ is a family of Markovian transition functions, which then generates a semigroup $\{P_t, t \geq 0\}$ of linear operators on $B_b(\mathbb{C}^d_{\mathbb{I}})$. Corollary 2.5.2 implies that this semigroup is Feller continuous.

Definition 2.6.3. Given a metric space $\mathbb{E}$, a Markovian semigroup $\{P_t, t \geq 0\}$ of linear operators on $B_b(\mathbb{E})$ is called *Feller continuous* if for any $f \in C_b(\mathbb{E})$ and $t \geq 0$, $P_t f(x)$ is a continuous function of x.

Corollary 2.6.2. *The semigroup induced by the SDDER (1.1) is Feller continuous.*

Remark. The argument in Theorem 1 of Section 2.3 of [8] can be used to show that any Feller continuous Markov process with continuous paths is also a strong Markov process.

Lemma 2.6.4. *The semigroup* $\{P_t, t \geq 0\}$ *induced by the SDDER (1.1) is stochastically continuous.*

Proof. By the definition of a solution, the solution X^x to (1.1) with initial condition $x \in \mathbb{C}^d_{\mathbb{I}}$ is continuous. For each $t \geq 0$, the function $p_t : \mathbb{C}^d_{\mathbb{J}} \to \mathbb{C}^d_{\mathbb{I}}$ (recall that $p_t(f) := f_t$) is also continuous (see [20], Lemma 4.2). Therefore the function $t \mapsto X^x_t(\omega)$ is continuous for each ω. So for any $f \in C_b(\mathbb{C}^d_{\mathbb{I}})$, the function $t \mapsto f(X_t(\omega))$ is continuous for each ω. It follows by the bounded convergence theorem that $\lim_{t\to 0+}(P_t f)(x) = \lim_{t \to 0+} E^x[f(X^x_t)] = f(x)$. The result follows by Proposition 2.6.1.

□

Chapter 3

Stationary Distributions

This chapter is devoted to defining stationary distributions for the SDDER (1.1), and to exhibiting a technique often used to prove that a stationary distribution exists. This involves precompactness of a sequence of averaging measures. A sufficient condition for this precompactness is provided in Theorem 3.3.1. More specific conditions on the coefficients b and σ are later given in Chapter 4. Assumptions 2.1.1 and 2.1.2 are assumed throughout this chapter, and $\{P_t(x,\Lambda) : x \in \mathbb{C}^d_{\mathbb{I}}, \Lambda \in \mathcal{M}_{\mathbb{I}}, t \geq 0\}$ is the family of Markovian transition functions induced by the SDDER (1.1).

3.1 Definition of a Stationary Distribution

For each $t \geq 0$ and probability measure μ on $(\mathbb{C}^d_{\mathbb{I}}, \mathcal{M}_{\mathbb{I}})$, consider the probability measure μP_t on $(\mathbb{C}^d_{\mathbb{I}}, \mathcal{M}_{\mathbb{I}})$ defined by

$$(\mu P_t)(\Lambda) = \int_{\mathbb{C}^d_{\mathbb{I}}} P_t(x,\Lambda)\mu(dx), \text{ for } \Lambda \in \mathcal{M}_{\mathbb{I}}.$$

Corollary 2.6.1 shows the required measurability for this integral to be meaningful.

Definition 3.1.1. A *stationary distribution for (1.1)* is a probability measure π on $(\mathbb{C}^d_{\mathbb{I}}, \mathcal{M}_{\mathbb{I}})$ such that $(\pi P_t)(\Lambda) = \pi(\Lambda)$ for all $t \geq 0$ and $\Lambda \in \mathcal{M}_{\mathbb{I}}$.

3.2 Krylov-Bogulyubov Measures

A common method for showing the *existence* of a stationary distribution for a Markov process is to exhibit a limit point of a sequence of Krylov-Bogulyubov measures ([3, 11, 20, 37]).

In light of that, given $x_o \in \mathbb{C}_{\mathbb{I}}^d$ and $T > 0$, we define the Krylov-Bogulyubov probability measure $Q_T^{x_o}$ on $(\mathbb{C}_{\mathbb{I}}^d, \mathcal{M}_{\mathbb{I}})$ by

$$Q_T^{x_o}(\Lambda) \quad := \quad \frac{1}{T}\int_0^T P_u(x_o, \Lambda)du \tag{3.1}$$

for all $\Lambda \in \mathcal{M}_{\mathbb{I}}$.

Remark. The integral in expression (3.1) is well-defined since the function $u \mapsto P_u(x_o, \Lambda)$ is measurable by Lemmas 2.6.1 and 2.6.4, and it is bounded by one.

The following proposition justifies a “Fubini theorem” (equality (3.3)), which will be used in the proof of Theorem 3.2.1.

Proposition 3.2.1. *For each $f \in B_b(\mathbb{C}_{\mathbb{I}}^d)$ and probability measure μ on $\mathbb{C}_{\mathbb{I}}^d$,*

$$\int_{\mathbb{C}_{\mathbb{I}}^d} f(x)(\mu P_t)(dx) \quad = \quad \int_{\mathbb{C}_{\mathbb{I}}^d} (P_t f)(y)\mu(dy). \tag{3.2}$$

Proof. If $f = 1_\Lambda$ for some $\Lambda \in \mathcal{M}_{\mathbb{I}}$, then (3.2) follows from the definitions of μP_t and $P_t f$. By the linearity of the integral, (3.2) also holds for simple f. The result now follows by invoking a monotone class theorem.

□

Therefore, for each $f \in B_b(\mathbb{C}_{\mathbb{I}}^d)$ and probability measure μ on $\mathbb{C}_{\mathbb{I}}^d$,

$$\begin{aligned}\int_{x\in\mathbb{C}_{\mathbb{I}}^d}\int_{y\in\mathbb{C}_{\mathbb{I}}^d} f(x)P_t(y,dx)\mu(dy) &= \int_{\mathbb{C}_{\mathbb{I}}^d} f(x)(\mu P_t)(dx) \\ &= \int_{\mathbb{C}_{\mathbb{I}}^d} (P_t f)(y)\mu(dy) \\ &= \int_{y\in\mathbb{C}_{\mathbb{I}}^d}\int_{x\in\mathbb{C}_{\mathbb{I}}^d} f(x)P_t(y,dx)\mu(dy). \end{aligned} \tag{3.3}$$

Theorem 3.2.1. *Assume that for some $x_o \in \mathbb{C}_{\mathbb{I}}^d$ and some sequence $\{T_n\}_{n=1}^\infty$ such that $T_n \nearrow \infty$ as $n \to \infty$, the sequence $\{Q_{T_n}^{x_o}\}_{n=1}^\infty$ converges weakly as $n \to \infty$ to some probability measure π^{x_o} on $(\mathbb{C}_{\mathbb{I}}^d, \mathcal{M}_{\mathbb{I}})$. Then π^{x_o} is a stationary distribution for the SDDER (1.1).*

Proof. We use a standard argument.

By Theorem 1.2 of [5], it suffices to show that for any bounded and continuous real-valued function f on $\mathbb{C}_{\mathbb{I}}^d$,

$$\int_{\mathbb{C}_{\mathbb{I}}^d} f(x)(\pi^{x_o}P_t)(dx) \quad = \quad \int_{\mathbb{C}_{\mathbb{I}}^d} f(x)\pi^{x_o}(dx) \text{ for all } t \geq 0. \tag{3.4}$$

So fix $t \geq 0$, and let $f : \mathbb{C}^d_{\mathbb{I}} \to \mathbb{R}$ be bounded and continuous. Then,

$$\begin{aligned}
\int_{\mathbb{C}^d_{\mathbb{I}}} f(x)(\pi^{x_o} P_t)(dx) &= \int_{\mathbb{C}^d_{\mathbb{I}}} (P_t f)(y)\pi^{x_o}(dy) \\
&= \lim_{n\to\infty} \int_{\mathbb{C}^d_{\mathbb{I}}} (P_t f)(y) Q^{x_o}_{T_n}(dy) \\
&= \lim_{n\to\infty} \frac{1}{T_n} \int_0^{T_n} \int_{\mathbb{C}^d_{\mathbb{I}}} (P_t f)(y) P_u\left(x_o, dy\right) du \\
&= \lim_{n\to\infty} \frac{1}{T_n} \int_0^{T_n} \int_{\mathbb{C}^d_{\mathbb{I}}} \int_{\mathbb{C}^d_{\mathbb{I}}} f(x) P_t\left(y, dx\right) P_u\left(x_o, dy\right) du \\
&= \lim_{n\to\infty} \frac{1}{T_n} \int_0^{T_n} \int_{\mathbb{C}^d_{\mathbb{I}}} f(x) P_{t+u}\left(x_o, dx\right) du \\
&= \lim_{n\to\infty} \frac{1}{T_n} \int_{v=t}^{t+T_n} \int_{\mathbb{C}^d_{\mathbb{I}}} f(x) P_v\left(x_o, dx\right) dv \\
&= \lim_{n\to\infty} \frac{1}{T_n} \int_0^{T_n} \int_{\mathbb{C}^d_{\mathbb{I}}} f(x) P_v\left(x_o, dx\right) dv \\
&\quad + \lim_{n\to\infty} \frac{1}{T_n} \int_{T_n}^{t+T_n} \int_{\mathbb{C}^d_{\mathbb{I}}} f(x) P_v\left(x_o, dx\right) dv \\
&\quad - \lim_{n\to\infty} \frac{1}{T_n} \int_0^{t} \int_{\mathbb{C}^d_{\mathbb{I}}} f(x) P_v\left(x_o, dx\right) dv \\
&= \lim_{n\to\infty} \int_{\mathbb{C}^d_{\mathbb{I}}} f(x) Q^{x_o}_{T_n}(dx) + 0 - 0 \\
&= \int_{\mathbb{C}^d_{\mathbb{I}}} f(x)\pi^{x_o}(dx),
\end{aligned}$$

thus proving (3.4). Here, the second equality follows since by the the Feller continuity of the semigroup (Corollary 2.6.2), $(P_t f)(\cdot)$ is a continuous function. Equality (3.3) and the Markov property (2.36) were used for the fifth equality.

□

3.3 Tightness Criterion for Krylov-Bogulyubov Measures

For each $x_o \in \mathbb{C}^d_{\mathbb{I}}$, let X^{x_o} together with Brownian motion W^{x_o} define a solution to (1.1) with initial condition $X^{x_o} \equiv x_o$ on some filtered probability space $(\Omega^{x_o}, \mathcal{F}^{x_o}, \{\mathcal{F}^{x_o}_t\}, P^{x_o})$.

The following theorem provides conditions guaranteeing tightness of the Krylov-Bogulyubov measures, and its proof uses Lemmas 2.5.1 and 2.5.2. Kushner ([25]) shows tightness of these measures under the assumption that b and σ are bounded. The linear growth conditions

(2.1) and (2.2) allow us to bound the oscillations of X^{x_o} by moments of $\|X_t^{x_o}\|_\infty$, assuming the latter are uniformly bounded.

Theorem 3.3.1. *Fix $x_o \in \mathbb{C}_{\mathbb{I}}^d$ and assume that $\sup_{t\geq 0} E^{x_o}[\|X_t^{x_o}\|_2^p] < \infty$ for some $p > 0$. Then for any sequence $\{T_n\}_{n=1}^\infty$ in $(0,\infty)$ increasing to ∞, the sequence $\{Q_{T_n}^{x_o}\}_{n=1}^\infty$ of Krylov-Bogulyubov measures is tight.*

Proof. Fix $\varepsilon, \lambda > 0$.

By Markov's inequality, for any $a > 0$,

$$\begin{aligned} Q_T^{x_o}\left(x \in \mathbb{C}_{\mathbb{I}}^d : |x(0)|_\infty > a\right) &= \frac{1}{T}\int_0^T P^{x_o}\left(|X^{x_o}(s)|_\infty > a\right) ds \\ &\leq \frac{1}{T}\int_0^T \frac{1}{a^p} E^{x_o}\left[|X^{x_o}(s)|_\infty^p\right] ds \\ &\leq \frac{1}{a^p}\sup_{t\geq 0} E^{x_o}\left[|X^{x_o}(t)|_\infty^p\right] \\ &\leq \frac{1}{a^p}\sup_{t\geq 0} E^{x_o}\left[|X^{x_o}(t)|_2^p\right]. \end{aligned}$$

The last term tends to zero as $a \to \infty$, independently of T. This establishes that condition (i) of Proposition 2.5.1 holds for $P_n = Q_{T_n}^{x_o}$, $n \geq 1$.

Fix $u \geq \tau$. Since $u - \tau \geq 0$, $w_{[(u-\tau)\wedge 0, u\wedge 0]}(X^{x_o}, \delta) = 0$, so Lemma 2.5.1 implies that for any $\delta > 0$ we have

$$\begin{aligned} P^{x_o}\left(w_{\mathbb{I}}(X_u^{x_o}, \delta) \geq \lambda\right) &= P^{x_o}\left(w_{[u-\tau,u]}(X^{x_o}, \delta) \geq \lambda\right) \\ &\leq P^{x_o}\left(\delta\left(C_1 + C_2\|X^{x_o}\|_{[u-2\tau,u],2}\right) \geq \frac{\lambda}{2}\right) \\ &\quad + P^{x_o}\left(\sup_{\substack{u-\tau\leq s<t\leq u \\ |s-t|<\delta}} \left|\int_s^t \sigma(X_r^{x_o}) dW^{x_o}(r)\right|_\infty \geq \frac{\lambda}{2}\right). \end{aligned} \tag{3.5}$$

By using Markov's inequality, the assumption that $\sup_{t\geq 0} E^{x_o}[\|X_t^{x_o}\|_2^p] < \infty$ implies that

$$\begin{aligned} \sup_{t\geq\tau} P^{x_o}\left(\|X^{x_o}\|_{[t-2\tau,t],2} > a\right) &\leq \frac{1}{a^p}\sup_{t\geq\tau} E^{x_o}\left[\|X^{x_o}\|_{[t-2\tau,t],2}^p\right] \\ &\leq \frac{1}{a^p}\sup_{t\geq\tau}\left(E^{x_o}\left[\|X^{x_o}\|_{[t-\tau,t],2}^p\right] + E^{x_o}\left[\|X^{x_o}\|_{[t-2\tau,t-\tau],2}^p\right]\right) \\ &\leq \frac{2}{a^p}\sup_{t\geq 0} E^{x_o}\left[\|X^{x_o}\|_{[t-\tau,t],2}^p\right] \\ &= \frac{2}{a^p}\sup_{t\geq 0} E^{x_o}\left[\|X_t^{x_o}\|_2^p\right], \end{aligned}$$

which approaches zero as $a \to \infty$, so that $\{\|X^{x_o}\|_{[t-2\tau,t],2}\,, t \geq \tau\}$ is stochastically bounded.

For $0 < \delta < \frac{\lambda}{4C_1}$, we have from inequality (B.7),

$$P^{x_o}\left(\delta\left(C_1 + C_2\|X^{x_o}\|_{[u-2\tau,u],2}\right) \geq \frac{\lambda}{2}\right) = P^{x_o}\left(\|X^{x_o}\|_{[u-2\tau,u],2} \geq \frac{\lambda - 2C_1\delta}{2C_2\delta}\right).$$

Then stochastic boundedness from above implies that there is a $\delta^{(1)}_{\varepsilon,\lambda} \in (0, \frac{\lambda}{2C_1})$ such that

$$\sup_{u\geq\tau} P^{x_o}\left(\delta\left(C_1 + C_2\|X^{x_o}\|_{[u-2\tau,u],2}\right) \geq \frac{\lambda}{2}\right) < \frac{\varepsilon}{4}$$

for all $\delta \in (0, \delta^{(1)}_{\varepsilon,\lambda}]$.

Lemma 2.5.2 implies that there is a $\delta^{(2)}_{\varepsilon,\lambda} > 0$ such that whenever $\delta \in (0, \delta^{(2)}_{\varepsilon,\lambda}]$, we have

$$\sup_{u\geq\tau} P^{x_o}\left(\sup_{\substack{u-\tau\leq s<t\leq u \\ |s-t|<\delta}} \left|\int_s^t \sigma(X^{x_o}_r)dW^{x_o}(r)\right|_\infty \geq \frac{\lambda}{2}\right) < \frac{\varepsilon}{4}. \tag{3.6}$$

It follows that

$$P^{x_o}\left(w_{\mathbb{I}}(X^{x_o}_u, \delta) \geq \lambda\right) < \frac{\varepsilon}{2}$$

whenever $0 < \delta < \delta_{\varepsilon,\lambda} := \delta^{(1)}_{\varepsilon,\lambda} \wedge \delta^{(2)}_{\varepsilon,\lambda}$ and $u \geq \tau$.

For any $T \geq \frac{2\tau}{\varepsilon} \vee \tau$ and $0 < \delta < \delta_{\varepsilon,\lambda}$, on combining the above we have

$$\begin{aligned}
Q^{x_o}_T\left(x \in \mathbb{C}^d_{\mathbb{I}} : w_{\mathbb{I}}(x,\delta) \geq \lambda\right) &= \frac{1}{T}\int_0^T P^{x_o}\left(w_{\mathbb{I}}(X^{x_o}_u, \delta) \geq \lambda\right) du \\
&= \frac{1}{T}\int_0^\tau P^{x_o}\left(w_{\mathbb{I}}(X^{x_o}_u, \delta) \geq \lambda\right) du \\
&\quad + \frac{1}{T}\int_\tau^T P^{x_o}\left(w_{\mathbb{I}}(X^{x_o}_u, \delta) \geq \lambda\right) du \\
&\leq \frac{\tau}{T} + \frac{1}{T}\int_\tau^T \frac{\varepsilon}{2} du \\
&\leq \frac{\varepsilon}{2} + \frac{T-\tau}{T}\frac{\varepsilon}{2} \\
&\leq \varepsilon.
\end{aligned} \tag{3.7}$$

It follows that condition (ii) of Proposition 2.5.1 holds, where $P_n = Q^{x_o}_{T_n}$, $n \geq 1$, for any $T_n \nearrow \infty$. Hence, $\{Q^{x_o}_{T_n}\}_{n=1}^\infty$ is tight.

□

Remark. Obvious modifications of the above proof yield the same result in the case that $\sup_{t\geq 0} E^{x_o}\left[\|X^{x_o}_t\|_2^p\right] < \infty$ is replaced by $\sup_{t\geq 0} E^{x_o}\left[f\left(\|X^{x_o}_t\|_2\right)\right] < \infty$, where $f : \mathbb{R}_+ \to \mathbb{R}_+$ is any strictly increasing function such that $\lim_{t\to\infty} f(t) = \infty$; e.g., a sufficient condition for tightness is that $\sup_{t\geq 0} E^{x_o}\left[(\log(\|X^{x_o}_t\|_2))^+\right] < \infty$.

3.4 Sufficient Conditions for Existence of a Stationary Distribution

We combine Theorems 3.2.1 and 3.3.1 to obtain the following corollary.

Corollary 3.4.1. *Assume that Assumptions 2.1.1 and 2.1.2 hold, and that there exist* $x_o \in \mathbb{C}^d_{\mathbb{I}}$ *and* $p > 0$ *such that* $\sup_{t \geq 0} E^{x_o}[\|X^{x_o}_t\|^p_2] < \infty$*. Then there exists a stationary distribution for the SDDER (1.1).*

Proof. Theorem 3.3.1 implies that for each sequence $\{T_n\}_{n=1}^{\infty}$ such that $T_n \nearrow \infty$ as $n \to \infty$, the sequence $\{Q^{x_o}_{T_n}\}_{n=1}^{\infty}$ is tight. Therefore, Prohorov's theorem implies that there is a subsequence $\{T_{n_k}\}_{k=1}^{\infty}$ such that $\{Q^{x_o}_{T_{n_k}}\}_{k=1}^{\infty}$ converges weakly as $k \to \infty$ to some probability measure π^{x_o} on $(\mathbb{C}^d_{\mathbb{I}}, \mathcal{M}_{\mathbb{I}})$. Theorem 3.2.1 then implies that π^{x_o} is a stationary distribution for the SDDER (1.1).

□

Thus, to ensure existence of a stationary distribution, we need only have a uniform (in $t \geq 0$) moment bound on $\|X_t\|_2^2$, and Chapter 4 has examples of different sets of assumptions on b and σ that are sufficient to guarantee such uniform moment bounds. One thing to notice about each of these sets of assumptions is that beyond the basic Assumptions 2.1.1 and 2.1.2, there are no restrictions on the coefficients b and σ on the set $\{x \in \mathbb{C}^d_{\mathbb{I}} : |x(0)|_2 < M\}$, where M is arbitrarily large. This freedom is possible because the uniform bound on the moments of $\|X_t\|_2^2$ just needs to be finite.

Chapter 4

Moment Bounds

Throughout this chapter, we assume that X is a solution of the SDDER (1.1) with a possibly random initial condition X_0. We give sufficient conditions on b and σ that yield moment bounds on $\|X_t\|_2$ uniformly in $t \geq 0$. Section 4.1 introduces an important auxiliary process, the overshoot process, and develops preliminary results on the "positive oscillation" of a path that are used for obtaining such bounds. Sections 4.2 and 4.3 develop moment bounds under the assumption that each component of b has a term providing a push in the negative direction (towards zero) on the set $\{x \in \mathbb{C}^d_{\mathbb{I}} : |x(0)|_2 \geq M\}$ for some $M > 0$. Sections 4.2 and 4.3 are distinguished by differences in the assumptions made on the restoring force and on the additional terms composing b and the assumptions on σ. Section 4.2 allows the additional terms to grow (in a sufficiently controlled manner) but requires the negative push at time t to be at least proportional to a value lying in the range of the segment X_t, all on $\{x \in \mathbb{C}^d_{\mathbb{I}} : |x(0)|_2 \geq M\}$. In Section 4.3, $\|\sigma\|_2$ and the components of b are bounded above and the negative push is strictly negative (uniformly), all on $\{x \in \mathbb{C}^d_{\mathbb{I}} : |x(0)|_2 \geq M\}$. This section also has stronger conclusions in the form of exponential moment bounds.

4.1 Overshoot and Positive Oscillation

In this section, we introduce two concepts that will be used frequently in what follows. The overshoot process will enable us to make use of conditions on b and σ that only hold on $\{x \in \mathbb{C}^d_{\mathbb{I}} : |x(0)|_\infty \geq M\}$ for some $M > 0$. The concept of positive oscillation is a convenient tool for studying the *increase* of each component of a reflected process.

4.1.1 Overshoot Process

Let $\tilde{M} \geq 0$. For each $i \in \{1, \ldots, d\}$, we ignore the dynamics of X^i when X_t is in the set $\{x \in \mathbb{C}^d_{\mathbb{I}} : x^i(0) \leq \tilde{M}\}$ by truncating X^i as follows. For each i, define the overshoot, Z^i, of X^i by

$$Z^i(t) \quad := \quad \left(X^i(t) - \tilde{M}\right)^+, \text{ for } t \geq -\tau. \tag{4.1}$$

Part (iv) of Definition 2.1.1 implies that $\int_0^t 1_{\{X^i(s) > \tilde{M}\}} dY^i(s) = 0$ for each $t \geq 0$ and $i = 1, \ldots, d$. Thus, by Tanaka's formula for continuous semimartingales (see, e.g., Theorem 1.2 of Chapter VI in [38]), we have that P-a.s., for all $t \geq 0$,

$$dZ^i(t) \quad = \quad 1_{\{X^i(t) > \tilde{M}\}} b^i(X_t)dt + 1_{\{X^i(t) > \tilde{M}\}} \sigma^i(X_t)dW(t) + dL^i(t), \tag{4.2}$$

where L^i is a constant multiple of the local time of X^i at $\tilde{M}$, which can increase only when $X^i(\cdot)$ is at $\tilde{M}$, and hence only when $Z^i(\cdot)$ is zero (see, e.g., Proposition VI.1.3 in [38]).

The following application of Itô's formula will be useful in Sections 4.2 and 4.3. For each $t \geq 0$,

$$\begin{aligned} d(Z^i(t))^2 &= 2Z^i(t)dZ^i(t) + d\langle Z^i \rangle(t) \\ &= 2Z^i(t)b^i(X_t)dt + 2Z^i(t)\sigma^i(X_t)dW(t) + 2Z^i(t)dL^i(t) \\ &\quad +1_{\{X^i(t) > \tilde{M}\}} \left|\sigma^i(X_t)\right|_2^2 dt \\ &= 2Z^i(t)b^i(X_t)dt + 2Z^i(t)\sigma^i(X_t)dW(t) + 1_{\{X^i(t) > \tilde{M}\}} \left|\sigma^i(X_t)\right|_2^2 dt, \end{aligned} \tag{4.3}$$

where $\langle Z^i \rangle$ denotes the quadratic variation process for of Z^i, and we have used the fact that L^i can increase only when Z^i is at zero. Thus

$$\begin{aligned} d\left(|Z(t)|_2^2\right) &= d\left((Z^1(t))^2 + \cdots + (Z^d(t))^2\right) \\ &= 2(Z(t))'b(X_t)dt + 2(Z(t))'\sigma(X_t)dW(t) \\ &\quad + \sum_{i=1}^d 1_{\{X^i(t) > \tilde{M}\}} \left|\sigma^i(X_t)\right|_2^2 dt. \end{aligned} \tag{4.4}$$

4.1.2 Positive Oscillation

We now introduce the notion of the positive oscillation (or largest increase) of a path over an interval. This refinement of the oscillation of a path (1.3) is well suited to our problem, and it still obeys an inequality analogous to part (i) of Proposition 2.3.1.

Definition 4.1.1. Given a path $x \in C([a_1, a_2], \mathbb{R})$, define the *positive oscillation of* x over $[a_1, a_2]$ by

$$\mathrm{Osc}^+(x, [a_1, a_2]) = \sup_{a_1 \le s \le t \le a_2} (x(t) - x(s)).$$

Remark. Note that there is no absolute value in the definition of Osc^+, so that we have the following obvious inequality:

$$\mathrm{Osc}^+(x, [a_1, a_2]) \le \mathrm{Osc}(x, [a_1, a_2]), \quad x \in C([a_1, a_2], \mathbb{R}).$$

Remark. We also have the following inequalities: for all $x \in \mathbb{C}_{\mathbb{I}}^d$ and $i = 1, \ldots, d$,

$$\begin{aligned} \mathrm{Osc}^+(x^i, \mathbb{I}) &\le \|x^i\|_{\mathbb{I}} \le \|x\|_2, \quad \text{and} \quad (4.5) \\ \|x^i\|_{\mathbb{I}} &\le x^i(-\tau) + \mathrm{Osc}^+(x^i, \mathbb{I}). \quad (4.6) \end{aligned}$$

We have the following property of Osc^+ when it is applied to a reflected path.

Lemma 4.1.1. *Fix* $0 \le t_1 < t_2 < \infty$. *Suppose that* $x, y, z \in C([t_1, t_2], \mathbb{R})$ *such that*

(i) $z(t) = x(t) + y(t) \in [0, \infty)$ *for all* $t \in [t_1, t_2]$,

(ii) $y(t_1) \ge 0$, *and* $y(\cdot)$ *is nondecreasing, and*

(iii) $y(\cdot)$ *can only increase when* z *is at zero:*

$$y(t) = y(t_1) + \int_{t_1}^{t} 1_{\{0\}}(z(s))dy(s), \quad \textit{for all} \ \ t \in [t_1, t_2].$$

Then,

$$\mathrm{Osc}^+(z, [t_1, t_2]) \le \mathrm{Osc}^+(x, [t_1, t_2]). \quad (4.7)$$

Proof. By continuity of z and compactness of the triangle $\{(s,t) : t_1 \le s \le t \le t_2\}$, there exist $s, t \in [t_1, t_2]$ such that $s \le t$ and $\mathrm{Osc}^+(z, [t_1, t_2]) = (z(t) - z(s))$. If $s = t$, then the inequality (4.7) is clear. So we suppose that $s < t$. Then there are two cases to consider.

Case 1: Assume that $y(s) = y(t)$. Then

$$\begin{aligned} z(t) - z(s) &= x(t) - x(s) \\ &\le \mathrm{Osc}^+(x, [t_1, t_2]). \quad (4.8) \end{aligned}$$

Case 2: Suppose that $y(s) < y(t)$. Then there is $u \in [s,t]$ such that $z(u) = 0$, by (iii). Let $u' = \sup\{v \le t : z(v) = 0\}$. Then $u' \in [u,t]$, $z(u') = 0$, and $z(v) > 0$ for all $v \in (u', t]$. Thus, y cannot increase on $(u', t]$ by (iii), and so by continuity, $y(u') = y(t)$. Then we have that

$$\begin{aligned} z(t) - z(s) &\le z(t) \\ &= z(t) - z(u') \\ &= x(t) - x(u') + y(t) - y(u') \\ &= x(t) - x(u') \\ &\le \text{Osc}^+(x, [t_1, t_2]), \end{aligned} \tag{4.9}$$

where we have used the facts that $z(s) \ge 0$, $z(u') = 0$, and $y(t) - y(u') = 0$.

□

We will also need the following technical lemma.

Lemma 4.1.2. *Suppose that* $X = \{X(t), t \in \mathbb{J}\}$ *is a solution of the SDDER (1.1). Then for each* $i = 1, \ldots, d$ *and* $\hat{M} \ge 0$, *for any* $0 \le t_1 < t_2 < \infty$, P*-a.s.,*

$$\begin{aligned} \text{Osc}^+(X^i, [t_1, t_2]) \le\ & \hat{M} + \int_{t_1}^{t_2} 1_{\{X^i(u) > \hat{M}\}} \left(b^i(X_u)\right)^+ du \\ & + \sup_{t_1 \le r < s \le t_2} \int_r^s 1_{\{X^i(u) > \hat{M}\}} \sigma^i(X_u) dW(u), \end{aligned} \tag{4.10}$$

and for any $t \ge 0$,

$$\begin{aligned} \text{Osc}^+(X^i, [t-\tau, t]) \le\ & \text{Osc}^+(X_0^i, \mathbb{I}) + \hat{M} + \int_{(t-\tau)^+}^t 1_{\{X^i(u) > \hat{M}\}} (b^i(X_u))^+ du \\ & + \sup_{(t-\tau)^+ \le r < s \le t} \int_r^s 1_{\{X^i(u) > \hat{M}\}} \sigma^i(X_u) dW(u). \end{aligned} \tag{4.11}$$

Proof. Fix $i \in \{1, \ldots, d\}$, $\hat{M} \ge 0$, $0 \le t_1 < t_2 < \infty$. In the definition of Z, set $\tilde{M} = \hat{M}$, so that $Z^i(\cdot) := (X^i(\cdot) - \hat{M})^+$. Then,

$$\text{Osc}^+(X^i, [t_1, t_2]) \le \hat{M} + \text{Osc}^+(Z^i, [t_1, t_2]). \tag{4.12}$$

The inequality (4.12) can be readily verified by considering $s \le t$ in $[t_1, t_2]$ such that the left hand side above is equal to $X^i(t) - X^i(s)$ and then considering the three cases: (a) $X^i(t) < \hat{M}$, (b) $X^i(t) \ge \hat{M}$ and $X^i(s) \ge \hat{M}$, and (c) $X^i(t) \ge \hat{M}$ and $X^i(s) < \hat{M}$. Thus, it suffices to estimate $\text{Osc}^+(Z^i, [t_1, t_2])$.

Since P-a.s. (4.2) holds and L^i can increase only when Z^i is zero, we may apply Lemma 4.1.1 to yield P-a.s.,

$$\text{Osc}^+(Z^i,[t_1,t_2]) \quad \leq \quad \text{Osc}^+(\mathcal{I}^i,[t_1,t_2]), \tag{4.13}$$

where

$$\mathcal{I}^i(t) \quad := \quad Z^i(0) + \int_0^t 1_{\{X^i(s)>\hat{M}\}} b^i(X_s)ds + \int_0^t 1_{\{X^i(s)>\hat{M}\}} \sigma^i(X_s)dW(s), \tag{4.14}$$

for $t \geq 0$. Now,

$$\begin{aligned} \text{Osc}^+(\mathcal{I}^i,[t_1,t_2]) \quad \leq \quad & \int_{t_1}^{t_2} 1_{\{X^i(u)>\hat{M}\}} (b^i(X_u))^+ du \\ & + \sup_{t_1\leq r<s\leq t_2} \left(\int_r^s 1_{\{X^i(u)>\hat{M}\}} \sigma^i(X_u)dW(u) \right). \end{aligned} \tag{4.15}$$

This establishes (4.10). Inequality (4.11) follows from (4.10) and the observation that for $t \geq 0$,

$$\text{Osc}^+(X^i,[t-\tau,t]) \quad \leq \quad \text{Osc}^+(X^i,\mathbb{I}) + \text{Osc}^+(X^i,[(t-\tau)^+,t]). \tag{4.16}$$

□

4.2 Bounded Moments when b and σ Satisfy an Integral Growth Condition

Throughout this section, we assume that the coefficients b, σ satisfy Assumption 4.2.1 below.

