

On Existence and Uniqueness of Stationary Distributions for Stochastic Delay Differential Equations with Positivity Constraints*

M. S. Kinnally and R. J. Williams

Department of Mathematics

University of California, San Diego

9500 Gilman Drive

La Jolla CA 92093-0112 USA

Email: williams@math.ucsd.edu

Webpage: <http://www.math.ucsd.edu/~williams>

Abstract

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

Keywords and phrases: stochastic differential equation, delay equation, stationary distribution, normal reflection, Lyapunov/Razumikhin-type argument, asymptotic coupling.

AMS 2000 subject classifications: 34K50, 37H10, 60H10, 60J25, 93E15.

Submitted to EJP on August 18, 2009. Final version accepted on April 9, 2010.

*Research supported in part by NSF grants DMS-0604537, DMS-0825686 and DMS-0906535.

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 [25, 37], neuronal models where the spatial distribution of neurons can result in delayed dynamics, epidemiological models where incubation periods result in delayed transmission of disease [5], and biochemical reactions in gene regulation where lengthy transcription and translation operations have been modeled with delayed dynamics [1, 4, 21]. There is an extensive literature, both theoretical and applied on ordinary delay differential equations. The book [13] by Hale and Lunel provides an introduction to this vast subject.

In some applications, the quantities of interest are naturally positive. For instance, rates and prices in Internet models are positive, concentrations of ions or chemical species and proportions of a population that are infected are all naturally positive quantities. In deterministic differential equation models for the delayed dynamics of such quantities, the dynamics may naturally keep the quantities positive or they may need to be adapted to be so, sometimes leading to piecewise continuous delay differential dynamics, see e.g., [25, 26, 27, 28, 29]. 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) [2, 7, 11, 15, 19, 20, 22, 23, 30, 34, 35, 36]. To obtain the analogue of such SDDE models with positivity constraints, in general, it is not simply a matter of adding a noise term to the ordinary differential equation dynamics, as this will frequently not lead to a solution respecting the state constraint, especially if the dispersion coefficient depends on a delayed state.

As described above, there is natural motivation for considering stochastic differential equations where all three features, delay, positivity 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., [17]), 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 positivity 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 positivity constraints of the form:

$$X(t) = X(0) + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW(s) + Y(t), \quad t \geq 0, \quad (1)$$

where $X(t)$ takes values in the closed 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, and Y is a vector-valued non-decreasing process which ensures that the positivity 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) as stochastic delay differential equations with reflection, where the action of Y is termed reflection (at the boundary of the orthant).

This paper is organized as follows. Our assumptions on the coefficients b and σ for well-posedness of (1), the rigorous definition of a solution of (1), and some properties of solutions are given in Section 2.1. Our main result giving sufficient conditions for existence and uniqueness of stationary distributions for (1) is stated in Section 2.2, and some examples of applications of the result are given in Section 2.3. In preparation for Section 3, a useful a priori moment bound on solutions to (1) is given in Section 2.4. Section 3 focuses on establishing sufficient conditions for existence of stationary distributions. A general condition guaranteeing existence is described in Section 3.1. This condition is in terms of uniform moment bounds, and it is fairly standard. Such bounds for second moments are shown to hold in Sections 3.4 and 3.5, under certain conditions on b and σ . The results of Sections 3.1, 3.4 and 3.5 are combined to give sufficient conditions for existence of a stationary distribution in Section 3.6. Our proofs of the moment bounds use stochastic Lyapunov/Razumikhin-type arguments applied to suitable functions of an overshoot process which is introduced in Section 3.2. For these arguments, the positive oscillation of a path, which is introduced in Section 3.3, proves to be a useful refinement of the usual notion of oscillation of a path. While our main results in Section 3 are new, we do use some results and adapt some techniques developed by Itô and Nisio [15] and Mao [20] for stochastic delay differential equations without reflection. Conditions for uniqueness of a stationary distribution are given in Section 4. Our proofs in that section are an adaptation of methods developed recently by Hairer, Mattingly and Scheutzow [12] for proving uniqueness of stationary distributions for stochastic delay differential equations without reflection. An important new aspect of the results in [12] 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, in contrast to prior results on uniqueness of stationary distributions for stochastic delay differential equations which often restricted to cases where the dispersion coefficient depended only on the current state $X(t)$ of the process [7, 17, 34, 36], with notable exceptions being [15, 30]. The important feature that distinguishes the results of [12] from those of [15, 30] is that the authors of [12] obtain uniqueness of the stationary distribution without requiring the existence of a unique random fixed point; see Section 4 for further discussion of this point. Appendix A states some well-known facts about reflection, and Appendix B covers some inequalities that appear frequently throughout this work.

1.2 Notation and Terminology

We shall use the following notation and terminology throughout this work.

For a real number a , we shall say that a is positive if $a \geq 0$ and we shall say that a is strictly positive if $a > 0$. For each strictly positive integer d , let \mathbb{R}^d denote d -dimensional Euclidean space, and let $\mathbb{R}_+^d = \{v \in \mathbb{R}^d : v_i \geq 0 \text{ for } i = 1, \dots, d\}$ 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 $i = 1, \dots, d$, 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 $i = 1, \dots, d$. 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., $v = (v^1, \dots, v^d)'$. 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 . Given $v \in \mathbb{R}^d$, $|v| = (v \cdot v)^{\frac{1}{2}}$, the Euclidean norm of 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_j^i 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. Given a $d \times m$ matrix A , $|A| := \left(\sum_{i=1}^d \sum_{j=1}^m (A_j^i)^2 \right)^{\frac{1}{2}}$ denotes the Frobenius norm of A .

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, continuous real-valued functions on \mathbb{E} will be denoted by $C_b(\mathbb{E})$.

For any two metric spaces $\mathbb{E}_1, \mathbb{E}_2$, let $C(\mathbb{E}_1, \mathbb{E}_2)$ denote the set 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 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 $\mathbb{C}_I^d = C(I, \mathbb{R}_+^d)$.

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

$$\text{Osc}(x, [a_1, a_2]) := \max_{i=1}^d \sup_{s, t \in [a_1, a_2]} |x^i(t) - x^i(s)|, \quad (2)$$

the modulus of continuity of x over I by

$$w_I(x, \delta) := \max_{i=1}^d \sup_{\substack{s, t \in I \\ |s-t| < \delta}} |x^i(t) - x^i(s)|, \quad \delta > 0, \quad (3)$$

and the supremum norm of x over I by

$$\|x\|_I = \sup_{t \in I} |x(t)|.$$

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 $C(\mathbb{I}, \mathbb{R}^d)$, $\mathbb{C}_{\mathbb{I}}^d$ has norm

$$\|x\| := \|x\|_{\mathbb{I}}, \quad x \in \mathbb{C}_{\mathbb{I}}^d,$$

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

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 outcome space Ω , P is a probability measure on the measurable space (Ω, \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., (Ω, \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 (Ω, \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 $\nu(\Lambda) = 0$. By a continuous process, we mean a process with all paths continuous.

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

- (i) $W(0) = 0$ a.s.,
- (ii) the coordinate processes, W^1, \dots, 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: $W^i(t_2) - W^i(t_1), W^i(t_3) - W^i(t_2), \dots, W^i(t_n) - W^i(t_{n-1})$, 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 $f : \{1, 2, \dots\} \rightarrow \mathbb{R}$ and $a \in (-\infty, \infty]$, the notation $f(n) \nearrow a$ as $n \rightarrow \infty$ means that $\lim_{n \rightarrow \infty} f(n) = a$ and $f(n) \leq f(n+1)$ for each $n = 1, 2, \dots$

2 Stochastic Delay Differential Equations with Reflection

In this section, we define our assumptions and the notion of a solution to equation (1) precisely. We state our main result and give some examples of its application. We also derive some useful properties of solutions to (1).

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 $\mathbb{I} = [-\tau, 0]$, $\mathbb{J} = [-\tau, \infty)$, $\mathbb{C}_{\mathbb{I}}^d = C(\mathbb{I}, \mathbb{R}_+^d)$ and $\mathbb{C}_{\mathbb{J}}^d = C(\mathbb{J}, \mathbb{R}_+^d)$. Furthermore, we fix positive integers d and m , and functions $b : \mathbb{C}_{\mathbb{I}}^d \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{C}_{\mathbb{I}}^d \rightarrow \mathbb{M}^{d \times m}$ that satisfy the following uniform Lipschitz assumption. Although we do not need as strong an assumption as this for well-posedness of (1), we will use this condition in proving uniqueness of stationary distributions. Accordingly, we shall *assume the following condition holds throughout this work*.

Assumption 2.1. *There exists a constant $\kappa_L > 0$ such that*

$$|b(x) - b(y)|^2 + |\sigma(x) - \sigma(y)|^2 \leq \kappa_L \|x - y\|^2 \quad \text{for all } x, y \in \mathbb{C}_{\mathbb{I}}^d. \quad (4)$$

Remark. A simple consequence of the Lipschitz condition (4) is that there exist strictly positive constants C_1, C_2, C_3 and C_4 such that for each $x \in \mathbb{C}_{\mathbb{I}}^d$,

$$|b(x)| \leq C_1 + C_2\|x\|, \text{ and} \quad (5)$$

$$|\sigma(x)|^2 \leq C_3 + C_4\|x\|^2. \quad (6)$$

Definition 2.1.1. Given a standard m -dimensional Brownian motion martingale $W = \{W(t), t \geq 0\}$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, P)$, a solution of the stochastic delay differential equation with reflection (SDDER) associated with (b, σ) is a d -dimensional continuous process $X = \{X(t), t \in \mathbb{J}\}$ on (Ω, \mathcal{F}, P) that P -a.s. satisfies (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) $Y = \{Y(t), t \geq 0\}$ 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) $\int_0^t X(s) \cdot dY(s) = 0$ for all $t \geq 0$, i.e., Y^i can increase only when X^i is at zero for $i = 1, \dots, d$.

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. 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 $t \rightarrow X_t(\omega)$ is continuous from \mathbb{R}_+ into $\mathbb{C}_{\mathbb{I}}^d$ for each $\omega \in \Omega$. It follows that the mapping $F : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{C}_{\mathbb{I}}^d$, where $F(t, \omega) = X_t(\omega)$, is progressively measurable, being continuous in t and adapted (see Lemma II.73.10 of [32]). Continuity of σ now implies that $\left\{ \int_0^t \sigma(X_s) dW(s), \mathcal{F}_t, t \geq 0 \right\}$ is a continuous d -dimensional local martingale; also, condition (ii) and continuity of b implies that $\{X(0) + \int_0^t b(X_s) ds + Y(t), \mathcal{F}_t, t \geq 0\}$ is a continuous adapted process that is locally of bounded variation. Therefore, $\{X(t), t \geq 0\}$ is a continuous semimartingale with respect to $\{\mathcal{F}_t, t \geq 0\}$.

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 , all defined on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, we define

$$\mathcal{I}(\xi)(t) := \xi(0) + \int_0^t b(\xi_s) ds + \int_0^t \sigma(\xi_s) dW(s), \quad t \geq 0. \quad (7)$$

For a solution X of the SDDER, $X(t) = \mathcal{I}(X)(t) + Y(t), t \geq 0$, where the regulator term, Y , has the following explicit formula in terms of $\mathcal{I}(X)$: for each $i = 1, \dots, d$,

$$Y^i(t) = \max_{s \in [0, t]} \left((\mathcal{I}(X))^i(s) \right)^-, \quad t \geq 0.$$

In the notation of Appendix A, $X = \phi(\mathcal{I}(X))$ and $Y = \psi(\mathcal{I}(X))$, because of the uniqueness of solutions to the Skorokhod problem; thus, Y is a function of X (cf. (108)). Then as a consequence of Proposition A.0.1(i), we have the following.

Proposition 2.1.1. For any $0 \leq a < b < \infty$

$$\text{Osc}(X, [a, b]) \leq \text{Osc}(\mathcal{I}(X), [a, b]). \quad (8)$$

Strong existence and uniqueness of a solution to (1) is a consequence of Assumption 2.1. We state this as a proposition. The proof is fairly standard and so we just sketch it.

Proposition 2.1.2. Given a Brownian motion martingale $\{W(t), t \geq 0\}$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, P)$ and an \mathcal{F}_0 -measurable $\mathbb{C}_{\mathbb{I}}^d$ -valued random element ξ , there exists a unique solution X to the SDDER (1) with initial condition $X_0 = \xi$ and driving Brownian motion W . Let X^x denote the solution with $X_0 = x$ in $\mathbb{C}_{\mathbb{I}}^d$. Then the associated family

$$\{P_t(x, \Lambda) := P(X_t^x \in \Lambda), t \geq 0, x \in \mathbb{C}_{\mathbb{I}}^d, \Lambda \in \mathcal{M}_{\mathbb{I}}\}$$

of transition functions is Markovian and Feller continuous.

Sketch of proof. As a consequence of Proposition 2.1.1 and the uniform Lipschitz properties of b , σ and ϕ , 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 for each $T > 0$ there exist constants $K_T, K'_T \geq 0$ such that for all stopping times η ,

$$E \left[\|X - X^\dagger\|_{[0, T \wedge \eta]}^2 \right] \leq K_T E \left[\|X_0 - X_0^\dagger\|^2 \right] + K'_T \int_0^T E \left[\|\tilde{X} - \tilde{X}^\dagger\|_{[0, r \wedge \eta]}^2 \right] dr. \quad (9)$$

By truncating the initial condition, we can reduce the proof of existence to the case where $E[\|\xi\|^2] < \infty$. Existence in this case then follows by a standard Picard iteration using (9) (see, e.g., Chapter 10 of [6]). Gronwall's inequality is used to prove uniqueness. Feller continuity follows from the standard argument that given a sequence of deterministic initial conditions $\{x^n\}_{n=1}^\infty$ such that $\lim_{n \rightarrow \infty} x^n = x \in \mathbb{C}_{\mathbb{I}}^d$, the sequence of distributions $\{P(X^{x^n} \in \cdot)\}_{n=1}^\infty$ on $(\mathbb{C}_{\mathbb{J}}^d, \mathcal{M}_{\mathbb{J}})$ is tight, and any weak limit point is the distribution of the solution X^x . The Markov property for the transition functions then follows from the uniqueness of solutions of (1). \square

Remark. It should be noted that global Lipschitz continuity is more than necessary to have a well-defined and Feller continuous family of Markovian transition functions associated with (1). One can obtain this same conclusion if the coefficients b and σ are continuous and obey (5) and (6), and weak existence and uniqueness in law of solutions to (1) holds.

2.2 Main Result

We begin by defining a stationary distribution.

Definition 2.2.1. A stationary distribution for (1) is a probability measure π on $(\mathbb{C}_{\mathbb{I}}^d, \mathcal{M}_{\mathbb{I}})$ such that $\pi(\Lambda) = (\pi P_t)(\Lambda) := \int_{\mathbb{C}_{\mathbb{I}}^d} P_t(x, \Lambda) \pi(dx)$ for all $t \geq 0$ and $\Lambda \in \mathcal{M}_{\mathbb{I}}$.

It is well-known that (non-delayed) Ornstein-Uhlenbeck processes have (unique) stationary distributions, and it is not hard to show that the reflected analogues also have stationary distributions. The following condition for delayed systems is motivated by these facts.

Assumption 2.2. *There exist positive constants $B_0, B_1, B_{1,1}, \dots, B_{1,d}, B_{2,1}, \dots, B_{2,d}, C_0, C_{2,1}, \dots, C_{2,d}, M$, constants $q_1 \in (0, 1], q_2 \in (0, 2]$, probability measures $\mu_1^1, \dots, \mu_1^d, \mu_2^1, \dots, \mu_2^d$ on $(\mathbb{I}, \mathcal{B}(\mathbb{I}))$, and a measurable function $\ell : \mathbb{C}_{\mathbb{I}}^d \rightarrow \mathbb{R}_+^d$, such that for each $x \in \mathbb{C}_{\mathbb{I}}^d$ and $i = 1, \dots, d$, $\ell^i(x) \in x^i(\mathbb{I}) := \{x^i(s), s \in \mathbb{I}\}$ for each i , and*

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

$$b^i(x) \leq B_0 - B_1 x^i(0) - B_{1,i} \ell^i(x) + B_{2,i} \int_{-\tau}^0 |x(r)|^{q_1} \mu_1^i(dr), \quad (10)$$

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

$$|\sigma^i(x)|^2 \leq C_0 + C_{2,i} \int_{-\tau}^0 |x(r)|^{q_2} \mu_2^i(dr), \quad (11)$$

(iii) *for $\underline{B}_1 := \min_i B_{1,i}$ and $\tilde{B}_2 := \left(\sum_{i=1}^d B_{2,i}^2 \right)^{\frac{1}{2}}$, we have*

$$B_1 + \underline{B}_1 > \left(\tau \left(\sum_{i=1}^d (B_{1,i} B_{2,i})^2 \right)^{\frac{1}{2}} + \tilde{B}_2 \right) \delta_{q_1,1} + \left(\frac{1}{2} \sum_{i=1}^d C_{2,i} + 4\sqrt{\tau} \left(\sum_{i=1}^d C_{2,i} B_{1,i}^2 \right)^{\frac{1}{2}} \right) \delta_{q_2,2}.$$

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 Lipschitz condition (4), which implies the linear growth bounds (5) and (6). These 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 growth control on b and $|\sigma|$ is weaker than the at-most-integral-linear growth imposed by parts (i) and (ii) of the above assumption.