4.2.1 Assumptions on b and σ

Assumption 4.2.1. *There exist non-negative constants B_0, B_1, $B_{1,1}, \ldots, B_{1,d}$, $B_{2,1}, \ldots, B_{2,d}$, C_0, $C_{2,1}, \ldots, C_{2,d}$, M, constants $q_1 \in (0,1], q_2 \in (0,2]$, probability measures $\mu_1^1, \ldots, \mu_1^d$, $\mu_2^1, \ldots, \mu_2^d$ on $(\mathbb{I}, \mathcal{B}(\mathbb{I}))$, and a measurable function $\ell : \mathbb{C}_{\mathbb{I}}^d \to \mathbb{R}_+^d$, such that for each $x \in \mathbb{C}_{\mathbb{I}}^d$ and $i = 1, \ldots, d$, $\ell^i(x) \in x^i(\mathbb{I}) := \{x^i(s), s \in \mathbb{I}\}$ for each i, and*

(i) whenever $x^i(0) \geq M$, we have

$$b^i(x) \;\leq\; B_0 - B_1 x^i(0) - B_{1,i}\ell^i(x) + B_{2,i} \int_{-\tau}^0 |x(r)|_2^{q_1} \mu_1^i(dr), \tag{4.17}$$

(ii) whenever $x^i(0) \geq M$*, we have*

$$\left|\sigma^i(x)\right|_2^2 \;\leq\; C_0 + C_{2,i}\int_{-\tau}^{0} |x(r)|_2^{q_2}\mu_2^i(dr), \tag{4.18}$$

(iii) for $\underline{B_1} := \min_i B_{1,i}$ *and* $\tilde{B}_2 := \left(\left(\sum_{i=1}^{d} B_{2,i}\right) \wedge \left(d\sum_{i=1}^{d} B_{2,i}^4\right)^{\frac{1}{4}}\right)$*, we have*

$$\begin{aligned} &B_1 + \underline{B_1} \\ &\quad > \left(\tau\left(\sum_{i=1}^{d}(B_{1,i}B_{2,i})^2\right)^{\frac{1}{2}} + \tilde{B}_2\right)\delta_{q_1,1} + \left(\frac{1}{2}\sum_{i=1}^{d} C_{2,i} + 4\sqrt{\tau}\left(\sum_{i=1}^{d} C_{2,i}B_{1,i}^2\right)^{\frac{1}{2}}\right)\delta_{q_2,2}. \end{aligned}$$

Remark. Note that parts (i) and (ii) restrict b^i and σ^i only on $\{x \in \mathbb{C}_{\mathbb{I}}^d : x^i(0) \geq M\}$, and the control on b^i is only one-sided. However, b and σ will always be required to satisfy the supremum linear growth bounds (2.1) and (2.2), which restrict the growth of b and σ for all $x \in \mathbb{C}_{\mathbb{I}}^d$, though, on $\bigcup_{i=1}^{d}\{x \in \mathbb{C}_{\mathbb{I}}^d : x^i(0) \geq M\}$, this supremum growth control on b and $\|\sigma\|_2$ is weaker than the at-most-integral-linear growth imposed by parts (i) and (ii) of the above assumption.

Throughout Section 4.2, we use $\tilde{M} = M + 1$ in the definition of the overshoot process Z in (4.1). The simple inequalities $X^i(\cdot) \leq Z^i(\cdot) + \tilde{M}$, for each i, reduce the problem of bounding the moments of $\|X_t\|_2$ to bounding the moments of $\|Z_t\|_2$. The intuition behind the following proofs is that the dynamics of the overshoot process Z^i when $Z^i(t) > 0$ are the same as the dynamics of X^i when $X^i(t) > \tilde{M}$, and since $\tilde{M} > M$, Assumption 4.2.1 is sufficient to get control over the growth of each component of Z.

4.2.2 Uniform Bound on $E\left[|X(t)|_2^2\right]$

The main result of this subsection is the following theorem, which holds under Assumption 4.2.1.

Theorem 4.2.1. *Suppose that* $E[\|X_0\|_2^2] < \infty$*. Then,* $\sup_{t \geq -\tau} E[|X(t)|_2^2] < \infty$*.*

The proof uses Lyapunov/Razumikhin-type arguments similar to those found in a theorem of Mao (Theorem 2.1 of [28]). Methods associated with the names Lyapunov and Razumikhin are often used to establish stability of dynamical systems. The proof is broken down into a series of supporting lemmas.

Lemma 4.2.1. *Assume* $E\left[\|X_0\|_2^2\right] < \infty$. *There exists a constant* $M_1 > 0$ *such that whenever* $t \geq \tau$ *is such that both*

$$E[|Z(t)|_2^2] \;\; \geq \;\; M_1, \tag{4.19}$$

and

$$E[|Z(r)|_2^2] \;\; \leq \;\; E[|Z(t)|_2^2] \text{ for all } r \in [t-2\tau, t], \tag{4.20}$$

then

$$E\left[2(Z(t))'b(X_t) + \sum_{i=1}^{d} 1_{\{X^i(t)>M\}} \left|\sigma^i(X_t)\right|_2^2\right] < 0.$$

Remark. We will refer to inequality (4.20) as the Razumikhin assumption.

Proof. Recall that we have set $\tilde{M} = M + 1$. Assume that we are given a $t \geq \tau$ such that (4.20) holds. For each $x \in \mathbb{C}_{\mathbb{I}}^d$, there is an $r_x \in \mathbb{I}^d$ such that for each $i = 1, \ldots, d$,

$$-\,\ell^i(x) \; = \; -x^i(r_x^i) \; \leq \; -x^i(0) + \text{Osc}^+(x^i, \mathbb{I}). \tag{4.21}$$

We note that for each $u \geq 0$ such that $Z^i(u) > 0$, we have $X^i(u) > \tilde{M} > M$ and so the inequalities (4.17) and (4.18) hold with $x = X_u$. Then,

$$\begin{aligned}
(Z(t))'b(X_t) &= \sum_{i=1}^{d} Z^i(t) b^i(X_t) \\
&\leq \sum_{i=1}^{d} Z^i(t) \left(B_0 - B_1 X^i(t) - B_{1,i}\ell^i(X_t) + B_{2,i} \int_{-\tau}^{0} |X(t+r)|_2^{q_1} \mu_1^i(dr) \right) \\
&\leq B_0 |Z(t)|_1 - (B_1 + \underline{B_1}) \sum_{i=1}^{d} Z^i(t) X^i(t) + \sum_{i=1}^{d} B_{1,i} Z^i(t) \text{Osc}^+(X^i, [t-\tau, t]) \\
&\quad + \sum_{i=1}^{d} B_{2,i} Z^i(t) \int_{-\tau}^{0} |X(t+r)|_2^{q_1} \mu_1^i(dr) \\
&\leq B_0 |Z(t)|_1 - (B_1 + \underline{B_1}) |Z(t)|_2^2 + \sum_{i=1}^{d} B_{1,i} Z^i(t) \left(\text{Osc}(X_0^i, \mathbb{I}) + M \right) \\
&\quad + \sum_{i=1}^{d} B_{1,i} Z^i(t) \int_{t-\tau}^{t} 1_{\{X^i(u)>M\}} (b^i(X_u))^+ du \\
&\quad + \sum_{i=1}^{d} B_{1,i} Z^i(t) \sup_{t-\tau \leq r < s \leq t} \left| \int_r^s 1_{\{X^i(u)>M\}} \sigma^i(X_u) dW(u) \right| \\
&\quad + |Z(t)|_2 \left(\sum_{i=1}^{d} B_{2,i}^2 \left(\int_{-\tau}^{0} |X(t+r)|_2^{q_1} \mu_1^i(dr) \right)^2 \right)^{\frac{1}{2}}.
\end{aligned} \tag{4.22}$$

Here, Assumption 4.2.1(i) and the non-negativity of the coordinates of Z were used for the first inequality, and the fact that $X(s) \geq Z(s)$ for all $s \geq -\tau$, Lemma 4.1.2 with $\hat{M} = M$, and the Cauchy-Schwarz inequality were used for the third inequality.

Using part (i) of Assumption 4.2.1, for each $i = 1, \ldots, d$, we have

$$\begin{aligned} \int_{t-\tau}^{t} 1_{\{X^i(u)>M\}}(b^i(X_u))^+ du &\leq \int_{t-\tau}^{t} \left(B_0 + B_{2,i} \int_{-\tau}^{0} |X(u+r)|_2^{q_1} \mu_1^i(dr) \right) du \\ &\leq B_0\tau + B_{2,i} \int_{t-\tau}^{t} \int_{-\tau}^{0} |X(u+r)|_2^{q_1} \mu_1^i(dr) du. \end{aligned} \quad (4.23)$$

Incorporating part (ii) of Assumption 4.2.1 with the above yields

$$\begin{aligned} &(Z(t))' b(X_t) + \frac{1}{2}\sum_{i=1}^{d} 1_{\{X^i(t)>M\}} \left|\sigma^i(X_t)\right|_2^2 \\ &\leq B_0|Z(t)|_1 - (B_1 + \underline{B_1})|Z(t)|_2^2 + \sum_{i=1}^{d} B_{1,i} Z^i(t) \left(\text{Osc}^+(X_0^i, \mathbb{I}) + M \right) \\ &\quad + \sum_{i=1}^{d} B_{1,i} Z^i(t) \left(B_0\tau + B_{2,i} \int_{t-\tau}^{t} \int_{-\tau}^{0} |X(u+r)|_2^{q_1} \mu_1^i(dr) du \right) \\ &\quad + \sum_{i=1}^{d} B_{1,i} Z^i(t) \sup_{t-\tau \leq r < s \leq t} \left| \int_{r}^{s} 1_{\{X^i(u)>M\}} \sigma^i(X_u) dW(u) \right| \\ &\quad + |Z(t)|_2 \left(\sum_{i=1}^{d} B_{2,i}^2 \left(\int_{-\tau}^{0} |X(t+r)|_2^{q_1} \mu_1^i(dr) \right)^2 \right)^{\frac{1}{2}} \\ &\quad + \frac{1}{2}\sum_{i=1}^{d} \left(C_0 + C_{2,i} \int_{-\tau}^{0} |X(t+r)|_2^{q_2} \mu_2^i(dr) \right). \end{aligned} \quad (4.24)$$

Define $\overline{B_1} := \max_{i=1}^{d} B_{1,i}$. Using the Cauchy-Schwarz inequality and taking expectations in inequality (4.24) yields

$$\begin{aligned}
E&\left[(Z(t))'b(X_t)+\frac{1}{2}\sum_{i=1}^{d}1_{\{X^i(t)>M\}}\left|\sigma^i(X_t)\right|_2^2\right]\\
&\le \left(B_0+M\overline{B_1}\right)E\left[|Z(t)|_1\right]-(B_1+\underline{B_1})E[|Z(t)|_2^2]\\
&\quad +\overline{B_1}E\left[|Z(t)|_2\left(\sum_{i=1}^{d}\left(\mathrm{Osc}^+(X_0^i,\mathbb{I})\right)^2\right)^{\frac{1}{2}}\right]\\
&\quad +E\left[\sum_{i=1}^{d}B_{1,i}Z^i(t)\left(B_0\tau+B_{2,i}\int_{t-\tau}^{t}\int_{-\tau}^{0}|X(u+r)|_2^{q_1}\mu_1^i(dr)du\right)\right]\\
&\quad +E\left[\sum_{i=1}^{d}B_{1,i}Z^i(t)\sup_{t-\tau\le r<s\le t}\left|\int_r^s 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right|\right]\\
&\quad +E\left[|Z(t)|_2\left(\sum_{i=1}^{d}B_{2,i}^2\left(\int_{-\tau}^{0}|X(t+r)|_2^{q_1}\mu_1^i(dr)\right)^2\right)^{\frac{1}{2}}\right]\\
&\quad +\frac{1}{2}dC_0+\frac{1}{2}\sum_{i=1}^{d}C_{2,i}E\left[\int_{-\tau}^{0}|X(t+r)|_2^{q_2}\mu_2^i(dr)\right]. \qquad (4.25)
\end{aligned}$$

We now separately develop estimates for the second and third to the last lines in (4.25). For each i,

$$\sup_{t-\tau\le r<s\le t}\left|\int_r^s 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right| \le 2\sup_{t-\tau\le s\le t}\left|\int_{t-\tau}^s 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right|. \qquad (4.26)$$

Part (ii) of Assumption 4.2.1, the assumption that $E[\|X_0\|_2^2]<\infty$, and Lemma 2.4.1 imply that for each i,

$$\left\{\int_{t-\tau}^s 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u),\mathcal{F}_s, s\ge t-\tau\right\}$$

is a square-integrable martingale. Then, Doob's submartingale inequality, the L^2 isometry for stochastic integrals, the independence of the coordinates of W, and (4.18) imply that

$$\begin{aligned}
E&\left[\sup_{t-\tau\le s\le t}\left|\int_{t-\tau}^s 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right|^2\right]\\
&\le 4E\left[\left|\int_{t-\tau}^t 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right|^2\right]\\
&= 4E\left[\int_{t-\tau}^t 1_{\{X^i(u)>M\}}\left|\sigma^i(X_t)\right|_2^2du\right]\\
&\le 4C_0\tau+4C_{2,i}\int_{t-\tau}^t\int_{-\tau}^0 E\left[|X(u+r)|_2^{q_2}\right]\mu_2^i(dr)du. \qquad (4.27)
\end{aligned}$$

We used Tonelli's theorem in the last inequality. Then, using the Cauchy-Schwarz inequality (twice) and inequality (B.1), we have

$$\begin{aligned}
& E\left[\sum_{i=1}^{d} B_{1,i} Z^i(t) \sup_{t-\tau\le r<s\le t}\left|\int_r^s 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right|\right] \\
\le\ & 2E\left[\sum_{i=1}^{d} B_{1,i} Z^i(t) \sup_{t-\tau\le s\le t}\left|\int_{t-\tau}^s 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right|\right] \\
\le\ & 2E\left[|Z(t)|_2\left(\sum_{i=1}^{d}\sup_{t-\tau\le s\le t}\left|\int_{t-\tau}^s B_{1,i}1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right|^2\right)^{\frac{1}{2}}\right] \\
\le\ & 2\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}\left(E\left[\sum_{i=1}^{d}\sup_{t-\tau\le s\le t}\left|\int_{t-\tau}^s B_{1,i}1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right|^2\right]\right)^{\frac{1}{2}} \\
\le\ & 4\left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}\left(\sum_{i=1}^{d}B_{1,i}^2C_0\tau+\sum_{i=1}^{d}C_{2,i}B_{1,i}^2\int_{t-\tau}^t\int_{-\tau}^0 E\left[|X(u+r)|_2^{q_2}\right]\mu_2^i(dr)du\right)^{\frac{1}{2}} \\
\le\ & 4\sqrt{C_0\tau}\left(\sum_{i=1}^{d}B_{1,i}^2\right)^{\frac{1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}} \\
& +4\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}\left(\sum_{i=1}^{d}C_{2,i}B_{1,i}^2\int_{t-\tau}^t\int_{-\tau}^0 E\left[|X(u+r)|_2^{q_2}\right]\mu_2^i(dr)du\right)^{\frac{1}{2}}. \qquad (4.28)
\end{aligned}$$

For the second last line in (4.25), the Cauchy-Schwarz inequality and Tonelli's theorem imply that

$$\begin{aligned}
& E\left[|Z(t)|_2\left(\sum_{i=1}^{d}B_{2,i}^2\left(\int_{-\tau}^0|X(t+r)|_2^{q_1}\mu_1^i(dr)\right)^2\right)^{\frac{1}{2}}\right] \\
& \quad\le\ \left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}\left(\sum_{i=1}^{d}B_{2,i}^2E\left[\left(\int_{-\tau}^0|X(t+r)|_2^{q_1}\mu_1^i(dr)\right)^2\right]\right)^{\frac{1}{2}} \\
& \quad\le\ \left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}\left(\sum_{i=1}^{d}B_{2,i}^4\right)^{\frac{1}{4}}\left(\sum_{i=1}^{d}\left(E\left[\left(\int_{-\tau}^0|X(t+r)|_2^{q_1}\mu_1^i(dr)\right)^2\right]\right)^2\right)^{\frac{1}{4}}. \qquad (4.29)
\end{aligned}$$

Proposition B.0.2 implies that for any $\gamma > 1$, and $s \ge -\tau$, there is a constant $K(d, \tilde{M}, \gamma, 2) \ge 0$ such that

$$\begin{aligned}
|X(s)|_2^2 &= (X^1(s))^2+\cdots+(X^d(s))^2 \\
&\le K(d,\tilde{M},\gamma,2)+\gamma\left((Z^1(s))^2+\cdots+(Z^d(s))^2\right) \\
&= K(d,\tilde{M},\gamma,2)+\gamma|Z(s)|_2^2. \qquad (4.30)
\end{aligned}$$

Two applications of Hölder's inequality, inequality (4.30), Fubini's theorem, and the Razumikhin assumption (4.20) imply that for each $\gamma > 1$,

$$\begin{aligned}
&\sum_{i=1}^{d}\left(E\left[\left(\int_{-\tau}^{0}|X(t+r)|_2^{q_1}\,\mu_1^i(dr)\right)^2\right]\right)^2 \\
&\quad\leq \sum_{i=1}^{d}\left(E\left[\left(\int_{-\tau}^{0}|X(t+r)|_2^{2}\,\mu_1^i(dr)\right)^{q_1}\right]\right)^2 \\
&\quad\leq \sum_{i=1}^{d}\left(E\left[\int_{-\tau}^{0}|X(t+r)|_2^{2}\,\mu_1^i(dr)\right]\right)^{2q_1} \\
&\quad\leq \sum_{i=1}^{d}\left(K(d,\tilde{M},\gamma,2)+\int_{-\tau}^{0}\gamma E\left[|Z(t+r)|_2^2\right]\mu_1^i(dr)\right)^{2q_1} \\
&\quad\leq \sum_{i=1}^{d}\left(K(d,\tilde{M},\gamma,2)+\int_{-\tau}^{0}\gamma E\left[|Z(t)|_2^2\right]\mu_1^i(dr)\right)^{2q_1} \\
&\quad\leq d\left(K(d,\tilde{M},\gamma,2)+\gamma E\left[|Z(t)|_2^2\right]\right)^{2q_1}, \qquad (4.31)
\end{aligned}$$

and thus

$$\begin{aligned}
&E\left[|Z(t)|_2\left(\sum_{i=1}^{d}B_{2,i}^2\left(\int_{-\tau}^{0}|X(t+r)|_2^{q_1}\,\mu_1^i(dr)\right)^2\right)^{\frac{1}{2}}\right] \\
&\quad\leq \left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}\left(\sum_{i=1}^{d}B_{2,i}^4\right)^{\frac{1}{4}}\left(d\left(K(d,\tilde{M},\gamma,2)+\gamma E\left[|Z(t)|_2^2\right]\right)^{2q_1}\right)^{\frac{1}{4}} \\
&\quad= \left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}\left(d\sum_{i=1}^{d}B_{2,i}^4\right)^{\frac{1}{4}}\left(K(d,\tilde{M},\gamma,2)+\gamma E\left[|Z(t)|_2^2\right]\right)^{\frac{q_1}{2}}. \qquad (4.32)
\end{aligned}$$

Alternatively, we could have used inequality (B.1), Hölder's inequality, (4.30), and (4.20), to yield for each $\gamma > 1$,

$$\begin{aligned}
& E\left[|Z(t)|_2\left(\sum_{i=1}^{d}B_{2,i}^2\left(\int_{-\tau}^{0}|X(t+r)|_2^{q_1}\mu_1^i(dr)\right)^2\right)^{\frac{1}{2}}\right] \\
& \quad \leq E\left[|Z(t)|_2\sum_{i=1}^{d}B_{2,i}\int_{-\tau}^{0}|X(t+r)|_2^{q_1}\mu_1^i(dr)\right] \\
& \quad = \sum_{i=1}^{d}B_{2,i}E\left[|Z(t)|_2\int_{-\tau}^{0}|X(t+r)|_2^{q_1}\mu_1^i(dr)\right] \\
& \quad \leq \left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}\sum_{i=1}^{d}B_{2,i}\left(E\left[\left(\int_{-\tau}^{0}|X(t+r)|_2^{q_1}\mu_1^i(dr)\right)^2\right]\right)^{\frac{1}{2}} \\
& \quad \leq \left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}\sum_{i=1}^{d}B_{2,i}\left(E\left[\left(\int_{-\tau}^{0}|X(t+r)|_2^{2}\mu_1^i(dr)\right)^{q_1}\right]\right)^{\frac{1}{2}} \\
& \quad \leq \left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}\sum_{i=1}^{d}B_{2,i}\left(E\left[\int_{-\tau}^{0}|X(t+r)|_2^{2}\mu_1^i(dr)\right]\right)^{\frac{q_1}{2}} \\
& \quad \leq \left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}\sum_{i=1}^{d}B_{2,i}\left(\int_{-\tau}^{0}\left(K(d,\tilde{M},\gamma,2)+\gamma E\left[|Z(t+r)|_2^2\right]\right)\mu_1^i(dr)\right)^{\frac{q_1}{2}} \\
& \quad \leq \left(\sum_{i=1}^{d}B_{2,i}\right)\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}\left(K(d,\tilde{M},\gamma,2)+\gamma E\left[|Z(t)|_2^2\right]\right)^{\frac{q_1}{2}}. \qquad (4.33)
\end{aligned}$$

Combining inequalities (4.32) and (4.33), and using (B.1), we have for each $\gamma > 1$,

$$\begin{aligned}
& E\left[|Z(t)|_2\left(\sum_{i=1}^{d}B_{2,i}^2\left(\int_{-\tau}^{0}|X(t+r)|_2^{q_1}\mu_1^i(dr)\right)^2\right)^{\frac{1}{2}}\right] \\
& \quad \leq \tilde{B}_2\left(\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_1}{2}}+\gamma^{\frac{q_1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{q_1}{2}}\right)\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}. \qquad (4.34)
\end{aligned}$$

Continuing from (4.25), by (B.6), the Cauchy-Schwarz inequality, (4.5), (4.28), and

(4.34), we have that

$$
\begin{aligned}
E & \left[(Z(t))' b(X_t) + \frac{1}{2}\sum_{i=1}^{d} 1_{\{X^i(t)>M\}} \left|\sigma^i(X_t)\right|_2^2\right] \\
\leq\ & (B_0 + (M + B_0\tau)\overline{B_1})\sqrt{d}\left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}} - (B_1 + \underline{B_1})E[|Z(t)|_2^2] \\
& + \overline{B_1}\sqrt{d}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}\left(E\left[\|X_0\|_2^2\right]\right)^{\frac{1}{2}} \\
& + E\left[\sum_{i=1}^{d} B_{1,i}B_{2,i}Z^i(t)\int_{t-\tau}^{t}\int_{-\tau}^{0}|X(u+r)|_2^{q_1}\mu_1^i(dr)du\right] \\
& + 4\sqrt{C_0\tau}\left(\sum_{i=1}^{d} B_{1,i}^2\right)^{\frac{1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}} \\
& + 4\left(\sum_{i=1}^{d} C_{2,i}B_{1,i}^2\int_{t-\tau}^{t}\int_{-\tau}^{0} E[|X(u+r)|_2^{q_2}]\mu_2^i(dr)du\right)^{\frac{1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}} \\
& + \tilde{B}_2\left(\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}} + \gamma^{\frac{q_1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{q_1+1}{2}}\right) \\
& + \frac{1}{2}dC_0 + \frac{1}{2}\sum_{i=1}^{d} C_{2,i}\int_{-\tau}^{0} E\left[|X(t+r)|_2^{q_2}\right]\mu_2^i(dr).
\end{aligned}
\tag{4.35}
$$

By Hölder's inequality (used twice), (4.30), and the Razumikhin assumption (4.20), we have for each $\gamma > 1$,

$$
\begin{aligned}
E & \left[\left(\int_{t-\tau}^{t}\int_{-\tau}^{0}|X(u+r)|_2^{q_1}\mu_1^i(dr)du\right)^2\right] \\
\leq\ & E\left[\tau\int_{t-\tau}^{t}\left(\int_{-\tau}^{0}|X(u+r)|_2^{q_1}\mu_1^i(dr)\right)^2 du\right] \\
\leq\ & \tau\int_{t-\tau}^{t} E\left[\left(\int_{-\tau}^{0}|X(u+r)|_2^2\mu_1^i(dr)\right)^{q_1}\right] du \\
\leq\ & \tau\int_{t-\tau}^{t}\left(E\left[\int_{-\tau}^{0}\left(K(d,\tilde{M},\gamma,2) + \gamma|Z(u+r)|_2^2\right)\mu_1^i(dr)\right]\right)^{q_1} du \\
\leq\ & \tau\int_{t-\tau}^{t}\left(K(d,\tilde{M},\gamma,2) + \gamma\int_{-\tau}^{0} E\left[|Z(u+r)|_2^2\right]\mu_1^i(dr)\right)^{q_1} du \\
\leq\ & \tau\int_{t-\tau}^{t}\left(K(d,\tilde{M},\gamma,2) + \gamma E\left[|Z(t)|_2^2\right]\right)^{q_1} du \\
\leq\ & \tau^2\left(K(d,\tilde{M},\gamma,2) + \gamma E\left[|Z(t)|_2^2\right]\right)^{q_1}.
\end{aligned}
\tag{4.36}
$$

Therefore, by the Cauchy-Schwarz inequality (twice), and (B.1),

$$
\begin{aligned}
& E\left[\sum_{i=1}^{d} B_{1,i}B_{2,i}Z^i(t)\int_{t-\tau}^{t}\int_{-\tau}^{0}|X(u+r)|_2^{q_1}\mu_1^i(dr)du\right] \\
\leq\ & \left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}\left(\sum_{i=1}^{d}(B_{1,i}B_{2,i})^2 E\left[\left(\int_{t-\tau}^{t}\int_{-\tau}^{0}|X(u+r)|_2^{q_1}\mu_1^i(dr)du\right)^2\right]\right)^{\frac{1}{2}} \\
\leq\ & \left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}\left(\tau^2\sum_{i=1}^{d}(B_{1,i}B_{2,i})^2\left(K(d,\tilde{M},\gamma,2)+\gamma E\left[|Z(t)|_2^2\right]\right)^{q_1}\right)^{\frac{1}{2}} \\
\leq\ & \tau\left(\sum_{i=1}^{d}(B_{1,i}B_{2,i})^2\right)^{\frac{1}{2}}\left(K(d,\tilde{M},\gamma,2)+\gamma E\left[|Z(t)|_2^2\right]\right)^{\frac{q_1}{2}}\left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}} \\
\leq\ & \tau\left(\sum_{i=1}^{d}(B_{1,i}B_{2,i})^2\right)^{\frac{1}{2}}\left(\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_1}{2}}+\gamma^{\frac{q_1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{q_1}{2}}\right)\left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}. \qquad (4.37)
\end{aligned}
$$

Hölder's inequality, (4.30), the Razumikhin assumption (4.20), and (B.1) also imply that for each $\gamma > 1$,

$$
\begin{aligned}
& \left(\sum_{i=1}^{d} C_{2,i}B_{1,i}^2\int_{t-\tau}^{t}\int_{-\tau}^{0}E[|X(u+r)|_2^{q_2}]\mu_2^i(dr)du\right)^{\frac{1}{2}} \\
&\quad\leq \left(\sum_{i=1}^{d} C_{2,i}B_{1,i}^2\int_{t-\tau}^{t}\int_{-\tau}^{0}\left(E\left[|X(u+r)|_2^2\right]\right)^{\frac{q_2}{2}}\mu_2^i(dr)du\right)^{\frac{1}{2}} \\
&\quad\leq \left(\sum_{i=1}^{d} C_{2,i}B_{1,i}^2\int_{t-\tau}^{t}\int_{-\tau}^{0}\left(K(d,\tilde{M},\gamma,2)+\gamma E\left[|Z(u+r)|_2^2\right]\right)^{\frac{q_2}{2}}\mu_2^i(dr)du\right)^{\frac{1}{2}} \\
&\quad\leq \sqrt{\tau}\left(\sum_{i=1}^{d} C_{2,i}B_{1,i}^2\right)^{\frac{1}{2}}\left(K(d,\tilde{M},\gamma,2)+\gamma E\left[|Z(t)|_2^2\right]\right)^{\frac{q_2}{4}} \\
&\quad\leq \sqrt{\tau}\left(\sum_{i=1}^{d} C_{2,i}B_{1,i}^2\right)^{\frac{1}{2}}\left(\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_2}{4}}+\gamma^{\frac{q_2}{4}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{q_2}{4}}\right). \qquad (4.38)
\end{aligned}
$$

By using Hölder's inequality, the fact that μ_2^i is a probability measure, inequalities (4.30) and (B.1), and the Razumikhin assumption (4.20), we obtain

$$
\begin{aligned}
\int_{-\tau}^{0}E\left[|X(t+r)|_2^{q_2}\right]\mu_2^i(dr) &\leq \int_{-\tau}^{0}\left(E\left[|X(t+r)|_2^2\right]\right)^{\frac{q_2}{2}}\mu_2^i(dr) \\
&\leq \int_{-\tau}^{0}\left(E\left[K(d,\tilde{M},\gamma,2)+\gamma|Z(t+r)|_2^2\right]\right)^{\frac{q_2}{2}}\mu_2^i(dr) \\
&\leq \left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_2}{2}}+\gamma^{\frac{q_2}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{q_2}{2}}. \qquad (4.39)
\end{aligned}
$$

Continuing from line (4.35) using inequalities (4.37), (4.38), and (4.39), we have

$$
\begin{aligned}
&E\left[(Z(t))'b(X_t)+\frac{1}{2}\sum_{i=1}^{d}1_{\{X^i(t)>M\}}\left|\sigma^i(X_t)\right|_2^2\right] \\
\leq\ & \left(B_0+\left(M+B_0\tau+\left(E\left[\|X_0\|_\infty^2\right]\right)^{\frac{1}{2}}\right)\overline{B_1}\right)\sqrt{d}\left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}} \\
& -(B_1+\underline{B_1})E[|Z(t)|_2^2] \\
& +\tau\left(\sum_{i=1}^{d}(B_{1,i}B_{2,i})^2\right)^{\frac{1}{2}}\left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}\left(\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_1}{2}}+\gamma^{\frac{q_1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{q_1}{2}}\right) \\
& +4\sqrt{C_0\tau}\left(\sum_{i=1}^{d}B_{1,i}^2\right)^{\frac{1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}} \\
& +4\sqrt{\tau}\left(\sum_{i=1}^{d}C_{2,i}B_{1,i}^2\right)^{\frac{1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}\left(\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_2}{4}}+\gamma^{\frac{q_2}{4}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{q_2}{4}}\right) \\
& +\tilde{B}_2\left(\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{1}{2}}+\gamma^{\frac{q_1}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{q_1+1}{2}}\right) \\
& +\frac{d}{2}C_0+\frac{1}{2}\left(\sum_{i=1}^{d}C_{2,i}\right)\left(\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_2}{2}}+\gamma^{\frac{q_2}{2}}\left(E\left[|Z(t)|_2^2\right]\right)^{\frac{q_2}{2}}\right) \\
=\ & K_1(\gamma)+K_2(\gamma)\left(E[|Z(t)|_2^2]\right)^{\frac{1}{2}}+K_3(\gamma)\left(E[|Z(t)|_2^2]\right)^{\frac{1+q_1}{2}}+K_4(\gamma)\left(E[|Z(t)|_2^2]\right)^{\frac{q_2}{2}} \\
& +K_5(\gamma)\left(E[|Z(t)|_2^2]\right)^{\frac{2+q_2}{4}}-(B_1+\underline{B_1})E[|Z(t)|_2^2], \qquad (4.40)
\end{aligned}
$$

where

$$\begin{aligned}
K_1(\gamma) &= \frac{d}{2}C_0 + \frac{1}{2}\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_2}{2}}\sum_{i=1}^{d} C_{2,i}, \\
K_2(\gamma) &= \left(B_0 + \left(M + B_0\tau + \left(E\left[\|X_0\|_\infty^2\right]\right)^{\frac{1}{2}}\right)\overline{B_1}\right)\sqrt{d} \\
&\quad +\tau\left(\sum_{i=1}^{d}(B_{1,i}B_{2,i})^2\right)^{\frac{1}{2}}\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_1}{2}} + 4\sqrt{C_0\tau}\left(\sum_{i=1}^{d}B_{1,i}^2\right)^{\frac{1}{2}} \\
&\quad +4\sqrt{\tau}\left(\sum_{i=1}^{d}C_{2,i}B_{1,i}^2\right)^{\frac{1}{2}}\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_2}{4}} + \tilde{B}_2\left(K(d,\tilde{M},\gamma,2)\right)^{\frac{q_1}{2}}, \\
K_3(\gamma) &= \left(\tau\left(\sum_{i=1}^{d}(B_{1,i}B_{2,i})^2\right)^{\frac{1}{2}} + \tilde{B}_2\right)\gamma^{\frac{q_1}{2}}, \\
K_4(\gamma) &= \frac{1}{2}\gamma^{\frac{q_2}{2}}\sum_{i=1}^{d}C_{2,i}, \text{ and} \\
K_5(\gamma) &= 4\sqrt{\tau}\left(\sum_{i=1}^{d}C_{2,i}B_{1,i}^2\right)^{\frac{1}{2}}\gamma^{\frac{q_2}{4}}.
\end{aligned}$$

By Assumption 4.2.1(iii), we can fix a $\gamma > 1$ such that $B_1 + \underline{B_1} > K_3(\gamma)\delta_{q_1,1} + (K_4(\gamma) + K_5(\gamma))\delta_{q_2,2}$. Therefore,

$$E\left[2(Z(t))'b(X_t) + \sum_{i=1}^{d} 1_{\{X^i(t)>M\}}\left|\sigma^i(X_t)\right|_2^2\right] < 0$$

whenever $E[|Z(t)|_2^2]$ is large enough. Indeed, define the function $f : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$f(r) := K_1(\gamma) + K_2(\gamma)r^{\frac{1}{2}} + K_3(\gamma)r^{\frac{1+q_1}{2}} + K_4(\gamma)r^{\frac{q_2}{2}} + K_5(\gamma)r^{\frac{2+q_2}{4}} - (B_1 + \underline{B_1})r.$$

All of the exponents on r are at most one, and the above shows that the constant in front of the highest degree term, namely,

$$-(B_1 + \underline{B_1}) + K_3(\gamma)\delta_{q_1,1} + (K_4(\gamma) + K_5(\gamma))\delta_{q_2,2},$$

is strictly negative, and this implies that

$$\lim_{r\to\infty} f(r) = -\infty,$$

so there exists a constant $M_1 > 0$ such that $r \geq M_1$ implies that $f(r) < 0$.

□

Lemma 4.2.2. *Assume $E[\|X_0\|_2^2] < \infty$. Let M_1 be defined as in the previous lemma, and assume that $t \geq \tau$ is such that both (4.19) and (4.20) hold. Then there exists an $h^* > 0$ such that*

$$E[|Z(t+s)|_2^2] \quad < \quad E[|Z(t)|_2^2], \qquad \textit{for each } s \in (0, h^*]. \tag{4.41}$$

Proof. Setting $\eta_n = t \vee \inf\{s \geq -\tau : |X(s)|_2 \geq n\}$ for each integer $n > 0$. Then the adapted process $\{1_{(t,\eta_n]}(s)(Z(s))'\sigma(X_s), s \geq t\}$ is bounded, so that the process

$$\left\{ \int_t^{(t+h)\wedge\eta_n} (Z(s))'\sigma(X_s)dW(s), \mathcal{F}_{t+h}, h \geq 0 \right\}$$

is a square-integrable martingale, and so

$$E\left[\int_t^{(t+h)\wedge\eta_n} (Z(s))'\sigma(X_s)dW(s)\right] = 0, \quad \text{for all } h \geq 0, n > 0.$$

We have from equality (4.4) that

$$\begin{aligned} &|Z((t+h)\wedge\eta_n)|_2^2 - |Z(t)|_2^2 \\ &\qquad = \int_t^{(t+h)\wedge\eta_n} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s)>\tilde{M}\}} \left|\sigma^i(X_s)\right|_2^2 \right) ds \\ &\qquad\quad + \int_t^{(t+h)\wedge\eta_n} 2(Z(s))'\sigma(X_s)dW(s), \end{aligned} \tag{4.42}$$

and since $E\left[\sup_{s\in[-\tau,t+h]} |Z(s)|_2^2\right] < \infty$ by Lemma 2.4.1 with $p = 2$, we have

$$\begin{aligned} &E\left[|Z((t+h)\wedge\eta_n)|_2^2\right] - E\left[|Z(t)|_2^2\right] \\ &\qquad = E\left[\int_t^{(t+h)\wedge\eta_n} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s)>\tilde{M}\}} \left|\sigma^i(X_s)\right|_2^2 \right) ds\right], \end{aligned} \tag{4.43}$$

and then for each $h > 0$, by the dominated convergence theorem,

$$\begin{aligned}
&E[|Z(t+h)|_2^2] - E[|Z(t)|_2^2] \\
&= E[\lim_{n\to\infty} |Z((t+h)\wedge\eta_n)|_2^2] - E[|Z(t)|_2^2] \\
&= \lim_{n\to\infty} E[|Z((t+h)\wedge\eta_n)|_2^2] - E[|Z(t)|_2^2] \\
&= \lim_{n\to\infty} E\left[\int_t^{(t+h)\wedge\eta_n} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s)>\tilde{M}\}} \left|\sigma^i(X_s)\right|_2^2\right) ds\right] \\
&= \lim_{n\to\infty} E\left[\int_t^{t+h} 1_{\{s\le\eta_n\}} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s)>\tilde{M}\}} \left|\sigma^i(X_s)\right|_2^2\right) ds\right] \\
&= E\left[\int_t^{t+h} \lim_{n\to\infty} 1_{\{s\le\eta_n\}} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s)>\tilde{M}\}} \left|\sigma^i(X_s)\right|_2^2\right) ds\right] \\
&= E\left[\int_t^{t+h} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s)>\tilde{M}\}} \left|\sigma^i(X_s)\right|_2^2\right) ds\right]. \qquad (4.44)
\end{aligned}$$

The second last equality uses dominated convergence conferred by Lemma 2.4.1and the linear growth bounds (2.1) and (2.2) on b and σ.

Define $f : \mathbb{R}_+ \to [0,1]$ by $f(r) = (r-M)^+ - (r-\tilde{M})^+$. Since $\tilde{M} = M+1$,

$$1_{(\tilde{M},\infty)}(r) \ \le\ f(r) \ \le\ 1_{(M,\infty)}(r) \quad \text{for all } r \ge 0. \qquad (4.45)$$

Then by (4.44), (4.45), dominated convergence, and Lebesgue's differentiation theorem, we have

$$\begin{aligned}
&\overline{\lim_{h\to 0+}} \frac{E[|Z(t+h)|_2^2] - E[|Z(t)|_2^2]}{h} \\
&= \overline{\lim_{h\to 0+}} E\left[\frac{1}{h}\int_t^{t+h} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s)>\tilde{M}\}} \left|\sigma^i(X_s)\right|_2^2\right) ds\right] \\
&\le \overline{\lim_{h\to 0+}} E\left[\lim_{h\to 0+}\frac{1}{h}\int_t^{t+h} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d f(X^i(s)) \left|\sigma^i(X_s)\right|_2^2\right) ds\right] \\
&= E\left[\lim_{h\to 0+}\frac{1}{h}\int_t^{t+h} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d f(X^i(s)) \left|\sigma^i(X_s)\right|_2^2\right) ds\right] \\
&= E\left[2(Z(t))'b(X_t) + \sum_{i=1}^d f(X^i(t)) \left|\sigma^i(X_t)\right|_2^2 ds\right] \\
&\le E\left[2(Z(t))'b(X_t) + \sum_{i}^d 1_{\{X^i(t)>M\}} \left|\sigma^i(X_t)\right|_2^2\right]. \qquad (4.46)
\end{aligned}$$

Here, we used the fact that the integrand in the third last line is a continuous function of s. According to Lemma 4.2.1, the last line above is strictly negative under the assumptions (4.19) and (4.20).

If there were no $h^* > 0$ such that $E[|Z(t+s)|_2^2] < E[|Z(t)|_2^2]$ for each $s \in (0, h^*]$, then we could construct a sequence $\{h_n\}_{n=1}^{\infty}$ of positive numbers decreasing to zero such that $E[|Z(t+h_n)|_2^2] \geq E[|Z(t)|_2^2]$ for all n. Then $\overline{\lim_{h \to 0+}} \frac{E[|Z(t+h)|_2^2] - E[|Z(t)|_2^2]}{h} \geq 0$, which is a contradiction to (4.46). Therefore there is an $h^* > 0$ such that (4.41) holds.

□

We now prove the main theorem of this subsection.

Proof of Theorem 4.2.1. First of all, as a consequence of Lemma 2.4.1, the continuity of Z, and the dominated convergence theorem, $E[|Z(s)|_2^2]$ is continuous as a function of $s \geq 0$.

Let $M_2 = \sup_{s \in [-\tau, \tau]} E[|Z(s)|_2^2] + M_1$, which is finite by Lemma 2.4.1 and the assumption that $E[\|X_0\|_2^2] < \infty$. If there was a $t_1 > \tau$ such that $E[|Z(t_1)|_2^2] > M_2$, then since $\sup_{s \in [-\tau, \tau]} E[|Z(s)|_2^2] < M_2$, $t := \inf\{s < t_1 : E[|Z(s)|_2^2] > M_2\}$ is a point of (τ, t_1). We also have $E[|Z(t)|_2^2] = M_2$ by continuity, and thus $E[|Z(r)|_2^2] \leq E[|Z(t)|_2^2]$ for all $r \in [t - 2\tau, t]$. Since $M_2 \geq M_1$, Lemma 4.2.2 implies that there is an $h^* > 0$ such that $E[|Z(s)|_2^2] < E[|Z(t)|_2^2] = M_2$ for all $s \in (t, t + h^*]$, but this contradicts the definition of t. Therefore, $\sup_{s \geq -\tau} E[|Z(s)|_2^2] \leq M_2$ which in turn implies that

$$\begin{aligned} \sup_{s \geq -\tau} E[|X(s)|_2^2] &\leq \sup_{s \geq -\tau} 2\left(E[|Z(s)|_2^2] + d\tilde{M}^2\right) \\ &\leq 2M_2 + 2d\tilde{M}^2. \end{aligned} \tag{4.47}$$

□

4.2.3 Uniform Bound on $E[\|X_t\|_2^2]$

Recall that we are assuming Assumption 4.2.1 holds.

Theorem 4.2.2. *Suppose that* $E[\|X_0\|_2^2] < \infty$. *Then,* $\sup_{t \geq 0} E\left[\|X_t\|_2^2\right] < \infty$.

Proof. Recall the definition of the overshoot process Z from Section 4.1.1 with $\tilde{M} = M + 1$. Theorem 4.2.1 implies that $\sup_{t \geq -\tau} E[|X(t)|_2^2] < \infty$.

For each $t \geq \tau$, by (4.6), (4.13)-(4.15), Proposition B.0.1, (4.17), and (4.26) with $\tilde{M}$

in place of M, we have

$$
\begin{aligned}
\|Z_t\|_2^2 &\leq \sum_{i=1}^{d} \left(Z^i(t-\tau) + \text{Osc}^+(Z^i,[t-\tau,t])\right)^2 \\
&\leq \sum_{i=1}^{d} \Bigg(Z^i(t-\tau) + \text{Osc}^+\left(\int_0^{\cdot} 1_{\{X^i(s)>\tilde{M}\}}(b^i(X_s))^+ ds, [t-\tau,t]\right) \\
&\qquad\qquad + \sup_{t-\tau\leq r<s\leq t}\left|\int_r^s 1_{\{X^i(s)>\tilde{M}\}}\sigma^i(X_u)dW(u)\right|\Bigg)^2 \\
&\leq 3\Bigg(\sum_{i=1}^{d}\left(Z^i(t-\tau)\right)^2 + \sum_{i=1}^{d}\left(\text{Osc}^+\left(\int_0^{\cdot} 1_{\{X^i(s)>\tilde{M}\}}(b^i(X_s))^+ ds, [t-\tau,t]\right)\right)^2 \\
&\qquad\qquad + \sum_{i=1}^{d}\left(\sup_{t-\tau\leq r<s\leq t}\left|\int_r^s 1_{\{X^i(s)>\tilde{M}\}}\sigma^i(X_u)dW(u)\right|\right)^2\Bigg) \\
&\leq 3\Bigg(|Z(t-\tau)|_2^2 + \sum_{i=1}^{d}\left(B_0\tau + B_{2,i}\int_{t-\tau}^{t}\int_{-\tau}^{0}|X(s+r)|_2^{q_1}\mu_1^i(dr)ds\right)^2 \\
&\qquad\qquad + \sum_{i=1}^{d} 4\sup_{t-\tau\leq s\leq t}\left|\int_{t-\tau}^{s} 1_{\{X^i(s)>\tilde{M}\}}\sigma^i(X_u)dW(u)\right|^2\Bigg). \qquad (4.48)
\end{aligned}
$$

Here, using Proposition B.0.1 and the Cauchy-Schwarz inequality, we have for each $i = 1, \ldots, d$,

$$
\begin{aligned}
&\left(B_0\tau + B_{2,i}\int_{t-\tau}^{t}\int_{-\tau}^{0}|X(s+r)|_2^{q_1}\mu_1^i(dr)ds\right)^2 \\
&\qquad \leq 2\left((B_0\tau)^2 + B_{2,i}^2\left(\int_{t-\tau}^{t}\int_{-\tau}^{0}|X(s+r)|_2^{q_1}\mu_1^i(dr)ds\right)^2\right) \\
&\qquad \leq 2\left((B_0\tau)^2 + B_{2,i}^2\tau\int_{t-\tau}^{t}\int_{-\tau}^{0}|X(s+r)|_2^{2q_1}\mu_1^i(dr)ds\right), \qquad (4.49)
\end{aligned}
$$

and by a similar argument to that for (4.27), using (4.18) we have for each i that

$$
\begin{aligned}
&E\left[\sup_{t-\tau\leq s\leq t}\left|\int_{t-\tau}^{s} 1_{\{X^i(s)>\tilde{M}\}}\sigma^i(X_u)dW(u)\right|^2\right] \\
&\qquad \leq 4\left(C_0\tau + C_{2,i}\int_{t-\tau}^{t}\int_{-\tau}^{0}E\left[|X(u+r)|_2^{q_2}\right]\mu_2^i(dr)du\right). \qquad (4.50)
\end{aligned}
$$

By Hölder's inequality, $E\left[|X(s)|_2^p\right] \leq E\left[|X(s)|_2^2\right]^{\frac{p}{2}}$ for all $s \geq -\tau$ and $0 < p \leq 2$. So by the hypotheses of the theorem and the fact that $r^{\frac{p}{2}} \leq 1 + r$ for all $r \geq 0$ and $0 < p \leq 2$, there is a constant $K > 0$ such that $\sup_{s\geq-\tau} E[|X(s)|_2^p] \leq K$ for all $0 < p \leq 2$. On combining the above and taking expectations in (4.48), we obtain for $t \geq \tau$,

$$\begin{aligned} E\left[\|Z_t\|_2^2\right] &\leq 3\Bigg(E[|Z(t-\tau)|_2^2] + 2d(B_0\tau)^2 \\ &\quad +2\sum_{i=1}^d B_{2,i}^2 \tau \int_{t-\tau}^t \int_{-\tau}^0 E\left[|X(s+r)|_2^{2q_1}\right]\mu_1^i(dr)ds \\ &\quad +16C_0\tau + 16\sum_{i=1}^d C_{2,i}\int_{t-\tau}^t \int_{-\tau}^0 E\left[|X(u+r)|_2^{q_2}\right]\mu_2^i(dr)du \Bigg) \\ &\leq 3\left(K + 2d(B_0\tau)^2 + 16C_0\tau d + 2\sum_{i=1}^d B_{2,i}^2\tau^2 K + 16C_0\tau + 16\sum_{i=1}^d C_{2,i}\tau K \right), \end{aligned}$$

which is a bound that is independent of $t \geq \tau$, so that

$$\sup_{t\geq\tau} E\left[\|Z_t\|_2^2\right] < \infty.$$

Therefore,

$$\begin{aligned} \sup_{t\geq\tau} E\left[\|X_t\|_2^2\right] &\leq 2\left(\sup_{t\geq\tau} E\left[\|Z_t\|_2^2\right] + d\tilde{M}^2 \right) \\ &< \infty, \end{aligned} \tag{4.51}$$

and thus for each $t \in [0, \tau]$,

$$E\left[\|X_t\|_2^2\right] \leq 2E\left[\|X_0\|_2^2 + \|X_\tau\|_2^2\right],$$

which is finite by (4.51) and the assumption that $E\left[\|X_0\|_2^2\right] < \infty$. □

4.3 Bounded Moments when b and σ Satisfy a Boundedness Assumption

Throughout this section, we assume that the coefficients b and σ satisfy Assumption 4.3.1 below.

4.3.1 Assumptions on b and σ

Assumption 4.3.1. *There exist non-negative constants K_u, M, strictly positive constants K_d, C_0, and a measurable function $\ell : \mathbb{C}_{\mathbb{I}}^d \to \mathbb{R}_+^d$, such that for each $x \in \mathbb{C}_{\mathbb{I}}^d$ and $i = 1, \ldots, d$, $\ell^i(x) \in x^i(\mathbb{I})$, and whenever $x^i(0) \geq M$, we have:*

(i) $b^i(x) \leq K_u 1_{[0,M]}(\ell^i(x)) - K_d 1_{[M,\infty)}(\ell^i(x))$, *and*

(ii) $|\sigma^i(x)|_2^2 \le C_0$.

Remark. Assumption 4.3.1 requires b^i and $|\sigma^i|_2$ to be bounded above on the set $\{x \in \mathbb{C}_{\mathbb{I}}^d : x^i(0) \ge M\}$, but this does not necessarily imply that they are bounded above on $\mathbb{C}_{\mathbb{I}}^d$. For instance, for $\tau = d = 1$, if we define the function $f : \mathbb{R}_+ \to [0,1]$ by $f(r) := (2-r)^+ - (1-r)^+$, then the drift coefficient

$$b(x) \;:=\; \|x\| f(x(0)) + (1 - x(-1))\,(1 - f(x(0)))$$

is continuous and satisfies part (i) with $M = 2$, $K_u = 1$, $K_d = 1$, and $\ell(x) = x(-1)$, yet it is unbounded on $\mathbb{C}_{\mathbb{I}}^d$: if for each $n \ge 1$, $x^{(n)} \in \mathbb{C}_{\mathbb{I}}^d$ is defined by $x^{(n)}(r) = 1 - nr$, $r \in [-1, 0]$, then $b(x^{(n)}) = n + 1$. Similarly, one may construct a continuous unbounded dispersion coefficient, σ, satisfying part (ii).

Remark. Note that Assumption 4.3.1 has no restrictions on the size of the constants M, K_u, K_d, C_0 (beyond strict positivity of K_d and C_0), cf. part (iii) of Assumption 4.2.1.

Recall the overshoot process Z defined in (4.1), where here we let $\tilde{M} = M$. In Theorem 4.3.1 below, for each $i = 1, \ldots, d$, we will use $E\left[(Z^i(t))^2 \exp(\alpha Z^i(t))\right]$ as a Lyapunov-type function to show that $E[\exp(\alpha X^i(t))]$ is bounded uniformly in $t \ge 0$, for some $\alpha > 0$.

4.3.2 Preliminary Lemma

A key role in the proof of Theorem 4.3.1 is played by the following proposition proved by Itô and Nisio ([20], Lemmas 8.1 and 8.2).

Proposition 4.3.1. *Assume that $f, g \in C([0,\infty), \mathbb{R})$ such that $f(0) \ge 0$ and $a_1, a_2, a_3 > 0$.*

(i) If $f(t) \le f(s) - a_1 \int_s^t f(u)du + \int_s^t g(u)du$ for all $0 \le s < t \le \infty$, then

$$f(t) \le f(0) + \int_0^t e^{-a_1(t-u)} g(u)du, \text{ for all } t \ge 0.$$

(ii) If $g(t) \le a_1 + a_2 \int_0^t e^{-a_3(t-u)} g(u)du$ for all $t \ge 0$, and $a_3 > a_2$, then

$$g(t) \le \frac{a_1 a_3}{a_3 - a_2} \text{ for all } t \ge 0.$$

Remark. This proposition allows us to use an analytical technique that Itô and Nisio developed. If for each $\varepsilon > 0$, there is a constant $K_\varepsilon \ge 0$ such that $g(t) \le K_\varepsilon + \varepsilon f(t)$, then the two parts of the proposition can be combined, and the collective result can be used in a manner similar to that in which Gronwall's inequality is often used. This suits our Lyapunov-type argument, where the role of $f(t)$ is played by $E[(Z^i(t))^2\exp(\alpha Z^i(t))]$, and that of $g(t)$ by $E[\exp(\alpha Z^i(t))]$. The details are in the proof of Theorem 4.3.1 below.

4.3.3 Uniform Bound on Exponential Moments of $X^i(t)$

The following theorem depends on some technical lemmas that are deferred until after the proof of the theorem.

Theorem 4.3.1. *Suppose that* $E[\exp(\kappa\|X_0\|_2)] < \infty$ *for each* $\kappa > 0$*. Then there exists* $\alpha > 0$ *such that* $\sup_{t\geq 0} E[\exp(\alpha X^i(t))] < \infty$ *for each* $i = 1, \ldots, d$*, and consequently,* $\sup_{t\geq -\tau} E[|X(t)|_p^p] < \infty$ *for all* $p \in [1, \infty)$.

Proof. Fix $i \in \{1, \ldots, d\}$. We derive the differential of the process $\{(Z^i(t))^2\exp(\alpha Z^i(t)), t \geq 0\}$. Itô's formula together with (4.2) yield for each $\alpha > 0$ and $t \geq 0$,

$$\begin{aligned} & d\left(\exp(\alpha Z^i(t))\right) \\ = \; & \alpha\exp\left(\alpha Z^i(t)\right) 1_{\{X^i(t)>M\}} b^i(X_t)dt + \alpha\exp\left(\alpha Z^i(t)\right) 1_{\{X^i(t)>M\}}\sigma^i(X_t)dW(t) \\ & + \frac{\alpha^2}{2}\exp\left(\alpha Z^i(t)\right) 1_{\{X^i(t)>M\}}\left|\sigma^i(X_t)\right|_2^2 dt + \alpha\exp\left(\alpha Z^i(t)\right) dL^i(t). \end{aligned} \tag{4.52}$$

Combining (4.52) with the differential (4.3) of $(Z^i(t))^2$, we obtain for each $i = 1, \ldots, d$ and $t \geq 0$,

$$\begin{aligned} & d\left((Z^i(t))^2\exp(\alpha Z^i(t))\right) \\ = \; & (Z^i(t))^2 d\left(\exp(\alpha Z^i(t))\right) + \exp(\alpha Z^i(t)) d\left((Z^i(t))^2\right) + d\left\langle (Z^i)^2, \exp(\alpha Z^i)\right\rangle(t) \\ = \; & \alpha\left(Z^i(t)\right)^2\exp\left(\alpha Z^i(t)\right) b^i(X_t)dt + \alpha\left(Z^i(t)\right)^2\exp\left(\alpha Z^i(t)\right)\sigma^i(X_t)dW(t) \\ & + \frac{\alpha^2}{2}(Z^i(t))^2\exp\left(\alpha Z^i(t)\right)\left|\sigma^i(X_t)\right|_2^2 dt + \alpha(Z^i(t))^2\exp\left(\alpha Z^i(t)\right) dL(t) \\ & + 2Z^i(t)\exp\left(\alpha Z^i(t)\right) b^i(X_t)dt \\ & + 2Z^i(t)\exp\left(\alpha Z^i(t)\right)\sigma^i(X_t)dW(t) \\ & + 1_{\{X^i(t)>M\}}\left|\sigma^i(X_t)\right|_2^2\exp\left(\alpha Z^i(t)\right) dt + 2Z^i(t)\alpha\exp\left(\alpha Z^i(t)\right)\left|\sigma^i(X_t)\right|_2^2 dt \\ = \; & \alpha(Z^i(t))^2\left(b^i(X_t) + \frac{\alpha\left|\sigma^i(X_t)\right|_2^2}{2}\right)\exp\left(\alpha Z^i(t)\right) dt \\ & + 2Z^i(t)\left(b^i(X_t) + \alpha\left|\sigma^i(X_t)\right|_2^2\right)\exp(\alpha Z^i(t))dt \\ & + 1_{\{X^i(t)>M\}}\left|\sigma^i(X_t)\right|_2^2\exp(\alpha Z^i(t))dt \\ & + \left(\alpha\left(Z^i(t)\right)^2 + 2Z^i(t)\right)\exp(\alpha Z^i(t))\sigma^i(X_t)dW(t). \end{aligned} \tag{4.53}$$

Here we have used the facts that $Z^i(t) = 0$ on $\{X^i(t) \le M\}$ and that L^i can increase only when Z^i is zero, which implies that the term $\alpha(Z^i(t))^2 \exp\left(\alpha Z^i(t)\right) dL(t)$ contributes zero. Assumption 4.3.1 can now be applied to the coefficients b^i and σ^i in this last expression because each appears only when multiplied by something that is zero when $X^i(t) < M$.

For $t \ge 0$, since $X^i(t) > M$ if and only if $Z^i(t) > 0$, using (4.53), Assumption 4.3.1, and Lemma 4.3.1 below, we have that for any choice of $\gamma > 0$,

$$\begin{aligned}
& d\left(\left(Z^i(t)\right)^2 \exp\left(\alpha Z^i(t)\right)\right) \\
\le\ & \alpha(Z^i(t))^2 \left((K_u + K_d)e^{\gamma M} e^{-\gamma X^i(t)+\gamma \operatorname{Osc}^+(X^i_t,\mathbb{I})} - K_d + \frac{\alpha C_0}{2}\right) \exp(\alpha Z^i(t))dt \\
& +2Z^i(t)\left((K_u + K_d)e^{\gamma M} e^{-\gamma X^i(t)+\gamma \operatorname{Osc}^+(X^i_t,\mathbb{I})} - K_d + \alpha C_0\right) \exp(\alpha Z^i(t))dt \\
& +C_0 \exp(\alpha Z^i(t))dt + \left(\alpha(Z^i(t))^2 + 2Z^i(t)\right) \exp(\alpha Z^i(t))\sigma^i(X_t)dW(t) \qquad (4.54) \\
\le\ & \alpha(K_u + K_d)e^{\gamma M}(Z^i(t))^2 \exp\left(\alpha X^i(t) - \gamma X^i(t) + \gamma \operatorname{Osc}^+(X^i_t,\mathbb{I})\right) dt \\
& + \left(\frac{\alpha C_0}{2} - K_d\right) \alpha(Z^i(t))^2 \exp(\alpha Z^i(t))dt \\
& +2(K_u + K_d)e^{\gamma M} Z^i(t) \exp\left(\alpha X^i(t) - \gamma X^i(t) + \gamma \operatorname{Osc}^+(X^i_t,\mathbb{I})\right) dt \\
& +2(\alpha C_0 - K_d)Z^i(t) \exp(\alpha Z^i(t))dt \\
& +C_0 \exp(\alpha Z^i(t))dt + \left(\alpha(Z^i(t))^2 + 2Z^i(t)\right) \exp(\alpha Z^i(t))\sigma^i(X_t)dW(t), \qquad (4.55)
\end{aligned}$$

where the second inequality follows from the facts that $X^i(t) \ge Z^i(t)$ for all $t \ge 0$ and that the exponential function is increasing.

Assume $\gamma > \alpha$, and consider the functions $f_1(r) = r\exp(-(\gamma-\alpha)r)$ and $f_2(r) = r^2 \exp(-(\gamma-\alpha)r)$. Both f_1 and f_2 are bounded on $\mathbb{R}_+$; in fact, $f_1(r) \le f_1\left(\frac{1}{\gamma-\alpha}\right) = \frac{1}{(\gamma-\alpha)e}$ and $f_2(r) \le f_2\left(\frac{2}{\gamma-\alpha}\right) = \frac{4}{(\gamma-\alpha)^2e^2}$, for all $r \ge 0$. Also, $Z^i(t) \le X^i(t)$ for all $t \ge 0$, so for $k = 1, 2$, $(Z^i(t))^k \exp(-(\gamma-\alpha)X^i(t)) \le f_k(X^i(t))$.

Therefore, for $t \ge 0$,

$$\begin{aligned}
& d\left(\left(Z^i(t)\right)^2 \exp\left(\alpha Z^i(t)\right)\right) \\
\le\ & \alpha\left(\frac{\alpha}{2}C_0 - K_d\right)(Z^i(t))^2 \exp(\alpha Z^i(t))dt \\
& +\alpha(K_u + K_d)e^{\gamma M}\frac{4}{(\gamma-\alpha)^2e^2} \exp\left(\gamma \operatorname{Osc}^+(X^i_t,\mathbb{I})\right) dt \\
& +2(\alpha C_0 - K_d)Z^i(t) \exp\left(\alpha Z^i(t)\right) dt \\
& +2(K_u + K_d)e^{\gamma M}\frac{1}{(\gamma-\alpha)e} \exp\left(\gamma \operatorname{Osc}^+(X^i_t,\mathbb{I})\right) dt \\
& +C_0 \exp\left(\alpha Z^i(t)\right) dt + \left(\alpha\left(Z^i(t)\right)^2 + 2Z^i(t)\right) \exp\left(\alpha Z^i(t)\right) \sigma^i(X_t)dW(t). \qquad (4.56)
\end{aligned}$$

Set

$$K_1 = \left(\frac{4\alpha(K_u+K_d)e^{\gamma M}}{(\gamma-\alpha)^2e^2} + \frac{2(K_u+K_d)e^{\gamma M}}{(\gamma-\alpha)e}\right) = \frac{2(K_u+K_d)e^{\gamma M}}{(\gamma-\alpha)e}\left(\frac{2\alpha}{(\gamma-\alpha)e}+1\right).$$

If we choose $\alpha \in \left(0, \frac{K_d}{C_0}\right)$, then $\beta := \alpha(K_d - \frac{\alpha}{2}C_0) > 0$, and we obtain

$$\begin{aligned} d\left(\left(Z^i(t)\right)^2 \exp\left(\alpha Z^i(t)\right)\right) \\ \leq\ & -\beta(Z^i(t))^2\exp(\alpha Z^i(t))dt + K_1\exp\left(\gamma \text{Osc}^+\left(X_t^i, \mathbb{I}\right)\right)dt \\ & + C_0\exp\left(\alpha Z^i(t)\right)dt + Z^i(t)\exp\left(\alpha Z^i(t)\right)\left(\alpha Z^i(t)+2\right)\sigma^i(X_t)dW(t). \quad (4.57) \end{aligned}$$

Here, we used the fact that $(\alpha C_0 - K_d)Z^i(t)\exp\left(\alpha Z^i(t)\right) \leq 0$ for all $t \geq 0$.