It is well-known that reflected Brownian motion on the half-line with strictly negative drift has a (unique) stationary distribution. The following assumption (which is distinct from Assumption 2.2) is sufficient for a stationary distribution for (1) to exist and the form of this condition is motivated by the aforementioned fact.

Assumption 2.3. *There exist positive constants K_u, M , strictly positive constants K_d, K_σ , and a measurable function $\ell : \mathbb{C}_{\mathbb{I}}^d \rightarrow \mathbb{R}_+^d$, such that for each $x \in \mathbb{C}_{\mathbb{I}}^d$ and $i = 1, \dots, d$, $\ell^i(x) \in x^i(\mathbb{I})$, 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) $|\sigma^i(x)|^2 \leq K_\sigma$.

Remark. Assumption 2.3 requires b^i and $|\sigma^i|$ to be bounded above on the set $\{x \in \mathbb{C}_{\mathbb{I}}^d : x^i(0) \geq M\}$, but this does not necessarily imply that they are bounded above on $\mathbb{C}_{\mathbb{I}}^d$. Also, note that unlike (iii) of Assumption 2.2, Assumption 2.3 has no restrictions on the size of the constants M , K_u , K_d , K_σ (beyond strict positivity of K_d and K_σ).

The following assumption is using in proving uniqueness of a stationary distribution.

Assumption 2.4. *The diffusion matrix $\sigma\sigma'$ is uniformly elliptic, i.e., there is a constant $a > 0$ such that $v'\sigma(x)(\sigma(x))'v \geq a|v|^2$ for all $x \in \mathbb{C}_{\mathbb{I}}^d$ and $v \in \mathbb{R}^d$.*

We shall use the following consequence of Assumption 2.4 in our proofs.

Proposition 2.2.1. *Under the Lipschitz condition (4) and Assumption 2.4, the dispersion coefficient σ has a continuous bounded right inverse, i.e., there is a constant $C_\sigma > 0$ and a continuous function $\sigma^\dagger : \mathbb{C}_{\mathbb{I}}^d \rightarrow \mathbb{M}^{m \times d}$ such that for all $x \in \mathbb{C}_{\mathbb{I}}^d$, $\sigma(x)\sigma^\dagger(x) = I_d$, and $|\sigma^\dagger(x)| \leq C_\sigma$ for all $x \in \mathbb{C}_{\mathbb{I}}^d$.*

Proof. Under the assumptions of the proposition, since $\sigma\sigma'$ is continuous and uniformly strictly positive definite, it follows by standard arguments that $(\sigma\sigma')^{-1}$ is a well-defined, continuous and bounded function on $\mathbb{C}_{\mathbb{I}}^d$. Then $\sigma^\dagger := \sigma'(\sigma\sigma')^{-1}$ is continuous and is a right inverse for σ . The (uniform) boundedness of σ^\dagger follows from the fact that

$$|\sigma^\dagger v|^2 = v'(\sigma\sigma')^{-1}\sigma\sigma'(\sigma\sigma')^{-1}v = v'(\sigma\sigma')^{-1}v \leq C_\sigma|v|^2,$$

where C_σ is a bound on the norm of $(\sigma\sigma')^{-1}$. □

Our main result is the following theorem.

Theorem 2.2.1. *Under Assumption 2.1, if Assumption 2.4 and either Assumption 2.2 or 2.3 hold, then there exists a unique stationary distribution for the SDDER (1).*

Proof. The result follows from Theorems 3.6.1 and 4.3.1 below. □

2.3 Examples

Example 2.3.1. *Differential delay equations with linear or affine coefficients are used often in engineering. We consider the following example of an SDDER with affine coefficients. For $x \in \mathbb{C}_{\mathbb{I}}$, let*

$$b(x) := b_0 - b_1x(0) - \sum_{i=2}^n b_i x(-r_i) + \sum_{i=n+1}^{n'} b_i x(-r_i), \quad (12)$$

and

$$\sigma(x) := a_0 + \sum_{i=1}^{n''} a_i x(-s_i), \quad (13)$$

where $0 \leq r_i \leq \tau$ and $0 \leq s_i \leq \tau$ for each i , $n' \geq n \geq 1$, $n'' \geq 0$, $b_0 \in \mathbb{R}$ and $b_1, \dots, b_{n'}, a_0, \dots, a_{n''} \geq 0$. 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 associated with (b, σ) has a unique stationary distribution. This follows from Theorem 2.2.1: the coefficients are clearly uniformly Lipschitz continuous, Assumption 2.2 holds with $M = 0$, $B_0 = (b_0)^+$, $B_1 = b_1$, $B_{1,1} = \sum_{i=2}^n b_i$, $B_{2,1} = \sum_{i=n+1}^{n'} b_i$, $q_1 = 1$,

$$q_2 = 2, \ell(x) = B_{1,1}^{-1} \sum_{i=2}^n b_i x(-r_i) \text{ if } B_{1,1} > 0,$$

$$\mu_1 = \frac{\sum_{i=n+1}^{n'} b_i \delta_{\{-r_i\}}}{B_{2,1}} \quad \text{if } B_{2,1} > 0, \quad \mu_2 = \frac{\sum_{i=1}^{n''} a_i \delta_{\{-s_i\}}}{\sum_{i=1}^{n''} a_i} \quad \text{if } C_{2,1} > 0, \quad (14)$$

where for $\gamma > 1$ sufficiently small, by (111) and the Cauchy-Schwarz inequality, we may take

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

and Assumption 2.4 holds because σ is uniformly positive definite when $a_0 > 0$.

Example 2.3.2. Fix $\alpha, \gamma, \varepsilon, C > 0$. For $x \in \mathbb{C}_{\mathbb{I}}$, define

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

The SDDER associated with this pair (b, σ) is a noisy version of a simple model used in the study of biochemical reaction systems [21]. In this model, a lengthy transcription/translation procedure leads to delayed negative feedback in the deterministic dynamics.

It is straightforward to verify that b, σ satisfy the uniform Lipschitz Assumption 2.1. 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 2.3 is satisfied with $\ell(x) = x(-\tau)$, $K_d = \frac{\gamma}{2}$, $K_u = \alpha$, $K_\sigma = \varepsilon^2(\alpha + \gamma)$ and $M = C\sqrt{\frac{2\alpha}{\gamma}}$. Also, σ is uniformly positive definite and so Assumption 2.4 holds. Therefore by Theorem 2.2.1, the SDDER associated with this (b, σ) has a unique stationary distribution.

Example 2.3.3. Deterministic delay differential equations have been used in modeling the dynamics of data transmission rates and prices in Internet congestion control [37]. In this context, the finiteness of transmission times results in delayed congestion signals and leads to differential dynamics with delayed negative feedback. There is a considerable body of work on obtaining sufficient conditions for stability of equilibrium points (see e.g., [8, 26, 27, 28, 29, 38, 39, 41]) for such models. It is natural to ask about the properties of noisy versions of these deterministic models and in particular to inquire about the existence and uniqueness of stationary distributions for such models. Here, as an illustration, we consider a noisy version of a model studied by Paganini and Wang [26], Peet and Lall [29], Papachristodolou

[27] and Papachristodolou, Doyle and Low [28]. This model has d links and d' sources. In the deterministic model the dynamics are given by

$$dX(t) = \hat{b}(X_t)dt, \quad (15)$$

where the i^{th} component of $X(t)$ represents the price at time t that link i charges for the transmission of a packet through it. The discontinuous drift \hat{b} is given for each $i = 1, \dots, d$ and $x \in \mathbb{C}_{\mathbb{I}}^d$ by

$$\hat{b}^i(x) = \begin{cases} -1 + \sum_{j=1}^{d'} A_{ij} \exp\left(-B_j \sum_{k=1}^d A_{kj} C_{kj} x^k (-r_{ijk})\right) & \text{if } x^i(0) > 0, \\ \left(-1 + \sum_{j=1}^{d'} A_{ij} \exp\left(-B_j \sum_{k=1}^d A_{kj} C_{kj} x^k (-r_{ijk})\right)\right)^+ & \text{if } x^i(0) = 0, \end{cases} \quad (16)$$

for some $B_1, \dots, B_d > 0$, $A_{ij} \geq 0$, $C_{kj} > 0$ and $r_{ijk} > 0$ for all $i, k \in \{1, \dots, d\}$ and $j \in \{1, \dots, d'\}$. The matrix A , which is related to routing in the network model, is assumed to have full row rank and to be such that for each i , $A_{ij} > 0$ for some j , indicating that each source must use at least one link. The solutions of (15) remain in the positive orthant by the form of \hat{b} (for the meaning of a solution with such a discontinuous right hand side, see, e.g., [10]). Various authors (see e.g., [26, 27, 28, 29]) have given sufficient conditions, principally in terms of smallness of the components of the gain parameter B , for there to be a unique globally asymptotically stable equilibrium point for (16). We now consider a noisy version of (16) and ask when it has a unique stationary distribution.

By uniqueness of solutions, the solutions of the SDDER associated with $\sigma \equiv 0$ coincide with the solutions of (15) when the drift b in (1) is defined by

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

Allowing σ to be non-zero yields a noisy version of (16). For this noisy version, we assume that $m \geq d$ and that $\sigma : \mathbb{C}_{\mathbb{I}}^d \rightarrow \mathbb{M}^{d \times m}$ is uniformly Lipschitz continuous and satisfies

$$a_1 |v|^2 \leq v' \sigma(x) \sigma(x)' v \leq a_2 |v|^2 \quad \text{for all } x \in \mathbb{C}_{\mathbb{I}}^d \text{ and } v \in \mathbb{R}^d, \quad (17)$$

for some $0 < a_1 < a_2 < \infty$. It is easily verified that b is uniformly Lipschitz continuous and for each $i = 1, \dots, d$,

$$b^i(x) \leq -1 + \sum_{j=1}^{d'} A_{ij} \exp(-B_j A_{ij} C_{ij} x^i (-r_{iji})) \leq -\frac{1}{2} \quad (18)$$

whenever $\sum_{j=1}^{d'} A_{ij} \exp(-B_j A_{ij} C_{ij} x^i (-r_{iji})) \leq \frac{1}{2}$. The latter holds if

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

(Recall that $\max_{j=1}^{d'} A_{ij} > 0$ by assumption.) It follows that b, σ satisfy Assumptions 2.1, 2.3 and 2.4 with $\tau := \max_{i,j,k} r_{ijk}$, $\ell^i(x) = \min_{j=1}^{d'} x^i(-r_{iji})$, $K_u = \max_i \sum_{j=1}^{d'} A_{ij}$, $K_d = \frac{1}{2}$, $K_\sigma = a_2$, $C_\sigma = \frac{1}{\sqrt{a_1}}$ and

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

Then, by Theorem 2.2.1, the SDDER with these coefficients (b, σ) has a unique stationary distribution. Thus, this noisy version of (16) has a unique stationary distribution without imposing additional restrictions on the parameters. This is in contrast to the known conditions for stability of the equilibrium solution in the deterministic equation (16).

Our results can be similarly applied to a slightly modified and noisy version of the Internet rate control model studied by Vinnecombe [38, 39] and Srikant et al. [8, 41] to yield existence and uniqueness of a stationary distribution for such a model without the strong conditions on the parameters used to obtain stability of the deterministic model. In particular, consider the deterministic delay model in equations (2)–(5) of [41] with $n_i, m_i \geq 0$ and replace $x_i^{n_i}(t)$ by $x_i^{n_i}(t) + c$ for some $c > 0$ (to prevent blowup of the drift when x_i reaches zero), replace $x_i^{m_i}(t)$ by $g(x_i^{m_i}(t))$ where $g(s) = s$ for $s \leq K$ and $g(s) = K$ for $s \geq K$ where K is a sufficiently large positive constant, and truncate the feedback functions f_l with an upper bound once the argument of these functions gets to a high level. Then with a dispersion coefficient of the same form as in (17) above one can prove that the associated SDDER has a unique stationary distribution by verifying that Assumptions 2.1, 2.2 and 2.4 hold.

2.4 Moment Bounds over Compact Time Intervals

Under Assumption 2.1, any solution X of the SDDER (1) satisfies the following supremum bound.

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

$$E \left[\|X\|_{[-\tau, T]}^p \right] \leq F_p(E[\|X_0\|^p], T) \text{ for each } T > 0. \quad (19)$$

In fact,

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

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 (5) and (6).

Sketch of proof. Inequality (110) and Proposition 2.1.1 can be used to obtain for any $T > 0$,

$$\begin{aligned} \|X\|_{[-\tau, T]}^p &\leq 2^{p-1} (\|X_0\|^p + (d \text{Osc}(\mathcal{I}(X), [0, T]))^p) \\ &\leq 2^{p-1} \|X_0\|^p + 2^{2p-2} d^p \left(\left(\int_0^T |b(X_t)| dt \right)^p + 2^p \sup_{s \in [0, T]} \left| \int_0^s \sigma(X_t) dW(t) \right|^p \right). \end{aligned}$$

The remainder of the proof follows from a standard argument (cf., Theorem 2.3 in Chapter 3 of [19]) using the linear growth conditions (5) and (6), the Burkholder-Davis-Gundy inequalities and a standard stopping argument allowing us to use Gronwall's inequality. \square

3 Existence of a Stationary Distribution

In this section, we prove that either Assumption 2.2 or 2.3 (in addition to Assumption 2.1) is sufficient to imply the existence of a stationary distribution for the SDDER (1). Throughout this section, we assume that X is a solution of the SDDER (1) with a possibly random initial condition X_0 . When the initial condition for X is deterministic, we will sometimes use the notation X^{x_o} for the unique solution with the initial condition x_o . We begin in Section 3.1 by describing a general sufficient condition for existence of a stationary distribution in terms of uniform (in t) moment bounds for $\|X_t\|$. We then use stochastic Lyapunov/Razumikhin-type arguments to verify that such bounds hold for second moments under either Assumption 2.2 or Assumption 2.3. Lyapunov-type functions are applied to an auxiliary process which we call the *overshoot process* and which we introduce in Section 3.2. In Section 3.3 we develop some preliminary results on the “positive oscillation” of a path. Sections 3.4 and 3.5 contain the key technical arguments for establishing the moment bounds under Assumption 2.2 and Assumption 2.3, respectively. Loosely speaking, each of these assumptions implies that, for each i , the i^{th} component of b has a term providing a push in the negative direction (towards zero) on the set $\{x \in \mathbb{C}_{\mathbb{I}}^d : |x^i(0)| \geq M\}$ for some $M > 0$. The two assumptions are distinguished by differences in the size of this “restoring force” and on the additional terms composing b and the assumptions on σ . Assumption 2.2 allows the additional terms in b to grow (in a sufficiently controlled manner) but requires the negative push in $b^i(x)$ to be at least proportional to a value lying in the range of x^i . For Assumption 2.3, $|\sigma|$ and the components of b are bounded above and the negative push is strictly negative and bounded away from zero. In Section 3.6 we combine the results of the preceding subsections to obtain the desired existence result.

Remark. Scrutiny of our proofs reveals that the results of this section still hold if Assumption 2.1 is replaced by the weaker assumptions that weak existence and uniqueness in law holds for (1), and that the coefficients b and σ are continuous and satisfy the linear growth conditions (5) and (6). As noted in the Remark following Proposition 2.1.2, under the latter conditions, the solutions of (1) define a Feller continuous Markov process. As explained in Section 2, we have assumed the stronger Assumption 2.1 throughout this paper because this assumption will be used critically in our uniqueness proof.

3.1 Sufficient Conditions for Existence of a Stationary Distribution

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-Bogolyubov measures [2, 7, 15, 30]. In light of that, given $x_o \in \mathbb{C}_{\mathbb{I}}^d$ and $T > 0$, we define the probability measure $Q_T^{x_o}$ on $(\mathbb{C}_{\mathbb{I}}^d, \mathcal{M}_{\mathbb{I}})$ by

$$Q_T^{x_o}(\Lambda) := \frac{1}{T} \int_0^T P_u(x_o, \Lambda) du \quad \text{for all } \Lambda \in \mathcal{M}_{\mathbb{I}}. \quad (20)$$

Remark. The function $u \rightarrow P_u(x_o, \Lambda)$ is measurable as a consequence of the stochastic continuity of the family $\{P_t(\cdot, \cdot), t \geq 0\}$, which follows from the continuity of the paths of X^{x_o} .