Using the fact that there exists a $\kappa_1 > \alpha$ such that $r^2(\alpha r+2)^2e^{2\alpha r} \leq e^{\kappa_1 r}$ for all $r \geq 0$, we obtain on using Assumption 4.3.1(ii) that for each $t \geq 0$,

$$\begin{aligned} E\left[\int_0^t (Z^i(s))^2\exp(2\alpha Z^i(s))(\alpha Z^i(s)+2)^2\left|\sigma^i(X_s)\right|_2^2 ds\right] &\leq E\left[\int_0^t C_0\exp(\kappa_1 Z^i(s))ds\right] \\ &\leq tC_0E\left[\sup_{0\leq s\leq t}\exp(\kappa_1 Z^i(s))\right], \end{aligned}$$

which is finite by Lemma 4.3.3 below and the assumption that $E\left[\exp(\kappa\|X_0\|_2)\right] < \infty$ for all $\kappa > 0$. It follows that $\left\{\int_0^t Z^i(s)\exp(\alpha Z^i(s))(\alpha Z^i(s)+2)\sigma^i(X_s)dW(s), \mathcal{F}_t, t \geq 0\right\}$ is a square-integrable martingale. Thus, integrating in time and taking expectations in (4.57), we obtain for each $t, h \geq 0$,

$$\begin{aligned} E\left[(Z^i(t+h))^2\exp(\alpha Z^i(t+h))\right] \leq\ & E\left[(Z^i(t))^2\exp(\alpha Z^i(t))\right] \\ & -\beta\int_t^{t+h} E\left[(Z^i(s))^2\exp(\alpha Z^i(s))\right]ds \\ & +\int_t^{t+h} C_0E[\exp(\alpha Z^i(s))]ds \\ & +\int_t^{t+h} K_1E\left[\exp(\gamma\text{Osc}^+(X_t^i, \mathbb{I}))\right]ds \\ \leq\ & E\left[(Z^i(t))^2\exp(\alpha Z^i(t))\right] \\ & -\beta\int_t^{t+h} E\left[(Z^i(s))^2\exp(\alpha Z^i(s))\right]ds \\ & +\int_t^{t+h}\left(K_1K(\gamma) + C_0E[\exp(\alpha Z^i(s))]\right)ds, \quad (4.58) \end{aligned}$$

where we have used Lemma 4.3.3 for the finiteness of $E[(Z^i(s))^2\exp(\alpha Z^i(s))]$ and Lemma 4.3.4 for the last inequality. By Lemma 4.3.3 and dominated convergence, the functions $f(t) :=$

$E[(Z^i(t))^2 \exp(\alpha Z^i(t))]$ and $g(t) := K_1 K(\gamma) + C_0 E[\exp(\alpha Z^i(t))]$ are continuous in $t \geq 0$. Then, part (i) of Proposition 4.3.1 implies that

$$\begin{aligned} & E\left[(Z^i(t))^2 \exp(\alpha Z^i(t))\right] \\ \leq \quad & E[(Z^i(0))^2 \exp(\alpha Z^i(0))] + \int_0^t \exp(-\beta(t-u)) \left(C_0 E\left[\exp(\alpha Z^i(u))\right] + K_1 K(\gamma)\right) du \\ \leq \quad & K_2 + C_0 \int_0^t \exp(-\beta(t-u)) E[\exp(\alpha Z^i(u))] du, \end{aligned} \tag{4.59}$$

where $K_2 = E[(Z^i(0))^2 \exp(\alpha Z^i(0))] + \frac{K_1 K(\gamma)}{\beta} \leq E[\exp((2+\alpha)Z^i(0))] + \frac{K_1 K(\gamma)}{\beta}$, by Proposition B.0.5, and this is finite by hypothesis. Using the fact that $\exp(ar) \leq \exp(\frac{a}{\sqrt{\varepsilon}}) + \varepsilon r^2 \exp(ar)$ for any $a, r \geq 0, \varepsilon > 0$ (as can be checked by considering the two cases when $r > \frac{1}{\sqrt{\varepsilon}}$ or $r \leq \frac{1}{\sqrt{\varepsilon}}$), we obtain that for each $t \geq 0$,

$$\begin{aligned} E[\exp(\alpha Z^i(t))] \quad & \leq \quad e^{\frac{\alpha}{\sqrt{\varepsilon}}} + \varepsilon E[(Z^i(t))^2 \exp(\alpha Z^i(t))] \\ & \leq \quad e^{\frac{\alpha}{\sqrt{\varepsilon}}} + \varepsilon K_2 + \varepsilon C_0 \int_0^t \exp(-\beta(t-u)) E[\exp(\alpha Z^i(u))] du. \end{aligned} \tag{4.60}$$

For $\varepsilon < \frac{\beta}{C_0}$, part (ii) of Proposition 4.3.1 yields

$$E[\exp(\alpha Z^i(t))] \quad \leq \quad \frac{(e^{\frac{\alpha}{\sqrt{\varepsilon}}} + \varepsilon K_2)\beta}{\beta - \varepsilon C_0}, \tag{4.61}$$

and therefore,

$$E[\exp(\alpha X^i(t))] \quad \leq \quad e^{\alpha M} \frac{(e^{\frac{\alpha}{\sqrt{\varepsilon}}} + \varepsilon K_2)\beta}{\beta - \varepsilon C_0}, \tag{4.62}$$

and hence $\sup_{t \geq 0} E[\exp(\alpha X^i(t))] < \infty$ for each $i = 1, \ldots, d$. By considering the Taylor expansion of $\exp(\alpha r)$, we can see that for each $r \in \mathbb{R}_+$ and positive integer n, $r^n \leq \frac{n!}{\alpha^n} \exp(\alpha r)$, and thus it follows from (4.62) and Hölder's inequality that for each $p \geq 1$ and $i = 1, \ldots, d$,

$$\sup_{t \geq 0} E[(X^i(t))^p] \quad < \quad \infty, \tag{4.63}$$

and the fact that $\sup_{t \geq -\tau} E[|X(t)|_p^p] < \infty$ follows.

□

4.3.4 Supporting Lemmas

We now prove the supporting lemmas.

Lemma 4.3.1. *For each* $\gamma > 0$ *and each* $x \in \mathbb{C}^d_{\mathbb{I}}$ *with* $x^i(0) \geq M$*, we have*

$$b^i(x) \quad \leq \quad (K_u + K_d)e^{\gamma M} \exp\left(-\gamma x^i(0) + \gamma \,\text{Osc}^+(x^i, \mathbb{I})\right) - K_d. \tag{4.64}$$

Proof. Let $x \in \mathbb{C}^d_{\mathbb{I}}$ with $x^i(0) \geq M$. Since $\ell^i(x) \in x^i(\mathbb{I})$, there is a $r_x \in \mathbb{I}$ such that $\ell^i(x) = x^i(r_x)$, and thus $x^i(0) \leq x^i(r_x) + \text{Osc}^+(x^i, \mathbb{I}) = \ell^i(x) + \text{Osc}^+(x^i, \mathbb{I})$ by the definition of Osc^+. Therefore,

$$-\ell^i(x) \quad \leq \quad -x^i(0) + \text{Osc}^+(x^i, \mathbb{I}). \tag{4.65}$$

From Assumption 4.3.1(i), it follows that for each $\gamma > 0$,

$$\begin{aligned}
b^i(x) \quad &\leq \quad K_u 1_{[0,M]}(\ell^i(x)) - K_d 1_{[M,\infty)}(\ell^i(x)) \\
&= \quad (K_u + K_d) 1_{[0,M]}(\ell^i(x)) - K_d \\
&\leq \quad (K_u + K_d) \exp\left(-\gamma(\ell^i(x) - M)\right) - K_d \\
&= \quad (K_u + K_d)e^{\gamma M} \exp(-\gamma \ell^i(x)) - K_d \\
&\leq \quad (K_u + K_d)e^{\gamma M} \exp\left(-\gamma x^i(0) + \gamma \,\text{Osc}^+(x^i, \mathbb{I})\right) - K_d.
\end{aligned} \tag{4.66}$$

□

The next lemma follows from basic growth estimates on solutions to stochastic (undelayed) differential equations with coefficients that grow at most linearly. This lemma only uses (ii) of Assumption 4.3.1.

Lemma 4.3.2. *For each* $t \geq 0$ *and* $i = 1, \ldots, d$*, define the process*

$$\left\{ \xi^{t,i}(s) := \exp\left(\int_{(t-\tau)^+}^{(t-\tau)^+ + s} 1_{\{X^i(u) > M\}} \sigma^i(X_u) dW(u) \right), s \geq 0 \right\},$$

where σ *satisfies part (ii) of Assumption 4.3.1. Then, there exists a function* $K : (0, \infty) \times \mathbb{R}_+ \to \mathbb{R}_+$ *independent of* t *and* i*, which can be chosen to be non-decreasing in each coordinate, such that for each* $p > 0$ *and* $T \geq 0$*,*

$$E\left[\|\xi^{t,i}\|^p_{[0,T]} \right] \vee E\left[\|(\xi^{t,i})^{-1}\|^p_{[0,T]} \right] \leq K(p, T).$$

Proof. For each $t \geq 0$, Itô's formula tells us that the process $\xi^{t,i}$ satisfies the stochastic differential equation (SDE)

$$\begin{aligned} d\xi^{t,i}(s) &= 1_{\{X^i((t-\tau)^+ +s)>M\}}\xi^{t,i}(s)\sigma^i(X_{(t-\tau)^+ +s})dW((t-\tau)^+ + s) \\ &\quad + \frac{|\sigma^i(X_{(t-\tau)^+ +s})|_2^2}{2} 1_{\{X^i((t-\tau)^+ +s)>M\}}\xi^{t,i}(s)ds, \quad s \geq 0, \end{aligned} \tag{4.67}$$

with initial condition $\xi^{t,i}(0) \equiv 1$. The coefficients satisfy linear growth conditions. Indeed,

$$\left|1_{\{X^i((t-\tau)^+ +s)>M\}}\xi^{t,i}(s)\sigma^i(X_{(t-\tau)^+ +s})\right|_2 \leq \sqrt{C_0}|\xi^{t,i}(s)|, \quad \text{and}$$

$$\left|\frac{1_{\{X^i((t-\tau)^+ +s)>M\}}|\sigma^i(X_{(t-\tau)^+ +s})|_2^2}{2}\xi^{t,i}(s)\right| \leq \frac{1}{2}C_0|\xi^{t,i}(s)|.$$

Therefore by the proof of Theorem V.12.1 in [40] (which is similar to the proof of Lemma 2.4.1), for each $p \geq 2$, $T \geq 0$, there is a $K(p,T) > 0$ such that $E\left[\|\xi^{t,i}\|_{[0,T]}^p\right] \leq K(p,T) < \infty$ for all t and i, since $K(p,T)$ depends only on the growth constants and initial condition, which are independent of t and i. Hölder's inequality can now be used to obtain the same result for $p \in (0,2]$. The process

$$\left\{\left(\xi^{t,i}(s)\right)^{-1} = \exp\left(\int_{(t-\tau)^+}^{(t-\tau)^+ +s} -1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right), s \geq 0\right\}$$

satisfies a similar SDE whose coefficients satisfy the same growth bounds, and thus for each $p > 0$ and $T \geq 0$,

$$E\left[\|(\xi^{t,i})^{-1}\|_{[0,T]}^p\right] \leq K(p,T) < \infty.$$

□

The following lemma is somewhat of an exponential analogue to Lemma 2.4.1 in the case that σ^i is bounded on $\{x \in \mathbb{C}_{\mathbb{I}}^d : x^i(0) \geq M\}$. It's validity should not be a surprise after consideration of Theorem 4.7 in [26].

Lemma 4.3.3. *Suppose* $E[\exp(\alpha\|X_0\|_2)] < \infty$ *for each* $\alpha > 0$. *Then for each* $T \geq 0$ *and* $\alpha > 0$,

$$E\left[\|\exp(\alpha|X(\cdot)|_1)\|_{[-\tau,T]}\right] < \infty.$$

Remark. This lemma remains valid without assuming part (i) of Assumption 4.3.1 as long as the linear growth condition (2.1) and Assumption 4.3.1(ii) hold.

Proof. Fix $\alpha > 0$. Define the stopping times $\eta_n := \inf\{t \geq 0 : \|X\|_{[-\tau,t],2} \geq n\}$, with the convention that $\inf \emptyset = \infty$.

Proposition B.0.3 implies that for each $T \geq 0$,

$$\begin{aligned}
&E\left[\|\exp(\alpha|X(\cdot)|_1)\|_{[-\tau,T\wedge\eta_n]}\right] \\
&\quad= E\left[\exp(\alpha\|X\|_{[-\tau,T\wedge\eta_n],1})\right] \\
&\quad\leq E\left[\exp\left(\alpha\|X^1\|_{[-\tau,T\wedge\eta_n]}\right)\cdots\exp\left(\alpha\|X^d\|_{[-\tau,T\wedge\eta_n]}\right)\right] \\
&\quad\leq \frac{1}{d}\left(E\left[\exp(\alpha d\|X^1\|_{[-\tau,T\wedge\eta_n]})\right]+\cdots+E\left[\exp(\alpha d\|X^d\|_{[-\tau,T\wedge\eta_n]})\right]\right). \qquad (4.68)
\end{aligned}$$

Since $X^i(t) \leq M + Z^i(t)$ for each $t \geq -\tau$ and $i = 1, \ldots, d$,

$$\begin{aligned}
E\left[\exp(\alpha d\|X^i\|_{[-\tau,T\wedge\eta_n]})\right] &= E\left[\|\exp(\alpha dX^i(\cdot))\|_{[-\tau,T\wedge\eta_n]}\right] \\
&\leq e^{\alpha dM}E\left[\|\exp(\alpha dZ^i(\cdot))\|_{[-\tau,T\wedge\eta_n]}\right] \\
&\leq e^{\alpha dM}\left(E\left[\exp(\alpha d\|Z_0^i\|)\right]+E\left[\|\exp(\alpha dZ^i(\cdot))\|_{[0,T\wedge\eta_n]}\right]\right).
\end{aligned}$$

Convexity of the exponential function implies that for any $a_1, a_2, a_3, a_4 \in \mathbb{R}$ and $\kappa \in (0, 1)$, we have

$$\begin{aligned}
&\exp\left(\frac{1-\kappa}{3}a_1+\kappa a_2+\frac{1-\kappa}{3}a_3+\frac{1-\kappa}{3}a_4\right) \\
&\qquad\leq \frac{1-\kappa}{3}\exp(a_1)+\kappa\exp(a_2)+\frac{1-\kappa}{3}\exp(a_3)+\frac{1-\kappa}{3}\exp(a_4),
\end{aligned}$$

which then implies that

$$\begin{aligned}
\exp(a_1+a_2+a_3+a_4) &\leq \frac{1-\kappa}{3}\exp\left(\frac{3}{1-\kappa}a_1\right)+\kappa\exp\left(\frac{a_2}{\kappa}\right) \\
&\quad+\frac{1-\kappa}{3}\exp\left(\frac{3}{1-\kappa}a_3\right)+\frac{1-\kappa}{3}\exp\left(\frac{3}{1-\kappa}a_4\right). \qquad (4.69)
\end{aligned}$$

Since (4.2) holds, as in the proof of Lemma 4.1.2, we use Lemma 4.1.1 to conclude that

$$\begin{aligned}
\mathrm{Osc}^+(Z^i,[0,T\wedge\eta_n]) &\leq \int_0^{T\wedge\eta_n} 1_{\{X^i(u)>M\}}(b^i(X_u))^+du \\
&\qquad+\sup_{0\leq r\leq s\leq T\wedge\eta_n}\int_r^s 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u) \\
&\leq \int_0^{T\wedge\eta_n}|b(X_u)|_2du+\sup_{0\leq s\leq T\wedge\eta_n}\int_0^s 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u) \\
&\qquad+\sup_{0\leq r\leq T\wedge\eta_n}\int_0^r -1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u), \qquad (4.70)
\end{aligned}$$

where we used the fact that $|b^i(X_u)| \leq |b(X_u)|_2$.

The linear growth condition (2.1) and Jensen's inequality imply that for any $\beta > 0$ and each $T > 0$,

$$\begin{aligned}
&\exp\left(\beta \int_0^T 1_{\{\|X\|_{[-\tau,t],2}<n\}}|b(X_t)|_2 dt\right) \\
&\quad \leq \exp\left(\beta \int_0^T 1_{\{\|X\|_{[-\tau,t],2}<n\}}(C_1 + C_2\|X_t\|_2)dt\right) \\
&\quad \leq \frac{1}{T}\int_0^T \exp\left(T\beta 1_{\{\|X\|_{[-\tau,t],2}<n\}}(C_1 + C_2\|X_t\|_2)\right) dt \\
&\quad \leq \frac{1}{T}\int_0^T \|\exp\left(T\beta(C_1 + C_2|X(\cdot)|_2)\right)\|_{[-\tau,t\wedge\eta_n]}\, dt. \qquad (4.71)
\end{aligned}$$

Therefore, for any $\kappa \in (0,1)$, $T > 0$, $i = 1, \ldots, d$,

$$\begin{aligned}
&E\left[\|\exp(\alpha d Z^i(\cdot))\|_{[0,T\wedge\eta_n]}\right] \\
&\leq E\left[\exp\left(\alpha d\left(Z^i(0) + \text{Osc}^+(Z^i,[0,T\wedge\eta_n])\right)\right)\right] \\
&\leq E\left[\exp\left(\alpha d\left(Z^i(0) + \int_0^{T\wedge\eta_n} |b(X_t)|_2 dt\right.\right.\right. \\
&\qquad \left.\left.\left. + \sup_{s\in[0,T\wedge\eta_n]}\int_0^s 1_{\{X^i(t)>M\}}\sigma^i(X_t)dW(t) + \sup_{s\in[0,T\wedge\eta_n]}\int_0^s -1_{\{X^i(t)>M\}}\sigma^i(X_t)dW(t)\right)\right)\right] \\
&\leq \frac{1-\kappa}{3}E\left[\exp\left(\frac{3\alpha d}{1-\kappa}Z^i(0)\right)\right] \\
&\quad +\kappa E\left[\exp\left(\frac{\alpha d}{\kappa}\int_0^T 1_{\{\|X\|_{[-\tau,t],2}<n\}}|b(X_t)|_2 dt\right)\right] \\
&\quad +\frac{1-\kappa}{3}E\left[\exp\left(\frac{3\alpha d}{1-\kappa}\sup_{s\in[0,T]}\int_0^s 1_{\{X^i(t)>M\}}\sigma^i(X_t)dW(t)\right)\right] \\
&\quad +\frac{1-\kappa}{3}E\left[\exp\left(\frac{3\alpha d}{1-\kappa}\sup_{s\in[0,T]}\int_0^s -1_{\{X^i(t)>M\}}\sigma^i(X_t)dW(t)\right)\right] \\
&\leq \frac{1-\kappa}{3}E\left[\exp\left(\frac{3\alpha d}{1-\kappa}\|X_0\|_2\right)\right] \\
&\quad +\kappa E\left[\frac{1}{T}\int_0^T \left\|\exp\left(T\frac{\alpha d}{\kappa}(C_1 + C_2|X(\cdot)|_2)\right)\right\|_{[-\tau,t\wedge\eta_n]} dt\right] \\
&\quad +\frac{1-\kappa}{3}E\left[\sup_{s\in[0,T]}\exp\left(\frac{3\alpha d}{1-\kappa}\int_0^s 1_{\{X^i(t)>M\}}\sigma^i(X_t)dW(t)\right)\right] \\
&\quad +\frac{1-\kappa}{3}E\left[\sup_{s\in[0,T]}\exp\left(\frac{-3\alpha d}{1-\kappa}\int_0^s 1_{\{X^i(t)>M\}}\sigma^i(X_t)dW(t)\right)\right]. \qquad (4.72)
\end{aligned}$$

Lemma 4.3.2 (with $t = 0$) along with (4.68) and (B.3) now imply that for each $T > 0$

and $\kappa \in (0,1)$,

$$
\begin{aligned}
& e^{-\alpha dM} E\left[\|\exp(\alpha|X(\cdot)|_1)\|_{[-\tau,T\wedge\eta_n]}\right] \\
& \quad \leq \; E[\exp(\alpha d\|X_0\|_2)] + \frac{1-\kappa}{3} E\left[\exp\left(\frac{3\alpha d}{1-\kappa}\|X_0\|_2\right)\right] \\
& \qquad + \frac{\kappa}{T}\exp\left(T\frac{\alpha d}{\kappa}C_1\right)\int_0^T E\left[\left\|\exp\left(T\frac{\alpha d C_2}{\kappa}|X(\cdot)|_1\right)\right\|_{[-\tau,t\wedge\eta_n]}\right] dt \\
& \qquad + \frac{1-\kappa}{3}\left(K\left(\frac{3\alpha d}{1-\kappa},T\right) + K\left(\frac{3\alpha d}{1-\kappa},T\right)\right). \qquad (4.73)
\end{aligned}
$$

If $T \in \left(0, \frac{1}{2dC_2}\right]$, we can set $\kappa = TdC_2 \in \left(0, \frac{1}{2}\right]$ and then we obtain for each $T \in \left(0, \frac{1}{2dC_2}\right]$,

$$
\begin{aligned}
& E\left[\|\exp(\alpha|X(\cdot)|_1)\|_{[-\tau,T\wedge\eta_n]}\right] \\
& \qquad \leq \; K_0(\alpha) + K_1(\alpha)\int_0^T E\left[\|\exp\left(\alpha|X(\cdot)|_1\right)\|_{[-\tau,t\wedge\eta_n]}\right] dt, \qquad (4.74)
\end{aligned}
$$

where

$$
\begin{aligned}
K_0(\alpha) &= e^{\alpha dM} E[\exp(\alpha d\|X_0\|_2)] + e^{\alpha dM}\frac{1}{3}E\left[\exp\left(6\alpha d\|X_0\|_2\right)\right] \\
&\quad + e^{\alpha dM}\frac{2}{3}K\left(6\alpha d, \frac{1}{2dC_2}\right), \text{ and} \\
K_1(\alpha) &= de^{\alpha dM}C_2\exp\left(\frac{\alpha C_1}{C_2}\right).
\end{aligned}
$$

Inequality (4.74) is obvious for $T=0$ because of inequality (B.6).

The assumptions imply $K_0(\alpha) < \infty$ and $K_1(\alpha) > 0$. Therefore, since

$$
\begin{aligned}
\|\exp(\alpha|X(\cdot)|_1)\|_{[-\tau,t\wedge\eta_n]} &\leq \left\|\exp(\alpha d^{\frac{1}{2}}|X(\cdot)|_2)\right\|_{[-\tau,t\wedge\eta_n]} \\
&\leq \exp(\alpha d^{\frac{1}{2}}\|X_0\|_2) + \exp(\alpha d^{\frac{1}{2}}n), \qquad (4.75)
\end{aligned}
$$

so that the expectation on the left of (4.74) is finite, Gronwall's inequality implies that

$$
E\left[\|\exp(\alpha|X(\cdot)|_1)\|_{\left[-\tau,\frac{1}{2dC_2}\wedge\eta_n\right]}\right] \leq K_0(\alpha)\exp\left(K_1(\alpha)\frac{1}{2dC_2}\right). \qquad (4.76)
$$

The monotone convergence theorem can then be applied to let $n\to\infty$ and obtain for each $\alpha > 0$,

$$
E\left[\|\exp(\alpha|X(\cdot)|_1)\|_{\left[-\tau,\frac{1}{2dC_2}\right]}\right] \leq K_0(\alpha)\exp\left(K_1(\alpha)\frac{1}{2dC_2}\right). \qquad (4.77)
$$

This procedure can be iterated to obtain a finite bound on $E\left[\|\exp(\alpha X(\cdot))\|_{[-\tau,T]}\right]$ for any $T>0$. Indeed, for each $k\geq 1$, set $T^{(k)}=\frac{k}{2dC_2}$. Assume that

$$E\left[\|\exp(\alpha X(\cdot))\|_{[-\tau,T^{(k)}]}\right]<\infty \text{ for each } \alpha>0,$$

which we have already shown to hold in the case when $k=1$. We can extend to $T^{(k+1)}$ as above with $T^{(k)}$ taking the place of 0, η_n replaced with $\eta_n^{(k)}:=\inf\{t\geq T^{(k)}:\|X\|_{[-\tau,t],2}\geq n\}$, $K_1(\alpha)$ unchanged, and $K_0(\alpha)$ replaced with

$$\begin{aligned} K_0^{(k)}(\alpha) \quad := \quad & e^{\alpha dM}E[\exp(\alpha d\|X\|_{[-\tau,T^{(k)}],2})]+e^{\alpha dM}\frac{1}{3}E\left[\exp\left(6\alpha d\|X\|_{[-\tau,T^{(k)}],2}\right)\right] \\ & +e^{\alpha dM}\frac{2}{3}K\left(6\alpha d,\frac{1}{2dC_2}\right). \end{aligned}$$

Then,

$$E\left[\|\exp(\alpha|X(\cdot)|_1)\|_{[-\tau,T^{(k+1)}]}\right] \quad\leq\quad K_0^{(k)}(\alpha)\exp\left(K_1(\alpha)\frac{1}{2dC_2}\right). \tag{4.78}$$

Recognizing that $T^{(k)}\to\infty$ as $k\to\infty$ completes the proof.

□

Lemma 4.3.4. *Fix a possibly random X_0 and assume that $E[\exp(\alpha\|X_0\|_2)]<\infty$ for each $\alpha>0$. Then for each $\gamma>0$, there is a constant $K(\gamma)>0$ such that for each $i\in\{1,\ldots,d\}$ and $t\geq 0$,*

$$E\left[\exp\left(\gamma\text{Osc}^+(X_t^i,\mathbb{I})\right)\right] \quad\leq\quad K(\gamma). \tag{4.79}$$

Proof. Lemma 4.1.2 with $\hat{M}=M$ and Assumption 4.3.1 imply that P-a.s.,

$$\begin{aligned} & \exp(\gamma\text{Osc}^+(X_t^i,\mathbb{I})) \\ & \quad\leq\quad e^{\gamma(M+\tau K_u)}\exp(\gamma\text{Osc}(X_0,\mathbb{I}))\exp\left(\sup_{(t-\tau)^+\leq s\leq t}\gamma\int_{(t-\tau)^+}^s\sigma^i(X_u)1_{\{X^i(u)>M\}}dW(u)\right) \\ & \qquad\qquad \times\exp\left(\sup_{(t-\tau)^+\leq s\leq t}\gamma\int_{(t-\tau)^+}^s -1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right). \end{aligned} \tag{4.80}$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&E\left[\exp\left(\sup_{(t-\tau)^+\le s\le t}\int_{(t-\tau)^+}^{s} 2\gamma 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right)\right.\\
&\qquad\qquad\left.\times\exp\left(\sup_{(t-\tau)^+\le s\le t}\int_{(t-\tau)^+}^{s} -2\gamma 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right)\right]\\
&\le \left(E\left[\exp\left(\sup_{(t-\tau)^+\le s\le t}\int_{(t-\tau)^+}^{s} 4\gamma 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right)\right]\right)^{\frac{1}{2}}\\
&\qquad\times\left(E\left[\exp\left(\sup_{(t-\tau)^+\le s\le t}\int_{(t-\tau)^+}^{s} -4\gamma 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right)\right]\right)^{\frac{1}{2}}\\
&= \left(E\left[\sup_{(t-\tau)^+\le s\le t}\exp\left(\int_{(t-\tau)^+}^{s} 4\gamma 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right)\right]\right)^{\frac{1}{2}}\\
&\qquad\times\left(E\left[\sup_{(t-\tau)^+\le s\le t}\exp\left(\int_{(t-\tau)^+}^{s} -4\gamma 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right)\right]\right)^{\frac{1}{2}}. \qquad (4.81)
\end{aligned}$$

Therefore, Lemma 4.3.2 implies that for each $t \ge 0$,

$$\begin{aligned}
&E\left[\exp\left(\sup_{(t-\tau)^+\le s\le t}\int_{(t-\tau)^+}^{s} 2\gamma 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right)\right.\\
&\qquad\qquad\left.\times\exp\left(\sup_{(t-\tau)^+\le s\le t}\int_{(t-\tau)^+}^{s} -2\gamma 1_{\{X^i(u)>M\}}\sigma^i(X_u)dW(u)\right)\right]\\
&\qquad\le K(4\gamma,\tau). \qquad (4.82)
\end{aligned}$$

Using the Cauchy-Schwarz inequality again we obtain that

$$
\begin{aligned}
&E\left[\exp(\gamma \text{Osc}^{+}(X_t^i, \mathbb{I}))\right] \\
&\leq \; e^{\gamma(M+\tau K_u)} E\Bigg[\exp(\gamma \text{Osc}(X_0, \mathbb{I})) \\
&\qquad \times \exp\left(\sup_{(t-\tau)^+ \leq s \leq t} \gamma \int_{(t-\tau)^+}^{s} 1_{\{X^i(u) \geq M\}} \sigma^i(X_u) dW(u)\right) \\
&\qquad \times \exp\left(\sup_{(t-\tau)^+ \leq s \leq t} \gamma \int_{(t-\tau)^+}^{s} -1_{\{X^i(u) \geq M\}} \sigma^i(X_u) dW(u)\right)\Bigg] \\
&\leq \; e^{\gamma(M+\tau K_u)} \left(E\left[\exp(2\gamma \text{Osc}(X_0, \mathbb{I}))\right]\right)^{\frac{1}{2}} \\
&\qquad \times \Bigg(E\Bigg[\exp\left(\sup_{(t-\tau)^+ \leq s \leq t} 2\gamma \int_{(t-\tau)^+}^{s} 1_{\{X^i(u) \geq M\}} \sigma^i(X_u) dW(u)\right) \\
&\qquad \times \exp\left(\sup_{(t-\tau)^+ \leq s \leq t} 2\gamma \int_{(t-\tau)^+}^{s} -1_{\{X^i(u) \geq M\}} \sigma^i(X_u) dW(u)\right)\Bigg]\Bigg)^{\frac{1}{2}} \\
&\leq \; e^{\gamma(M+\tau K_u)} \left(E\left[\exp(2\gamma \|X_0\|_2)\right]\right)^{\frac{1}{2}} \left(K(4\gamma, \tau)\right)^{\frac{1}{2}},
\end{aligned} \tag{4.83}
$$

which is finite by assumption.

□

4.3.5 Uniform Bound on $E[\|X_t\|_2^2]$

Lemma 4.3.5. *Assume that* $\sup_{t \geq -\tau} E[|X(t)|_2^2] < \infty$ *and* $E[\|X_0\|_2^2] < \infty$. *Then* $\sup_{t \geq 0} E[\|X_t\|_2^2] < \infty$.

Proof. After replacing B_0 by K_u and setting $B_{2,i} = C_{2,i} = 0$ for each i, the proof is identical to the proof of Theorem 4.2.2.

□

Combining Theorem 4.3.1 with $p = 2$ and Lemma 4.3.5 yields the following.

Corollary 4.3.1. *Under Assumption 4.3.1, if* $E\left[\exp\left(\kappa \|X_0\|_2\right)\right] < \infty$ *for all* $\kappa > 0$*, then* $\sup_{t \geq 0} E\left[\|X_t\|_2^2\right] < \infty$.

Chapter 5

Existence and Uniqueness of Stationary Distributions

5.1 Existence of Stationary Distributions

For simplicity of exposition, we introduce the following assumption.

Assumption 5.1.1. *Either Assumption 4.2.1 or 4.3.1 holds.*

The following is obtained by combining our results from Chapters 3 and 4.

Theorem 5.1.1. *Under Assumptions 2.1.1, 2.1.2, and 5.1.1, there exists a stationary distribution for the SDDER (1.1).*

Proof. For each $x_o \in \mathbb{C}^d_{\mathbb{I}}$, the hypotheses on the initial conditions of either Theorem 4.2.2 or Corollary 4.3.1 are met, so that $\sup_{t \geq 0} E[\|X_t^{x_o}\|_2^2] < \infty$. The result now follows from Corollary 3.4.1.

□

5.2 Uniqueness of Stationary Distributions

In this section, we prove uniqueness of a stationary distribution for the SDDER under the following Assumption 5.2.1 on b and σ. We use an asymptotic coupling argument that is an adaptation to the situation with reflection of a novel argument recently introduced by Hairer, Mattingley, and Scheutzow [16] for stochastic delay differential equations without reflection.

Assumption 5.2.1.

(i) there exists a bounded right inverse for σ, i.e., there is a constant $C_6 > 0$ and a measurable function $\sigma^\dagger : \mathbb{C}^d_{\mathbb{I}} \to \mathbb{M}^{m\times d}$ such that for all $x \in \mathbb{C}^d_{\mathbb{I}}$, $\sigma(x)\sigma^\dagger(x) = I_d$, and $\|\sigma^\dagger(x)\|_2 \leq C_6$ for all $x \in \mathbb{C}^d_{\mathbb{I}}$, and

(ii) the coefficients b and σ are globally Lipschitz continuous, i.e., there exists a constant $\kappa_L > 0$ such that

$$|b(x) - b(y)|_2^2 + \|\sigma(x) - \sigma(y)\|_2^2 \quad \leq \quad \kappa_L \|x - y\|_2^2 \quad \text{for all } x, y \in \mathbb{C}^d_{\mathbb{I}}. \tag{5.1}$$

Remark. Part (i) implies that $m \geq d$, since the rank of a product of matrices cannot exceed the rank of either factor. If $\sigma\sigma'$ is uniformly elliptic (or uniformly positive definite), i.e., there is a constant $a > 0$ such that $v\sigma(x)(\sigma(x))'v \geq a|v|_2^2$ for all $x \in \mathbb{C}^d_{\mathbb{I}}$ and $v \in \mathbb{R}^d_+$, then σ has a bounded right inverse. Indeed, in this case for each $x \in \mathbb{C}^d_{\mathbb{I}}$, let a singular value decomposition be $\sigma(x) = U(x)\Lambda(x)V(x)$, where $U(x) \in \mathbb{M}^{d\times d}$ and $V(x) \in \mathbb{M}^{m\times m}$ are unitary matrices, $\Lambda(x) \in \mathbb{M}^{d\times m}$, $\Lambda^i_j(x) = 0$ whenever $i \neq j$, and $\Lambda^i_i(x) \geq 0$ for each $i = 1, \ldots, d$. We shall drop the x in what follows. Then $\sigma\sigma' = U\Lambda\Lambda' U'$, and uniform ellipticity implies that the diagonal entries of the $d \times d$ diagonal matrix $\Lambda\Lambda'$ are at least $a > 0$, so that the diagonal entries of $(\Lambda\Lambda')^{-1}$ are at most $\frac{1}{a}$. Therefore, the maximal diagonal entry of the Moore-Penrose pseudoinverse $\Lambda^\dagger = \Lambda'(\Lambda\Lambda')^{-1}$ of Λ is at most $\frac{1}{\sqrt{a}}$. Therefore,

$$\sigma^\dagger \; := \; \sigma'(\sigma\sigma')^{-1} \; = \; V'\Lambda'(\Lambda\Lambda')^{-1}U' \; = \; V'\Lambda^\dagger U',$$

is a right inverse for σ with spectral radius bounded above by $\frac{1}{\sqrt{a}}$, since U and V are unitary. Also, $\sigma^\dagger(\cdot)$ is continuous since $\sigma(\cdot)$ and $\sigma(\cdot)(\sigma(\cdot))'$ are continuous, and taking the inverse of a non-singular matrix is a continuous operation.

Remark. Part (ii) of Assumption 5.2.1 implies that Assumptions 2.1.1 and 2.1.2 hold (see Appendix C).

The main result of this is section is the following theorem.

Theorem 5.2.1. *Under Assumption 5.2.1, there exists at most one stationary distribution for the SDDER (1.1).*

A key element to our proof is the following proposition, which is adapted to our situation from Corollary 2.2 of [16]. Before stating it, we introduce some notation. Denote the space of sequences $\{x^n\}_{n=0}^\infty$ in $\mathbb{C}^d_{\mathbb{I}}$ by $(\mathbb{C}^d_{\mathbb{I}})^\infty$, and endow this with the product topology and associated

Borel σ-algebra. Denote the space of pairs of sequences with values in $\mathbb{C}_{\mathbb{I}}^d$ by $(\mathbb{C}_{\mathbb{I}}^d)^\infty \times (\mathbb{C}_{\mathbb{I}}^d)^\infty$, again with the product topology and associated σ-algebra. Denote the space of probability measures on $(\mathbb{C}_{\mathbb{I}}^d)^\infty \times (\mathbb{C}_{\mathbb{I}}^d)^\infty$ by $\mathcal{P}((\mathbb{C}_{\mathbb{I}}^d)^\infty \times (\mathbb{C}_{\mathbb{I}}^d)^\infty)$. Let $\{P_t(x,\Lambda), x \in \mathbb{C}_{\mathbb{I}}^d, \Lambda \in \mathcal{M}_{\mathbb{I}}, t \geq 0\}$ be the family of Markovian transition functions associated with the SDDER (1.1) and let μ be a probability measure on $\mathbb{C}_{\mathbb{I}}^d$. Define the probability measure P_∞^μ on $(\mathbb{C}_{\mathbb{I}}^d)^\infty$ as follows. A cylinder set $A \subset (\mathbb{C}_{\mathbb{I}}^d)^\infty$ has the form:

$$A = \left\{ \{x^n\}_{n=0}^\infty \in (\mathbb{C}_{\mathbb{I}}^d)^\infty : x^n \in A_n \text{ for all } n \right\},$$

where $A_n \in \mathcal{M}_{\mathbb{I}}$ for each n, and there is a non-negative integer N such that $A_n = \mathbb{C}_{\mathbb{I}}^d$ for all $n \geq N+1$. Then, P_∞^μ is defined on such a set A by

$$P_\infty^\mu(A) = \int_{x_0 \in A_0} \int_{x_1 \in A_1} \cdots \int_{x_N \in A_N} P_\tau(x_{N-1}, dx_N) \cdots P_\tau(x_0, dx_1)\mu(dx_0).$$

Kolmogorov's extension theorem (see, e.g., [11] or [39]) ensures that P_∞^μ extends uniquely to a probability measure on $(\mathbb{C}_{\mathbb{I}}^d)^\infty$. Thus, P_∞^μ is the distribution of the sequence $\{X_{n\tau}\}_{n=0}^\infty$ when X is a solution of (1.1) started with distribution μ. Recall that the symbol $\sim$ between two probability measures means that they are mutually absolutely continuous.

The following proposition follows immediately from Corollary 2.2 of [16] by setting $A = \mathbb{C}_{\mathbb{I}}^d$ there.

Proposition 5.2.1. *Assume that there is a family* $\left\{ \tilde{P}_{x,y} : (x,y) \in \mathbb{C}_{\mathbb{I}}^d \times \mathbb{C}_{\mathbb{I}}^d \right\}$ *of probability measures on* $(\mathbb{C}_{\mathbb{I}}^d)^\infty \times (\mathbb{C}_{\mathbb{I}}^d)^\infty$ *such that for each* $x, y \in \mathbb{C}_{\mathbb{I}}^d$,

(i) $\tilde{P}_{x,y}(\cdot \times (\mathbb{C}_{\mathbb{I}}^d)^\infty) \sim P_\infty^{\delta_x}(\cdot)$ *and* $\tilde{P}_{x,y}((\mathbb{C}_{\mathbb{I}}^d)^\infty \times \cdot) \sim P_\infty^{\delta_y}(\cdot)$,

(ii) for each $x, y \in \mathbb{C}_{\mathbb{I}}^d$,

$$\tilde{P}_{x,y}\left((\{x^n\}_{n=0}^\infty, \{y^n\}_{n=0}^\infty) \in (\mathbb{C}_{\mathbb{I}}^d)^\infty \times (\mathbb{C}_{\mathbb{I}}^d)^\infty : \lim_{n\to\infty} \|x^n - y^n\|_2 = 0 \right) > 0,$$

and for each $\Gamma \in \mathcal{B}((\mathbb{C}_{\mathbb{I}}^d)^\infty \times (\mathbb{C}_{\mathbb{I}}^d)^\infty)$,

(iii) the mapping $(x,y) \mapsto \tilde{P}_{x,y}(\Gamma)$ *is measurable on* $\mathbb{C}_{\mathbb{I}}^d \times \mathbb{C}_{\mathbb{I}}^d$.

Then there exists at most one stationary distribution for the semigroup that is associated with $\{P_t(\cdot,\cdot), t \geq 0\}$.

We need to develop some preliminary results before giving the proof of Theorem 5.2.1. We begin with a stochastic variation of constants formula.

Proposition 5.2.2. *Assume that on some filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, $\{\xi^{(1)}(t), t \geq 0\}$ *is an adapted process satisfying the following stochastic differential equation:*

$$d\xi^{(1)}(t) = \alpha \xi^{(1)}(t)dt + d\xi^{(2)}(t), \tag{5.2}$$

for some $\alpha \in \mathbb{R}$ *and some continuous semimartingale* $\{\xi^{(2)}(t), t \geq 0\}$. *Then*

$$\xi^{(1)}(t) = e^{\alpha t}\xi^{(1)}(0) + \int_0^t e^{\alpha(t-s)} d\xi^{(2)}(s), \quad t \geq 0, \tag{5.3}$$

and thus for each $t \geq s \geq 0$,

$$\xi^{(1)}(t) = e^{\alpha(t-s)}\xi^{(1)}(s) + \int_s^t e^{\alpha(t-r)} d\xi^{(2)}(r).$$

Proof. Denote the right-hand-side of (5.3) by $\xi^{(3)}(t)$. Then $\xi^{(3)}(0) = \xi^{(1)}(0)$, and

$$\begin{aligned} d\xi^{(3)}(t) &= \alpha e^{\alpha t}\xi^{(1)}(0)dt + \alpha e^{\alpha t}\int_0^t e^{-\alpha s} d\xi^{(2)}(s)dt \\ &\quad + e^{\alpha(t-t)} d\xi^{(2)}(t) + d\left\langle e^{\alpha \cdot}, \int_0^{\cdot} e^{-\alpha s} d\xi^{(2)}(s) \right\rangle (t) \\ &= \alpha \xi^{(3)}(t)dt + d\xi^{(2)}(t). \end{aligned} \tag{5.4}$$

Thus,

$$d\left(\xi^{(1)}(t) - \xi^{(3)}(t)\right) = \alpha \left(\xi^{(1)}(t) - \xi^{(3)}(t)\right) dt,$$

and it follows that

$$\begin{aligned} \xi^{(1)}(t) - \xi^{(3)}(t) &= \left(\xi^{(1)}(0) - \xi^{(3)}(0)\right) e^{\alpha t} \\ &= 0, \end{aligned}$$

and so

$$P\left(\xi^{(1)}(t) = \xi^{(3)}(t) \text{ for all } t \geq 0\right) = 1.$$

□

We assume that an m-dimensional Brownian motion martingale $\{W(t), t \geq 0\}$ is given on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, P)$. For each $\lambda > 0$, consider the system of SDDERs

$$dX(t) = b(X_t)dt + \sigma(X_t)dW(t) + dY(t), \tag{5.5}$$

$$d\tilde{X}^\lambda(t) = b(\tilde{X}_t^\lambda)dt + \lambda(X(t) - \tilde{X}^\lambda(t))dt + \sigma(\tilde{X}_t^\lambda)dW(t) + d\tilde{Y}^\lambda(t), \tag{5.6}$$

where P-a.s., $(X(t), \tilde{X}^\lambda(t)) \in \mathbb{R}_+^{2d}$ for all $t \geq -\tau$, and where $(Y, \tilde{Y}^\lambda)$ is a continuous adapted process such that P-a.s., $Y(0) = \tilde{Y}^\lambda(0) = 0$ and Y^i (resp. $\tilde{Y}^{\lambda,i}$) can increase only when X^i (resp. $\tilde{X}^{\lambda,i}$) is zero. We allow possibly random initial conditions $X_0 \equiv \xi$ and $\tilde{X}_0^\lambda \equiv \tilde{\xi}$. This is a $2d$-dimensional system with globally Lipschitz coefficients, and thus Appendix C implies that there exists a (pathwise) unique strong solution for any pair of square-integrable initial conditions:

$$E[\|\xi\|_2^2] < \infty \quad \text{and} \quad E[\|\tilde{\xi}\|_2^2] < \infty. \tag{5.7}$$

We consider a solution pair $(X, \tilde{X}^\lambda)$ with initial conditions satisfying (5.7). The difference $U^\lambda(t) := X(t) - \tilde{X}^\lambda(t)$ satisfies

$$\begin{aligned} dU^\lambda(t) &= \left(b(X_t) - b(\tilde{X}_t^\lambda)\right) dt - \lambda U^\lambda(t) dt \\ &\quad + \left(\sigma(X_t) - \sigma(\tilde{X}_t^\lambda)\right) dW(t) + d\left(Y - \tilde{Y}^\lambda\right)(t), \quad t \geq 0. \end{aligned} \tag{5.8}$$

The following lemma is a modified version of Lemma 3.5 of [16], where here we have equations with reflection. Inequality (5.11) is the reason that this lemma remains true in the reflected case. Our proof is very similar to that in [16] from (5.14) onwards.

Lemma 5.2.1. *For each $\alpha > 0$, there exist $\lambda > 0$ and $K > 0$ such that*

$$E\left[\sup_{t\geq 0} e^{\alpha t}\|U_t^\lambda\|_2^8\right] \leq KE\left[\|U_0^\lambda\|_2^8\right],$$

whenever $E[\|U_0^\lambda\|_2^8\|] < \infty$.

The proof uses the following proposition, which is a slight generalization of Lemma 3.4 in [16] for the case where W is m-dimensional, and specializes to the case where h is continuous. The proof is nearly identical, and so we omit it. The proof uses the representation $V^\alpha(t) = e^{-\alpha t}\int_0^t e^{\alpha s}h(s)dW(s)$, the Burkholder-Davis-Gundy inequality, an integration by parts, and estimates on V^α on the segments $\left[\frac{kT}{N}, \frac{(k+1)T}{N}\right]$, $k = 0, \ldots, N-1$, for large enough integers N.

Proposition 5.2.3. *On a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, let W be an m-dimensional Brownian motion martingale, let $\{h(s), s \geq 0\}$ be a continuous adapted process taking values in $\mathbb{M}^{1\times m}$, and assume that for each $\alpha > 0$, we have an adapted continuous real-valued process $\{V^\alpha(t), t \geq 0\}$ satisfying the stochastic differential equation*

$$dV^\alpha(t) = -\alpha V^\alpha(t)dt + h(t)dW(t), \quad t \geq 0, \tag{5.9}$$

with the initial condition $V^\alpha(0) = 0$*. Then for each* $T > 0$ *and* $p > 2$*, there exists a function* $\nu_{T,p} : \mathbb{R}_+ \to \mathbb{R}_+$ *satisfying* $\lim_{\alpha\to\infty} \nu_{T,p}(\alpha) = 0$ *such that for any stopping time* η*,*

$$E\left[\sup_{0\le t\le T\wedge\eta} |V^\alpha(t)|^p\right] \le \nu_{T,p}(\alpha) E\left[\sup_{0\le t\le T\wedge\eta} |h(t)|_2^p\right].$$

Proof of Lemma 5.2.1. Fix $\alpha > 0$. Without loss of generality, we may assume that $\lambda > \frac{\alpha}{2}$. From equation (5.8), we have

$$\begin{aligned} d\left(|U^\lambda(t)|_2^2\right) &= 2\left(U^\lambda(t)\right)'\left(b(X_t) - b(\tilde{X}_t^\lambda)\right)dt - 2\lambda|U^\lambda(t)|_2^2 dt \\ &\quad +2\left(U^\lambda(t)\right)'\left(\sigma(X_t) - \sigma(\tilde{X}_t^\lambda)\right)dW(t) + 2\left(U^\lambda(t)\right)' d\left(Y - \tilde{Y}^\lambda\right)(t) \\ &\quad + \left\|\sigma(X_t) - \sigma(\tilde{X}_t^\lambda)\right\|_2^2 dt. \end{aligned} \tag{5.10}$$

The constraints on where Y and $\tilde{Y}^\lambda$ can increase and the non-negativity of X and $\tilde{X}^\lambda$ imply that for each $t \ge s \ge 0$ and $\beta \in \mathbb{R}$,

$$\begin{aligned} \int_s^t e^{\beta r}(U^\lambda(r))' d(Y - \tilde{Y}^\lambda)(r) &= \sum_{i=1}^d \int_s^t e^{\beta r} U^{\lambda,i}(r) d\left(Y^i - \tilde{Y}^{\lambda,i}\right)(r) \\ &= -\sum_{i=1}^d \left(\int_s^t e^{\beta r} X^i(r) d\tilde{Y}^{\lambda,i}(r) \right. \\ &\qquad \left. + \int_s^t e^{\beta r} \tilde{X}^{\lambda,i}(r) dY^i(r)\right) \\ &\le 0. \end{aligned} \tag{5.11}$$

The Lipschitz continuity condition (5.1) on b and σ implies that for any $x, y \in \mathbb{C}_{\mathbb{I}}^d$,

$$\begin{aligned} &2(x(0) - y(0))'(b(x) - b(y)) + \|\sigma(x) - \sigma(y)\|_2^2 \\ &\quad \le 2|x(0) - y(0)|_2 |b(x) - b(y)|_2 + \|\sigma(x) - \sigma(y)\|_2^2 \\ &\quad \le |x(0) - y(0)|_2^2 + |b(x) - b(y)|_2^2 + \|\sigma(x) - \sigma(y)\|_2^2 \\ &\quad \le (1 + \kappa_L)\|x - y\|_2^2. \end{aligned} \tag{5.12}$$

Itô's formula and equality (5.10) yield

$$\begin{aligned} d\left(e^{\alpha t}|U^\lambda(t)|_2^2\right) &= \alpha e^{\alpha t}|U^\lambda(t)|_2^2 dt + e^{\alpha t} d\left(|U^\lambda(t)|_2^2\right) \\ &= (\alpha - 2\lambda)e^{\alpha t}|U^\lambda(t)|_2^2 dt + 2e^{\alpha t}\left(U^\lambda(t)\right)'\left(b(X_t) - b(\tilde{X}_t^\lambda)\right)dt \\ &\quad +2e^{\alpha t}\left(U^\lambda(t)\right)'\left(\sigma(X_t) - \sigma(\tilde{X}_t^\lambda)\right)dW(t) + 2e^{\alpha t}\left(U^\lambda(t)\right)' d\left(Y - \tilde{Y}^\lambda\right)(t) \\ &\quad +e^{\alpha t}\left\|\sigma(X_t) - \sigma(\tilde{X}_t^\lambda)\right\|_2^2 dt. \end{aligned} \tag{5.13}$$

Then Proposition 5.2.2 and inequalities (5.12) and (5.11) imply that for each fixed $t_1 \geq 0$, and $t \geq t_1$,

$$\begin{aligned} e^{\alpha t}|U^\lambda(t)|_2^2 &= e^{(\alpha-2\lambda)(t-t_1)}|U^\lambda(t_1)|_2^2 + 2\int_{t_1}^t e^{(\alpha-2\lambda)(t-r)}e^{\alpha r}\left(U^\lambda(r)\right)'\left(b(X_r)-b(\tilde{X}_r^\lambda)\right)dr \\ &\quad +2\int_{t_1}^t e^{(\alpha-2\lambda)(t-r)}e^{\alpha r}\left(U^\lambda(r)\right)'\left(\sigma(X_r)-\sigma(\tilde{X}_r^\lambda)\right)dW(r) \\ &\quad +2\int_{t_1}^t e^{(\alpha-2\lambda)(t-r)}e^{\alpha r}\left(U^\lambda(r)\right)' d\left(Y-\tilde{Y}^\lambda\right)(r) \\ &\quad +\int_{t_1}^t e^{(\alpha-2\lambda)(t-r)}e^{\alpha r}\left\|\sigma(X_r)-\sigma(\tilde{X}_r^\lambda)\right\|_2^2 dr \\ &\leq e^{(\alpha-2\lambda)(t-t_1)}|U^\lambda(t_1)|_2^2 + (1+\kappa_L)\int_{t_1}^t e^{(\alpha-2\lambda)(t-r)}e^{\alpha r}\left\|U_r^\lambda\right\|_2^2 dr \\ &\quad +2\int_{t_1}^t e^{(\alpha-2\lambda)(t-r)}e^{\alpha r}\left(U^\lambda(r)\right)'\left(\sigma(X_r)-\sigma(\tilde{X}_r^\lambda)\right)dW(r). \end{aligned} \tag{5.14}$$

The remainder of the proof is very similar to that of Lemma 3.5 in [16] (following on from the top of page 15, or the third display of the proof), but for completeness, we provide the details and correct a few minor oversights.

For each $\beta > 0$, $p \geq 1$, integer $n \geq 0$, and $t_2 \geq t_1$, we have

$$\sup_{s\in[t_1,t_2]} e^{\beta s}\|U_s^\lambda\|_2^p \leq e^{\beta\tau}\left(\sup_{s\in[t_1-\tau,t_1]} e^{\beta s}|U^\lambda(s)|_2^p + \sup_{s\in[t_1,t_2]} e^{\beta s}|U^\lambda(s)|_2^p\right). \tag{5.15}$$

Therefore, continuing from (5.14)

$$\begin{aligned} &e^{\alpha t}|U^\lambda(t)|_2^2 \\ &\quad\leq |U^\lambda(t_1)|_2^2 \\ &\qquad +(1+\kappa_L)e^{\alpha\tau}\left(\int_{t_1}^t e^{(\alpha-2\lambda)(t-r)}dr\right)\left(\sup_{r\in[t_1-\tau,t_1]} e^{\alpha r}|U^\lambda(r)|_2^2 + \sup_{r\in[t_1,t]} e^{\alpha r}|U^\lambda(r)|_2^2\right) \\ &\qquad +2\sup_{r\in[t_1,t]}\left|\int_{t_1}^r e^{(\alpha-2\lambda)(r-u)}e^{\alpha u}\left(U^\lambda(u)\right)'\left(\sigma(X_u)-\sigma(\tilde{X}_u^\lambda)\right)dW(u)\right|. \end{aligned} \tag{5.16}$$

In a similar manner to that in which inequality (4.69) was derived, using the convexity of $r \to r^4$ we have for each $\gamma > 1$ and $a_1, a_2, a_3, a_4 \in \mathbb{R}_+$,

$$(a_1+a_2+a_3+a_4)^4 \leq \gamma^3 a_1^4 + \left(\frac{3\gamma}{\gamma-1}\right)^3 a_2^4 + \left(\frac{3\gamma}{\gamma-1}\right)^3 a_3^4 + \left(\frac{3\gamma}{\gamma-1}\right)^3 a_4^4.$$

Then, since $\int_{t_1}^{t_2} e^{-\zeta(t_2-u)}du = \frac{1}{\zeta}(1-e^{\zeta(t_1-t_2)}) \leq \frac{1}{\zeta}$ for any $\zeta > 0$ and $t_2 \geq t_1$, raising both sides of (5.16) to the fourth power yields for $t \geq t_1$,

$$\begin{aligned} e^{4\alpha t}|U^\lambda(t)|_2^8 & \\ \leq\ & \gamma^3|U^\lambda(t_1)|_2^8 \\ & + \left(\frac{3\gamma}{\gamma-1}\right)^3 \frac{(1+\kappa_L)^4 e^{4\alpha\tau}}{(2\lambda-\alpha)^4}\left(\sup_{r\in[t_1-\tau,t_1]} e^{4\alpha r}|U^\lambda(r)|_2^8 + \sup_{r\in[t_1,t]} e^{4\alpha r}|U^\lambda(r)|_2^8\right) \\ & + \left(\frac{3\gamma}{\gamma-1}\right)^3 2^4 \sup_{r\in[t_1,t]}\left|\int_{t_1}^r e^{(\alpha-2\lambda)(r-u)}e^{\alpha u}\Big(U^\lambda(u)\Big)'\Big(\sigma(X_u)-\sigma(\tilde{X}_u^\lambda)\Big)dW(u)\right|^4. \quad (5.17) \end{aligned}$$

By Itô's formula, the differential of

$$V(t) := \int_{t_1}^{t_1+t} e^{(\alpha-2\lambda)((t_1+t)-u)}e^{\alpha u}\left(U^\lambda(u)\right)'\left(\sigma(X_u)-\sigma(\tilde{X}_u^\lambda)\right)dW(u), \quad t\geq 0,$$

is

$$\begin{aligned} dV(t) &= d\left(e^{(\alpha-2\lambda)(t_1+t)}\int_{t_1}^{t_1+t} e^{(\alpha-2\lambda)(-u)}e^{\alpha u}\left(U^\lambda(u)\right)'\left(\sigma(X_u)-\sigma(\tilde{X}_u^\lambda)\right)dW(u)\right) \\ &= (\alpha-2\lambda)e^{(\alpha-2\lambda)(t_1+t)}\int_{t_1}^{t_1+t} e^{2\lambda u}\left(U^\lambda(u)\right)'\left(\sigma(X_u)-\sigma(\tilde{X}_u^\lambda)\right)dW(u)dt \\ &\quad + e^{(\alpha-2\lambda)(t_1+t)}e^{2\lambda(t_1+t)}\left(U^\lambda(t_1+t)\right)'\left(\sigma(X_{t_1+t})-\sigma(\tilde{X}_{t_1+t}^\lambda)\right)dW(t_1+t) \\ &= -(2\lambda-\alpha)V(t)dt \\ &\quad + e^{\alpha(t_1+t)}\left(U^\lambda(t_1+t)\right)'\left(\sigma(X_{t_1+t})-\sigma(\tilde{X}_{t_1+t}^\lambda)\right)dW^{t_1}(t), \quad (5.18) \end{aligned}$$

where $\{W^{t_1}(t) := W(t_1+t)-W(t_1), t\geq 0\}$ is a Brownian motion martingale relative to the filtration $\{\mathcal{F}_t^{t_1} := \mathcal{F}_{t_1+t}, t\geq 0\}$. If we define $\eta_n^\lambda = \inf\{r\geq t_1 : \left|U^\lambda(r)\right|_2 \geq n\}$, then $\eta_n^\lambda - t_1$ is a stopping time relative to $\{\mathcal{F}_t^{t_1}\}$, so Proposition 5.2.3 and Assumption 5.2.1(ii) imply that

$$\begin{aligned} & E\left[\sup_{r\in[t_1,(t_1+\tau)\wedge\eta_n^\lambda]}\left|\int_{t_1}^r e^{(\alpha-2\lambda)(r-u)}e^{\alpha u}\left(U^\lambda(u)\right)'\left(\sigma(X_u)-\sigma(\tilde{X}_u^\lambda)\right)dW(u)\right|^4\right] \\ &= E\left[\sup_{r\in[0,\tau\wedge(\eta_n^\lambda-t_1)]}(V(r))^4\right] \\ &\leq \nu_{\tau,4}(2\lambda-\alpha)E\left[\sup_{r\in[0,\tau(\wedge\eta_n^\lambda-t_1)]} e^{4\alpha(t_1+r)}\left|\left(U^\lambda(t_1+r)\right)'\left(\sigma(X_{(t_1+r)})-\sigma(\tilde{X}_{(t_1+r)}^\lambda)\right)\right|^4\right] \\ &\leq \nu_{\tau,4}(2\lambda-\alpha)E\left[\sup_{r\in[t_1,(t_1+\tau)\wedge\eta_n^\lambda]}\left|e^{\alpha r}\left(U^\lambda(r)\right)'\left(\sigma(X_r)-\sigma(\tilde{X}_r^\lambda)\right)\right|_2^4\right] \\ &\leq \nu_{\tau,4}(2\lambda-\alpha)\kappa_L^2 E\left[\sup_{r\in[t_1,(t_1+\tau)\wedge\eta_n^\lambda]} e^{4\alpha r}\left\|U_r^\lambda\right\|_2^8\right]. \quad (5.19) \end{aligned}$$

Taking the supremum up to time $(t_1+\tau)\wedge\eta_n^\lambda$ and then the expectation on both sides of inequality (5.17), and using (5.15) and (5.19), we obtain

$$\begin{aligned}
&E\left[\sup_{t\in[t_1,(t_1+\tau)\wedge\eta_n^\lambda]} e^{4\alpha t}|U^\lambda(t)|_2^8\right] \\
&\quad\le \gamma^3 E\left[|U^\lambda(t_1)|_2^8\right] \\
&\qquad + \left(\frac{3\gamma}{\gamma-1}\right)^3 \frac{(1+\kappa_L)^4 e^{4\alpha\tau}}{(2\lambda-\alpha)^4} E\left[\sup_{t\in[t_1-\tau,t_1]} e^{4\alpha t}|U^\lambda(t)|_2^8 + \sup_{t\in[t_1,(t_1+\tau)\wedge\eta_n^\lambda]} e^{4\alpha t}\left|U^\lambda(t)\right|_2^8\right] \\
&\qquad + \left(\frac{3\gamma}{\gamma-1}\right)^3 2^4\nu_{\tau,4}(2\lambda-\alpha)\kappa_L^2 e^{4\alpha\tau} \\
&\qquad\qquad \times\left(E\left[\sup_{t_1-\tau\le t\le t_1} e^{4\alpha t}|U^\lambda(t)|_2^8 + \sup_{t\in[t_1,(t_1+\tau)\wedge\eta_n^\lambda]} e^{4\alpha t}\left|U^\lambda(t)\right|_2^8\right]\right) \\
&\quad\le \left(\gamma^3 + \left(\frac{3\gamma}{\gamma-1}\right)^3\delta(\lambda,\alpha)\right) E\left[\sup_{t\in[t_1-\tau,t_1]} e^{4\alpha t}|U^\lambda(t)|_2^8\right] \\
&\qquad + \left(\frac{3\gamma}{\gamma-1}\right)^3 \delta(\lambda,\alpha) E\left[\sup_{t\in[t_1,(t_1+\tau)\wedge\eta_n^\lambda]} e^{4\alpha t}\left|U^\lambda(t)\right|_2^8\right],
\end{aligned} \tag{5.20}$$

where $\delta(\lambda,\alpha) := \frac{(1+\kappa_L)^4 e^{4\alpha\tau}}{(2\lambda-\alpha)^4} + 2^4\nu_{\tau,4}(2\lambda-\alpha)\kappa_L^2 e^{4\alpha\tau} > 0$. Note that $\lim\limits_{\lambda\to\infty}\delta(\lambda,\alpha) = 0$ for each fixed $\alpha > 0$. By the definition of η_n^λ,

$$\begin{aligned}
&E\left[\sup_{t\in[t_1,(t_1+\tau)\wedge\eta_n^\lambda]} e^{4\alpha t}|U^\lambda(t)|_2^8\right] \\
&\quad\le e^{4\alpha(t_1+\tau)} E\left[1_{\{|U^\lambda(t_1)|_2\ge n\}}|U^\lambda(t_1)|_2^8 + 1_{\{|U^\lambda(t_1)|_2< n\}}\sup_{t\in[t_1,t_1+\tau]}\left(n\wedge|U^\lambda(t)|_2\right)^8\right] \\
&\quad\le e^{4\alpha(t_1+\tau)}\left(E\left[|U^\lambda(t_1)|_2^8\right] + n^8\right),
\end{aligned} \tag{5.21}$$

which is finite if $E[|U^\lambda(t_1)|_2^8] < \infty$. Thus inequality (5.20) implies that for all λ sufficiently large, $\delta(\lambda,\alpha) < \left(\frac{\gamma-1}{3\gamma}\right)^3$ and provided $E[|U^\lambda(t_1)|_2^8] < \infty$, we will have

$$E\left[\sup_{t\in[t_1,(t_1+\tau)\wedge\eta_n^\lambda]} e^{4\alpha t}|U^\lambda(t)|_2^8\right] \le \frac{\gamma^3 + \left(\frac{3\gamma}{\gamma-1}\right)^3\delta(\lambda,\alpha)}{1-\left(\frac{3\gamma}{\gamma-1}\right)^3\delta(\lambda,\alpha)} E\left[\sup_{t\in[t_1-\tau,t_1]} e^{4\alpha t}|U^\lambda(t)|_2^8\right]. \tag{5.22}$$

Letting $n\to\infty$, monotone convergence yields

$$E\left[\sup_{t\in[t_1,t_1+\tau]} e^{4\alpha t}|U^\lambda(t)|_2^8\right] \le \frac{\gamma^3 + \left(\frac{3\gamma}{\gamma-1}\right)^3\delta(\lambda,\alpha)}{1-\left(\frac{3\gamma}{\gamma-1}\right)^3\delta(\lambda,\alpha)} E\left[\sup_{t\in[t_1-\tau,t_1]} e^{4\alpha t}|U^\lambda(t)|_2^8\right]. \tag{5.23}$$

Define

$$\beta(\gamma,\lambda,\alpha) := \frac{\gamma^3 + \left(\frac{3\gamma}{\gamma-1}\right)^3 \delta(\lambda,\alpha)}{1 - \left(\frac{3\gamma}{\gamma-1}\right)^3 \delta(\lambda,\alpha)}.$$

Then $\lim_{\lambda\to\infty} \beta(\gamma,\lambda,\alpha) = \gamma^3$. Consider λ large enough that $\delta(\lambda,\alpha) < \left(\frac{\gamma-1}{3\gamma}\right)^3$. We now prove by induction that provided $E[\|U_0^\lambda\|_2^8] < \infty$, we have

$$E\left[\sup_{(n-1)\tau\le t\le n\tau} e^{4\alpha t}|U^\lambda(t)|_2^8\right] \;\le\; (\beta(\gamma,\lambda,\alpha))^n\, E\left[\|U_0^\lambda\|_2^8\right], \tag{5.24}$$

for $n = 0,1,2,\ldots$. For $n=0$, this holds trivially since $e^{4\alpha t} \le 1$ for all $t \in \mathbb{I}$. Now suppose that (5.24) holds for $n = k-1 \ge 0$. Since $E[|U^\lambda((k-1)\tau)|_2^8] \le (\beta(\gamma,\lambda,\alpha))^k\, E\left[\|U_0^\lambda\|_2^8\right] < \infty$ by the induction hypothesis, setting $t_1 = (k-1)\tau$ in (5.23) we obtain

$$\begin{aligned} E\left[\sup_{(k-1)\tau\le t\le k\tau} e^{4\alpha t}|U^\lambda(t)|_2^8\right] &\le \beta(\gamma,\lambda,\alpha) E\left[\sup_{(k-2)\tau\le t\le (k-1)\tau} e^{4\alpha t}|U^\lambda(t)|_2^8\right] \\ &\le (\beta(\gamma,\lambda,\alpha))^k\, E\left[\left\|U_0^\lambda\right\|_2^8\right]. \end{aligned}$$

This completes the inductive step. Now,

$$\begin{aligned} E\left[\sup_{t\ge 0} e^{\alpha t}\|U_t^\lambda\|_2^8\right] &\le e^{\alpha\tau}\left(E\left[\sup_{t\in\mathbb{I}} e^{\alpha t}|U^\lambda(t)|_2^8\right] + E\left[\sup_{t\ge 0} e^{\alpha t}|U^\lambda(t)|_2^8\right]\right) \\ &\le e^{\alpha\tau}\left(E\left[\|U_0^\lambda\|_2^8\right] + E\left[\sum_{n=1}^\infty \sup_{(n-1)\tau\le t\le n\tau} e^{\alpha t}|U^\lambda(t)|_2^8\right]\right) \\ &\le e^{\alpha\tau}\left(E\left[\|U_0^\lambda\|_2^8\right] + E\left[\sum_{n=1}^\infty e^{-3\alpha(n-1)\tau} \sup_{(n-1)\tau\le t\le n\tau} e^{4\alpha t}|U^\lambda(t)|_2^8\right]\right) \\ &\le e^{\alpha\tau}\left(E\left[\|U_0^\lambda\|_2^8\right] + \sum_{n=1}^\infty e^{-3\alpha(n-1)\tau}\,(\beta(\gamma,\lambda,\alpha))^n\, E\left[\left\|U_0^\lambda\right\|_2^8\right]\right) \\ &= e^{\alpha\tau} E\left[\|U_0^\lambda\|_2^8\right]\left(1 + e^{3\alpha\tau}\sum_{n=1}^\infty e^{-3\alpha n\tau}\,(\beta(\gamma,\lambda,\alpha))^n\right), \end{aligned} \tag{5.25}$$

which is finite provided $\beta(\gamma,\lambda,\alpha) < e^{3\alpha\tau}$. To accomplish this, we may first choose $\gamma > 1$ small enough that $\gamma^3 < \frac{1+e^{3\alpha\tau}}{2}$, and then choose $\lambda > \frac{\alpha}{2}$ large enough that both $\delta(\lambda,\alpha) < \left(\frac{\gamma-1}{3\gamma}\right)^3$ and $\beta(\gamma,\lambda,\alpha) < \gamma^3 + \left(\frac{1+e^{3\alpha\tau}}{2} - \gamma^3\right) < e^{3\alpha\tau}$. Thus, the lemma is proved with $K = e^{\alpha\tau}\left(1 + e^{3\alpha\tau}\sum_{n=1}^\infty e^{-3\alpha n\tau}\,(\beta(\gamma,\lambda,\alpha))^n\right)$.