The following theorem gives sufficient conditions for the existence of a stationary distribution for the SDDER (1). Although we only use this result with $p = 2$ in this work, we give the result for general $p > 0$ as the proof is similar for all p .

Theorem 3.1.1. Fix $x_o \in \mathbb{C}_{\mathbb{I}}^d$ and assume that $\sup_{t \geq 0} E[\|X_t^{x_o}\|^p] < \infty$ for some $p > 0$. Then for any sequence $\{T_n\}_{n=1}^\infty$ in $(0, \infty)$ that increases to ∞ , the sequence $\{Q_{T_n}^{x_o}\}_{n=1}^\infty$ of probability measures is tight and any weak limit point is a stationary distribution for the SDDER (1).

Proof. By Markov's inequality, for any $T > 0$ and $a > 0$,

$$Q_T^{x_o} \left(x \in \mathbb{C}_{\mathbb{I}}^d : |x(0)| > a \right) = \frac{1}{T} \int_0^T P(|X^{x_o}(s)| > a) ds \leq \frac{1}{a^p} \sup_{t \geq 0} E[|X^{x_o}(t)|^p]. \quad (21)$$

The last term tends to zero as $a \rightarrow \infty$, independently of T .

Fix $\varepsilon, \lambda > 0$ and $u \geq \tau$, and recall the notation for the modulus of continuity from (3). The linear growth condition (5) and Proposition 2.1.1 imply that for any $\delta > 0$:

$$\begin{aligned} P(w_{\mathbb{I}}(X_u^{x_o}, \delta) \geq \lambda) &\leq P\left(\delta(C_1 + C_2\|X^{x_o}\|_{[u-2\tau, u]}) \geq \frac{\lambda}{2}\right) \\ &+ P\left(\sup_{\substack{u-\tau \leq s < t \leq u \\ |s-t| < \delta}} \max_{i=1, \dots, d} \left| \int_s^t \sigma^i(X_r^{x_o}) dW(r) \right| \geq \frac{\lambda}{2}\right). \end{aligned} \quad (22)$$

By Markov's inequality,

$$\sup_{t \geq \tau} P(\|X^{x_o}\|_{[t-2\tau, t]} > a) \leq \frac{2}{a^p} \sup_{t \geq 0} E[\|X_t^{x_o}\|^p],$$

which approaches zero as $a \rightarrow \infty$. This implies that there is $\delta_{\varepsilon, \lambda}^{(1)} \in (0, \frac{\lambda}{4C_1})$ such that

$$\sup_{u \geq \tau} P\left(\delta(C_1 + C_2\|X^{x_o}\|_{[u-2\tau, u]}) \geq \frac{\lambda}{2}\right) < \frac{\varepsilon}{4} \quad \text{for all } \delta \in (0, \delta_{\varepsilon, \lambda}^{(1)}].$$

Since σ grows at most linearly, a standard time-change argument (see, e.g., Theorem 3.4.6 and Problem 3.4.7 of [16], or V.1.7 of [31]) implies that there is $\delta_{\varepsilon, \lambda}^{(2)} > 0$ such that whenever $\delta \in (0, \delta_{\varepsilon, \lambda}^{(2)}]$ we have

$$\sup_{u \geq \tau} P\left(\sup_{\substack{u-\tau \leq s < t \leq u \\ |s-t| < \delta}} \max_{i=1, \dots, d} \left| \int_s^t \sigma^i(X_r^{x_o}) dW(r) \right| \geq \frac{\lambda}{2}\right) < \frac{\varepsilon}{4}. \quad (23)$$

It follows that

$$P(w_{\mathbb{I}}(X_u^{x_o}, \delta) \geq \lambda) < \frac{\varepsilon}{2} \quad \text{whenever } 0 < \delta < \delta_{\varepsilon, \lambda} := \delta_{\varepsilon, \lambda}^{(1)} \wedge \delta_{\varepsilon, \lambda}^{(2)} \text{ and } u \geq \tau.$$

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

$$\begin{aligned} Q_T^{x_o} \left(x \in \mathbb{C}_{\mathbb{I}}^d : w_{\mathbb{I}}(x, \delta) \geq \lambda \right) &= \frac{1}{T} \int_0^\tau P(w_{\mathbb{I}}(X_u^{x_o}, \delta) \geq \lambda) du + \frac{1}{T} \int_\tau^T P(w_{\mathbb{I}}(X_u^{x_o}, \delta) \geq \lambda) du \\ &< \varepsilon. \end{aligned} \quad (24)$$

Tightness of $\{Q_{T_n}^{x_o}\}_{n=1}^\infty$ follows from (21) and (24), by Theorem 7.3 of [3].

The fact that any weak limit point of $\{Q_{T_n}^{x_o}\}_{n=1}^\infty$ is a stationary distribution is a consequence of Theorem 1.2 of [3] and the Feller continuity of the associated family $\{P_t(\cdot, \cdot), t \geq 0\}$ of Markovian transition functions. \square

3.2 Overshoot Process

Let $\tilde{M} \geq 0$. For each $i = 1, \dots, d$, define the overshoot, Z^i , of X^i by

$$Z^i(t) := \left(X^i(t) - \tilde{M} \right)^+, \text{ for } t \geq -\tau. \quad (25)$$

Part (iii) of Definition 2.1.1 implies that $\int_0^t \mathbf{1}_{\{X^i(s) > \tilde{M}\}} dY^i(s) = 0$ for each $t \geq 0$. Thus, by (1) and Tanaka's formula for continuous semimartingales (see, e.g., Theorem 1.2 of Chapter VI in [31]), we have that P -a.s., for all $t \geq 0$,

$$dZ^i(t) = \mathbf{1}_{\{X^i(t) > \tilde{M}\}} b^i(X_t) dt + \mathbf{1}_{\{X^i(t) > \tilde{M}\}} \sigma^i(X_t) dW(t) + dL^i(t), \quad (26)$$

where L^i is a constant multiple of the local time of X^i at \tilde{M} , which can increase only when X^i is at \tilde{M} and hence only when Z^i is at zero.

The following consequence of Itô's formula will be useful in Sections 3.4 and 3.5. For each $t \geq 0$,

$$\begin{aligned} d(Z^i(t))^2 &= 2Z^i(t)b^i(X_t)dt + 2Z^i(t)\sigma^i(X_t)dW(t) + 2Z^i(t)dL^i(t) + \mathbf{1}_{\{X^i(t) > \tilde{M}\}} |\sigma^i(X_t)|^2 dt \\ &= 2Z^i(t)b^i(X_t)dt + 2Z^i(t)\sigma^i(X_t)dW(t) + \mathbf{1}_{\{X^i(t) > \tilde{M}\}} |\sigma^i(X_t)|^2 dt, \end{aligned} \quad (27)$$

where we have used the facts that $Z^i(t) = 0$ when $X^i(t) \leq \tilde{M}$ and L^i can increase only when Z^i is at zero. Thus we have

$$d(|Z(t)|^2) = 2(Z(t))'b(X_t)dt + 2(Z(t))'\sigma(X_t)dW(t) + \sum_{i=1}^d \mathbf{1}_{\{X^i(t) > \tilde{M}\}} |\sigma^i(X_t)|^2 dt. \quad (28)$$

3.3 Positive Oscillation

We now introduce the notion of the positive oscillation of a real-valued path over a given time interval. This refinement of the oscillation of a path (2) is well suited to our problem, and it still obeys an inequality analogous to (107).

Definition 3.3.1. Given a path $x \in C([a_1, a_2], \mathbb{R})$, define the *positive oscillation* of x over $[a_1, a_2]$ by

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

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

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

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

$$\text{Osc}^+(x^i, \mathbb{I}) \leq \|x^i\| \leq \|x\|, \quad \text{and} \quad (29)$$

$$\|x^i\| \leq x^i(-\tau) + \text{Osc}^+(x^i, \mathbb{I}). \quad (30)$$

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

Lemma 3.3.1. Fix $0 \leq t_1 < t_2 < \infty$. 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) $y(t_1) \geq 0$, $y(\cdot)$ is non-decreasing, and
- (iii) $y(\cdot)$ can only increase when z is at zero, i.e., $\int_{t_1}^{t_2} z(s) dy(s) = 0$.

Then,

$$\text{Osc}^+(z, [t_1, t_2]) \leq \text{Osc}^+(x, [t_1, t_2]). \quad (31)$$

Proof. By continuity of z and compactness of the triangle $\{(s, t) : t_1 \leq s \leq t \leq t_2\}$, there exist $s, t \in [t_1, t_2]$ such that $s \leq t$ and $\text{Osc}^+(z, [t_1, t_2]) = z(t) - z(s)$. If $s = t$, then the inequality (31) is clear. So we suppose that $s < t$. Then there are two cases to consider.

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

$$z(t) - z(s) = x(t) - x(s) \leq \text{Osc}^+(x, [t_1, t_2]). \quad (32)$$

Case 2: Suppose that $y(s) < y(t)$. Then there is $u \in [s, t]$ such that $z(u) = 0$, by (iii). Let $u' = \sup\{v \leq t : z(v) = 0\}$. 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

$$\begin{aligned} z(t) - z(s) &\leq z(t) = z(t) - z(u') = x(t) - x(u') + y(t) - y(u') = x(t) - x(u') \\ &\leq \text{Osc}^+(x, [t_1, t_2]), \end{aligned} \quad (33)$$

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

We will also use the following technical lemma.

Lemma 3.3.2. For each $i = 1, \dots, d$ and $\hat{M} \geq 0$, for any $0 \leq t_1 < t_2 < \infty$, P -a.s.,

$$\begin{aligned} \text{Osc}^+(X^i, [t_1, t_2]) &\leq \hat{M} + \int_{t_1}^{t_2} 1_{\{X^i(u) > \hat{M}\}} (b^i(X_u))^+ du \\ &\quad + \sup_{t_1 \leq r < s \leq t_2} \int_r^s 1_{\{X^i(u) > \hat{M}\}} \sigma^i(X_u) dW(u). \end{aligned} \quad (34)$$

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

$$\text{Osc}^+(X^i, [t_1, t_2]) \leq \hat{M} + \text{Osc}^+(Z^i, [t_1, t_2]). \quad (35)$$

The inequality (35) can be readily verified by considering $s \leq t$ 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) $X^i(t) \geq \hat{M}$ and $X^i(s) \geq \hat{M}$, and (c) $X^i(t) \geq \hat{M}$ and $X^i(s) < \hat{M}$. Thus, it suffices to estimate $\text{Osc}^+(Z^i, [t_1, t_2])$.

Since P -a.s., (26) holds and L^i can increase only when Z^i is zero, we may apply Lemma 3.3.1 to obtain that P -a.s.:

$$\text{Osc}^+(Z^i, [t_1, t_2]) \leq \text{Osc}^+(\mathcal{I}^i, [t_1, t_2]), \quad (36)$$

where, for each $t \geq 0$,

$$\mathcal{I}^i(t) := Z^i(0) + \int_0^t 1_{\{X^i(s) > \hat{M}\}} b^i(X_s) ds + \int_0^t 1_{\{X^i(s) > \hat{M}\}} \sigma^i(X_s) dW(s), \quad (37)$$

and

$$\text{Osc}^+(\mathcal{I}^i, [t_1, t_2]) \leq \int_{t_1}^{t_2} 1_{\{X^i(u) > \hat{M}\}} (b^i(X_u))^+ du + \sup_{t_1 \leq r < s \leq t_2} \left(\int_r^s 1_{\{X^i(u) > \hat{M}\}} \sigma^i(X_u) dW(u) \right). \quad (38)$$

Thus, (34) holds. \square

3.4 Bounded Second Moments when b and σ Satisfy an Integral Growth Condition

Throughout this subsection, we assume that the coefficients b, σ satisfy Assumption 2.2 (in addition to Assumption 2.1) and we set $\tilde{M} = M + 1$ in the definition of the overshoot process Z in (25). The simple inequalities $X^i(\cdot) \leq Z^i(\cdot) + \tilde{M}$, for each i , reduce the problem of bounding the second moment of $\|X_t\|$ to that of bounding the second moment of $\|Z_t\|$.

3.4.1 Uniform Bound on $E[|X(t)|^2]$

Theorem 3.4.1. *Suppose that $E[\|X_0\|^2] < \infty$. Then, $\sup_{t \geq -\tau} E[|X(t)|^2] < \infty$.*

Our proof of this theorem uses stochastic Lyapunov/Razumikhin-type arguments similar to those found in a theorem of Mao (Theorem 2.1 of [20]). We first prove two technical lemmas and then the proof of the theorem is given. To simplify notation, in the following we let

$$\mathcal{Z}(t) := E[|Z(t)|^2] \quad \text{and} \quad \mathcal{I}^i(t) := 1_{\{X^i(t) > M\}} \quad (39)$$

for $i = 1, \dots, d$ and $t \geq -\tau$.

Lemma 3.4.1. *Suppose that $E[\|X_0\|^2] < \infty$. There exists a constant $M_1 > 0$ such that whenever $t \geq \tau$ is such that*

$$\mathcal{Z}(t) \geq M_1 \quad \text{and} \quad \mathcal{Z}(r) \leq \mathcal{Z}(t) \quad \text{for all } r \in [t - 2\tau, t], \quad (40)$$

then

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

Remark. We will refer to the second inequality in (40) as the Razumikhin assumption.

Proof. Suppose that $t \geq \tau$ is such that (40) holds. For each $x \in \mathbb{C}_{\mathbb{I}}^d$, there is an $r_x \in \mathbb{I}^d$ such that for each $i = 1, \dots, d$,

$$-\ell^i(x) = -x^i(r_x^i) \leq -x^i(0) + \text{Osc}^+(x^i, \mathbb{I}). \quad (41)$$

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 (10) and (11) hold with $x = X_u$. Then,

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

Here, Assumption 2.2(i) and the positivity 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 3.3.2 with $\hat{M} = M$ and the Cauchy-Schwarz inequality were used for the second inequality. Combining the above with parts (i) and (ii) of Assumption 2.2, on taking expectations and setting $\overline{B}_1 := \max_{i=1}^d B_{1,i}$ we obtain

$$\begin{aligned} & E \left[(Z(t))'b(X_t) + \frac{1}{2} \sum_{i=1}^d \mathcal{I}^i(t) |\sigma^i(X_t)|^2 \right] \\ & \leq (B_0 + M\overline{B}_1) E \left[\sum_{i=1}^d Z^i(t) \right] - (B_1 + \underline{B}_1) \mathcal{Z}(t) \\ & \quad + E \left[\sum_{i=1}^d B_{1,i} Z^i(t) \left(B_0 \tau + B_{2,i} \int_{t-\tau}^t \int_{-\tau}^0 |X_u(r)|^{q_1} \mu_1^i(dr) du \right) \right] \\ & \quad + E \left[\sum_{i=1}^d B_{1,i} Z^i(t) \sup_{t-\tau \leq r < s \leq t} \left| \int_r^s \mathcal{I}^i(u) \sigma^i(X_u) dW(u) \right| \right] \\ & \quad + E \left[|Z(t)| \left(\sum_{i=1}^d B_{2,i}^2 \left(\int_{-\tau}^0 |X_t(r)|^{q_1} \mu_1^i(dr) \right)^2 \right)^{\frac{1}{2}} \right] \\ & \quad + \frac{1}{2} dC_0 + \frac{1}{2} \sum_{i=1}^d C_{2,i} E \left[\int_{-\tau}^0 |X_t(r)|^{q_2} \mu_2^i(dr) \right]. \quad (42) \end{aligned}$$

We now separately develop estimates for the second and third to the last lines in (42). For each i ,

$$\sup_{t-\tau \leq r < s \leq t} \left| \int_r^s \mathcal{I}^i(u) \sigma^i(X_u) dW(u) \right| \leq 2 \sup_{t-\tau \leq s \leq t} \left| \int_{t-\tau}^s \mathcal{I}^i(u) \sigma^i(X_u) dW(u) \right|. \quad (43)$$

Part (ii) of Assumption 2.2, the assumption that $E[\|X_0\|^2] < \infty$ and Lemma 2.4.1 imply that for each i ,

$$\left\{ \int_{t-\tau}^s \mathcal{I}^i(u) \sigma^i(X_u) dW(u), \mathcal{F}_s, s \geq t - \tau \right\}$$

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 (11) imply that

$$E \left[\sup_{t-\tau \leq s \leq t} \left| \int_{t-\tau}^s \mathcal{I}^i(u) \sigma^i(X_u) dW(u) \right|^2 \right] \leq 4C_0\tau + 4C_{2,i} \int_{t-\tau}^t \int_{-\tau}^0 E[|X_u(r)|^{q_2}] \mu_2^i(dr) du. \quad (44)$$