□

Lemma 5.2.1 provides us with strong asymptotic convergence of the paths of X and $\tilde{X}^\lambda$ for large enough λ. From this point on, we shall fix a $\lambda > 0$ such that the result of Lemma

5.2.1 holds. For two segments $x, y \in \mathbb{C}^d_{\mathbb{I}}$, let $(X^x, \tilde{X}^{x,y})$ be the unique strong solution to (5.5) and (5.6) with the initial condition $(X_0, \tilde{X}^\lambda_0) \equiv (x, y)$. For each $n \geq 1$ and $x, y \in \mathbb{C}^d_{\mathbb{I}}$, define the stopping time

$$\eta^{x,y,n} := \inf\left\{t \geq 0 : \int_0^t \lambda^2 \left|\sigma^\dagger(\tilde{X}^{x,y}_s)(X^x(s) - \tilde{X}^{x,y}(s))\right|_2^2 ds \geq n\right\}.$$

Lemma 5.2.2. *For each* $x, y \in \mathbb{C}^d_{\mathbb{I}}$*,*

$$P\left(\lim_{t\to\infty} |X^x(t) - \tilde{X}^{x,y}(t)|_2 = 0\right) = 1, \quad \textit{and}$$

$$\lim_{n\to\infty} P\left(\eta^{x,y,n} = \infty\right) = 1.$$

Proof. The first claim is a direct consequence of Lemma 5.2.1.

Define the random variable $\Upsilon := \sup_{t\geq 0} e^t \|U^\lambda_t\|_2^2$, which is P-a.s. finite by Lemma 5.2.1. Then, by Assumption 5.2.1(i) and the fact that $\|U^\lambda_t\|_2^2 \leq e^{-t}\Upsilon$ for each $t \geq 0$,

$$\begin{aligned} \int_0^\infty \lambda^2 \left|\sigma^\dagger(\tilde{X}^\lambda_s)U^\lambda(s)\right|_2^2 ds &\leq \lambda^2 C_6^2 \int_0^\infty \left|U^\lambda(s)\right|_2^2 ds \\ &\leq \lambda^2 C_6^2 \int_0^\infty e^{-s}\Upsilon ds \\ &\leq \lambda^2 C_6^2 \Upsilon. \end{aligned} \tag{5.26}$$

Therefore,

$$\begin{aligned} P\left(\eta^{x,y,n} = \infty\right) &= P\left(\int_0^\infty \lambda^2 \left|\sigma^\dagger(\tilde{X}^\lambda_s)U^\lambda(s)\right|_2^2 ds < n\right) \\ &\geq P\left(\lambda^2 C_6^2 \Upsilon < n\right), \end{aligned} \tag{5.27}$$

which increases to one as $n \to \infty$ since $\Upsilon < \infty$, P-a.s..

□

Proof of Theorem 5.2.1. Define the function $N : \mathbb{C}^d_{\mathbb{I}} \times \mathbb{C}^d_{\mathbb{I}} \to \{1, 2, \dots\}$ by

$$N(x, y) := \inf\left\{n \geq 1 : P(\eta^{x,y,n} = \infty) \geq \frac{1}{2}\right\},$$

which is finite by Lemma 5.2.2. The map

$$\mathbb{C}^d_{\mathbb{J}} \times \mathbb{C}^d_{\mathbb{J}} \ni (\omega^{(1)}, \omega^{(2)}) \to \int_0^\infty \lambda^2 \left|\sigma^\dagger(\omega^{(2)}_s)(\omega^{(1)}(s) - \omega^{(2)}(s))\right|_2^2 ds$$

is measurable because it is the limit in n of the measurable maps

$$\mathbb{C}^d_{\mathbb{J}} \times \mathbb{C}^d_{\mathbb{J}} \ni (\omega^{(1)}, \omega^{(2)}) \to \int_0^n \lambda^2 \left|\sigma^\dagger(\omega^{(2)}_s)(\omega^{(1)}(s) - \omega^{(2)}(s))\right|_2^2 ds.$$

It follows from the result of Lemma 2.6.2 that the map $(x,y) \mapsto P((X^x, \tilde{X}^{x,y}) \in \Gamma)$ is measurable for each $\Gamma \in \mathcal{M}_{\mathbb{J}} \otimes \mathcal{M}_{\mathbb{J}}$. Therefore, $N(\cdot,\cdot)$ is measurable. Henceforth, we abbreviate

$$\eta^{x,y} := \eta^{x,y,N(x,y)}.$$

For each $x, y \in \mathbb{C}^d_{\mathbb{I}}$, let $v^{x,y}(t) = 1_{\{t \le \eta^{x,y}\}} \lambda \sigma^{\dagger}(\tilde{X}^{x,y}_t) \left(X^x(t) - \tilde{X}^{x,y}(t) \right)$ for $t \ge 0$. Define the process

$$\tilde{W}^{x,y}(t) := W(t) + \int_0^t v^{x,y}(s) ds, \quad t \ge 0.$$

By construction of $\eta^{x,y}$ and $v^{x,y}$,

$$\int_0^\infty |v^{x,y}(s)|_2^2 ds \ \le \ N(x,y),$$

so by Novikov's criterion (see, e.g., Proposition VIII.1.15 of [38]),

$$\rho^{x,y}(t) \ := \ \exp\left(- \int_0^t (v^{x,y}(s))' \, dW(s) - \frac{1}{2} \int_0^t |v^{x,y}(s)|_2^2 ds \right), \quad t \ge 0,$$

defines a uniformly integrable martingale. Let $\rho^{x,y}(\infty)$ denote the P-a.s. strictly positive limit of $\rho^{x,y}(t)$ as $t \to \infty$. It then follows from Girsanov's theorem (see, e.g., Section 1 of Chapter VIII of [38]) that the probability measure $Q^{x,y}$, defined by $Q^{x,y}(A) = E^P[\rho^{x,y}(\infty) 1_A]$ for all $A \in \mathcal{F}$, is equivalent to P, and under $Q^{x,y}$, $\tilde{W}^{x,y}$ is a Brownian motion $\{\mathcal{F}_t\}$-martingale. Let $\bar{X}^{x,y}$ be the unique solution under $Q^{x,y}$ to the SDDER

$$d\bar{X}(t) \quad = \quad b(\bar{X}_t)dt + \sigma(\bar{X}_t) d\tilde{W}^{x,y}(t) + d\bar{Y}(t), \tag{5.28}$$

with initial condition $\bar{X}_0 = y$. Then, P-a.s.,

$$d\bar{X}^{x,y}(t) \ = \ b(\bar{X}^{x,y}_t)dt + 1_{\{t \le \eta^{x,y}\}} \lambda \left(X^x(t) - \tilde{X}^{x,y}(t) \right) dt + \sigma(\bar{X}^{x,y}_t) dW(t) + d\bar{Y}(t), \tag{5.29}$$

where W is the Brownian motion under P. For (5.29), we used the facts that $\sigma\sigma^{\dagger} = I_d$ and

$$\begin{aligned} P&\left(\sigma^{\dagger}(\bar{X}^{x,y}_t) = \sigma^{\dagger}(\tilde{X}^{x,y}_t) \text{ for all } t \in [0, \eta^{x,y}] \cap \mathbb{R} \right) \\ &\ge \ P\left(\bar{X}^{x,y}(t) = \tilde{X}^{x,y}(t) \text{ for all } t \in [-\tau, \eta^{x,y}] \cap \mathbb{R} \right) \\ &= \ 1. \end{aligned} \tag{5.30}$$

The equality above follows by a very similar proof to that for the pathwise uniqueness for the SDDER with Lipschitz coefficients (see the proof of Theorem C.0.2 with $\eta^{x,y}$ in place of η_n).

Since uniqueness in law holds for solutions of (5.5), the distribution of $\bar{X}^{x,y}$ under $Q^{x,y}$ is the same as that of the solution X^y to (5.5) under P with initial condition $X_0 = y$. Then, since $Q^{x,y} \sim P$, the distribution of $\bar{X}^{x,y}$ under P is equivalent to that of X^y under P. Thus, if we let $\tilde{P}_{x,y}$ be the probability measure on $(\mathbb{C}^d_{\mathbb{I}})^\infty \times (\mathbb{C}^d_{\mathbb{I}})^\infty$ that is the law of $\left(\{X^x_{n\tau}\}_{n=0}^\infty, \{\bar{X}^{x,y}_{n\tau}\}_{n=0}^\infty\right)$ under P, then $\tilde{P}_{x,y}$ satisfies condition (i) of Proposition 5.2.1.

On the set $\{\eta^{x,y} = \infty\}$, we have $\bar{X}^{x,y} = \tilde{X}^{x,y}$ P-a.s. by (5.30). Thus, on $\{\eta^{x,y} = \infty\}$, P-a.s.,

$$\begin{aligned} \lim_{t\to\infty} |X^x(t) - \bar{X}^{x,y}(t)|_2 &= \lim_{t\to\infty} |X^x(t) - \tilde{X}^{x,y}(t)|_2 \\ &= 0 \end{aligned} \tag{5.31}$$

as was shown in Lemma 5.2.2. Therefore,

$$\begin{aligned} \tilde{P}_{x,y}\left((\{x^n\}_{n=0}^\infty, \{y^n\}_{n=0}^\infty) \in (\mathbb{C}^d_{\mathbb{I}})^\infty \times (\mathbb{C}^d_{\mathbb{I}})^\infty : \lim_{n\to\infty} \|x^n - y^n\| = 0\right) &\geq P(\eta^{x,y} = \infty) \\ &\geq \frac{1}{2}, \end{aligned} \tag{5.32}$$

so that $\tilde{P}_{x,y}$ also satisfies condition (ii) of Proposition 5.2.1.

All that remains to show is the measurability of $(x, y) \to \tilde{P}_{x,y}(\Gamma)$ for each $\Gamma \in \mathcal{B}((\mathbb{C}^d_{\mathbb{I}})^\infty \times (\mathbb{C}^d_{\mathbb{I}})^\infty)$, which would follow from the measurability of $(x, y) \to \bar{P}_{x,y}(B)$ for each Borel measurable $B \subset C(\mathbb{R}_+, \mathbb{C}^d_{\mathbb{I}}) \times C(\mathbb{R}_+, \mathbb{C}^d_{\mathbb{I}})$, where $\bar{P}_{x,y}$ is the law of $(X^x_\cdot, \bar{X}^{x,y}_\cdot)$ under P. This is proved in the lemma below, and completes the proof.

□

Lemma 5.2.3. *Using the notation in the proof of Theorem 5.2.1, for each measurable* $B \subset C(\mathbb{R}_+, \mathbb{C}^d_{\mathbb{I}}) \times C(\mathbb{R}_+, \mathbb{C}^d_{\mathbb{I}})$*, the map* $(x, y) \to \bar{P}_{x,y}(B)$ *is measurable.*

Proof. Using monotone class theorem arguments (see, e.g., Theorem II.3.2 in [39]), it suffices to prove the measurability of $(x, y) \mapsto E\left[g_1(X^x_{t_1}, \bar{X}^{x,y}_{t_1}) \cdots g_k(X^x_{t_k}, \bar{X}^{x,y}_{t_k})\right]$ for each collection of times $0 \leq t_1 < \cdots < t_k < \infty$ and functions $g_1, \ldots, g_k \in C_b(\mathbb{C}^{2d}_{\mathbb{I}})$ for $k = 1, 2, \ldots$.

Fix an integer $k \geq 1$. Define the sets $A_0 := \{0 \leq \eta^{x,y} < t_1\}$, $A_k := \{t_k \leq \eta^{x,y}\}$, and for each $j = 1, \ldots, k-1$, $A_j := \{t_j \leq \eta^{x,y} < t_{j+1}\}$, so that Ω is the disjoint union of

$A_0, \ldots, A_k$. Then

$$\begin{aligned}
&E\left[g_1(X^x_{t_1}, \bar{X}^{x,y}_{t_1}) \cdots g_k(X^x_{t_k}, \bar{X}^{x,y}_{t_k})\right] \\
&= \sum_{j=0}^{k} E\left[\prod_{i=1}^{k} g_i(X^x_{t_i}, \bar{X}^{x,y}_{t_i}) 1_{A_j}\right] \\
&= \sum_{j=0}^{k} E\left[\left(\prod_{i=1}^{j} g_i(X^x_{t_i}, \bar{X}^{x,y}_{t_i})\right) 1_{A_j} \left(\prod_{i=j+1}^{k} g_i(X^x_{t_i}, \bar{X}^{x,y}_{t_i})\right)\right] \\
&= \sum_{j=0}^{k} E\left[\left(\prod_{i=1}^{j} g_i(X^x_{t_i}, \bar{X}^{x,y}_{t_i})\right) 1_{A_j} E\left[\left(\prod_{i=j+1}^{k} g_i(X^x_{t_i}, \bar{X}^{x,y}_{t_i})\right) \Big| \mathcal{F}_{\eta^{x,y}}\right]\right], \qquad (5.33)
\end{aligned}$$

where we use the standard convention that $\prod_{i=1}^{0} a_i = \prod_{i=k+1}^{k} a_i = 1$ for any real numbers a_i.

By (5.30),

$$\left(\prod_{i=1}^{j} g_i(X^x_{t_i}, \bar{X}^{x,y}_{t_i})\right) 1_{A_j} = \left(\prod_{i=1}^{j} g_i(X^x_{t_i}, \tilde{X}^{x,y}_{t_i})\right) 1_{A_j}.$$

For each $\check{x}, \hat{x} \in \mathbb{C}^d_{\mathbb{I}}$, let $P^{\check{x},\hat{x}}$ denote the law induced on $C(\mathbb{R}_+, \mathbb{C}^{2d}_{\mathbb{I}})$ by the pair of strong solutions to (5.5) with the two initial conditions $\check{x}, \hat{x}$ and the same driving Brownian motion W. Now, on $\{\eta^{x,y} < \infty\}$, we define

$$\check{X}^{x,y}(t) := X^x(\eta^{x,y} + t), \quad \text{and} \quad \hat{X}^{x,y}(t) := \bar{X}^{x,y}(\eta^{x,y} + t), \quad t \geq -\tau,$$

which satisfy

$$\check{X}^{x,y}(t) = X^x(\eta^{x,y}) + \int_0^t b(\check{X}^{x,y}_s) ds + \int_0^t \sigma(\check{X}^{x,y}_s) d\hat{W}^{x,y}(s) + \check{Y}^{x,y}(t), \qquad (5.34)$$

and

$$\hat{X}^{x,y}(t) = \bar{X}^{x,y}(\eta^{x,y}) + \int_0^t b(\check{X}^{x,y}_s) ds + \int_0^t \sigma(\hat{X}^{x,y}_s) d\hat{W}^{x,y}(s) + \hat{Y}^{x,y}(t), \qquad (5.35)$$

where $\check{X}^{x,y}_0 = X^x_{\eta^{x,y}}$, $\hat{X}^{x,y}_0 = \bar{X}^{x,y}_{\eta^{x,y}}$, $\hat{W}^{x,y}(t) := W(\eta^{x,y}+t) - W(\eta^{x,y})$, $\check{Y}^{x,y}(t) := Y^x(\eta^{x,y}+t) - Y^x(\eta^{x,y})$, and $\hat{Y}^{x,y}(t) := \bar{Y}^{x,y}(\eta^{x,y}+t) - \bar{Y}^{x,y}(\eta^{x,y})$ (here, Y^x and $\bar{Y}^{x,y}$ are the regulator processes of X^x and $\bar{X}^{x,y}$). Now, on $\{\eta^{x,y} < \infty\}$, conditioned on $\mathcal{F}_{\eta^{x,y}}$, $\hat{W}^{x,y}$ is a standard m-dimensional Brownian motion independent of $\mathcal{F}_{\eta^{x,y}}$. It follows from strong uniqueness for the pair of solutions to the SDDER that the conditional law of $(\check{X}^{x,y}, \hat{X}^{x,y})$ is given by $P^{\check{x},\hat{x}}$

where $\check{x} = X^x_{\eta^{x,y}}$ and $\hat{x} = \bar{X}^{x,y}_{\eta^{x,y}}$. Then on $\{\eta^{x,y} < \infty\}$, for $j \le k$,

$$\begin{aligned} & 1_{A_j} E\left[\left(\prod_{i=j+1}^k g_i(X^x_{t_i}, \bar{X}^{x,y}_{t_i})\right) \Big| \mathcal{F}_{\eta^{x,y}}\right] \\ &= 1_{A_j} E\left[\left(\prod_{i=j+1}^k g_i(\check{X}^{x,y}_{t_i-\eta^{x,y}}, \hat{X}^{x,y}_{t_i-\eta^{x,y}})\right) \Big| \mathcal{F}_{\eta^{x,y}}\right] \\ &= 1_{A_j} \int\limits_{\left(C(\mathbb{R}_+,\mathbb{C}^d_{\mathbb{I}})\right)^2} \prod_{i=j+1}^k g_i\left(\check{w}(t_i - \eta^{x,y}), \hat{w}(t_i - \eta^{x,y})\right) P^{X^x_{\eta^{x,y}}, \tilde{X}^{x,y}_{\eta^{x,y}}}(d\check{w}, d\hat{w}). \quad (5.36) \end{aligned}$$

We note that the above expression equals 1_{A_j} if $j = k$. Let $h_k(\check{x}, \hat{x}, t) \equiv 1$ for all $(\check{x}, \hat{x}, t) \in \mathbb{C}^d_{\mathbb{I}} \times \mathbb{C}^d_{\mathbb{I}} \times \mathbb{R}_+$. For $j = 0, 1, \ldots, k-1$, we define the functions $h_j : \mathbb{C}^d_{\mathbb{I}} \times \mathbb{C}^d_{\mathbb{I}} \times [0, \infty)$ by

$$h_j(\check{x}, \hat{x}, t) := \int\limits_{\left(C(\mathbb{R}_+,\mathbb{C}^d_{\mathbb{I}})\right)^2} \prod_{i=j+1}^k g_i\left(\check{w}(t_i - t), \hat{w}(t_i - t)\right) P^{\check{x},\hat{x}}(d\check{w}, d\hat{w}), \quad \text{for } t < t_{j+1},$$

and $h_j(\check{x}, \hat{x}, t) = 0$ for $t \ge t_{j+1}$. The last line in (5.36) equals $1_{A_j} h_j(X^x_{\eta^{x,y}}, \tilde{X}^{x,y}_{\eta^{x,y}}, \eta^{x,y})$. The function h_j is measurable in $(\check{x}, \hat{x})$ for each fixed t by Lemma 2.6.2, and it is piecewise continuous in t for each fixed pair $(\check{x}, \hat{x})$. Hence, h_j is measurable (see, e.g., exercise 11 of Section 2.1 of [14]).

For each $x, y \in \mathbb{C}^d_{\mathbb{I}}$, define $\hat{P}^{x,y}$ to be the distribution of $(X^x, \tilde{X}^{x,y})$ under P on $(\tilde{\Omega} := (\mathbb{C}^d_{\mathbb{J}})^2, \tilde{\mathcal{F}} := \mathcal{B}((\mathbb{C}^d_{\mathbb{J}})^2))$, and denote by $(w, \tilde{w})$ the coordinate mapping process on $(\tilde{\Omega}, \tilde{\mathcal{F}})$. Also, define on $(\tilde{\Omega}, \tilde{\mathcal{F}})$

$$\eta := \inf\left\{t \ge 0 : \int_0^t \lambda \left|\sigma^\dagger(\tilde{w}_s)(w(s) - \tilde{w}(s))\right|_2^2 ds \ge N(w_0, \tilde{w}_0)\right\},$$

which is measurable since $N(\cdot,\cdot)$ is measurable. The random time η is also a stopping time with respect to $\left\{\tilde{\mathcal{F}}_t := \sigma\left(w_s, \tilde{w}_s, s \le t\right), t \ge 0\right\}$. Therefore $(w_\eta, \tilde{w}_\eta, \eta)$ is measurable. Combining all of the above,

$$\begin{aligned} & E\left[g_1(X^x_{t_1}, \bar{X}^{x,y}_{t_1}) \cdots g_k(X^x_{t_k}, \bar{X}^{x,y}_{t_k})\right] \\ &= \sum_{j=0}^k E\left[\left(\prod_{i=1}^j g_i(X^x_{t_i}, \tilde{X}^{x,y}_{t_i})\right) 1_{A_j} h_j(X^x_{\eta^{x,y}}, \tilde{X}^{x,y}_{\eta^{x,y}}, \eta^{x,y})\right] \\ &= \sum_{j=0}^k E^{\hat{P}^{x,y}}\left[\left(\prod_{i=1}^j g_i(w_{t_i}, \tilde{w}_{t_i})\right) 1_{B_j} h_j(w_\eta, \tilde{w}_\eta, \eta)\right], \quad (5.37) \end{aligned}$$

where B_j is defined in the same manner as A_j, but with $\eta^{x,y}$ replaced with η.

Lemma 2.6.2 implies that the mapping $(x, y) \mapsto \hat{P}^{x,y}(\Gamma)$ is measurable for each $\Gamma \in \mathcal{B}(\mathbb{C}^d_{\mathbb{J}} \times \mathbb{C}^d_{\mathbb{J}})$, and thus the measurability of the integrand implies that the expression in (5.37) is a measurable function of (x, y).

□

Corollary 5.2.1. *Under Assumptions 5.1.1 and 5.2.1, there exists a unique stationary distribution for the SDDER (1.1).*

Proof. The second remark after Assumption 5.2.1 implies that Assumptions 2.1.1 and 2.1.2 hold, so the result follows from Theorems 5.1.1 and 5.2.1. □

Chapter 6

Applications

We now apply the previous results to the Examples 2.2.1, 2.2.2, and 2.2.3.

6.1 Example 2.2.1: Biochemical Reaction System

Fix $\alpha, \gamma, \varepsilon, C > 0$.

Lemma 6.1.1. *The functions*

$$\begin{aligned} b(x) &:= \frac{\alpha}{\left(1+\frac{x(-\tau)}{C}\right)^2} - \gamma, \ \textit{for } x \in \mathbb{C}_{\mathbb{I}}, \ \textit{and} \\ \sigma(x) &:= \varepsilon \left(\frac{\alpha}{\left(1+\frac{x(-\tau)}{C}\right)^2} + \gamma \right)^{\frac{1}{2}}, \ \textit{for } x \in \mathbb{C}_{\mathbb{I}}, \end{aligned}$$

satisfy the Lipschitz condition in Assumption 5.2.1(ii).

Proof. Since the derivative of $r \mapsto \frac{1}{r^2}$ is bounded by 1 for $r \geq 1$, we have $\left|\frac{1}{r^2} - \frac{1}{s^2}\right| \leq |r-s|$ for $r, s \geq 1$. Therefore, for any $x, y \in \mathbb{C}_{\mathbb{I}}$,

$$\begin{aligned} \left| \frac{\alpha}{\left(1+\frac{x(-\tau)}{C}\right)^2} - \frac{\alpha}{\left(1+\frac{y(-\tau)}{C}\right)^2} \right| &\leq \alpha \left| \left(1+\frac{x(-\tau)}{C}\right) - \left(1+\frac{y(-\tau)}{C}\right) \right| \\ &= \frac{\alpha}{C} \left| x(-\tau) - y(-\tau) \right|. \end{aligned} \tag{6.1}$$

Therefore,

$$\begin{aligned} (b(x)-b(y))^2 &\leq \left(\frac{\alpha}{C}\right)^2 |x(-\tau)-y(-\tau)|^2 \\ &\leq \left(\frac{\alpha}{C}\right)^2 \|x-y\|_2^2. \end{aligned} \tag{6.2}$$

Using the equality $|\sqrt{r}-\sqrt{s}| = \frac{|r-s|}{\sqrt{r}+\sqrt{s}}$, we have

$$\begin{aligned}(\sigma(x)-\sigma(y))^2 &\leq \varepsilon^2 \frac{\left(\frac{\alpha}{C}\right)^2 \|x-y\|_2^2}{\left(\sqrt{\frac{\alpha}{\left(1+\frac{x(-\tau)}{C}\right)^2}+\gamma}+\sqrt{\frac{\alpha}{\left(1+\frac{y(-\tau)}{C}\right)^2}+\gamma}\right)^2} \\ &\leq \frac{\varepsilon^2}{4\gamma}\left(\frac{\alpha}{C}\right)^2 \|x-y\|_2^2. \qquad (6.3)\end{aligned}$$

□

If $x \in \mathbb{C}_{\mathbb{I}}$ such that $x(-\tau) \geq C\sqrt{\frac{2\alpha}{\gamma}}$, then $b(x) \leq -\frac{\gamma}{2}$. The dispersion coefficient is bounded by $\varepsilon\sqrt{\alpha+\gamma}$. Therefore, Assumption 4.3.1 is satisfied with $\ell(x) = x(-\tau)$, $K_d = \frac{\gamma}{2}$, $K_u = \alpha$, $C_0 = \varepsilon^2(\alpha+\gamma)$, and $M = C\sqrt{\frac{2\alpha}{\gamma}}$. Also, σ has a measurable right inverse bounded by $\frac{1}{\varepsilon\sqrt{\gamma}}$. Therefore, the SDDER associated with this (b,σ) has a unique stationary distribution.

6.2 Example 2.2.2: Affine Coefficients

Set

$$b(x) := b_0 - b_1 x(0) - \sum_{i=2}^{n} b_i x(-r_i) + \sum_{i=n+1}^{n'} b_i x(-r_i), \qquad (6.4)$$

and

$$\sigma(x) := a_0 + \sum_{i=1}^{n''} a_i x(-s_i), \qquad (6.5)$$

where $2 \leq n < n'$, $0 \leq r_i \leq \tau$ for each $i = 1, \ldots, n'$, $0 \leq s_i \leq \tau$ and $a_i \geq 0$ for each $i = 0, \ldots, n''$, $b_0 \in \mathbb{R}$, and $b_1, \ldots, b_{n'} \geq 0$. If $a_0 > 0$ and

$$\sum_{i=1}^{n} b_i > \left(\sum_{i=n+1}^{n'} b_i\right)\left(1+\tau\sum_{i=2}^{n} b_i\right) + \frac{1}{2}\left(\sum_{i=1}^{n''} a_i\right)^2 + 4\sqrt{\tau}\sum_{i=1}^{n''} a_i \sum_{i=2}^{n} b_i,$$

then the one-dimensional SDDER

$$\begin{aligned} dX(t) &= \left(b_0 - b_1 X(t) - \sum_{i=2}^{n} b_i X(t-r_i) + \sum_{i=n+1}^{n'} b_i X(t-r_i)\right) dt \\ &\quad + \left(a_0 + \sum_{i=1}^{n''} a_i X(t-s_i)\right) dW(t) + dY(t) \end{aligned}$$

will have a unique stationary distribution. The example (1.2) in Chapter 1 is a special case of this result.

Existence follows because the coefficients b and σ are uniformly Lipschitz continuous and satisfy Assumption 4.2.1 with

$$M = 0, \qquad B_0 = (b_0)^+, \qquad B_1 = b_1, \qquad B_{1,1} = \sum_{i=2}^{n} b_i,$$
$$B_{2,1} = \sum_{i=n+1}^{n'} b_i, \qquad q_1 = 1, \qquad q_2 = 2,$$
$$\ell(x) = \frac{\sum_{i=2}^{n} b_i x(-r_i)}{B_{1,1}}, \qquad \mu_1 = \frac{\sum_{i=n+1}^{n'} b_i \delta_{\{-r_i\}}}{B_{2,1}}, \qquad \mu_2 = \frac{\sum_{i=1}^{n''} a_i \delta_{\{-s_i\}}}{\sum_{i=1}^{n''} a_i},$$

where $\frac{0}{0} \equiv 0$, and for any $\gamma > 1$,

$$C_{2,1} = \gamma \left(\sum_{i=1}^{n''} a_i \right)^2 \quad \text{and} \quad C_0 = K(a_0, \gamma, 2)$$

by using Proposition B.0.2. Uniqueness follows because σ has a measurable right inverse bounded by $\frac{1}{a_0}$, so that Assumption 5.2.1 is also satisfied.

6.3 Example 2.2.3: Internet Congestion Control

Recall, the drift from Example 2.2.3 was given by

$$b^i(x) := -1 + \sum_{j=1}^{d'} A_{ij} \exp \left(-B_j \sum_{k=1}^{d} A_{kj} C_{kj} x^k(-r_{ijk}) \right), \quad i = 1, \ldots, d,$$

for some $B_1, \ldots, B_d > 0$, and $A_{ij} \geq 0$, $C_{kj} > 0$, and $r_{ijk} > 0$ for all $i, k \in \{1, \ldots, d\}$ and $j \in \{1, \ldots, d'\}$. Let us assume that $m \geq d$, and that there are $0 < a_1 < a_2$ such that $\sigma : \mathbb{C}^d_{\mathbb{I}} \to \mathbb{M}^{d \times m}$ is uniformly Lipschitz continuous and satisfies

$$a_1 |v|_2^2 \leq v' \sigma(x) \sigma(x)' v \leq a_2 |v|_2^2 \text{ for all } x \in \mathbb{C}^d_{\mathbb{I}} \text{ and } v \in \mathbb{R}^d. \tag{6.6}$$

Inequality (6.6) implies that $\sigma\sigma'$ is invertible, and thus $\sigma'(\sigma\sigma')^{-1}$ is a right-inverse for σ that is bounded above by $\frac{1}{\sqrt{a_1}}$. Inequality (6.6) also implies that $\|\sigma^i\|_2^2 \leq a_2$ on all of $\mathbb{C}^d_{\mathbb{I}}$ for all i.

For each $i = 1, \ldots, d$,

$$\begin{aligned} b^i(x) &= -1 + \sum_{j=1}^{d'} A_{ij} \exp \left(-B_j \sum_{k=1}^{d} A_{kj} C_{kj} x^k(-r_{ijk}) \right) \\ &\leq -1 + \sum_{j=1}^{d'} A_{ij} \exp \left(-B_j A_{ij} C_{ij} x^i(-r_{iji}) \right) \\ &\leq -\frac{1}{2} \end{aligned} \tag{6.7}$$

if $\sum_{j=1}^{d'} A_{ij} \exp\left(-B_j A_{ij} C_{ij} x^i(-r_{iji})\right) \leq \frac{1}{2}$, which will be true whenever

$$A_{ij} \exp\left(-B_j A_{ij} C_{ij} x^i(-r_{iji})\right) \leq \frac{1}{2d'} \text{ for each } j = 1, \ldots, d'.$$

The latter will be true if $A_{ij} = 0$ for each j; otherwise, $\max_{j=1}^{d'} A_{ij} > 0$ and $\min_{j:A_{ij}\neq 0} B_j A_{ij} C_{ij} > 0$, and (6.7) will be true whenever

$$\min_{j=1}^{d'} x^i(-r_{iji}) \geq \frac{\ln\left(2d' \max_{j=1}^{d'} A_{ij}\right)}{\min_{j:A_{ij}\neq 0} B_j A_{ij} C_{ij}}.$$

Therefore, b and σ satisfy Assumptions 4.3.1 and 5.2.1 with $\tau := \max_{i,j,k} r_{ijk}$, $\ell^i(x) = \min_{j=1}^{d'} x^i(-r_{iji})$, $K_u = \max_i \sum_{j=1}^{d'} A_{ij}$, $M = \max_{i=1}^{d} 1_{\{\max_{j=1}^{d'} A_{ij} > 0\}} \frac{\ln\left(2d' \max_{j=1}^{d'} A_{ij}\right)}{\min_{j:A_{ij}\neq 0} B_j A_{ij} C_{ij}}$, $K_d = \frac{1}{2}$, $C_0 = a_2$, and $C_6 = \frac{1}{\sqrt{a_1}}$.

Therefore, the SDDER

$$\begin{aligned} dX^i(t) \quad = \quad & \left(-1 + \sum_{j=1}^{d'} A_{ij} \exp\left(-B_j \sum_{k=1}^{d} A_{kj} C_{kj} x^k(-r_{ijk})\right)\right) dt \\ & + \sigma^i(X_t) dW(t) + dY^i(t), \quad i = 1, \ldots, d, \end{aligned}$$

has a unique stationary distribution.

Appendix A

Notation List

For the convenience of the reader, we summarize here some notation used in this paper.

- $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}_+ = [0, \infty)$, $\mathbb{R}^d = (-\infty, \infty)^d$, and $\mathbb{R}_+ = [0, \infty)^d$
- $\tau \in (0, \infty)$, $\mathbb{I} = [-\tau, 0]$, and $\mathbb{J} = [-\tau, \infty)$
- for any real numbers r, s, $\delta_{r,s}$ denotes the Kronecker delta, i.e., it is one if $r = s$ and zero otherwise
- for any metric space $(\mathbb{E}, \rho)$,
  - $B(x, r) = \{y \in \mathbb{E} : \rho(x, y) < r\}$, for each $x \in \mathbb{E}$ and $r > 0$
  - $\mathcal{B}(\mathbb{E})$ denotes the associated Borel σ-algebra
  - $C_b(\mathbb{E})$ denotes the space of continuous bounded real-valued functions on $\mathbb{E}$
  - $B_b(\mathbb{E})$ denotes the space of bounded Borel-measurable real-valued functions defined on $\mathbb{E}$ with norm $\|f\| := \sup\limits_{x \in \mathbb{E}} |f(x)|$
- for metric spaces $\mathbb{E}_1, \mathbb{E}_2$, $C(\mathbb{E}_1, \mathbb{E}_2)$ is the space of continuous functions $x : \mathbb{E}_1 \to \mathbb{E}_2$
- for each positive integer d, $C_+(\mathbb{R}_+, \mathbb{R}^d) := \{x \in C(\mathbb{R}_+, \mathbb{R}^d) : x(0) \in \mathbb{R}^d_+\}$
- given a vector $v = (v^1, \ldots, v^d)' \in \mathbb{R}^d$, $|v|_\infty = \max\limits_{i=1,\ldots d} |v^i|$, and for each $p \in [1, \infty)$, $|v|_p = \sqrt[p]{|v^i|^p + \cdots + |v^d|^p}$
- I_d denotes the $(d \times d)$-identity matrix
- given a matrix $A = (A^i_j) \in \mathbb{M}^{d \times m} := \{(d \times m)\text{-matrices with real entries}\}$

- A' is the transpose of A
- A^i denotes the i^{th} row and A_j denotes the j^{th} column of A
- $\|A\|_\infty := \max_{i,j}|A^i_j|$ denotes the l^∞-norm
- $\|A\|_2 := \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{m}(A^i_j)^2}$ denotes the Frobenius norm

- $C_0(\mathbb{R}_+, \mathbb{R}^m) = \{x \in C(\mathbb{R}_+, \mathbb{R}^m) : x(0) = 0\}$
- for any closed interval $F \subset \mathbb{R}$
  - $\mathbb{C}_F = C(F, \mathbb{R}_+)$ and $\mathbb{C}^d_F = C(F, \mathbb{R}^d_+)$
  - $\mathcal{M}_F = \mathcal{B}(\mathbb{C}^d_F)$
  - given $x = (x^1, \ldots, x^d)' \in C(F, \mathbb{R}^d)$,
    * for each $p \geq 1$ and $G \subset F$, $\|x\|_{G,p} = \sup_{r\in G}|x(r)|_p$
    * for any $\delta > 0$, $w_F(x, \delta) = \max_i \sup_{\substack{s,t\in F \\ |s-t|<\delta}} |x^i(s) - x^i(t)|$
    * for any $[a, b] \subset F$, $\text{Osc}(x, [a, b]) := \max_i \sup_{s,t\in[a,b]} |x^i(s) - x^i(t)|$
  - for each $x \in C(F, \mathbb{R})$, $\|x\|_F = \sup_{r\in F}|x(r)| = \|x\|_{F,p}$ for any $p \in [1, \infty]$
- for each $x \in \mathbb{C}_{\mathbb{I}}$, $\|x\| = \|x\|_{\mathbb{I}}$
- for each $x \in \mathbb{C}^d_{\mathbb{I}}$ and $p \in [1, \infty]$, $\|x\|_p = \|x\|_{\mathbb{I},p}$
- for any $x \in C(\mathbb{J}, \mathbb{E})$ and $t \geq 0$, the segment $x_t \in C(\mathbb{I}, \mathbb{E})$ is defined by $x_t(r) = x(t + r), r \in \mathbb{I}$
- for any $t \in F$, $e_t : C(F, \mathbb{E}) \to \mathbb{E}$ is defined by $e_t(x) = x(t)$
- for any $t \geq 0$, $p_t : C(\mathbb{J}, \mathbb{E}) \to C(\mathbb{I}, \mathbb{E})$ is defined by $p_t(x) = x_t$

Appendix B

Useful Inequalities

For referencing purposes, we state here several inequalities that are used in this paper. For any $a_1, a_2 \geq 0$, we have the inequality

$$(a_1 + a_2)^q \leq a_1^q + a_2^q, \text{ for all } q \in [0,1], \tag{B.1}$$

which is obvious if $a_1 = a_2$ or if either is 0, and if $a_1 > a_2 > 0$ then $(a_1 + a_2)^q - a_1^q \leq q a_1^{q-1} a_2 < a_2^q$. A consequence of (B.1) is that for any $a_1, a_2 \geq 0$, we have

$$(a_1 + a_2)^q \geq a_1^q + a_2^q, \text{ for all } q \geq 1. \tag{B.2}$$

From (B.1), it follows that for any $1 \leq p_1 < p_2 < \infty$, and $v \in \mathbb{R}^d$, we have

$$\begin{aligned} |v|_{p_2}^{p_1} &= \left(|v^1|^{p_2} + \cdots + |v^d|^{p_2}\right)^{\frac{p_1}{p_2}} \\ &\leq \left(|v^1|^{p_1} + \cdots + |v^d|^{p_1}\right) \\ &= |v|_{p_1}^{p_1}. \end{aligned} \tag{B.3}$$

We also have by Hölder's inequality that for any $\infty > p_1 \geq p_2 \geq 1$,

$$\begin{aligned} |v|_{p_2}^{p_1} &= \left(|v^1|^{p_2} + \cdots + |v^d|^{p_2}\right)^{\frac{p_1}{p_2}} \\ &\leq d^{\frac{p_1 - p_2}{p_2}} \left(|v^1|^{p_1} + \cdots + |v^d|^{p_1}\right) \\ &= d^{\frac{p_1 - p_2}{p_2}} |v|_{p_1}^{p_1}. \end{aligned} \tag{B.4}$$

Combining inequalities (B.3) and (B.4), we obtain for any $p_1, p_2 \in [1, \infty)$, $v \in \mathbb{R}^d$,

$$|v|_{p_2}^{p_1} \leq d^{\left(\frac{p_1}{p_2} - 1\right)^+} |v|_{p_1}^{p_1}. \tag{B.5}$$

Finally, after raising both sides of (B.5) to the power $\frac{1}{p_1}$, we obtain

$$|v|_{p_2} \leq d^{\frac{(p_1-p_2)^+}{p_1 p_2}} |v|_{p_1}. \tag{B.6}$$

We also have the following comparisons with the maximum norm: for any $p \in [1, \infty)$,

$$\begin{aligned} |v|_\infty = \max_i |v^i| = (\max_i |v^i|^p)^{\frac{1}{p}} &\leq \left(\sum_{i=1}^{d} |v^i|^p \right)^{\frac{1}{p}} = |v|_p \\ &\leq \left(d \max_i |v^i|^p \right)^{\frac{1}{p}} = d^{\frac{1}{p}} |v|_\infty, \end{aligned} \tag{B.7}$$

which can also be seen as the limiting case of (B.6).

The following is a well-known fact that follows from the convexity of power functions.

Proposition B.0.1. *For any $p > 1$, $a_1, \ldots, a_n \in \mathbb{R}$, we have*

$$|a_1 + \cdots + a_n|^p \leq n^{p-1}(|a_1|^p + \cdots + |a_n|^p).$$

Proof. Since $f(x) := x^p$ is a convex function on $\mathbb{R}$, we obtain that

$$\begin{aligned} |a_1 + \cdots + a_n|^p &= n^p \left| \frac{a_1 + \cdots + a_n}{n} \right|^p \\ &\leq n^p \left(\frac{|a_1|^p + \cdots + |a_n|^p}{n} \right) \\ &= n^{p-1}(|a_1|^p + \cdots + |a_n|^p). \end{aligned} \tag{B.8}$$

□

Sometimes n^{p-1} is too big for our needs, and we will use the following alternative.

Proposition B.0.2. *For any $\gamma > 1$, and $a, q \geq 0$, there is a $K = K(a, \gamma, q) \geq 0$ such that*

$$(a+t)^q \leq K + \gamma t^q \quad \textit{for all } t \in \mathbb{R}_+.$$

Proof. If $q \leq 1$, then we can choose $K = a^q$ because for $a_1, a_2 \geq 0$, we have inequality (B.1).

If $q > 1$, then for each $K \geq 0$, define $f_K(t) = \gamma t^q + K - (a+t)^q$ for $t \geq 0$. We must find K so that f_K is nonnegative. Since $\gamma > 1$, $\lim_{t \to \infty} f_K(t) = \infty$, and thus f_K has a global minimum on $\mathbb{R}_+$. We have that $f_K'(t) = q\gamma t^{q-1} - q(a+t)^{q-1}$ is zero only at $t = \frac{a}{\gamma^{\frac{1}{q-1}}-1}$, so that $f_K(t) \geq 0$ whenever $K \geq \left(a + \frac{a}{\gamma^{\frac{1}{q-1}}-1}\right)^q - \gamma\left(\frac{a}{\gamma^{\frac{1}{q-1}}-1}\right)^q$.

Therefore we can choose

$$K(a,\gamma,q) = a^q \vee \left(\left(a + \frac{a}{\gamma^{\frac{1}{q-1}}-1}\right)^q - \gamma\left(\frac{a}{\gamma^{\frac{1}{q-1}}-1}\right)^q\right).$$

□

The following inequality is related to Artin's inequality and the arithmetic-geometric mean inequality, and it is due to the concavity of the logarithm function.

Proposition B.0.3. *For any* $d \in \mathbb{N}$ *and* $a_1, \ldots, a_d \in \mathbb{R}_+$,

$$a_1 a_2 \cdots a_d \leq \frac{1}{d}\left(a_1^d + \cdots + a_d^d\right). \tag{B.9}$$

Proof. The function $t \mapsto \log(t)$ is concave (i.e., $t \mapsto -\log(t)$ is convex), so

$$\begin{aligned} \log(\sqrt[d]{a_1 \cdots a_d}) &= \frac{1}{d}\log(a_1 \cdots a_d) \\ &= \frac{1}{d}\left(\log(a_1) + \ldots \log(a_d)\right) \\ &\leq \log\left(\frac{a_1 + \cdots + a_d}{d}\right). \end{aligned} \tag{B.10}$$

Since the function $t \mapsto \exp(t)$ is increasing,

$$\sqrt[d]{a_1 \cdots a_d} \leq \frac{a_1 + \cdots + a_d}{d},$$

from which it follows that

$$\begin{aligned} a_1 \cdots a_d &\leq \left(\frac{a_1 + \cdots + a_d}{d}\right)^d \\ &\leq \frac{1}{d}\left(a_1^d + \cdots + a_d^d\right), \end{aligned} \tag{B.11}$$

where we've used Proposition B.0.1 for the last inequality. □

We now state Gronwall's inequality, which is used frequently. For a proof, consult [9], p. 250 and p. 262, or [22], pp. 287-288 and pp. 387-388.

Proposition B.0.4. *Fix $T > 0$. Assume that f, g are Borel measurable, nonnegative, integrable functions defined on $[0, T]$. Suppose that there is a $K > 0$ such that*

$$f(t) \leq g(t) + K \int_0^t f(s)ds, \quad \textit{for all } t \in [0, T].$$

Then

$$f(t) \leq g(t) + K \int_0^t e^{K(t-s)} g(s)ds, \quad \textit{for all } t \in [0, T].$$

If g is constant, then $f(t) \leq ge^{Kt}$.

The following simple estimate is used in Section 4.3.

Proposition B.0.5. *For each $\alpha > 0$, we have $t^2 e^{\alpha t} \leq e^{(\alpha+2)t}$ for all $t \geq 0$.*

Proof. Define $f(t) = e^{(\alpha+2)t} - t^2 e^{\alpha t}, t \geq 0$. Then, $f(0) = 1$ and $f'(t) = (\alpha + 2)e^{(\alpha+2)t} - 2te^{\alpha t} - \alpha t^2 e^{\alpha t}$, so $f'(t) \geq 0$ for all $t > 0$ if

$$e^{2t} \geq \frac{\alpha}{\alpha+2} t^2 + \frac{2}{\alpha+2} t, \text{ for all } t > 0. \tag{B.12}$$

Inequality (B.12) is true because it is true for $t = 0$, and the slope of the left side is at least the slope of the right side for $t > 0$, since $e^{2t} \geq 2t + 1$ for all $t \geq 0$.

□

Appendix C

Sufficient Conditions for Strong Existence and Uniqueness of Solutions

In this appendix, we provide specific assumptions on b and σ that imply that Assumption 2.1.2 holds. Assume that there exists a positive constant $\kappa_L < \infty$ such that inequality (5.1) holds, i.e., for each $x, y \in \mathbb{C}_{\mathbb{I}}^d$,

$$|b(x) - b(y)|_2^2 + \|\sigma(x) - \sigma(y)\|_2^2 \quad \leq \quad \kappa_L \|x - y\|_2^2.$$

The equivalence of all matrix norms can be used to show that Assumption 2.1.1 is a consequence of this Lipschitz condition.

We consider the SDDER (1.1) and show that strong existence and pathwise uniqueness hold under the Lipschitz condition (5.1). We assume $\{\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, P, \{W(t), t \geq 0\}\}$ is given with the properties stated in Definition 2.1.1. First, we state a few remarks for referencing purposes.

Remark. The Lipschitz condition (5.1) and Proposition B.0.1 imply that

$$\begin{aligned} |b(x)|_2^2 &= |b(\mathbf{0}) + b(x) - b(\mathbf{0})|_2^2 \\ &\leq 2\,|b(\mathbf{0})|_2^2 + 2\kappa_L \|x\|_2^2, \quad \text{and} \qquad \text{(C.1)} \\ \|\sigma(x)\|_2^2 &= \|\sigma(\mathbf{0}) + \sigma(x) - \sigma(\mathbf{0})\|_2^2 \\ &\leq 2\|\sigma(\mathbf{0})\|_2^2 + 2\kappa_L \|x\|_2^2 \qquad \text{(C.2)} \end{aligned}$$

for each $x \in \mathbb{C}_{\mathbb{I}}$, where **0** here stands for the element of $\mathbb{C}_{\mathbb{I}}^d$ that is identically 0 on all of $\mathbb{I}$.

Remark. The Lipschitz condition (5.1) also implies that for each $f, g \in \mathbb{C}^d_{\mathbb{J}}$,

$$\int_0^t \left(|b(f_s) - b(g_s)|_2^2 + \|\sigma(f_s) - \sigma(g_s)\|_2^2 \right) ds \leq \int_0^t \kappa_L \|f_s - g_s\|_2^2 ds \leq \int_0^t \kappa_L \|f - g\|^2_{[-\tau,s],2} ds \tag{C.3}$$

Remark. It follows from the definition of the oscillation of x that

$$\text{Osc}(x, [a,b]) \leq 2\|x\|_{[a,b],\infty}. \tag{C.4}$$

We begin with a lemma. Recall the notation $\mathcal{I}(\cdot)$ from line (2.7).

Lemma C.0.1. *If $X(t), \tilde{X}(t), X^\dagger(t), \tilde{X}^\dagger(t), t \geq -\tau$ and $Y(t), Y^\dagger(t), t \geq 0$ are continuous $\mathbb{R}^d_+$-valued processes such that $X_0 = \tilde{X}_0$, $X_0^\dagger = \tilde{X}_0^\dagger$, and $(X|_{\mathbb{R}_+}, Y)$ solves the Skorokhod problem for $\mathcal{I}(\tilde{X})$ and $(X^\dagger|_{\mathbb{R}_+}, Y^\dagger)$ solves the Skorokhod problem for $\mathcal{I}(\tilde{X}^\dagger)$, then assuming $E\left[\|\tilde{X}\|^2_{[-\tau,t],2}\right] \vee E\left[\|\tilde{X}^\dagger\|^2_{[-\tau,t],2}\right] < \infty$, we have*

$$E\left[\|X\|^2_{[-\tau,t],2}\right] \leq (24d + 1 + 2d)E\left[\|X_0\|_2^2\right] + 192d\|\sigma(\mathbf{0})\|_2^2 t + 48d|b(\mathbf{0})|_2^2 t^2 + (48t + 192)\kappa_L d \int_0^t E\left[\|\tilde{X}_r\|_2^2\right] dr, \tag{C.5}$$

and

$$E\left[\|X - X^\dagger\|^2_{[0,t],2}\right] \leq 3K_\ell^2(1 + \kappa_L t(t+4))E\left[\|X_0 - X_0^\dagger\|_2^2\right] + 3K_\ell^2 \kappa_L (t+4) \int_0^t E\left[\|\tilde{X} - \tilde{X}^\dagger\|^2_{[0,r],2}\right] dr. \tag{C.6}$$

Proof. By inequality (B.7) and the definitions, for each $t \geq 0$,

$$\begin{aligned}
\|X\|^2_{[0,t],2} &= \sup_{s\in[0,t]} |X(s)|_2^2 \\
&\leq d \sup_{s\in[0,t]} |X(s)|_\infty^2 \\
&\leq d\left(|X(0)|_\infty + \sup_{r,s\in[0,t]} |X(s) - X(r)|_\infty \right)^2 \\
&\leq 2d|X(0)|_2^2 + 2d(\text{Osc}(X, [0,t]))^2.
\end{aligned} \tag{C.7}$$

It follows from inequality (C.4) and Proposition 2.3.1(i) that for each $t \geq 0$,

$$\begin{aligned}
\|X\|^2_{[-\tau,t],2} &\leq \|X_0\|_2^2 + \|X\|^2_{[0,t],2} \\
&\leq \|X_0\|_2^2 + 2d|X_0(0)|_2^2 + 2d(\text{Osc}(X, [0,t]))^2 \\
&\leq (1+2d)\|X_0\|_2^2 + 2d(\text{Osc}(\mathcal{I}(\tilde{X}), [0,t]))^2 \\
&\leq (1+2d)\|X_0\|_2^2 + 8d\|\mathcal{I}(\tilde{X})\|^2_{[0,t],\infty}.
\end{aligned} \tag{C.8}$$

After fixing $t \geq 0$, for each $r \in [0,t]$, it follows from inequality (C.2) that

$$\begin{aligned} E\left[\|\sigma(\tilde{X}_r)\|_2^2\right] &\leq 2\|\sigma(\mathbf{0})\|_2^2 + 2\kappa_L E\left[\|\tilde{X}_r\|_2^2\right] &\text{(C.9)} \\ &\leq 2\|\sigma(\mathbf{0})\|_2^2 + 2\kappa_L E\left[\|\tilde{X}\|_{[-\tau,t],2}^2\right], &\text{(C.10)} \end{aligned}$$

which is finite by assumption.

Therefore, Itô's isometry implies that for each $s \in [0,t]$,

$$\begin{aligned} E\left[\left|\int_0^s \sigma(\tilde{X}_r)dW(r)\right|_2^2\right] &= E\left[\sum_{i=1}^d \left|\int_0^s \sigma^i(\tilde{X}_r)dW(r)\right|^2\right] \\ &= \sum_{i=1}^d E\left[\int_0^s \sum_{j=1}^m \left(\sigma_j^i(\tilde{X}_r)\right)^2 dr\right] \\ &= E\left[\int_0^s \|\sigma(\tilde{X}_r)\|_2^2 dr\right] \\ &\leq E\left[\int_0^s \left(C_3 + C_4\|\tilde{X}_r\|_2^2\right) dr\right] \\ &\leq tC_3 + tC_4 E\left[\|\tilde{X}\|_{[-\tau,t],2}^2\right], &\text{(C.11)} \end{aligned}$$

which is finite by assumption.

It follows from (C.11) that $\left\{\int_0^s \sigma(\tilde{X}_r)dW(r), \mathcal{F}_s, s \in [0,t]\right\}$ is a square-integrable martingale. Therefore Doob's inequality and the Itô isometry imply that

$$E\left[\sup_{0\leq s\leq t}\left|\int_0^s \sigma(\tilde{X}_r)dW(r)\right|_2^2\right] \leq 4E\left[\int_0^t \left\|\sigma(\tilde{X}_r)\right\|_2^2 dr\right]. \tag{C.12}$$

For each $t \geq 0$, we have

$$\begin{aligned} E\left[\|\mathcal{I}(\tilde{X})\|_{[0,t],2}^2\right] &\leq 3E\left[|X(0)|_2^2\right] + 3E\left[\sup_{0\leq s\leq t}\left|\int_0^s b(\tilde{X}_r)dr\right|_2^2\right] \\ &\quad +3E\left[\sup_{0\leq s\leq t}\left|\int_0^s \sigma(\tilde{X}_r)dW(r)\right|_2^2\right] \\ &\leq 3E\left[\|X_0\|_2^2\right] + 3tE\left[\int_0^t |b(\tilde{X}_r)|_2^2 dr\right] \\ &\quad +12E\left[\int_0^t \|\sigma(\tilde{X}_r)\|_2^2 dr\right] \\ &\leq 3E\left[\|X_0\|^2\right] + 3t\int_0^t E\left[2|b(\mathbf{0})|_2^2 + 2\kappa_L\|\tilde{X}_r\|_2^2\right] dr \\ &\quad +12\int_0^t E\left[2\|\sigma(\mathbf{0})\|_2^2 + 2\kappa_L\|\tilde{X}_r\|_2^2\right] dr. &\text{(C.13)} \end{aligned}$$

The first inequality is a consequence of convexity, the second inequality follows from the Cauchy-Schwarz inequality and inequality (C.12), and the third inequality used inequalities (C.1) and (C.9). Combining (C.8) with (C.13) yields the first result.

Using Proposition 2.3.1,

$$\begin{aligned}
E\left[\|X - X^{\dagger}\|^2_{[0,t],2}\right] &\leq K_\ell^2 E\left[\|\mathcal{I}(\tilde{X}) - \mathcal{I}(\tilde{X}^{\dagger})\|^2_{[0,t],2}\right] \\
&\leq 3K_\ell^2 E\left[|X(0) - X^{\dagger}(0)|_2^2\right] + 3K_\ell^2 E\left[\sup_{0\leq s\leq t}\left|\int_0^s \left(b(\tilde{X}_r) - b(\tilde{X}_r^{\dagger})\right) dr\right|_2^2\right] \\
&\quad +3K_\ell^2 E\left[\sup_{0\leq s\leq t}\left|\int_0^s \left(\sigma(\tilde{X}_r) - \sigma(\tilde{X}_r^{\dagger})\right) dW(r)\right|_2^2\right] \\
&\leq 3K_\ell^2 E\left[\|X_0 - X_0^{\dagger}\|_2^2\right] + 3K_\ell^2 E\left[\sup_{0\leq s\leq t} s\int_0^s \left|b(\tilde{X}_r) - b(\tilde{X}_r^{\dagger})\right|_2^2 dr\right] \\
&\quad +12K_\ell^2 E\left[\int_0^t \left\|\sigma(\tilde{X}_r) - \sigma(\tilde{X}_r^{\dagger})\right\|_2^2 dr\right] \\
&\leq 3K_\ell^2 E\left[\|X_0 - X_0^{\dagger}\|_2^2\right] + 3K_\ell^2 \kappa_L (t+4)\int_0^t E\left[\|\tilde{X} - \tilde{X}^{\dagger}\|^2_{[-\tau,r],2}\right] dr \\
&\leq 3K_\ell^2 (1+\kappa_L t(t+4)) E\left[\|X_0 - X_0^{\dagger}\|_2^2\right] + 3K_\ell^2 \kappa_L (t+4)\int_0^t E\left[\|\tilde{X} - \tilde{X}^{\dagger}\|^2_{[0,r],2}\right] dr
\end{aligned}$$

The second inequality used the Cauchy-Schwarz inequality and Doob's inequality, which we can use by (C.10) and the analogous inequality involving $\tilde{X}^{\dagger}$, and the third inequality follows from inequality (C.3).

□

Theorem C.0.1. *Under the global Lipschitz condition (5.1), given any $\mathbb{C}_{\mathbb{I}}^d$-valued, $\mathcal{F}_0$-measurable random variable ξ such that $E[\|\xi\|_2^2] < \infty$, there exists a strong solution X to (1.1) with initial condition $X_0 = \xi$.*

The following existence proof is a standard argument using Picard's iteration technique. It was adapted from the proof of Theorem 11 in [20].

Proof. Define the processes $\{X^{(n)}(t), t \geq -\tau\}, n = 0, 1, \dots$ inductively by

$$X^{(0)}(t) = \begin{cases} \xi(t) & \text{for } t \in \mathbb{I} \\ \xi(0) & \text{for } t \geq 0, \end{cases}$$

and for $n \in \{1, 2, \dots\}$,

$$X^{(n+1)}(t) = \begin{cases} \xi(t) & \text{for } t \in \mathbb{I} \\ \mathcal{I}(X^{(n)})(t) + Y^{(n+1)}(t) & \text{for } t \geq 0, \end{cases} \tag{C.14}$$

i.e., $\left\{\left(X^{(n+1)}(t), Y^{(n+1)}(t)\right), t \geq 0\right\}$ solves the one-dimensional Skorokhod problem (with normal reflection) for $\mathcal{I}(X^{(n)})$. Recall that

$$\mathcal{I}(X^{(n)})(t) = \xi(0) + \int_0^t b(X_s^{(n)})ds + \int_0^t \sigma(X_s^{(n)})dW(s) \quad \text{for } t \geq 0. \tag{C.15}$$

By definition, for each $t \geq 0$, $E\left[\|X^{(0)}(s)\|^2_{[-\tau,t],2}\right] = E\left[\|\xi\|_2^2\right] < \infty$. Assume for some integer $n \geq 0$, that $E\left[\|X^{(n)}\|^2_{[-\tau,t],2}\right] < \infty$ for each $t \geq 0$. Then by Lemma C.0.1,

$$\begin{aligned} E\left[\|X^{(n+1)}\|^2_{[-\tau,t],2}\right] \quad &\leq \quad (24d + 1 + 2d)E[\|\xi\|_2^2] + 192d\|\sigma(\mathbf{0})\|_2^2 t \\ &\quad +48d|b(\mathbf{0})|_2^2 t^2 \\ &\quad +(48t + 192)\kappa_L d \int_0^t E\left[\|X_r^{(n)}\|^2_{[-\tau,t],2}\right] dr, \end{aligned} \tag{C.16}$$

which is finite by assumption. So by induction, $E\left[\|X^{(n)}\|^2_{[-\tau,t],2}\right] < \infty$ for every integer $n \geq 0$.

Lemma C.0.1 also implies that for each $n \geq 1$,

$$E\left[\|X^{(n+1)} - X^{(n)}\|^2_{[0,t],2}\right] \leq 3K_\ell^2\kappa_L(t+4)\int_0^t E\left[\|X^{(n)} - X^{(n-1)}\|^2_{[0,r],2}\right] dr, \tag{C.17}$$

since $X_0^{(n+1)} = X_0^{(n)} = \xi$.

For each $n \in \{0, 1, \dots\}$, define

$$f_n(t) := E\left[\|X^{(n+1)} - X^{(n)}\|^2_{[0,t],2}\right], \text{ for } t \geq 0.$$

It follows from Lemma C.0.1, the fact that $E\left[\|X^{(0)}\|^2_{[-\tau,t],2}\right] = E[\|\xi\|^2]$, and inequality (C.16) with $n = 0$ that

$$\begin{aligned} f_0(t) \quad &= \quad E\left[\sup_{0\leq s\leq t}\left|X^{(1)}(s) - X^{(0)}(s)\right|_2^2\right] \\ &\leq \quad 2\left(E[|\xi(0)|_2^2] + E\left[\|X^{(1)}\|^2_{[0,t],2}\right]\right) \\ &\leq \quad K(t), \end{aligned}$$

where $K(t)$ is the following quadratic polynomial in t:

$$\begin{aligned} K(t) \quad &= \quad 96d\left(|b(\mathbf{0})|_2^2 + \kappa_L E[\|\xi\|_2^2]\right)t^2 + 384d\left(\|\sigma(\mathbf{0})\|_2^2 + \kappa_L E[\|\xi\|_2^2]\right)t \\ &\quad +2(24d + 2 + d)E[\|\xi\|_2^2]. \end{aligned} \tag{C.18}$$

By induction, we'll show that for each integer $n \geq 0$,

$$f_n(t) \quad \leq \quad K(t)\frac{\left(3K_\ell^2\kappa_L\right)^n(t+4)^n t^n}{n!}, \text{ for all } t \geq 0, \tag{C.19}$$

which is evident for $n = 0$ (we use the convention that $0! = 1$). So suppose (C.19) holds for some $n \geq 0$ (we have already shown it holds for $n = 0$). Using inequality (C.17) and the fact that $K(t)$ is increasing on $\mathbb{R}_+$, we obtain

$$\begin{aligned} f_{n+1}(t) &\leq 3K_\ell^2\kappa_L(t+4)\int_0^t f_n(s)ds \\ &\leq 3K_\ell^2\kappa_L(t+4)\int_0^t K(s)\frac{\left(3K_\ell^2\kappa_L\right)^n (s+4)^n s^n}{n!}ds \\ &\leq K(t)\left(3K_\ell^2\kappa_L\right)^{n+1}(t+4)^{n+1}\int_0^t \frac{s^n}{n!}ds \\ &= K(t)\left(3K_\ell^2\kappa_L\right)^{n+1}(t+4)^{n+1}\frac{t^{n+1}}{(n+1)!}, \end{aligned}$$

so (C.19) holds for all $n \geq 0$.

Therefore by Chebychev's inequality, for each $\varepsilon > 0$ and $t \geq 0$,

$$\begin{aligned} P\left(\left\|X^{(n+1)} - X^{(n)}\right\|_{[0,t],2} > \varepsilon\right) &\leq \varepsilon^{-2}E\left[\left\|X^{(n+1)} - X^{(n)}\right\|_{[0,t],2}^2\right] \\ &\leq \varepsilon^{-2}f_n(t) \\ &\leq \varepsilon^{-2}K(t)\frac{\left(3K_\ell^2\kappa_L\right)^n (t+4)^n t^n}{n!}. \end{aligned}$$

Setting $\varepsilon = \frac{1}{n}$ and $t = \log(n)$ and using the fact that $\|X_0^{(n+1)} - X_0^{(n)}\| = 0$ for all $n \geq 0$, we have

$$\begin{aligned} &P\left(\left\|X^{(n+1)} - X^{(n)}\right\|_{[-\tau,\log(n)],2} > \frac{1}{n}\right) \\ &\leq n^2 K(\log(n))\frac{\left(3K_\ell^2\kappa_L\right)^n (\log(n)+4)^n(\log(n))^n}{n!}. \end{aligned} \tag{C.20}$$

Using the ratio test and the facts that $\lim_{n\to\infty}\frac{K(\log(n))}{K(\log(n+1))} = \lim_{n\to\infty}\left(\frac{\log(n)+k}{\log(n+1)+k}\right)^n = 1$ for all $k \geq 0$, the terms in the right member of (C.20) sum to a finite quantity. Therefore the Borel-Cantelli lemma implies that the sequence of processes $\{X^{(n)}\}$ is P-a.s. Cauchy in $\mathbb{C}_\mathbb{J}$ for each $t \geq 0$, and hence uniformly convergent on any compact interval, and thus P-a.s. has a continuous limit $\{X(t), t \geq -\tau\}$ in the topology of $\mathbb{C}_\mathbb{J}$. Thus, for each $n \geq 0$ and $t \geq 0$, we have

$$\begin{aligned}
\left| \|X - X^{(n)}\|_{[-\tau,t],2} - \left\| \sum_{k=n}^{\infty} \left(X^{(k+1)} - X^{(k)} \right) \right\|_{[-\tau,t],2} \right| & \\
\leq & \left\| X - X^{(n)} - \sum_{k=n}^{\infty} \left(X^{(k+1)} - X^{(k)} \right) \right\|_{[-\tau,t],2} \\
= & \left\| X - X^{(n)} - \lim_{N\to\infty} \sum_{k=n}^{N} \left(X^{(k+1)} - X^{(k)} \right) \right\|_{[-\tau,t],2} \\
= & \left\| X - \lim_{N\to\infty} X^{(N)} \right\|_{[-\tau,t],2} \\
= & \ 0. \qquad \text{(C.21)}
\end{aligned}$$

Because of (C.19) and the definition of f_n, for each $t \geq 0$,

$$\left(X^{(n+1)}(t) - X^{(0)}(t) \right) = \sum_{k=0}^{n} \left(X^{(k+1)}(t) - X^{(k)}(t) \right)$$

converges absolutely in $L^2(\Omega, P)$. Therefore $\{X^{(n)}(t)\}_{n=0}^{\infty}$ converges in $L^2(\Omega, P)$ for each $t \geq 0$. Since $X^{(n)}(t) \to X(t)$ P-a.s. for each $t \geq 0$, we obtain

$$E\left[|X^{(n)}(t) - X(t)|_2^2 \right] \to 0 \quad \text{as} \quad n \to \infty. \qquad \text{(C.22)}$$

Also, by (C.21), for each $t \geq 0$ and $n \geq 1$,

$$\begin{aligned}
E\left[\|X - X^{(n)}\|^2_{[-\tau,t],2}\right]^{\frac{1}{2}} &= E\left[\sup_{-\tau\leq s\leq t}\left|\sum_{k=n}^{\infty}\left(X^{(k+1)}(s) - X^{(k)}(s)\right)\right|_2^2\right]^{\frac{1}{2}} \\
&\leq E\left[\sup_{-\tau\leq s\leq t}\left(\sum_{k=n}^{\infty}\left|X^{(k+1)}(s) - X^{(k)}(s)\right|_2\right)^2\right]^{\frac{1}{2}} \\
&= E\left[\left(\sup_{-\tau\leq s\leq t}\sum_{k=n}^{\infty}\left|X^{(k+1)}(s) - X^{(k)}(s)\right|_2\right)^2\right]^{\frac{1}{2}} \\
&\leq E\left[\left(\sum_{k=n}^{\infty}\sup_{-\tau\leq s\leq t}\left|X^{(k+1)}(s) - X^{(k)}(s)\right|_2\right)^2\right]^{\frac{1}{2}} \\
&= E\left[\varliminf_{N\to\infty}\left(\sum_{k=n}^{N}\sup_{-\tau\leq s\leq t}\left|X^{(k+1)}(s) - X^{(k)}(s)\right|_2\right)^2\right]^{\frac{1}{2}} \\
&\leq \varliminf_{N\to\infty} E\left[\left(\sum_{k=n}^{N}\sup_{-\tau\leq s\leq t}\left|X^{(k+1)}(s) - X^{(k)}(s)\right|_2\right)^2\right]^{\frac{1}{2}} \\
&\leq \varliminf_{N\to\infty}\sum_{k=n}^{N} E\left[\sup_{-\tau\leq s\leq t}\left|X^{(k+1)}(s) - X^{(k)}(s)\right|_2^2\right]^{\frac{1}{2}} \\
&\leq \sum_{k=n}^{\infty}\left(K(t)\frac{\left(3K_\ell^2\kappa_L\right)^k (t+4)^k t^k}{k!}\right)^{\frac{1}{2}}, \qquad \text{(C.23)}
\end{aligned}$$

which approaches 0 as $n \to \infty$ by the ratio test. Fatou's Lemma was used for the third inequality, and the triangle inequality for the $\mathscr{L}^2(\Omega)$-norm (with finite sums) was used for the fourth inequality.