Then, using the Cauchy-Schwarz inequality and inequality (109), we have

$$\begin{aligned} & E \left[\sum_{i=1}^d B_{1,i} Z^i(t) \sup_{t-\tau \leq r < s \leq t} \left| \int_r^s \mathcal{I}^i(u) \sigma^i(X_u) dW(u) \right| \right] \\ & \leq 4(\mathcal{Z}(t))^{\frac{1}{2}} \left(\sqrt{C_0\tau} \left(\sum_{i=1}^d B_{1,i}^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^d C_{2,i} B_{1,i}^2 \int_{t-\tau}^t \int_{-\tau}^0 E[|X_u(r)|^{q_2}] \mu_2^i(dr) du \right)^{\frac{1}{2}} \right). \quad (45) \end{aligned}$$

For the second last line in (42), using the Cauchy-Schwarz inequality we obtain that

$$\begin{aligned} & E \left[|Z(t)| \left(\sum_{i=1}^d B_{2,i}^2 \left(\int_{-\tau}^0 |X_t(r)|^{q_1} \mu_1^i(dr) \right)^2 \right)^{\frac{1}{2}} \right] \\ & \leq (\mathcal{Z}(t))^{\frac{1}{2}} \left(\sum_{i=1}^d B_{2,i}^2 E \left[\left(\int_{-\tau}^0 |X_t(r)|^{q_1} \mu_1^i(dr) \right)^2 \right] \right)^{\frac{1}{2}}. \quad (46) \end{aligned}$$

Inequality (111) implies that for any $\gamma > 1$ and $s \geq -\tau$ there is a constant $K_\gamma \geq 0$, which depends on d and M in addition to γ , such that

$$|X(s)|^2 \leq K_\gamma + \gamma |Z(s)|^2. \quad (47)$$

Then Jensen's inequality, the fact that each μ_1^i is a probability measure, Fubini's theorem, inequality (109) and the Razumikhin assumption in (40) can be used to obtain, for each $\gamma > 1$ and each i ,

$$\begin{aligned} E \left[\left(\int_{-\tau}^0 |X_t(r)|^{q_1} \mu_1^i(dr) \right)^2 \right] & \leq E \left[\left(\int_{-\tau}^0 \left(K_{\frac{1}{\gamma}} + \gamma^{\frac{1}{\gamma}} |Z_t(r)|^2 \right)^{\frac{q_1}{2}} \mu_1^i(dr) \right)^2 \right] \\ & \leq \int_{-\tau}^0 E \left[\left(K_{\frac{1}{\gamma}} + \gamma^{\frac{1}{\gamma}} |Z_t(r)|^2 \right)^{q_1} \right] \mu_1^i(dr) \\ & \leq C_\gamma + \gamma \int_{-\tau}^0 (\mathcal{Z}(t+r))^{q_1} \mu_1^i(dr) \\ & \leq C_\gamma + \gamma (\mathcal{Z}(t))^{q_1}, \quad (48) \end{aligned}$$

where $C_\gamma := \left(K_{\gamma^{\frac{1}{q_1}}}\right)^{q_1}$. Thus, we obtain

$$E \left[|Z(t)| \left(\sum_{i=1}^d B_{2,i}^2 \left(\int_{-\tau}^0 |X_t(r)|^{q_1} \mu_1^i(dr) \right)^2 \right)^{\frac{1}{2}} \right] \leq \tilde{B}_2 (\mathcal{Z}(t))^{\frac{1}{2}} (C_\gamma + \gamma (\mathcal{Z}(t))^{q_1})^{\frac{1}{2}}. \quad (49)$$

Continuing on from (42), by the Cauchy-Schwarz inequality, (29), (45), (49) and (109) we have that

$$\begin{aligned} & E \left[(Z(t))' b(X_t) + \frac{1}{2} \sum_{i=1}^d \mathcal{I}^i(t) |\sigma^i(X_t)|^2 \right] \\ & \leq (B_0 + (M + B_0 \tau) \overline{B}_1) \sqrt{d} (\mathcal{Z}(t))^{\frac{1}{2}} - (B_1 + \underline{B}_1) \mathcal{Z}(t) \\ & \quad + E \left[\sum_{i=1}^d B_{1,i} B_{2,i} Z^i(t) \int_{t-\tau}^t \int_{-\tau}^0 |X_u(r)|^{q_1} \mu_1^i(dr) du \right] \\ & \quad + 4 \sqrt{C_0 \tau} \left(\sum_{i=1}^d B_{1,i}^2 \right)^{\frac{1}{2}} (\mathcal{Z}(t))^{\frac{1}{2}} \\ & \quad + 4 \left(\sum_{i=1}^d C_{2,i} B_{1,i}^2 \int_{t-\tau}^t \int_{-\tau}^0 E[|X_u(r)|^{q_2}] \mu_2^i(dr) du \right)^{\frac{1}{2}} (\mathcal{Z}(t))^{\frac{1}{2}} \\ & \quad + \tilde{B}_2 \left(C_\gamma (\mathcal{Z}(t))^{\frac{1}{2}} + \gamma (\mathcal{Z}(t))^{\frac{q_1+1}{2}} \right) \\ & \quad + \frac{1}{2} d C_0 + \frac{1}{2} \sum_{i=1}^d C_{2,i} \int_{-\tau}^0 E[|X_t(r)|^{q_2}] \mu_2^i(dr). \end{aligned} \quad (50)$$

We now examine the fifth last, third last and last lines in (50) more closely. By Hölder's inequality, Fubini's theorem, (47), the Razumikhin assumption and Jensen's inequality, we have for each $\gamma > 1$,

$$\begin{aligned} E \left[\left(\int_{t-\tau}^t \int_{-\tau}^0 |X_u(r)|^{q_1} \mu_1^i(dr) du \right)^2 \right] & \leq E \left[\tau \int_{t-\tau}^t \int_{-\tau}^0 |X_u(r)|^{2q_1} \mu_1^i(dr) du \right] \\ & \leq \tau \int_{t-\tau}^t \int_{-\tau}^0 E[(K_\gamma + \gamma |Z_u(r)|^2)^{q_1}] \mu_1^i(dr) du \\ & \leq \tau \int_{t-\tau}^t \int_{-\tau}^0 (E[K_\gamma + \gamma |Z_u(r)|^2])^{q_1} \mu_1^i(dr) du \\ & \leq \tau^2 (K_\gamma + \gamma \mathcal{Z}(t))^{q_1}. \end{aligned} \quad (51)$$

Therefore, by the Cauchy-Schwarz inequality and (109), we have

$$\begin{aligned} & E \left[\sum_{i=1}^d B_{1,i} B_{2,i} Z^i(t) \int_{t-\tau}^t \int_{-\tau}^0 |X_u(r)|^{q_1} \mu_1^i(dr) du \right] \\ & \leq \tau \left(\sum_{i=1}^d (B_{1,i} B_{2,i})^2 \right)^{\frac{1}{2}} \left(K_\gamma^{\frac{q_1}{2}} + \gamma^{\frac{q_1}{2}} (\mathcal{Z}(t))^{\frac{q_1}{2}} \right) (\mathcal{Z}(t))^{\frac{1}{2}}. \end{aligned} \quad (52)$$

Similarly, Hölder's inequality, Jensen's inequality, (47), the Razumikhin assumption and (109) also imply that for each $\gamma > 1$:

$$\left(\sum_{i=1}^d C_{2,i} B_{1,i}^2 \int_{t-\tau}^t \int_{-\tau}^0 E[|X_u(r)|^{q_2}] \mu_2^i(dr) du \right)^{\frac{1}{2}} \leq \sqrt{\tau} \left(\sum_{i=1}^d C_{2,i} B_{1,i}^2 \right)^{\frac{1}{2}} \left(K_{\gamma^{\frac{q_2}{4}}} + \gamma^{\frac{q_2}{4}} (\mathcal{Z}(t))^{\frac{q_2}{4}} \right), \quad (53)$$

and

$$\int_{-\tau}^0 E[|X_t(r)|^{q_2}] \mu_2^i(dr) \leq K_{\gamma^{\frac{q_2}{2}}} + \gamma^{\frac{q_2}{2}} (\mathcal{Z}(t))^{\frac{q_2}{2}}. \quad (54)$$

Continuing on from line (50), using inequalities (52), (53) and (54) we have

$$\begin{aligned} E \left[(Z(t))' b(X_t) + \frac{1}{2} \sum_{i=1}^d \mathcal{I}^i(t) |\sigma^i(X_t)|^2 \right] \\ \leq K_1(\gamma) + K_2(\gamma) (\mathcal{Z}(t))^{\frac{1}{2}} + K_3(\gamma) (\mathcal{Z}(t))^{\frac{1+q_1}{2}} + K_4(\gamma) (\mathcal{Z}(t))^{\frac{q_2}{2}} \\ + K_5(\gamma) (\mathcal{Z}(t))^{\frac{2+q_2}{4}} - (B_1 + \underline{B}_1) \mathcal{Z}(t), \end{aligned} \quad (55)$$

where K_1, K_2 are real-valued functions on $(1, \infty)$, and

$$\begin{aligned} K_3(\gamma) &= \tau \left(\sum_{i=1}^d (B_{1,i} B_{2,i})^2 \right)^{\frac{1}{2}} \gamma^{\frac{q_1}{2}} + \tilde{B}_2 \gamma, \\ K_4(\gamma) &= \frac{1}{2} \gamma^{\frac{q_2}{2}} \sum_{i=1}^d C_{2,i} \quad \text{and} \quad K_5(\gamma) = 4\sqrt{\tau} \left(\sum_{i=1}^d C_{2,i} B_{1,i}^2 \right)^{\frac{1}{2}} \gamma^{\frac{q_2}{4}}. \end{aligned}$$

By Assumption 2.2(iii), we can fix a $\gamma > 1$ sufficiently small such that $B_1 + \underline{B}_1 > K_3(\gamma) \delta_{q_1,1} + (K_4(\gamma) + K_5(\gamma)) \delta_{q_2,2}$. For such a γ ,

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

whenever $\mathcal{Z}(t)$ is large enough. Indeed, define the function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ 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 for r are at most one. By the choice of γ , the constant in front of the highest degree term is strictly negative and this implies that $\lim_{r \rightarrow \infty} f(r) = -\infty$. Thus, there exists a constant $M_1 > 0$ such that $r \geq M_1$ implies that $f(r) < 0$. \square

Lemma 3.4.2. *Suppose that $E[\|X_0\|^2] < \infty$. Let M_1 be defined as in the previous lemma and assume that $t \geq \tau$ is such that (40) holds. Then there exists an $h^* > 0$ such that*

$$\mathcal{Z}(t+s) < \mathcal{Z}(t), \quad \text{for each } s \in (0, h^*]. \quad (56)$$

Proof. Let $\eta_n = t \vee \inf\{s \geq -\tau : |X(s)| \geq n\}$ for each integer $n > 0$. We have from equality (28) that for each $n > 0$ and $h \geq 0$,

$$\begin{aligned} |Z((t+h) \wedge \eta_n)|^2 - |Z(t)|^2 &= \int_t^{(t+h) \wedge \eta_n} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s) > \tilde{M}\}} |\sigma^i(X_s)|^2 \right) ds \\ &\quad + \int_t^{(t+h) \wedge \eta_n} 2(Z(s))'\sigma(X_s)dW(s). \end{aligned} \quad (57)$$

By the definition of η_n , the stochastic integral with respect to W in the above has zero mean since it defines a square integrable martingale as a function of h . Since $E \left[\sup_{s \in [-\tau, t+h]} |Z(s)|^2 \right] < \infty$ by Lemma 2.4.1 with $p = 2$, and b and σ satisfy the linear growth bounds (5) and (6), we can take expectations in (57) and apply the dominated convergence theorem to conclude that

$$\mathcal{Z}(t+h) - \mathcal{Z}(t) = E \left[\int_t^{t+h} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s) > \tilde{M}\}} |\sigma^i(X_s)|^2 \right) ds \right]. \quad (58)$$

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

$$1_{(\tilde{M}, \infty)}(r) \leq f(r) \leq 1_{(M, \infty)}(r) \quad \text{for all } r \geq 0. \quad (59)$$

Then by (58), (59), dominated convergence and Lebesgue's differentiation theorem we have

$$\begin{aligned} \overline{\lim}_{h \rightarrow 0+} \frac{\mathcal{Z}(t+h) - \mathcal{Z}(t)}{h} &= \overline{\lim}_{h \rightarrow 0+} E \left[\frac{1}{h} \int_t^{t+h} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d 1_{\{X^i(s) > \tilde{M}\}} |\sigma^i(X_s)|^2 \right) ds \right] \\ &\leq \overline{\lim}_{h \rightarrow 0+} E \left[\frac{1}{h} \int_t^{t+h} \left(2(Z(s))'b(X_s) + \sum_{i=1}^d f(X^i(s)) |\sigma^i(X_s)|^2 \right) ds \right] \\ &\leq E \left[2(Z(t))'b(X_t) + \sum_{i=1}^d \mathcal{I}^i(t) |\sigma^i(X_t)|^2 \right]. \end{aligned} \quad (60)$$

Here we used the fact that the integrand in the second last line is a continuous function of s , and the fact that $f(X^i(t)) \leq 1_{(M, \infty)}(X^i(t)) = \mathcal{I}^i(t)$. According to Lemma 3.4.1, the last line above is strictly negative under the assumption (40).

If there is no $h^* > 0$ such that $\mathcal{Z}(t+s) < \mathcal{Z}(t)$ for each $s \in (0, h^*]$, then we can construct a sequence $\{h_n\}_{n=1}^\infty$ of positive numbers decreasing to zero such that $\mathcal{Z}(t+h_n) \geq \mathcal{Z}(t)$ for all n . Then $\overline{\lim}_{h \rightarrow 0+} \frac{\mathcal{Z}(t+h) - \mathcal{Z}(t)}{h} \geq 0$, which is a contradiction to (60). Therefore there is an $h^* > 0$ such that (56) holds. \square

Proof of Theorem 3.4.1. By Lemma 2.4.1, the continuity of Z and the dominated convergence theorem, $\mathcal{Z}(s)$ is continuous as a function of $s \geq 0$. Let $M_1 > 0$ be as in Lemma 3.4.1. Then $M_2 := \sup_{s \in [-\tau, \tau]} \mathcal{Z}(s) + M_1$ is finite. If there is a $t_1 > \tau$ such that $\mathcal{Z}(t_1) > M_2$, then since

$\sup_{s \in [-\tau, \tau]} \mathcal{Z}(s) < M_2$, the time $t := \inf\{s \geq -\tau : \mathcal{Z}(s) > M_2\}$ is in (τ, t_1) . We also have $\mathcal{Z}(t) = M_2$ by continuity, and thus $\mathcal{Z}(r) \leq \mathcal{Z}(t)$ for all $r \in [t-2\tau, t]$. Since $M_2 \geq M_1$, Lemma 3.4.2 implies that there is an $h^* > 0$ such that $\mathcal{Z}(s) < \mathcal{Z}(t) = M_2$ for all $s \in (t, t+h^*]$. However, this last conclusion contradicts the definition of t . Thus, there cannot be such a t_1 and consequently $\sup_{s \geq -\tau} \mathcal{Z}(s) \leq M_2$. This in turn implies that

$$\sup_{s \geq -\tau} E[|X(s)|^2] \leq \sup_{s \geq -\tau} 2 \left(\mathcal{Z}(s) + d\tilde{M}^2 \right) \leq 2M_2 + 2d\tilde{M}^2. \quad (61)$$

□

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

Theorem 3.4.2. *Suppose that $E[\|X_0\|^2] < \infty$. Then, $\sup_{t \geq 0} E[\|X_t\|^2] < \infty$.*

Proof. Recall that we are assuming that Assumption 2.2 holds and that the overshoot process Z is defined by (25) with $\tilde{M} = M + 1$. For each $t \geq \tau$, by (30), (36)-(38), (43), (110) and (10) with \tilde{M} in place of \hat{M} and M , we have

$$\begin{aligned} \|Z_t\|^2 &\leq \sum_{i=1}^d \left(Z^i(t-\tau) + \int_{t-\tau}^t 1_{\{X^i(s) > \tilde{M}\}} (b^i(X_s))^+ ds \right. \\ &\quad \left. + 2 \sup_{t-\tau \leq s \leq t} \left| \int_{t-\tau}^s 1_{\{X^i(u) > \tilde{M}\}} \sigma^i(X_u) dW(u) \right| \right)^2 \\ &\leq 3 \left(|Z(t-\tau)|^2 + \sum_{i=1}^d \left(B_0\tau + B_{2,i} \int_{t-\tau}^t \int_{-\tau}^0 |X(s+r)|^{q_1} \mu_1^i(dr) ds \right)^2 \right. \\ &\quad \left. + \sum_{i=1}^d 4 \sup_{t-\tau \leq s \leq t} \left| \int_{t-\tau}^s 1_{\{X^i(u) > \tilde{M}\}} \sigma^i(X_u) dW(u) \right|^2 \right). \quad (62) \end{aligned}$$