Using inequality (C.3), we obtain for $t \geq 0$ and $n \geq 1$ fixed,

$$\begin{aligned}
&E\left[\sup_{0\leq s\leq t}\left|\int_0^s\left(b(X_r) - b(X_r^{(n)})\right)dr\right|_2^2\right] \\
&\qquad\leq E\left[t\int_0^t\left|b(X_s) - b(X_s^{(n)})\right|_2^2 ds\right] \\
&\qquad\leq tE\left[\int_0^t \kappa_L\|X - X^{(n)}\|^2_{[-\tau,s],2}ds\right] \\
&\qquad\leq t^2\kappa_L E\left[\|X - X^{(n)}\|^2_{[-\tau,t],2}\right] \\
&\qquad\leq t^2\kappa_L\left(\sum_{k=n}^{\infty}\left(K(t)\frac{\left(3K_\ell^2\kappa_L\right)^k (t+4)^k t^k}{k!}\right)^{\frac{1}{2}}\right)^2. \qquad \text{(C.24)}
\end{aligned}$$

In particular, for each $t \geq 0$,

$$\int_0^t b(X_s^{(n)})ds \to \int_0^t b(X_s)ds \text{ in } L^2(\Omega) \text{ as } n \to \infty. \tag{C.25}$$

The Itô isometry yields for each $t \geq 0$,

$$\begin{aligned}
E\left[\left|\int_0^t \left(\sigma(X_s) - \sigma(X_s^{(n)})\right) dW(s)\right|_2^2\right] & \\
&= E\left[\int_0^t \left\|\sigma(X_s) - \sigma(X_s^{(n)})\right\|_2^2 ds\right] \\
&\leq E\left[\int_0^t \kappa_L \|X - X^{(n)}\|^2_{[-\tau,s],2} ds\right] \\
&\leq t\kappa_L E\left[\|X - X^{(n)}\|^2_{[-\tau,t],2}\right] \\
&\leq t\kappa_L \left(\sum_{k=n}^{\infty}\left(K(t)\frac{\left(3K_\ell^2\kappa_L\right)^k (t+4)^k t^k}{k!}\right)^{\frac{1}{2}}\right)^2, \qquad \text{(C.26)}
\end{aligned}$$

so that for each $n \geq 1$, $\left\{\int_0^t \left(\sigma(X_s) - \sigma(X_s^{(n)})\right) dW(s), \mathcal{F}_t, t \geq 0\right\}$ is a square-integrable martingale, and for each fixed $t \geq 0$,

$$\int_0^t \sigma(X_s^{(n)})dW(s) \to \int_0^t \sigma(X_s)dW(s) \text{ in } L^2(\Omega) \text{ as } n \to \infty. \tag{C.27}$$

Lines (C.25) and (C.27) imply that for each $t \geq 0$,

$$\mathcal{I}(X^{(n)})(t) \to \mathcal{I}(X)(t) \text{ in } L^2(\Omega) \text{ as } n \to \infty. \tag{C.28}$$

By (C.26), we may apply Doob's inequality to obtain

$$\begin{aligned}
E\left[\sup_{0\leq s\leq t}\left|\int_0^s \left(\sigma(X_r)dr - \sigma(X_r^{(n)})\right) dW(r)\right|_2^2\right] & \\
&\leq 4E\left[\int_0^t \left\|\sigma(X_r) - \sigma(X_r^{(n)})\right\|_2^2 dr\right] \\
&\leq 4t\kappa_L \left(\sum_{k=n}^{\infty}\left(K(t)\frac{\left(3K_\ell^2\kappa_L\right)^k (t+4)^k t^k}{k!}\right)^{\frac{1}{2}}\right)^2. \qquad \text{(C.29)}
\end{aligned}$$

Setting $Y(t) = \psi\left(\mathcal{I}(X)\right)(t)$ for $t \geq 0$, Proposition 2.3.1 and inequalities (C.24) and (C.29) imply that

$$E\left[|Y^{(n)}(t) - Y(t)|_2^2\right] \leq K_\ell^2 E\left[\|\mathcal{I}(X^{(n)}) - \mathcal{I}(X)\|^2_{[0,t],2}\right] \to 0 \text{ as } n \to \infty. \tag{C.30}$$

Taking the limit as $n \to \infty$ in line (C.14), the facts (C.22), (C.28), and (C.30) imply that for each $t \geq 0$, $X(t)$ satisfies equation (1.1) P-a.s.. Since $X_0^{(n)} = \xi$ for all $n \geq 0$, X satisfies the initial condition $X_0 = \xi$.

By induction, $X^{(n)}(t)$ and $Y^{(n)}(t)$ are $\mathcal{F}_t$-measurable for each $t \geq 0$. Therefore for each $t \geq 0$, $X(t)$ and $Y(t)$ are $\mathcal{F}_t$-measurable, as they are the P-a.s. limits of the $\mathcal{F}_t$-measurable random variables $X^{(n)}(t)$ and $Y^{(n)}(t)$, respectively, and by assumption, $\mathcal{F}_t$ is P-complete.

□

Theorem C.0.2. *Fix a $\mathbb{C}_{\mathbb{I}}^d$-valued $\mathcal{F}_0$-measurable random variable ξ. Under the Lipschitz condition (5.1), any solution X to (1.1) with the initial condition $X_0 = \xi$ is unique up to indistinguishability.*

Proof. Suppose that X and X' both solve (1.1) with the same initial condition ξ, so that for any $t \geq 0$,

$$\begin{aligned} X(t) - X'(t) &= \int_0^t \big(b(X_s) - b(X'_s)\big)\, ds \\ &\quad + \int_0^t \big(\sigma(X_s) - \sigma(X'_s)\big)\, dW(s) + (Y(t) - Y'(t)). \end{aligned}$$

Here $Y = \psi(\mathcal{I}(X))$ and $Y' = \psi(\mathcal{I}(X'))$. For $t \geq 0$,

$$\begin{aligned} |\mathcal{I}(X)(t) - \mathcal{I}(X')(t)|_2^2 &\leq 2\left|\int_0^t (b(X_s) - b(X'_s))ds\right|_2^2 \\ &\quad + 2\left|\int_0^t (\sigma(X_s) - \sigma(X'_s))dW(s)\right|_2^2, \end{aligned} \tag{C.31}$$

from which it follows via the Cauchy-Schwarz inequality that

$$\begin{aligned} \|\mathcal{I}(X) - \mathcal{I}(X')\|_{[0,t],2}^2 &\leq 2t\int_0^t \left|b(X_r) - b(X'_r)\right|_2^2 dr \\ &\quad + \sup_{0\leq s\leq t} 2\left|\int_0^s \big(\sigma(X_r) - \sigma(X'_r)\big)\, dW(r)\right|_2^2. \end{aligned} \tag{C.32}$$

Since the coefficients b and σ are bounded on bounded subsets of $\mathbb{C}_{\mathbb{I}}^d$, by stopping at the stopping time $\eta_n = 0 \vee \inf\{t \geq -\tau : |X(t)|_2 \vee |X'(t)|_2 \geq n\}$, we obtain the square-integrable d-dimensional martingale

$$\left\{\int_0^{t\wedge\eta_n} \big(\sigma(X_s) - \sigma(X'_s)\big)\, dW(s), \mathcal{F}_t, t \geq 0\right\}.$$

Using Doob's inequality and the fact that $X_0 \equiv X_0'$, we have for $t \geq 0$,

$$
\begin{aligned}
E&\left[\sup_{0\leq s\leq t}\left|\mathcal{I}(X)(s\wedge\eta_n)-\mathcal{I}(X')(s\wedge\eta_n)\right|_2^2\right] \\
&\leq 2E\left[t\int_0^{t\wedge\eta_n}\left|b(X_s)-b(X_s')\right|_2^2 ds\right] \\
&\qquad +8E\left[\int_0^{t\wedge\eta_n}\left\|\sigma(X_s)-\sigma(X_s')\right\|_2^2 ds\right] \\
&\leq 2\kappa_L(t+4)\int_0^t E\left[\sup_{0\leq s\leq r}\left|X(s\wedge\eta_n)-X'(s\wedge\eta_n)\right|_2^2\right] dr \\
&\leq 2K_\ell^2\kappa_L(t+4)\int_0^t E\left[\sup_{0\leq s\leq r}\left|\mathcal{I}(X)(s\wedge\eta_n)-\mathcal{I}(X')(s\wedge\eta_n)\right|_2^2\right] dr, \qquad \text{(C.33)}
\end{aligned}
$$

where the second inequality follows from the remark (C.3), and the third inequality follows from Proposition 2.3.1. Gronwall's inequality then implies that

$$
E\left[\sup_{0\leq s\leq t}\left|\mathcal{I}(X)(s\wedge\eta_n)-\mathcal{I}(X')(s\wedge\eta_n)\right|_2^2\right] = 0, \text{ for all } t\geq 0. \qquad \text{(C.34)}
$$

Proposition 2.3.1 and Fatou's lemma then imply that for each $t \geq 0$,

$$
\begin{aligned}
E\left[\sup_{0\leq s\leq t}\left|X(s)-X'(s)\right|_2^2\right] &\leq K_\ell^2 E\left[\sup_{0\leq s\leq t}\left|\mathcal{I}(X)(s)-\mathcal{I}(X')(s)\right|_2^2\right] \\
&\leq \varliminf_{n\to\infty} K_\ell^2 E\left[\sup_{0\leq s\leq t}\left|\mathcal{I}(X)(s\wedge\eta_n)-\mathcal{I}(X')(s\wedge\eta_n)\right|_2^2\right] \\
&= 0. \qquad \text{(C.35)}
\end{aligned}
$$

Since $X_0 = X_0'$, it follows that $P\left(X(t) = X'(t) \text{ for all } t \geq -\tau\right) = 1$.

□

Definition C.0.1. The uniqueness proved in Theorem C.0.2 is called *pathwise uniqueness of solutions for (1.1)* (see Definition IV.1.5 of [21], or Definition 5.3.2 of [22]). If instead we had that for any two weak solutions $\{(\Omega,\mathcal{F},P),\{\mathcal{F}_t\},X,W\}$ and $\{(\Omega',\mathcal{F}',P'),\{\mathcal{F}_t'\},X',W'\}$ to (1.1) with the same initial distribution (i.e., $P(X_0 \in \Lambda) = P'(X_0' \in \Lambda)$ for each $\Lambda \in \mathcal{M}_{\mathbb{I}}$) that $P(X \in \Gamma) = P'(X' \in \Gamma)$ for each $\Gamma \in \mathcal{M}_{\mathbb{J}}$, then we say that *uniqueness in law* holds for (1.1).

Since the spaces $\mathbb{C}_{\mathbb{I}}^d, \mathbb{C}_{\mathbb{J}}^d$ and $C_0(\mathbb{R}_+,\mathbb{R}^m)$ are all Polish spaces, the techniques of Yamada and Watanabe (see, e.g., the Corollary to Theorem IV.1.1 in [21], or Proposition 5.3.20 in [22]) can be applied to our situation to yield the following corollary.

Corollary C.0.1. *The Lipschitz condition (5.1) implies that uniqueness in law holds.*

The following corollary will be used to relax the assumption of global Lipschitz continuity in Theorems C.0.1 and C.0.2.

Corollary C.0.2. *Let $\Lambda \subset \mathbb{C}_{\mathbb{I}}^d$ be open, and assume that $b(x) = b'(x)$ and $\sigma(x) = \sigma'(x)$ for each $x \in \Lambda$, and that all four satisfy the global Lipschitz condition (5.1). Denote the unique strong solutions to the SDDER (1.1) with coefficients b, σ and b', σ' starting with identical initial conditions $X_0 = X_0'$, by X and X', respectively. Then the exit times $T_\Lambda(X) := \inf\{t \geq 0 : X_t \in \Lambda^c\}$ and $T_\Lambda(X')$ are identical off a null set, and*

$$P\left(X(t \wedge T_\Lambda(X)) = X'(t \wedge T_\Lambda(X')), \text{ for all } t \geq -\tau\right) = 1.$$

Proof. The assumptions on the coefficients imply that

$$\mathcal{I}'(X')(s \wedge T_\Lambda(X') \wedge T_\Lambda(X)) = \mathcal{I}(X')(s \wedge T_\Lambda(X') \wedge T_\Lambda(X)) \text{ for each } s \geq 0,$$

where $\mathcal{I}'$ is defined as $\mathcal{I}$ but with the coefficients b', σ' in place of b, σ. Therefore,

$$\begin{aligned} & E\left[\sup_{0\leq s\leq t} \left|X(s \wedge T_\Lambda(X') \wedge T_\Lambda(X)) - X'(s \wedge T_\Lambda(X') \wedge T_\Lambda(X))\right|_2^2\right] \\ \leq \quad & K_\ell^2 E\left[\sup_{0\leq s\leq t} \left|\mathcal{I}(X)(s \wedge T_\Lambda(X') \wedge T_\Lambda(X)) - \mathcal{I}'(X')(s \wedge T_\Lambda(X') \wedge T_\Lambda(X))\right|_2^2\right] \\ = \quad & K_\ell^2 E\left[\sup_{0\leq s\leq t} \left|\mathcal{I}(X)(s \wedge T_\Lambda(X') \wedge T_\Lambda(X)) - \mathcal{I}(X')(s \wedge T_\Lambda(X') \wedge T_\Lambda(X))\right|_2^2\right]. \end{aligned} \tag{C.36}$$

First we assume that Λ is bounded. The coefficients b, σ are bounded on bounded sets, so that

$$\left\{\int_0^{s \wedge T_\Lambda(X') \wedge T_\Lambda(X)} \left(\sigma(X_r) - \sigma(X_r')\right) dW(r), \mathcal{F}_s, s \geq 0\right\}$$

is a square-integrable martingale. Then as in the proof of Theorem C.0.2, we can use Gronwall's inequality to obtain equality (C.34) with $T_\Lambda(X') \wedge T_\Lambda(X)$ in place of η_n (in fact, they are the same stopping times when $\Lambda = B(\mathbf{0}, n)$, where $B(\mathbf{0}, n) := \{x \in \mathbb{C}_{\mathbb{I}}^d : \|x\|_2 \leq n\}$). Thus the set

$$F := \left\{X(t \wedge T_\Lambda(X') \wedge T_\Lambda(X)) = X'(t \wedge T_\Lambda(X') \wedge T_\Lambda(X)), \text{ for all } t \geq -\tau\right\}$$

has full measure. On $F \cap \{T_\Lambda(X) \leq T_\Lambda(X') < \infty\}$ we have that $X'(T_\Lambda(X') \wedge T_\Lambda(X)) = X(T_\Lambda(X') \wedge T_\Lambda(X)) = X(T_\Lambda(X)) \in \Lambda^c$ and thus $T_\Lambda(X') = T_\Lambda(X)$. Similarly, $T_\Lambda(X') = T_\Lambda(X)$ on $F \cap \{T_\Lambda(X') \leq T_\Lambda(X) < \infty\}$. On $F \cap \{T_\Lambda(X) = \infty\}$, $X(t) \in \Lambda$ for all $t \geq 0$, so that $X'(t \wedge T_\Lambda(X')) = X'(t \wedge T_\Lambda(X') \wedge T_\Lambda(X)) = X(t \wedge T_\Lambda(X') \wedge T_\Lambda(X)) \in \Lambda$, which implies that $t < T_\Lambda(X')$ for all $t \geq 0$, and thus $T_\Lambda(X') = \infty = T_\Lambda(X)$. Similarly, $T_\Lambda(X') = T_\Lambda(X)$ on $F \cap \{T_\Lambda(X') = \infty\}$, and thus $T_\Lambda(X') = T_\Lambda(X)$ on all of F, and the result follows.

If Λ is not bounded, then consider $\Lambda_n := \Lambda \cap B(\mathbf{0}, n)$ for each $n \geq 1$. It was shown that $P(X(t \wedge T_{\Lambda_n}(X)) = X'(t \wedge T_{\Lambda_n}(X')),$ for all $t \geq -\tau) = 1$ for each $n \geq 1$. By tightness of a finite number of distributions, for any $\varepsilon > 0$ and $t \geq 0$, there exists $N > 0$ such that

$$\begin{aligned} 1 - \varepsilon &< P\left(\|X\|_{[-\tau, t \wedge T_\Lambda(X)]} \vee \|X'\|_{[-\tau, t\wedge T_\Lambda(X)]} < N\right) \\ &\leq P\left(t \wedge T_\Lambda(X) = t \wedge T_{\Lambda_N}(X), t \wedge T_\Lambda(X') = t \wedge T_{\Lambda_N}(X')\right) \\ &\leq P\left(X(t \wedge T_\Lambda(X)) = X'(t \wedge T_\Lambda(X'))\right). \end{aligned} \tag{C.37}$$

Since $\varepsilon > 0$ was arbitrary, $P(X(t \wedge T_\Lambda(X)) \neq X'(t \wedge T_\Lambda(X'))) = 0$. Since $t \geq 0$ was arbitrary and $X_0 = X_0'$, continuity of X and X' implies that

$$\begin{aligned} & P(X(t \wedge T_\Lambda(X)) \neq X'(t \wedge T_\Lambda(X')) \text{ for some } t \geq -\tau) \\ &= P(X(t \wedge T_\Lambda(X)) \neq X'(t \wedge T_\Lambda(X')) \text{ for some } t \in \mathbb{Q} \cap \mathbb{R}_+) \\ &\leq \sum_{t \in \mathbb{Q} \cap \mathbb{R}_+} P(X(t \wedge T_\Lambda(X)) \neq X'(t \wedge T_\Lambda(X'))) \\ &= 0. \end{aligned} \tag{C.38}$$

The result follows.

□

We now have a tool that allows us to weaken the conditions of Theorem C.0.1 by means of a standard technique (see [40], Section V.12) under the assumption that b and σ are locally Lipschitz.

Assumption C.0.1. *The coefficients b, σ are locally Lipschitz, i.e., for each $N > 0$, there is a κ_N such that for all x, y with $\|x\|_2, \|y\|_2 \leq N$, we have*

$$|b(x) - b(y)|_2^2 + \|\sigma(x) - \sigma(y)\|_2^2 \leq \kappa_N \|x - y\|_2^2. \tag{C.39}$$

Clearly, Assumption C.0.1 follows if the global Lipschitz condition (5.1) holds. When the local Lipschitz condition does not hold, there are still occasions when there will exist a weak solution to (1.1) that is unique in law, but often times the linear growth conditions (2.1) and (2.2) will be required to prevent explosion in finite time.

Theorem C.0.3. *Under Assumptions 2.1.1 and C.0.1, there exists a unique strong solution to the SDDER (1.1) for any $\mathcal{F}_0$-measurable initial condition X_0.*

Proof. Define for each $n \geq 1$, the functions $b_n(x) = c_n(\|x\|_2)b(x)$ and $\sigma_n(x) = c_n(\|x\|_2)\sigma(x)$, where $c_n : \mathbb{R}_+ \to [0,1]$ is defined as $c_n(r) = 1 - (r-n)^+ + (r-(n+1))^+$, so that $c_n(r) = 0$ whenever $r \geq n+1$, and $c_n(r) = 1$ whenever $r \leq n$. Then b_n, σ_n are globally Lipschitz. Indeed, without loss of generality, we can assume that $\|x\|_2 \leq \|y\|_2$, so that $c_n(\|x\|_2) \geq c_n(\|y\|_2)$. If $\|x\|_2 \geq n+1$, then $b_n(x) = b_n(y) = 0$. Otherwise, using the triangle inequality and the Lipschitz continuity of c_n, we obtain

$$\begin{aligned} |b_n(x) - b_n(y)|_2^2 &= |(c_n(\|x\|_2) - c_n(\|y\|_2))b(x) + c_n(\|y\|_2)(b(x) - b(y))|_2^2 \\ &\leq 2\,|(c_n(\|x\|_2) - c_n(\|y\|_2))b(x)|_2^2 + 2\,|c_n(\|y\|_2)(b(x)-b(y))|_2^2 \\ &\leq 2\,|\|x\|_2 - \|y\|_2|^2\,(C_1 + C_2(n+1))^2 + 2\kappa_{n+1}\|x-y\|_2^2 \\ &\leq 2\left((C_1 + C_2(n+1))^2 + \kappa_{n+1}\right)\|x-y\|_2^2. \end{aligned} \tag{C.40}$$

A similar inequality holds for σ_n, so that for all $x, y \in \mathbb{C}_{\mathbb{I}}^d$,

$$\begin{aligned} &|b_n(x) - b_n(y)|_2^2 + \|\sigma_n(x) - \sigma_n(y)\|_2^2 \\ &\quad\leq \left(4\kappa_{n+1} + 2\big(C_1 + C_2(n+1)\big)^2 + 2\big(C_3 + C_4(n+1)^2\big)\right)\|x-y\|^2. \end{aligned} \tag{C.41}$$

It follows from Assumption 2.1.1 that b_n, σ_n grow at most linearly with the same growth constants C_1, C_2, C_3, C_4, since $c_n(\|x\|_2) \leq 1$ for all $x \in \mathbb{C}_{\mathbb{I}}^d$.

Thus for each $n \geq 1$ and $\mathbb{C}_{\mathbb{I}}^d$-valued random variable ξ such that $E[\|\xi\|_2^2] < \infty$, there is a unique strong solution $X^{(n)}$ to (1.1) with the coefficients b_n, σ_n in place of b, σ with initial condition $X_0^{(n)} = \xi$. Corollary C.0.2 implies that these solutions are consistent in the sense that $X^{(n+1)}(t \wedge T_{B(\mathbf{0},n)}(X^{(n+1)})) = X^{(n+1)}(t \wedge T_{B(\mathbf{0},n)}(X^{(n)})) = X^{(n)}(t \wedge T_{B(\mathbf{0},n)}(X^{(n)}))$ for all $t \geq -\tau$, P-a.s.. Therefore, we can define a solution X to (1.1) (until the explosion time $T^{(\infty)}(X) := \lim_{n\to\infty} T_{B(\mathbf{0},n)}(X) = \lim_{n\to\infty} T_{B(\mathbf{0},n)}(X^{(n)})$) via localization by setting $X(t) = X^{(n)}(t)$ whenever $t \in [-\tau, T_{B(\mathbf{0},n)}(X^{(n)})]$. Lemma 2.4.1 implies that

$$E\left[\|X^{(n)}\|_{[-\tau,t],2}^2\right] \leq F_2(E[\|\xi\|_2^2], t), \qquad \text{for each } n \geq 1,$$

since F_2 depends only on the linear growth constants of the coefficients, which in our case are the same for each n. Thus for each $t \geq 0$,

$$\begin{aligned} P\left(t > T_{B(\mathbf{0},n)}(X^{(n)})\right) &= P\left(\|X^{(n)}\|_{[-\tau,t],2} > n\right) \\ &\leq \frac{1}{n^2}E\left[\|X^{(n)}\|_{[-\tau,t],2}^2\right] \leq \frac{1}{n^2}F_2(E[\|\xi\|_2^2], t) \\ &\to 0 \quad \text{as} \quad n \to \infty. \end{aligned}$$

Therefore X does not explode, i.e., $T^{(\infty)} = \infty$ P-a.s., and we can define the solution on all of $\mathbb{R}_+$. Since each $X^{(n)}(t)$ was the unique solution to (1.1) for $t \in [0, T_{B(\mathbf{0},n)}(X^{(n)})]$, it follows that X is unique.

$\square$

Bibliography

[1] Anderson, D. F. (2007). A modified next reaction method for simulating chemical systems with time dependent propensities and delays, *J. Chem. Phys.* **127** 214107.

[2] Atar R., Budhiraja, A., and Dupuis P. (2001). On positive recurrence of constrained diffusion processes. *Ann. Prob.* Vol. **29**, No. 2 979-1000.

[3] Bakhtin, Y. and Mattingly, J. C. (2005). Stationary solutions of stochastic differential equation with memory and stochastic partial differential equations. *Commun. Comtemp. Math.* 7 (5), 553-582.

[4] Berman, A. and Plemmons, R. J. (1979). *Nonnegative Matrices in the Mathematical Sciences.* Academic Press, New York.

[5] Billingsley, P. (1999). *Convergence of Probability Measures* (2nd Edition). J. Wiley & Sons, New York.

[6] Bratsun, D., Volfson, D., Tsimring, L. S. and Hasty, J. (2005). Delay-induced stochastic oscillations in gene regulation, *PNAS,* **102** 14593–14598.

[7] Brauer, F. and Castillo-Chávez C. (2001). *Mathematical Models in Population Biology and Epidemiology,* Springer, New York.

[8] Chung, K. L. (1982). *Lectures from Markov Processes to Brownian Motion.* Springer-Verlag, New York.

[9] Chung, K. L. and Williams, R. J. (1990). *Introduction to Stochastic Integration* (2nd Edition). Birkhäuser, Boston.

[10] Conway, J. B. (1978). *Functions of One Complex Variable I.* Springer-Verlag, New York

[11] Da Prato, G. and Zabczyk, J. (1996). *Ergodicity for Infinite Dimensional Systems.* Cambridge University Press, Cambridge.

[12] Dupuis, P. and Ishii, H. (1991). On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications. *Stochastics and Stochastics Reports* **35**, 31-62.

[13] Filippov, A. F. (1988). *Differential Equations with Discontinuous Righthand Sides.* Kluwer Academic Publishers, Dordrecht.

[14] Folland, G. B. (1999). *Real Analysis: Modern Techniques and Their Applications* (2nd edition). Wiley-Interscience, New York. Prentice-Hall, Englewood Cliffs, New Jersey.

[15] Hairer, M. and Ohashi, A. (2007). Ergodic theory for SDEs with extrinsic memory, *Ann. Prob.* **35** 1950–1977.

[16] Hairer, M., Mattingley, J. C., and Scheutzow, M. (2009). Asymptotic coupling and a weak form of Harris' theorem with applications to stochastic delay equations, preprint, http://arxiv.org/abs/0902.4495.

[17] Hale, J. K. and Lunel, S. M. (1993). *Introduction to Functional Differential Equations.* Applied Mathematical Sciences. Springer-Verlag New York.

[18] Harrison, J. M. and Reiman, M. I. (1981). Reflected Brownian motion on an orthant. *Ann. Prob.* **9** 302-308.

[19] Harrison, J. M. and Williams, R. J. (1987). Brownian models of open queueing networks with homogeneous customer populations. *Stochastics* **22** 77-115.

[20] Itô, K. and Nisio, M. (1964). On stationary solutions of a stochastic differential equation. *J. Math. Kyoto Univ.* **4** 1-75.

[21] Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* Elsevier-North Holland, New York.

[22] Karatzas, I. and Shreve, S. (1998). *Brownian Motion and Stochastic Calculus* (2nd Edition). Springer, New York.

[23] Kurtz, T. G. and Protter, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Prob.* **19** 3 1035-1070.

[24] Kushner, H. J. (2001). *Heavy Traffic Analysis of Controlled Queueing and Communication Networks.* Springer-Verlag, New York.

[25] Kushner, H. J. (2008). *Numerical methods for controlled stochastic delay systems.* Birkhäuser, Boston.

[26] Liptser, R. S., and Shiryayev A. N. (1977). *Statistics of Random Processes I General Theory*. Springer-Verlag, New York.

[27] Mao, X. (1994). *Exponential Stability of Stochastic Differential Equations.* Marcel Dekker.

[28] Mao, X. (1996). Razumikhin-type theorems on exponential stability of stochastic functional-differential equations. *Stochastic Process. Appl.* **65** 2 233-250.

[29] Mather, W., Bennett, M. R., Hasty, J., and Tsimring L. S. (2009). Delay-induced degrade-and-fire oscillations in small genetic circuits. *Physical Review Letters*, **102** 068105, pp. 1-4.

[30] Mohammed, S-E. A. (1984). *Stochastic Functional Differential Equations.* Pitman Advanced Publishing Program, Boston, London, Melbourne.

[31] Mohammed, S-E. A. (1986). Stability of linear delay equations under a small noise. *Proc. Edinburgh Math. Soc.* **29** 2 233-254.

[32] Paganini, F., Doyle, J., and Low, S. (2001). Scalable Laws for stable network congestion control. *Proceedings of the 40th IEEE Conference on Decision and Control*. Orlando, Fl.

[33] Paganini, F. and Wang, Z. (2002). Global stability with time-delay in network congestion control. *Proceedings of the 41st IEEE Conference on Decision and Control* December 2002. Las Vegas, NV. 3632-3637.

[34] Papachristodoulou, A. (2004). Global stability analysis of a TCP/AQM protocol for arbitrary networks with delay. *43rd IEEE Conference on Decision and Control*, December 14-17, 2004. Altlantis, Paradise Island, Bahamas. 1029-1034.

[35] Papchristodoulou, A., Doyle J. and Low, S. (2004). Analysis of nonlinear delay differential equation models of TCP/AQM protocols using sums of squares. *43rd IEEE Conference on Decision and Control*, December 14-17, 2004. Atlantis, Paradise Island, Bahamas.

[36] Peet, M. and Lall, S (2007). Global stability analysis of a nonlinear model of internet congestion control with delay. *IEEE Transactions on Automatic Control* **52** 3 553-559.

[37] Reiß, M., Riedle, M., and van Gaans, O. (2006). Delay differential equations driven by Lévy processes: Stationarity and Feller properties. *Stochastic Processes and their Applications* **116** 1409-1432.

[38] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion* (3rd Edition). Springer-Verlag, Berlin, Heidelberg, New York.

[39] Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes and Martingales Volume 1: Foundations* (2nd Edition). Cambridge University Press, Cambridge.

[40] Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes and Martingales Volume 2: Itô Calculus* (2nd Edition). Cambridge University Press, Cambridge.

[41] Scheutzow, M. (1983). Qualitative behaviour of stochastic delay equations with a bounded memory. *Stochastics* **12** 41-80.

[42] Scheutzow, M. (2000). Stability and instability of routings through a queueing network. *Queueing Systems Theory Appl.* **35** 117-128.

[43] Scheutzow, M. (2004). Exponential growth rates for stochastic delay differential equations. *Stochastics and Dynamics* **5** 2 163-174.

[44] Srikant, R. (2004). *The Mathematics of Internet Congestion Control.* Birkhäuser, Boston.

[45] Williams, R. J. (1995). Semimartingale reflecting Brownian motions in the orthant. In *Stochastic Networks*, eds. F.P. Kelly and R.J. Williams, IMA Vol. Math. Appl., **71**, Springer, New York, p. 125–137.

[46] Williams, R.J. (1998). An invariance principle for semimartingale reflecting Brownian motions in an orthant. *Queueing Systems Theory Appl.* **30** 5-25.