Using (110) and the Cauchy-Schwarz inequality, we have for each $i = 1, \dots, d$,

$$\begin{aligned} &\left(B_0\tau + B_{2,i} \int_{t-\tau}^t \int_{-\tau}^0 |X(s+r)|^{q_1} \mu_1^i(dr) ds \right)^2 \\ &\leq 2 \left((B_0\tau)^2 + B_{2,i}^2 \tau \int_{t-\tau}^t \int_{-\tau}^0 |X(s+r)|^{2q_1} \mu_1^i(dr) ds \right), \quad (63) \end{aligned}$$

and by a similar argument to that used for (44), by (11) we have for each i :

$$\begin{aligned} &E \left[\sup_{t-\tau \leq s \leq t} \left| \int_{t-\tau}^s 1_{\{X^i(u) > \tilde{M}\}} \sigma^i(X_u) dW(u) \right|^2 \right] \\ &\leq 4 \left(C_0\tau + C_{2,i} \int_{t-\tau}^t \int_{-\tau}^0 E[|X(u+r)|^{q_2}] \mu_2^i(dr) du \right). \quad (64) \end{aligned}$$

By Hölder's inequality, $E[|X(s)|^p] \leq E[|X(s)|^2]^{\frac{p}{2}}$ for all $s \geq -\tau$ and $0 < p \leq 2$. Then, by Theorem 3.4.1 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 $\sup_{s \geq -\tau} E[|X(s)|^p] \leq K$ for all $0 < p \leq 2$. On combining the above and taking expectations in (62), we obtain for all $t \geq \tau$:

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

where we have used the facts that $|Z(s)| \leq |X(s)|$ for all $s \geq -\tau$, $0 < q_1 \leq 1$ and $0 < q_2 \leq 2$. The last line above is independent of $t \geq \tau$ and so we have

$$\sup_{t \geq \tau} E[\|X_t\|^2] \leq 2 \left(\sup_{t \geq \tau} E[\|Z_t\|^2] + d\tilde{M}^2 \right) < \infty. \quad (65)$$

Combining this with the hypothesis of the theorem and the fact that

$$E[\|X_t\|^2] \leq 2E[\|X_0\|^2 + \|X_\tau\|^2] \quad \text{for each } t \in [0, \tau],$$

yields the desired result. \square

3.5 Bounded Second Moments when b and σ Satisfy a Boundedness Assumption

Throughout this subsection, we assume that the coefficients b and σ satisfy Assumption 2.3 (in addition to Assumption 2.1) and we set $\tilde{M} = M$ in the definition of the overshoot process Z . In Theorem 3.5.1 below, for each $i = 1, \dots, d$, we show that $E[\exp(\alpha X^i(t))]$ is bounded uniformly in $t \geq 0$, for some $\alpha > 0$, provided that suitable initial bounds hold. In turn, this will be used to bound $E[\|X_t\|^2]$ uniformly for all $t \geq 0$.

3.5.1 Uniform Bound on an Exponential Moment of $X^i(t)$

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

Theorem 3.5.1. *Suppose that $E[\exp(\kappa\|X_0\|)] < \infty$ for each $\kappa > 0$. Then there exists $\alpha > 0$ such that $\sup_{t \geq 0} E[\exp(\alpha X^i(t))] < \infty$ for each $i = 1, \dots, d$, and consequently, $\sup_{t \geq -\tau} E[|X(t)|^p] < \infty$ for all $p > 0$.*

Proof. Fix $i \in \{1, \dots, d\}$. Let $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a twice continuously differentiable non-decreasing function such that $f(r) = 0$ for $r \leq M$, $f(r) \leq \exp(\alpha r)$ for $M \leq r \leq M+1$, and $f(r) = \exp(\alpha r)$ for $r \geq M+1$. Then there exist positive constants C' and C'' , depending on α , such that

$$f'(r) \leq C' + \alpha f(r) \quad \text{and} \quad f''(r) \leq C'' + \alpha^2 f(r) \quad \text{for all } r \in \mathbb{R}_+.$$

Then the fact that Y^i can increase only when X^i is at zero, the form of f , and Lemma 3.5.1 below imply that for any $\gamma > 0$, the differential of $f(X^i(t))$ satisfies

$$\begin{aligned}
d(f(X^i(t))) &= f'(X^i(t))b^i(X_t)dt + f'(X^i(t))\sigma^i(X_t)dW(t) + \frac{1}{2}f''(X^i(t))|\sigma^i(X_t)|^2dt \\
&\leq (C' + \alpha f(X^i(t)))((K_u + K_d)e^{\gamma M} \exp(-\gamma X^i(t) + \gamma \text{Osc}^+(X_t^i, \mathbb{I})) - K_d) dt \\
&\quad + f'(X^i(t))\sigma^i(X_t)dW(t) + (C'' + \alpha^2 f(X^i(t)))K_\sigma dt \\
&\leq f'(X^i(t))\sigma^i(X_t)dW(t) + (-K_d\alpha + \alpha^2 K_\sigma) f(X^i(t)) dt + (C''K_\sigma - C'K_d)dt \\
&\quad + (K_u + K_d)e^{\gamma M} (C' \exp(-\gamma X^i(t)) + \alpha \exp((\alpha - \gamma)X^i(t))) \exp(\gamma \text{Osc}^+(X_t^i, \mathbb{I}))dt.
\end{aligned}$$

Set $\alpha = \frac{K_d}{2K_\sigma}$, $\beta = \frac{K_d^2}{2K_\sigma}$ and $\gamma = 2\alpha$. Then

$$d(f(X^i(t))) \leq f'(X^i(t))\sigma^i(X_t)dW(t) - \frac{\beta}{2}f(X^i(t)) dt + C(1 + \exp(\gamma \text{Osc}^+(X_t^i, \mathbb{I})))dt,$$

where $C > 0$ is an appropriately chosen constant (depending on α). Therefore,

$$\begin{aligned}
d\left(e^{\beta t} f(X^i(t))\right) &\leq \frac{\beta}{2}e^{\beta t} f(X^i(t)) dt + e^{\beta t} f'(X^i(t)) \sigma^i(X_t)dW(t) \\
&\quad + Ce^{\beta t}(1 + \exp(\gamma \text{Osc}^+(X_t^i, \mathbb{I})))dt.
\end{aligned} \tag{66}$$

Since Lemma 3.5.3 implies that for each $t \geq 0$,

$$E\left[e^{2\beta t} (f'(X^i(t)))^2 |\sigma^i(X_t)|^2\right] \leq 2K_\sigma e^{2\beta t} ((C')^2 + \alpha^2 E[\exp(2\alpha X^i(t))]) < \infty,$$

we have $E\left[\int_0^t e^{\beta s} f'(X^i(s)) \sigma^i(X_s)dW(s)\right] = 0$, which in turn implies that on taking expectations in (66) we have

$$\begin{aligned}
e^{\beta t} E[f(X^i(t))] &\leq \frac{\beta}{2} \int_0^t e^{\beta s} E[f(X^i(s))] ds + C \int_0^t e^{\beta s} (1 + E[\exp(\gamma \text{Osc}^+(X_s^i, \mathbb{I}))]) ds \\
&\leq \frac{\beta}{2} \int_0^t e^{\beta s} E[f(X^i(s))] ds + C \frac{e^{\beta t}}{\beta} (1 + K(\gamma)),
\end{aligned}$$

where $K(\cdot)$ is defined in Lemma 3.5.4 below. Gronwall's inequality now implies that

$$e^{\beta t} E[f(X^i(t))] \leq C \frac{e^{\beta t}}{\beta} (1 + K(\gamma)) + \frac{C}{2} (1 + K(\gamma)) \int_0^t e^{\frac{\beta}{2}(t-s)} e^{\beta s} ds,$$

and thus for all $t \geq 0$,

$$E[f(X^i(t))] \leq \frac{2C(1 + K(\gamma))}{\beta}.$$

The form of f then implies that

$$\sup_{t \geq 0} E[\exp(\alpha X^i(t))] \leq \exp(\alpha(M + 1)) + \frac{2C(1 + K(\gamma))}{\beta}. \tag{67}$$

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 (67), the hypothesis of the theorem and Hölder's inequality that for each $p > 0$ and $i = 1, \dots, d$,

$$\sup_{t \geq -\tau} E[(X^i(t))^p] < \infty.$$

□

3.5.2 Supporting Lemmas

We now prove the additional lemmas used in the proof of Theorem 3.5.1. We will again use the notation \mathcal{S} from (39).

Lemma 3.5.1. *For each $\gamma > 0$ and each $x \in \mathbb{C}_{\mathbb{I}}^d$ with $x^i(0) \geq M$, we have*

$$b^i(x) \leq (K_u + K_d)e^{\gamma M} \exp(-\gamma x^i(0) + \gamma \text{Osc}^+(x^i, \mathbb{I})) - K_d. \quad (68)$$

Proof. Let $x \in \mathbb{C}_{\mathbb{I}}^d$ with $x^i(0) \geq M$. Since $\ell^i(x) \in x^i(\mathbb{I})$, there is $r_x \in \mathbb{I}$ 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,

$$0 \leq \ell^i(x) - x^i(0) + \text{Osc}^+(x^i, \mathbb{I}). \quad (69)$$

From Assumption 2.3(i), it follows that for each $\gamma > 0$,

$$\begin{aligned} b^i(x) &\leq K_u 1_{[0, M]}(\ell^i(x)) - K_d 1_{[M, \infty)}(\ell^i(x)) \leq (K_u + K_d) 1_{[0, M]}(\ell^i(x)) - K_d \\ &\leq (K_u + K_d)e^{\gamma M} \exp(-\gamma x^i(0) + \gamma \text{Osc}^+(x^i, \mathbb{I})) - K_d. \end{aligned} \quad (70)$$

□

Lemma 3.5.2. *For each $t \geq 0$ and $i = 1, \dots, d$, define the process*

$$\left\{ \xi^{t,i}(s) := \exp \left(\int_{(t-\tau)^+}^{(t-\tau)^++s} \mathcal{S}^i(u) \sigma^i(X_u) dW(u) \right), s \geq 0 \right\}.$$

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

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

Proof. By Hölder's inequality, it suffices to prove this result for $p > 1$. Fix i, t and consider $q \in \mathbb{R}$ satisfying $|q| > 1$. Since $\int_{(t-\tau)^+}^{(t-\tau)^++s} \mathcal{S}^i(u) |\sigma^i(X_u)|^2 du \leq K_\sigma s$ for all $s \geq 0$, by Novikov's condition for exponential martingales, the process

$$\left\{ \chi_q(s) := (\xi^{t,i}(s))^q \exp \left(-\frac{q^2}{2} \int_{(t-\tau)^+}^{(t-\tau)^++s} \mathcal{S}^i(u) |\sigma^i(X_u)|^2 du \right), \mathcal{F}_{(t-\tau)^++s}, s \geq 0 \right\}$$

is a martingale. Thus, for any stopping time η , for each $s \geq 0$,

$$E \left[(\xi^{t,i}(s \wedge \eta))^q \right] \leq E [\chi_q(s \wedge \eta)] \exp \left(\frac{q^2}{2} K_\sigma s \right) = \exp \left(\frac{q^2}{2} K_\sigma s \right).$$

Now $\xi^{t,i}$ and $(\xi^{t,i})^{-1}$ are local submartingales and so there is a sequence of stopping times $\{\eta_n\}$ tending to infinity as $n \rightarrow \infty$ and such that for each n , the stopped processes, $\xi^{t,i}(\cdot \wedge \eta_n)$ and $(\xi^{t,i}(\cdot \wedge \eta_n))^{-1}$ are submartingales. Setting $q = p$ and $q = -p$ for $p > 1$, using Doob's inequality we obtain for each $T \geq 0$,

$$E \left[\sup_{u \in [0, T]} (\xi^{t,i}(u \wedge \eta_n))^q \right] \leq C_p E [(\xi^{t,i}(T \wedge \eta_n))^q] \leq C_p \exp \left(\frac{p^2}{2} K_\sigma T \right),$$

for a constant C_p depending only on p and which can be chosen to be increasing with p . Letting $n \rightarrow \infty$ and using monotone convergence completes the proof. \square

The following is an analogue of Lemma 2.4.1 for exponential moments in the case that σ^i is bounded on $\{x \in \mathbb{C}_\mathbb{I}^d : x^i(0) \geq M\}$. Although this result can be viewed as a delayed and constrained analogue to Theorem 4.7 of [18], we include a proof for completeness. Scrutiny of our proof reveals that the result does not need Assumption 2.3 (i), although the proof uses Assumption 2.3 (ii) and the linear growth condition (5).

Lemma 3.5.3. *If $E[\exp(\kappa \|X_0\|)] < \infty$ for each $\kappa > 0$, then for each $T \geq 0$ and $\kappa > 0$,*

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

Proof. Fix $\kappa > 0$. For each positive integer n , define the stopping time $\eta_n := \inf\{t \geq 0 : \|X\|_{[-\tau, t]} \geq n\}$, with the convention that $\inf \emptyset = \infty$. Convexity of the exponential function implies that for each $T \geq 0$,

$$\begin{aligned} E \left[\|\exp(\kappa |X(\cdot)|)\|_{[-\tau, T \wedge \eta_n]} \right] \\ \leq \frac{1}{d} \left(E [\exp(\kappa d \|X^1\|_{[-\tau, T \wedge \eta_n]})] + \cdots + E [\exp(\kappa d \|X^d\|_{[-\tau, T \wedge \eta_n]})] \right). \end{aligned} \quad (71)$$

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

$$E [\exp(\kappa d \|X^i\|_{[-\tau, T \wedge \eta_n]})] \leq e^{\kappa d M} (E [\exp(\kappa d \|Z_0^i\|)] + E [\|\exp(\kappa d Z^i(\cdot))\|_{[0, T \wedge \eta_n]}]). \quad (72)$$

Since (26) holds with $\tilde{M} = M$, as in the proof of Lemma 3.3.2, we can use Lemma 3.3.1 to conclude that

$$\begin{aligned} \text{Osc}^+(Z^i, [0, T \wedge \eta_n]) &\leq \int_0^T \mathbf{1}_{\{\|X\|_{[-\tau, t]} < n\}} |b(X_t)| dt + \sup_{0 \leq s \leq T \wedge \eta_n} \int_0^s \mathcal{I}^i(t) \sigma^i(X_t) dW(t) \\ &\quad + \sup_{0 \leq s \leq T \wedge \eta_n} \int_0^s -\mathcal{I}^i(t) \sigma^i(X_t) dW(t), \end{aligned} \quad (73)$$

where $\mathcal{I}^i(t)$ is given by (39). The linear growth condition (5) and Jensen's inequality imply that for any $\beta > 0$ and each $T > 0$,

$$\begin{aligned} \exp\left(\beta \int_0^T 1_{\{\|X\|_{[-\tau,t]} < n\}} |b(X_t)| dt\right) &\leq \frac{1}{T} \int_0^T \exp\left(T\beta 1_{\{\|X\|_{[-\tau,t]} < n\}} (C_1 + C_2 \|X_t\|)\right) dt \\ &\leq \frac{1}{T} \int_0^T \|\exp(T\beta(C_1 + C_2 |X(\cdot)|))\|_{[-\tau,t \wedge \eta_n]} dt. \end{aligned} \quad (74)$$

Then (73)–(74) (with $\beta = \kappa d/\lambda$) together with inequality (112) (with $n = 4$ and $a_1 = \kappa d \int_0^T 1_{\{\|X\|_{[-\tau,t]} < n\}} |b(X_t)| dt$) imply that for any $\lambda \in (0, 1)$, $T > 0$ and $i = 1, \dots, d$:

$$\begin{aligned} &E \left[\|\exp(\kappa d Z^i(\cdot))\|_{[0, T \wedge \eta_n]} \right] \\ &\leq E \left[\exp \left(\kappa d \left(Z^i(0) + \int_0^T 1_{\{\|X\|_{[-\tau,t]} < n\}} |b(X_t)| dt \right. \right. \right. \\ &\quad \left. \left. + \sup_{s \in [0, T \wedge \eta_n]} \int_0^s \mathcal{I}^i(t) \sigma^i(X_t) dW(t) + \sup_{s \in [0, T \wedge \eta_n]} \int_0^s -\mathcal{I}^i(t) \sigma^i(X_t) dW(t) \right) \right) \Bigg] \\ &\leq \frac{1-\lambda}{3} E \left[\exp \left(\frac{3\kappa d}{1-\lambda} \|X_0\| \right) \right] \\ &\quad + \lambda E \left[\frac{1}{T} \int_0^T \left\| \exp \left(T \frac{\kappa d}{\lambda} (C_1 + C_2 |X(\cdot)|) \right) \right\|_{[-\tau, t \wedge \eta_n]} dt \right] \\ &\quad + \frac{1-\lambda}{3} E \left[\sup_{s \in [0, T]} \exp \left(\frac{3\kappa d}{1-\lambda} \int_0^s \mathcal{I}^i(t) \sigma^i(X_t) dW(t) \right) \right] \\ &\quad + \frac{1-\lambda}{3} E \left[\sup_{s \in [0, T]} \exp \left(\frac{-3\kappa d}{1-\lambda} \int_0^s \mathcal{I}^i(t) \sigma^i(X_t) dW(t) \right) \right]. \end{aligned} \quad (75)$$

Lemma 3.5.2 (with $t = 0$) along with (71), (72) and (75) now imply that for each $T > 0$ and $\lambda \in (0, 1)$,

$$\begin{aligned} &e^{-\kappa d M} E \left[\|\exp(\kappa |X(\cdot)|)\|_{[-\tau, T \wedge \eta_n]} \right] \\ &\leq E[\exp(\kappa d \|X_0\|)] + \frac{1-\lambda}{3} E \left[\exp \left(\frac{3\kappa d}{1-\lambda} \|X_0\| \right) \right] \\ &\quad + \frac{\lambda}{T} \exp \left(\frac{T\kappa d C_1}{\lambda} \right) \int_0^T E \left[\left\| \exp \left(\frac{T\kappa d C_2}{\lambda} |X(\cdot)| \right) \right\|_{[-\tau, t \wedge \eta_n]} \right] dt \\ &\quad + \frac{2(1-\lambda)}{3} K \left(\frac{3\kappa d}{1-\lambda}, T \right). \end{aligned} \quad (76)$$

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

$$E \left[\|\exp(\kappa |X(\cdot)|)\|_{[-\tau, T \wedge \eta_n]} \right] \leq K_0(\kappa) + K_1(\kappa) \int_0^T E \left[\|\exp(\kappa |X(\cdot)|)\|_{[-\tau, t \wedge \eta_n]} \right] dt, \quad (77)$$

where

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

Inequality (77) is obvious for $T = 0$. Our assumptions imply that $K_0(\kappa) < \infty$ and $K_1(\kappa) > 0$. Therefore, since

$$\|\exp(\kappa |X(\cdot)|)\|_{[-\tau, t \wedge \eta_n]} \leq \exp(\kappa(\|X_0\| + n)), \quad (78)$$

so that the expectation on the left of (77) is finite, Gronwall's inequality implies that

$$E \left[\|\exp(\kappa |X(\cdot)|)\|_{[-\tau, \frac{1}{2dC_2} \wedge \eta_n]} \right] \leq K_0(\kappa) \exp \left(K_1(\kappa) \frac{1}{2dC_2} \right). \quad (79)$$

The monotone convergence theorem can then be applied to obtain for each $\kappa > 0$,

$$E \left[\|\exp(\kappa |X(\cdot)|)\|_{[-\tau, \frac{1}{2dC_2}]} \right] \leq K_0(\kappa) \exp \left(K_1(\kappa) \frac{1}{2dC_2} \right). \quad (80)$$

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

$$E \left[\|\exp(\kappa |X(\cdot)|)\|_{[-\tau, T^{(k)}]} \right] < \infty \quad \text{for each } \kappa > 0. \quad (81)$$

We can show that this holds with $k+1$ in place of k by applying the above procedure with 0 replaced by $T^{(k)}$, η_n replaced with $\eta_n^{(k)} := \inf\{t \geq T^{(k)} : \|X\|_{[-\tau, t]} \geq n\}$, $K_1(\kappa)$ unchanged, and $K_0(\kappa)$ replaced with $K_0^{(k)}(\kappa)$ which is given by the expression for $K_0(\kappa)$ with $\|X\|_{[-\tau, T^{(k)}]}$ in place of $\|X_0\|$. Then,

$$E \left[\|\exp(\kappa |X(\cdot)|)\|_{[-\tau, T^{(k+1)}]} \right] \leq K_0^{(k)}(\kappa) \exp \left(K_1(\kappa) \frac{1}{2dC_2} \right). \quad (82)$$

By induction, (81) holds for all k . Since $T^{(k)} \rightarrow \infty$ as $k \rightarrow \infty$, the proof is complete. \square

Lemma 3.5.4. *Fix a possibly random X_0 . Suppose that $\gamma > 0$ and $E[\exp(2\gamma\|X_0\|)] < \infty$. Then for each $i \in \{1, \dots, d\}$ and $t \geq 0$,*

$$E[\exp(\gamma \text{Osc}^+(X_t^i, \mathbb{I}))] \leq K(\gamma) := e^{\gamma(M+K_u)} (E[\exp(2\gamma\|X_0\|)])^{\frac{1}{2}} (K(4\gamma, \tau))^{\frac{1}{2}}, \quad (83)$$

where $K(\cdot, \cdot)$ is specified in Lemma 3.5.2.

Proof. Lemma 3.3.2 with $\hat{M} = M$ and Assumption 2.3 imply that P -a.s. for any $t \geq 0$,

$$\begin{aligned} \exp(\gamma \text{Osc}^+(X_t^i, \mathbb{I})) &\leq e^{\gamma(M+\tau K_u)} \exp(\gamma \text{Osc}^+(X_0^i, \mathbb{I})) \exp \left(\sup_{(t-\tau)^+ \leq s \leq t} \gamma \int_{(t-\tau)^+}^s \sigma^i(X_u) \mathcal{I}^i(u) dW(u) \right) \\ &\quad \times \exp \left(\sup_{(t-\tau)^+ \leq s \leq t} \gamma \int_{(t-\tau)^+}^s -\mathcal{I}^i(u) \sigma^i(X_u) dW(u) \right) \\ &\leq e^{\gamma(M+\tau K_u)} (E[\exp(2\gamma\|X_0\|)])^{\frac{1}{2}} (K(4\gamma, \tau))^{\frac{1}{2}}, \end{aligned} \quad (84)$$

which is finite by assumption. Here we have used the Cauchy-Schwarz inequality twice, Lemma 3.5.2, and (29) to obtain the second inequality. \square

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

Lemma 3.5.5. *Assume $\sup_{t \geq -\tau} E[|X(t)|^2] < \infty$ and $E[\|X_0\|^2] < \infty$. Then $\sup_{t \geq 0} E[\|X_t\|^2] < \infty$.*

Proof. After replacing B_0 by K_u , setting $B_{2,i} = C_{2,i} = 0$ for each i , and using the hypothesis of the lemma in place of Theorem 3.4.1, the proof is identical to the proof of Theorem 3.4.2. \square

Combining Theorem 3.5.1 with $p = 2$ and Lemma 3.5.5 yields the following. Recall for this that we are assuming that Assumption 2.3 holds.

Corollary 3.5.1. *Suppose $E[\exp(\kappa\|X_0\|)] < \infty$ for all $\kappa > 0$. Then $\sup_{t \geq 0} E[\|X_t\|^2] < \infty$.*

3.6 Existence Theorem

The following is obtained by combining the results from Section 3.1 and either Section 3.4 or Section 3.5.

Theorem 3.6.1. *If either Assumption 2.2 or 2.3 holds (in addition to Assumption 2.1), then there exists a stationary distribution for the SDDER (1).*

Proof. For each $x_o \in \mathbb{C}_{\mathbb{I}}^d$, the hypotheses on the initial conditions of either Theorem 3.4.2 or Corollary 3.5.1 are met, so that $\sup_{t \geq 0} E[\|X_t^{x_o}\|^2] < \infty$. The result now follows from Theorem 3.1.1. \square

4 Uniqueness of Stationary Distributions

Throughout this section, we assume that Assumption 2.4 holds (in addition to Assumption 2.1). We will prove uniqueness of any stationary distribution for the SDDER. For this proof, we adapt to equations with reflection a clever asymptotic coupling argument recently introduced by Hairer, Mattingly and Scheutzow [12] for stochastic delay differential equations without reflection. Of particular note is the fact that this argument applies to equations with a dispersion coefficient that depends on the history of the process over the delay period. Most previous work on proving uniqueness of stationary distributions relied on showing the mutual equivalence of distributions of X_t at some time $t > 0$ for all starting states, and then applying either Doob's theorem (see Theorem 4.2.1 in [7]) as in [30], or the techniques of Döblin (see, e.g., [24]) as in [17, 34]. However, these arguments cannot be easily extended to situations where σ depends on past states, because of the potential for reconstruction of the initial condition from the quadratic variation process (see [30, 36]). In [12], this potential difficulty is avoided by use of a different ergodic argument (see Theorem 1.1 of [12]). The main idea of this argument is to introduce a change of probability measure under which, with strictly positive probability, two solutions of the SDDER starting from different initial conditions are driven towards one another as time goes to infinity.

Although our general line of argument is very similar to that in [12], there are some differences due to the presence of reflection in the dynamics and we also provide more details for some steps. We begin in Section 4.1 by stating an abstract uniqueness result proved in

[12]. The key technical section is Section 4.2 where the novel asymptotic coupling introduced in [12] is adapted to our setting. The uniqueness result is then stated and proved in Section 4.3.

4.1 An Abstract Uniqueness Result

A key element for our proof is the following proposition, which is adapted to our situation from Corollary 2.2 of [12]. Before stating it, we introduce some notation. Denote the space of sequences $\{x^n\}_{n=0}^\infty$ in $\mathbb{C}_\mathbb{I}^d$ by $(\mathbb{C}_\mathbb{I}^d)^\infty$, and endow this with the product topology and associated Borel σ -algebra. For $x \in \mathbb{C}_\mathbb{I}^d$, let P_∞^x denote the probability measure on $(\mathbb{C}_\mathbb{I}^d)^\infty$ that is the distribution of the sequence $\{X_{n\tau}\}_{n=0}^\infty$ when X is a solution of (1) started from x . 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 [12] by setting $A = \mathbb{C}_\mathbb{I}^d$ there.

Proposition 4.1.1. *Assume that there is a family $\{\tilde{P}^{x,y}, (x,y) \in \mathbb{C}_\mathbb{I}^d \times \mathbb{C}_\mathbb{I}^d\}$ 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) $\tilde{P}^{x,y}(\cdot \times (\mathbb{C}_\mathbb{I}^d)^\infty) \sim P_\infty^x(\cdot)$ and $\tilde{P}^{x,y}((\mathbb{C}_\mathbb{I}^d)^\infty \times \cdot) \sim P_\infty^y(\cdot)$,
- (ii) for each $x, y \in \mathbb{C}_\mathbb{I}^d$,

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

and

- (iii) for each Borel set Γ in $(\mathbb{C}_\mathbb{I}^d)^\infty \times (\mathbb{C}_\mathbb{I}^d)^\infty$, the mapping $(x, y) \rightarrow \tilde{P}^{x,y}(\Gamma)$ is measurable on $\mathbb{C}_\mathbb{I}^d \times \mathbb{C}_\mathbb{I}^d$.

Then there exists at most one stationary distribution for the SDDER (1).

4.2 Asymptotic Coupling of a Pair of Processes

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

$$dX(t) = b(X_t)dt + \sigma(X_t)dW(t) + dY(t), \quad (85)$$

$$d\tilde{X}^\lambda(t) = b(\tilde{X}_t^\lambda)dt + \lambda(X(t) - \tilde{X}^\lambda(t))dt + \sigma(\tilde{X}_t^\lambda)dW(t) + d\tilde{Y}^\lambda(t), \quad (86)$$

where P -a.s., $(X(t), \tilde{X}^\lambda(t)) \in \mathbb{R}_+^{2d}$ for all $t \geq -\tau$, and where (Y, \tilde{Y}^λ) 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 random \mathcal{F}_0 -measurable initial conditions for X_0 and \tilde{X}_0^λ . This is a $2d$ -dimensional system with globally Lipschitz coefficients, and thus Proposition 2.1.2 implies that there exists a unique strong solution for any pair of initial conditions. Consider such a solution pair (X, \tilde{X}^λ) , and let $U^\lambda(t) := X(t) - \tilde{X}^\lambda(t)$ for $t \geq -\tau$. Then, for $t \geq 0$,

$$dU^\lambda(t) = \left(b(X_t) - b(\tilde{X}_t^\lambda) \right) dt - \lambda U^\lambda(t)dt + \left(\sigma(X_t) - \sigma(\tilde{X}_t^\lambda) \right) dW(t) + d\left(Y - \tilde{Y}^\lambda \right)(t). \quad (87)$$

The following lemma is a modified version of Lemma 3.5 of [12], where here we have equations with reflection and we give the result for all $q > 4$ rather than just for $q = 8$. Inequality (97) is the reason that this lemma remains true in the reflected case. Our proof of this lemma is very similar to that in [12], but for completeness, we provide the details. We have also extracted Lemma 4.2.2 as a separate preliminary result. In [12] the analogous result for $q = 4$ is included within the proof of their Lemma 4.2.1.

Lemma 4.2.1. *For each $\alpha > 0$ and $q > 4$, there exist $\lambda > 0$ and $K > 0$ depending only on $\alpha, q, \tau, \kappa_L$ such that*

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

Before proving this lemma, we give two propositions and a preliminary lemma. The first proposition is a simple stochastic variation of constants formula.

Proposition 4.2.1. *Assume that on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, $\{\xi(t), t \geq 0\}$ is a continuous adapted process satisfying the following stochastic differential equation:*

$$d\xi(t) = \gamma \xi(t) dt + d\chi(t), \quad (88)$$

for some $\gamma \in \mathbb{R}$ and some continuous semimartingale $\{\chi(t), t \geq 0\}$. Then

$$\xi(t) = e^{\gamma t} \xi(0) + \int_0^t e^{\gamma(t-s)} d\chi(s), \quad t \geq 0, \quad (89)$$

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

$$\xi(t) = e^{\gamma(t-s)} \xi(s) + \int_s^t e^{\gamma(t-r)} d\chi(r).$$

Proof. It is straight-forward to verify that the right member of (89) satisfies (88). By the uniqueness of solutions for this equation given the initial state $\xi(0)$, the result follows. \square

The next proposition is a slight generalization of Lemma 3.4 in [12] to the case where W is m -dimensional, and specializes to the case where h is continuous. The proof of the proposition is nearly identical to that in [12], and so we omit it. In brief, this proof uses the representation $V^\beta(t) = e^{-\beta t} \int_0^t e^{\beta s} h(s) dW(s)$, the Burkholder-Davis-Gundy inequality, an integration by parts and estimates on V^β on the segments $\left[\frac{kT}{N}, \frac{(k+1)T}{N} \right]$, $k = 0, \dots, N-1$, for sufficiently large integers N .

Proposition 4.2.2. *Suppose that $\{h(s), s \geq 0\}$ is a continuous adapted process taking values in $\mathbb{M}^{1 \times m}$, and assume that for each $\beta > 0$ we have an adapted continuous real-valued process $\{V^\beta(t), t \geq 0\}$ satisfying the stochastic differential equation*

$$dV^\beta(t) = -\beta V^\beta(t) dt + h(t) dW(t), \quad t \geq 0, \quad (90)$$

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

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

For $p = 4$, the next lemma is implicitly proved in Lemma 3.5 of [12]. We state and prove this result for general $p > 2$ here since the result may be of independent interest; however, our proof for the general case uses the same line of argument as in [12].

Lemma 4.2.2. *For each $\lambda > 0$, let ζ^λ be a continuous process defined on the time interval $[-\tau, \infty)$, taking values in $[0, \infty)$ such that $\zeta^\lambda(t) \in \mathcal{F}_0$ for all $t \in [-\tau, 0]$, $\zeta^\lambda(t) \in \mathcal{F}_t$ for all $t \geq 0$, and*

$$\zeta^\lambda(t) = \zeta^\lambda(0) + \int_0^t k^\lambda(s) ds + \int_0^t \ell^\lambda(s) dW(s) + v^\lambda(t) \quad \text{for all } t \geq 0, \quad (91)$$

where k^λ, ℓ^λ and v^λ are continuous adapted processes taking values in $\mathbb{R}, \mathbb{M}^{1 \times m}$ and $(-\infty, 0]$, respectively, and v^λ is a non-increasing process satisfying $v^\lambda(0) = 0$. Suppose that there are strictly positive constants K_1, K_2 and K_3 such that for each $\lambda > 0$, P -a.s. for all $s \in [0, \infty)$,

$$k^\lambda(s) \leq -\lambda K_1 \zeta^\lambda(s) + K_2 \|\zeta_s^\lambda\|, \quad (92)$$

$$|\ell^\lambda(s)| \leq K_3 \|\zeta_s^\lambda\|. \quad (93)$$

Then for each $p \in (2, \infty)$ and $\alpha > 0$, there are positive constants λ and C (depending only on $p, \tau, \alpha, K_1, K_2, K_3$) such that

$$E \left[\sup_{t \geq 0} e^{\alpha t} \|\zeta_t^\lambda\|^p \right] \leq C E \left[\|\zeta_0^\lambda\|^p \right]. \quad (94)$$

Proof. Fix $p > 2$, $\alpha > 0$ and $\kappa > 2\alpha/p$ such that $e^{\kappa p \tau/2} > 2 \cdot 3^{p-1}$. For each $\lambda > 0$ let

$$\xi^\lambda(t) = e^{\kappa t} \zeta^\lambda(t) \quad \text{for all } t \in [-\tau, \infty). \quad (95)$$

Then for all $t \geq 0$:

$$\begin{aligned} d\xi^\lambda(t) &= \kappa \xi^\lambda(t) dt + e^{\kappa t} d\zeta^\lambda(t) \\ &= -\beta \xi^\lambda(t) dt + e^{\kappa t} \tilde{k}^\lambda(t) dt + e^{\kappa t} \ell^\lambda(t) dW(t) + e^{\kappa t} dv^\lambda(t), \end{aligned}$$

where $\beta = -\kappa + \lambda K_1$ and $\tilde{k}^\lambda(t) = \lambda K_1 \zeta^\lambda(t) + k^\lambda(t)$ for all $t \geq 0$. It follows from Proposition 4.2.1 (with $\gamma = -\beta$) that for any $0 \leq s \leq t$,

$$\begin{aligned} \xi^\lambda(t) &= e^{-\beta(t-s)} \xi^\lambda(s) + \int_s^t e^{-\beta(t-r)} e^{\kappa r} \tilde{k}^\lambda(r) dr \\ &\quad + \int_s^t e^{-\beta(t-r)} e^{\kappa r} \ell^\lambda(r) dW(r) + \int_s^t e^{-\beta(t-r)} e^{\kappa r} dv^\lambda(r) \\ &\leq e^{-\beta(t-s)} \xi^\lambda(s) + K_2 \int_s^t e^{-\beta(t-r)} e^{\kappa r} \|\zeta_r^\lambda\| dr + V^\beta(t) - V^\beta(s), \\ &\leq e^{-\beta(t-s)} \xi^\lambda(s) + K_2 e^{\kappa \tau} \int_s^t e^{-\beta(t-r)} \|\xi_r^\lambda\| dr + V^\beta(t) - V^\beta(s), \end{aligned}$$

where $V^\beta(t) = \int_0^t e^{-\beta(t-r)} e^{\kappa r} \ell^\lambda(r) dW(r)$ for all $t \geq 0$, and we have used the fact that v^λ is non-increasing together with (92) for the first inequality above.

Fix an integer $n \geq 0$. For each positive integer m , let $\eta_m^\lambda = \inf\{r \geq n\tau : |\xi^\lambda(r)| \geq m\}$. Then by applying Proposition 4.2.2 with $\{\mathcal{F}_{n\tau+t}, t \geq 0\}$, $V^\beta(\cdot + n\tau) - V^\beta(n\tau)$, $W(\cdot + n\tau) - W(n\tau)$ and η_m^λ in place of $\{\mathcal{F}_t, t \geq 0\}$, $V^\beta(\cdot)$, $W(\cdot)$ and η , respectively, and then using (93), we obtain the following:

$$\begin{aligned} E \left[\sup_{r \in [n\tau, \eta_m^\lambda \wedge (n+1)\tau]} |V^\beta(r) - V^\beta(n\tau)|^p \right] &\leq \nu_{\tau,p}(\beta) E \left[\sup_{r \in [n\tau, \eta_m^\lambda \wedge (n+1)\tau]} |e^{\kappa r} \ell^\lambda(r)|^p \right] \\ &\leq \nu_{\tau,p}(\beta) E \left[\sup_{r \in [n\tau, \eta_m^\lambda \wedge (n+1)\tau]} e^{p\kappa r} K_3^p \|\xi_r^\lambda\|^p \right] \\ &\leq \nu_{\tau,p}(\beta) K_3^p e^{p\kappa\tau} E \left[\sup_{r \in [n\tau, \eta_m^\lambda \wedge (n+1)\tau]} \|\xi_r^\lambda\|^p \right], \end{aligned}$$

whenever λ is sufficiently large that $\beta > 0$. Let

$$\Gamma(\lambda, n, m) = E \left[\sup_{r \in [n\tau, \eta_m^\lambda \wedge (n+1)\tau]} \|\xi_r^\lambda\|^p \right].$$

On combining the above with (110) and the fact that ξ^λ takes values in $[0, \infty)$, for all λ sufficiently large that $\beta > 0$, we obtain that

$$\Gamma(\lambda, n, m) \leq 3^{p-1} \left(E \left[\|\xi_{n\tau}^\lambda\|^p \right] + \frac{K_2^p e^{p\kappa\tau} \Gamma(\lambda, n, m)}{\beta^p} + \nu_{\tau,p}(\beta) K_3^p e^{p\kappa\tau} \Gamma(\lambda, n, m) \right).$$

Now by choosing λ sufficiently large (depending only on $p, \tau, \alpha, \kappa, K_1, K_2, K_3$) we can guarantee that $\beta > 0$ and that β is sufficiently large that

$$\frac{3^{p-1} K_2^p e^{p\kappa\tau}}{\beta^p} \leq \frac{1}{4} \quad \text{and} \quad 3^{p-1} \nu_{\tau,p}(\beta) K_3^p e^{p\kappa\tau} \leq \frac{1}{4}.$$

Fix such a λ . Then for all n, m ,

$$\Gamma(\lambda, n, m) \leq 2 \cdot 3^{p-1} E \left[\|\xi_{n\tau}^\lambda\|^p \right].$$

(We note that by the definition of η_m^λ , this holds even if $\Gamma(\lambda, n, m)$ is infinite, since this will only occur when the right hand side above is also infinite.) Now, on letting $m \rightarrow \infty$, we obtain by monotone convergence that

$$E \left[\sup_{r \in [n\tau, (n+1)\tau]} \|\xi_r^\lambda\|^p \right] \leq 2 \cdot 3^{p-1} E \left[\|\xi_{n\tau}^\lambda\|^p \right].$$

It follows that for each integer $n \geq 0$,

$$E \left[\|\xi_{n\tau}^\lambda\|^p \right] \leq 2^n 3^{n(p-1)} E \left[\|\xi_0^\lambda\|^p \right].$$

Then,

$$\begin{aligned}
E \left[\sup_{t \geq 0} e^{\alpha t} |\zeta^\lambda(t)|^p \right] &\leq E \left[\sup_{t \geq 0} e^{\kappa p t / 2} |\zeta^\lambda(t)|^p \right] \leq \sum_{n=1}^{\infty} E \left[\sup_{t \in [(n-1)\tau, n\tau]} e^{\kappa p t / 2} |\zeta^\lambda(t)|^p \right] \\
&\leq \sum_{n=1}^{\infty} E \left[e^{-\kappa(n-1)p\tau/2} \|\xi_{n\tau}^\lambda\|^p \right] \leq \sum_{n=1}^{\infty} e^{-\kappa(n-1)p\tau/2} 2^n 3^{n(p-1)} E \left[\|\xi_0^\lambda\|^p \right] \\
&\leq C_1 E \left[\|\xi_0^\lambda\|^p \right] \leq C_1 E \left[\|\zeta_0^\lambda\|^p \right],
\end{aligned}$$

where $C_1 = \sum_{n=1}^{\infty} e^{-\kappa(n-1)p\tau/2} 2^n 3^{n(p-1)}$ is finite by the choice of κ . The desired result then follows by the observation that

$$E \left[\sup_{t \geq 0} e^{\alpha t} \|\zeta_t^\lambda\|^p \right] \leq e^{\alpha\tau} \left(E \left[\sup_{t \geq 0} e^{\alpha t} |\zeta^\lambda(t)|^p \right] + E \left[\|\zeta_0^\lambda\|^p \right] \right).$$

□

Proof of Lemma 4.2.1. Fix $\alpha > 0$ and $q > 4$. Let $p = q/2$. From equation (87), we have by Itô's formula that

$$\begin{aligned}
d|U^\lambda(t)|^2 &= -2\lambda|U^\lambda(t)|^2 dt + 2 \left(U^\lambda(t) \right)' \left(b(X_t) - b(\tilde{X}_t^\lambda) \right) dt \\
&\quad + 2 \left(U^\lambda(t) \right)' \left(\sigma(X_t) - \sigma(\tilde{X}_t^\lambda) \right) dW(t) + 2 \left(U^\lambda(t) \right)' d \left(Y - \tilde{Y}^\lambda \right) (t) \\
&\quad + \left| \sigma(X_t) - \sigma(\tilde{X}_t^\lambda) \right|^2 dt.
\end{aligned} \tag{96}$$

The constraints on where Y and \tilde{Y}^λ can increase and the positivity of X and \tilde{X}^λ imply that for each $0 \leq s \leq t$,

$$\begin{aligned}
\int_s^t (U^\lambda(r))' d(Y - \tilde{Y}^\lambda)(r) &= \sum_{i=1}^d \int_s^t U^{\lambda,i}(r) d \left(Y^i - \tilde{Y}^{\lambda,i} \right) (r) \\
&= -\sum_{i=1}^d \left(\int_s^t X^i(r) d\tilde{Y}^{\lambda,i}(r) + \int_s^t \tilde{X}^{\lambda,i}(r) dY^i(r) \right) \\
&\leq 0.
\end{aligned} \tag{97}$$

The Lipschitz continuity condition (4) on b and σ implies that for any $x, y \in \mathbb{C}_\mathbb{1}^d$,

$$2(x(0) - y(0))'(b(x) - b(y)) + |\sigma(x) - \sigma(y)|^2 \leq (1 + \kappa_L) \|x - y\|^2. \tag{98}$$

Thus, on setting $\zeta^\lambda(t) = |U^\lambda(t)|^2$ for $t \in [-\tau, \infty)$, we see that the hypotheses of Lemma 4.2.2 are satisfied with

$$\begin{aligned}
k^\lambda(s) &= -2\lambda\zeta^\lambda(s) + 2(U^\lambda(s))'(b(X_s) - b(\tilde{X}_s^\lambda)) + |\sigma(X_s) - \sigma(\tilde{X}_s^\lambda)|^2, \\
\ell^\lambda(s) &= 2(U^\lambda(s))'(\sigma(X_s) - \sigma(\tilde{X}_s^\lambda)),
\end{aligned}$$

and $K_1 = 2$, $K_2 = 1 + \kappa_L$ and $K_3 = 2\sqrt{1 + \kappa_L}$. The desired result then follows immediately from Lemma 4.2.2. □

From this point on, we shall fix a $\lambda > 0$ such that the result of Lemma 4.2.1 holds for $\alpha = 4$ and $q = 8$ there. For two segments $x, y \in \mathbb{C}_{\mathbb{I}}^d$, let $(X^x, \tilde{X}^{x,y})$ be the unique strong solution to (85)–(86) with the initial condition $(X_0, \tilde{X}_0^\lambda) \equiv (x, y)$. Recall that we are assuming that Assumptions 2.1 and 2.4 hold. Then by Proposition 2.2.1, there is a continuous inverse σ^\dagger for σ that is uniformly bounded by some finite constant C_σ . For each integer $n \geq 1$ and $x, y \in \mathbb{C}_{\mathbb{I}}^d$, define the stopping time

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

Lemma 4.2.3. *For each $x, y \in \mathbb{C}_{\mathbb{I}}^d$,*

$$P \left(\limsup_{t \rightarrow \infty} |X^x(t) - \tilde{X}^{x,y}(t)| = 0 \right) = 1 \quad \text{and} \quad \lim_{n \rightarrow \infty} P(\eta^{x,y,n} = \infty) = 1.$$

Proof. Fix $x, y \in \mathbb{C}_{\mathbb{I}}^d$. Let $U^\lambda(t) = X^x(t) - \tilde{X}^{x,y}(t)$ for all $t \in [-\tau, \infty)$ and let $\Upsilon := \sup_{t \geq 0} e^t \|U_t^\lambda\|^2$. Then by Lemma 4.2.1, $E[\Upsilon^4] \leq K \|x - y\|^8$. Thus, $\Upsilon < \infty$ P -a.s., and the first conclusion of the lemma follows immediately.

Then by the bound on σ^\dagger and the fact that $\|U_t^\lambda\|^2 \leq e^{-t} \Upsilon$ for each $t \geq 0$, we have

$$\int_0^\infty \lambda^2 \left| \sigma^\dagger(\tilde{X}_s^{x,y}) U^\lambda(s) \right|^2 ds \leq \lambda^2 C_\sigma^2 \int_0^\infty |U^\lambda(s)|^2 ds \leq \lambda^2 C_\sigma^2 \int_0^\infty e^{-s} \Upsilon ds \leq \lambda^2 C_\sigma^2 \Upsilon.$$

Therefore,

$$P(\eta^{x,y,n} = \infty) \geq P \left(\int_0^\infty \lambda^2 \left| \sigma^\dagger(\tilde{X}_s^{x,y}) U^\lambda(s) \right|^2 ds < n \right) \geq P(\lambda^2 C_\sigma^2 \Upsilon < n), \quad (99)$$

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

4.3 Uniqueness Theorem

Theorem 4.3.1. *Under Assumption 2.4 (in addition to Assumption 2.1), there exists at most one stationary distribution for the SDDER (1).*

Given the results of the previous section, our proof of the uniqueness theorem has the same general structure as the proof of Theorem 3.1 in [12], although we include a few more details, especially with regard to certain measurability properties.

Proof. We verify the hypotheses of Proposition 4.1.1. Define the function $N : \mathbb{C}_{\mathbb{I}}^d \times \mathbb{C}_{\mathbb{I}}^d \rightarrow \{1, 2, \dots\}$ by

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

which is finite valued by Lemma 4.2.3. The following maps are measurable:

$$\mathbb{C}_{\mathbb{J}}^d \times \mathbb{C}_{\mathbb{J}}^d \ni (\omega, \tilde{\omega}) \rightarrow \int_0^k \lambda^2 \left| \sigma^\dagger(\tilde{\omega}_s)(\omega(s) - \tilde{\omega}(s)) \right|^2 ds, \quad k = 1, 2, \dots$$

It follows that for each $n = 1, 2, \dots$,

$$\Gamma_n := \left\{ (\omega, \tilde{\omega}) : \int_0^k \lambda^2 \left| \sigma^\dagger(\tilde{\omega}_s) (\omega(s) - \tilde{\omega}(s)) \right|^2 ds < n \text{ for } k = 1, 2, \dots \right\}$$

is in $\mathcal{M}_{\mathbb{J}} \otimes \mathcal{M}_{\mathbb{J}}$. Then Feller continuity of the Markovian transition functions implies that the map $(x, y) \rightarrow P(\eta^{x,y,n} = \infty) = P((X^x, \tilde{X}^{x,y}) \in \Gamma_n)$ is measurable for each n , and so the measurability of $N(\cdot, \cdot)$ follows. Henceforth, we abbreviate

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

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

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

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

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

so by Novikov's condition (see, e.g., Proposition VIII.1.15 of [31]),

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

defines a uniformly integrable martingale. Let $\rho^{x,y}(\infty)$ denote the P -a.s. strictly positive limit of $\rho^{x,y}(t)$ as $t \rightarrow \infty$. It then follows from Girsanov's theorem (see, e.g., Section 1 of Chapter VIII of [31]) 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

$$d\bar{X}(t) = b(\bar{X}_t)dt + \sigma(\bar{X}_t)d\tilde{W}^{x,y}(t) + d\bar{Y}(t), \quad (100)$$

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

$$d\bar{X}^{x,y}(t) = b(\bar{X}_t^{x,y})dt + 1_{\{t \leq \eta^{x,y}\}} \lambda \left(X^x(t) - \tilde{X}^{x,y}(t) \right) dt + \sigma(\bar{X}_t^{x,y})dW(t) + d\bar{Y}(t), \quad (101)$$

where W is a Brownian motion under P . For (101), we used the facts that $\sigma\sigma^\dagger = I_d$ and

$$\begin{aligned} P \left(\sigma^\dagger(\bar{X}_t^{x,y}) = \sigma^\dagger(\tilde{X}_t^{x,y}) \text{ for all } t \in [0, \eta^{x,y}] \cap [0, \infty) \right) \\ \geq P \left(\bar{X}^{x,y}(t) = \tilde{X}^{x,y}(t) \text{ for all } t \in [-\tau, \eta^{x,y}] \cap [-\tau, \infty) \right) = 1. \end{aligned} \quad (102)$$

The equality above follows by a very similar proof to that for the strong uniqueness of solutions for the SDDER with Lipschitz coefficients using Gronwall's inequality.

Since uniqueness in law holds for solutions of (85), the distribution of $\bar{X}^{x,y}$ under $Q^{x,y}$ is the same as that of the solution X^y to (85) under P with initial condition $X_0 = y$. Then, since $Q^{x,y} \sim P$, the distribution of $\bar{X}^{x,y}$ under P and the distribution of X^y under P are mutually

absolutely continuous. Thus, if we let $\tilde{P}^{x,y}$ be the probability measure on $(\mathbb{C}_{\mathbb{I}}^d)^\infty \times (\mathbb{C}_{\mathbb{I}}^d)^\infty$ that is the law of $(\{X_{n\tau}^x\}_{n=0}^\infty, \{\tilde{X}_{n\tau}^{x,y}\}_{n=0}^\infty)$ under P , then $\tilde{P}^{x,y}$ satisfies condition (i) of Proposition 4.1.1.

On the set $\{\eta^{x,y} = \infty\}$, we have $\bar{X}^{x,y} = \tilde{X}^{x,y}$ P -a.s. by (102). Thus, on $\{\eta^{x,y} = \infty\}$, by Lemma 4.2.3 we have P -a.s.:

$$\lim_{t \rightarrow \infty} |X^x(t) - \bar{X}^{x,y}(t)| = \lim_{t \rightarrow \infty} |X^x(t) - \tilde{X}^{x,y}(t)| = 0. \quad (103)$$

Therefore,

$$\begin{aligned} \tilde{P}^{x,y} \left((\{x^n\}_{n=0}^\infty, \{y^n\}_{n=0}^\infty) \in (\mathbb{C}_{\mathbb{I}}^d)^\infty \times (\mathbb{C}_{\mathbb{I}}^d)^\infty : \lim_{n \rightarrow \infty} \|x^n - y^n\| = 0 \right) &\geq P(\eta^{x,y} = \infty) \\ &\geq \frac{1}{2}, \end{aligned} \quad (104)$$

so that $\tilde{P}^{x,y}$ also satisfies condition (ii) of Proposition 4.1.1.

All that remains to be shown is the measurability property (iii) in Proposition 4.1.1. This will follow from the measurability of $(x, y) \rightarrow \bar{P}^{x,y}(B)$ for each Borel measurable set $B \subset C(\mathbb{R}_+, \mathbb{C}_{\mathbb{I}}^d) \times C(\mathbb{R}_+, \mathbb{C}_{\mathbb{I}}^d)$, where $\bar{P}^{x,y}$ is the law of $(X^x, \bar{X}^{x,y})$ under P . We establish the latter below.

By a monotone class argument, it suffices to prove that for each $k = 1, 2, \dots$, $0 \leq t_1 < t_2 < \dots < t_k < \infty$, and g_1, \dots, g_k in $C_b(\mathbb{C}_{\mathbb{I}}^{2d})$, the mapping $(x, y) \rightarrow E[g_1(X_{t_1}^x, \bar{X}_{t_1}^{x,y}) \cdots g_k(X_{t_k}^x, \bar{X}_{t_k}^{x,y})]$ is measurable. For the proof of this, let $A_0 := \{0 \leq \eta^{x,y} < t_1\}$, $A_k := \{t_k \leq \eta^{x,y}\}$, and for each $j = 1, \dots, k-1$ let $A_j := \{t_j \leq \eta^{x,y} < t_{j+1}\}$. Then, Ω is the disjoint union of A_0, \dots, A_k and

$$\begin{aligned} &E[g_1(X_{t_1}^x, \bar{X}_{t_1}^{x,y}) \cdots g_k(X_{t_k}^x, \bar{X}_{t_k}^{x,y})] \\ &= \sum_{j=0}^k E \left[\left(\prod_{i=1}^j g_i(X_{t_i}^x, \bar{X}_{t_i}^{x,y}) \right) 1_{A_j} E \left[\left(\prod_{i=j+1}^k g_i(X_{t_i}^x, \bar{X}_{t_i}^{x,y}) \right) \middle| \mathcal{F}_{\eta^{x,y}} \right] \right], \end{aligned} \quad (105)$$

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

For each $\tilde{x}, \hat{x} \in \mathbb{C}_{\mathbb{I}}^d$, let $P^{\tilde{x}, \hat{x}}$ denote the law induced on $C(\mathbb{R}_+, \mathbb{C}_{\mathbb{I}}^{2d})$ by the pair of strong solutions to (85) with the two initial conditions \tilde{x}, \hat{x} and the same driving Brownian motion W . Then, by the strong uniqueness for this pair, P -a.s., on $\{\eta^{x,y} < \infty\}$, the law of $(X_{\eta^{x,y}+}^x, \bar{X}_{\eta^{x,y}+}^{x,y})$ under P conditioned on $\mathcal{F}_{\eta^{x,y}}$ is given by $P^{X_{\eta^{x,y}}^x, \bar{X}_{\eta^{x,y}}^{x,y}}$. Let $h_k(\tilde{x}, \hat{x}, t) \equiv 1$ for all $(\tilde{x}, \hat{x}, t) \in \mathbb{C}_{\mathbb{I}}^d \times \mathbb{C}_{\mathbb{I}}^d \times [0, \infty]$. For $j = 0, 1, \dots, k-1$, define the measurable real-valued functions h_j on $\mathbb{C}_{\mathbb{I}}^d \times \mathbb{C}_{\mathbb{I}}^d \times [0, \infty)$ by

$$h_j(\tilde{x}, \hat{x}, t) := E^{P^{\tilde{x}, \hat{x}}} \left[\prod_{i=j+1}^k g_i(\tilde{w}(t_i - t), \hat{w}(t_i - t)) \right], \quad \text{for } t < t_{j+1},$$

and $h_j(\tilde{x}, \hat{x}, t) = 0$ for $t \geq t_{j+1}$, where (\tilde{w}, \hat{w}) denotes a generic point in $C(\mathbb{R}_+, \mathbb{C}_{\mathbb{I}}^{2d})$. From the above and using the fact (see (102)) that P -a.s., $\bar{X}^{x,y}(\cdot)$ and $\tilde{X}^{x,y}(\cdot)$ agree on $[-\tau, \eta^{x,y}] \cap$

$[-\tau, \infty)$, we see that the right hand side of (105) is equal to

$$\begin{aligned} & \sum_{j=0}^k E \left[\left(\prod_{i=1}^j g_i(X_{t_i}^x, \tilde{X}_{t_i}^{x,y}) \right) 1_{A_j} h_j(X_{\eta^{x,y}}^x, \tilde{X}_{\eta^{x,y}}^{x,y}, \eta^{x,y}) \right] \\ &= \sum_{j=0}^k E^{\hat{P}^{x,y}} \left[\left(\prod_{i=1}^j g_i(w_{t_i}, \tilde{w}_{t_i}) \right) 1_{B_j} h_j(w_\eta, \tilde{w}_\eta, \eta) \right], \end{aligned} \quad (106)$$

where for each $x, y \in \mathbb{C}_{\mathbb{I}}^d$, $\hat{P}^{x,y}$ is the law of $(X^x(\cdot), \tilde{X}^{x,y}(\cdot))$ under P ,

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

and B_j is defined in the same manner as A_j , but with η in place of $\eta^{x,y}$. The desired measurability in (x, y) of the expression in (106) then follows directly from the Feller continuity of the transition functions associated with $\{(X^x, \tilde{X}^{x,y}) : x, y \in \mathbb{C}_{\mathbb{I}}^d\}$, since the fact that $N(\cdot, \cdot)$ is measurable can be used to show that η is a stopping time with respect to $\{\tilde{\mathcal{F}}_t := \sigma((w_s, \tilde{w}_s) : 0 \leq s \leq t), t \geq 0\}$. \square

Acknowledgments. We thank Martin Hairer, Jonathan Mattingly and Michael Scheutzow for conversations concerning uniqueness of stationary solutions for stochastic delay differential equations and for sending us a preprint of their important recent work [12] on this topic. We also thank R. Srikant for discussions on Internet congestion control models and Jeff Hasty and Lev Tsimring for discussions on dynamic models with delay arising in systems biology. We are also grateful to an anonymous referee for several helpful suggestions, especially with regard to refining some of the proofs in Section 3.

A Reflection

To ensure that a solution of (1) remains positive, we have employed Skorokhod's well-known mapping for constraining a continuous real-valued function to be positive by means of reflection at the origin. This mapping was applied 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 A.0.1. 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 $i = 1, \dots, d$, $y^i(0) = 0$ and y^i is non-decreasing,
- (iii) for each $i = 1, \dots, d$, $\int_0^t z^i(s) dy^i(s) = 0$ 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 .

We summarize some basic facts about the Skorokhod problem in the next proposition.

Proposition A.0.1. *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 $(\phi, \psi) : C_+(\mathbb{R}_+, \mathbb{R}^d) \rightarrow C_+(\mathbb{R}_+, \mathbb{R}^{2d})$ defined by $(\phi(x), \psi(x)) = (z, y)$. The pair (ϕ, ψ) satisfies the following:*

(i)

$$\text{Osc}(\phi(x), [a, b]) \leq \text{Osc}(x, [a, b]). \quad (107)$$

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

$$\begin{aligned} \|\psi(x) - \psi(y)\|_{[0, t]} &\leq K_\ell \|x - y\|_{[0, t]}, \text{ and} \\ \|\phi(x) - \phi(y)\|_{[0, t]} &\leq K_\ell \|x - y\|_{[0, t]}. \end{aligned}$$

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

$$y^i(t) = \max_{0 \leq s \leq t} (x^i(s))^- , \quad i = 1, \dots, d. \quad (108)$$

For more details, see [9, 14, 40]. We note that $K_\ell \leq 2$, but we keep the notation K_ℓ for convenience. \square

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

$$(a_1 + a_2)^q \leq a_1^q + a_2^q, \text{ for all } q \in [0, 1], \quad (109)$$

which is obvious if $a_1 = a_2$ or if either is 0, and if $a_1 > a_2 > 0$ then $(a_1 + a_2)^q - a_1^q \leq qa_1^{q-1}a_2 < a_2^q$.

The following is a well-known fact that follows from the convexity of power functions. For any $p > 1$, $a_1, \dots, a_n \in \mathbb{R}$, we have

$$|a_1 + \dots + a_n|^p \leq n^{p-1}(|a_1|^p + \dots + |a_n|^p). \quad (110)$$

Sometimes n^{p-1} is too big for our needs, and we will use the following alternative, which can be proved with standard optimization techniques. For any $\gamma > 1$, and $a, q \geq 0$, there is a $K = K(a, \gamma, q) \geq 0$ such that

$$(a + t)^q \leq K + \gamma t^q \quad \text{for all } t \geq 0. \quad (111)$$

Convexity of the exponential function implies that for any integer $n \geq 2$, numbers $a_1, a_2, \dots, a_n \in \mathbb{R}_+$, and $\lambda \in (0, 1)$, we have

$$\exp\left(\lambda a_1 + \frac{1-\lambda}{n-1} a_2 + \dots + \frac{1-\lambda}{n-1} a_n\right) \leq \lambda \exp(a_1) + \frac{1-\lambda}{n-1} \exp(a_2) + \dots + \frac{1-\lambda}{n-1} \exp(a_n),$$

which then implies that

$$\exp(a_1 + a_2 + \cdots + a_n) \leq \lambda \exp\left(\frac{1}{\lambda}a_1\right) + \frac{1-\lambda}{n-1} \exp\left(\frac{n-1}{1-\lambda}a_2\right) + \cdots + \frac{1-\lambda}{n-1} \exp\left(\frac{n-1}{1-\lambda}a_n\right). \quad (112)$$

We now state Gronwall's inequality. For a proof, see for example [6], p. 250 and p. 262.

Proposition B.0.2. *Fix $T > 0$. Assume that f, g are Borel measurable, integrable functions defined on $[0, T]$ and taking values in \mathbb{R}_+ . Suppose that there is a $K > 0$ such that*

$$f(t) \leq g(t) + K \int_0^t f(s) ds, \quad \text{for all } t \in [0, T].$$

Then

$$f(t) \leq g(t) + K \int_0^t e^{K(t-s)} g(s) ds, \quad \text{for all } t \in [0, T].$$

If g is constant, then $f(t) \leq ge^{Kt}$.

References

- [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] Bakhtin, Y. and Mattingly, J. C. (2005). Stationary solutions of stochastic differential equation with memory and stochastic partial differential equations. *Commun. Contemp. Math.* **7** 553–582.
- [3] Billingsley, P. (1999). *Convergence of Probability Measures* (2nd Edition). J. Wiley & Sons, New York.
- [4] Bratsun, D., Volfson, D., Tsimring, L. S. and Hasty, J. (2005). Delay-induced stochastic oscillations in gene regulation. *PNAS*, **102** 14593–14598.
- [5] Brauer, F. and Castillo-Chávez C. (2001). *Mathematical Models in Population Biology and Epidemiology*. Springer, New York.
- [6] Chung, K. L. and Williams, R. J. (1990). *Introduction to Stochastic Integration* (2nd Edition). Birkhäuser, Boston.
- [7] Da Prato, G. and Zabczyk, J. (1996). *Ergodicity for Infinite Dimensional Systems*. Cambridge University Press, Cambridge, UK.
- [8] Deb, S. and Srikant, R. (2003). Global stability of congestion controllers for the Internet. *IEEE Transactions on Automatic Control*. **48** 1055–1060.
- [9] 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.
- [10] Filippov, A. F. (1988). *Differential Equations with Discontinuous Righthand Sides*. Kluwer Academic Publishers, Dordrecht.

- [11] Hairer, M. and Ohashi, A. (2007). Ergodic theory for SDEs with extrinsic memory. *Ann. Prob.* **35** 1950–1977.
- [12] Hairer, M., Mattingly, J. C., and Scheutzow, M. (2009). Asymptotic coupling and a general form of Harris’ theorem with applications to stochastic delay equations. *Probability Theory and Related Fields*, DOI 10.1007/s00440-009-0250-6.
- [13] Hale, J. K. and Lunel, S. M. (1993). *Introduction to Functional Differential Equations*. Applied Mathematical Sciences. Springer-Verlag, New York.
- [14] Harrison, J. M. and Reiman, M. I. (1981). Reflected Brownian motion on an orthant. *Ann. Prob.* **9** 302–308.
- [15] Itô, K. and Nisio, M. (1964). On stationary solutions of a stochastic differential equation. *J. Math. Kyoto Univ.* **4** 1–75.
- [16] Karatzas, I. and Shreve, S. (1998). *Brownian Motion and Stochastic Calculus* (2nd Edition). Springer, New York.
- [17] Kushner, H. J. (2008). *Numerical Methods for Controlled Stochastic Delay Systems*. Birkhäuser, Boston.
- [18] Liptser, R. S., and Shiriyayev A. N. (1977). *Statistics of Random Processes I General Theory*. Springer-Verlag, New York.
- [19] Mao, X. (1994). *Exponential Stability of Stochastic Differential Equations*. Marcel Dekker.
- [20] Mao, X. (1996). Razumikhin-type theorems on exponential stability of stochastic functional-differential equations. *Stochastic Processes and their Applications* **65** 233–250.
- [21] 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.
- [22] Mohammed, S-E. A. (1984). *Stochastic Functional Differential Equations*. Pitman, Boston.
- [23] Mohammed, S-E. A. (1986). Stability of linear delay equations under a small noise. *Proc. Edinburgh Math. Soc.* **29** 233–254.
- [24] Orey, S. (1971). *Lecture Notes on Limit Theorems for Markov Chain Transition Probabilities*. Van Nostrand Reinhold Company, London.
- [25] 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*, 185–190.
- [26] 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*, 3632–3637.
- [27] Papachristodoulou, A. (2004). Global stability analysis of a TCP/AQM protocol for arbitrary networks with delay. *43rd IEEE Conference on Decision and Control*, 1029–1034.

- [28] Papachristodoulou, 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*, 4684–4689.
- [29] 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** 553–559.
- [30] 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.
- [31] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion* (3rd Edition). Springer-Verlag, Berlin.
- [32] Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes and Martingales Volume 1: Foundations* (2nd Edition). Cambridge University Press, Cambridge, UK.
- [33] Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes and Martingales Volume 2: Itô Calculus* (2nd Edition). Cambridge University Press, Cambridge, UK.
- [34] Scheutzow, M. (1983). Qualitative behaviour of stochastic delay equations with a bounded memory. *Stochastics* **12** 41–80.
- [35] Scheutzow, M. (2000). Stability and instability of routings through a queueing network. *Queueing Systems Theory Appl.* **35** 117–128.
- [36] Scheutzow, M. (2004). Exponential growth rates for stochastic delay differential equations. *Stochastics and Dynamics* **5** 163–174.
- [37] Srikant, R. (2004). *The Mathematics of Internet Congestion Control*. Birkhäuser, Boston.
- [38] Vinnicombe, G. (2002). On the stability of networks operating TCP-like congestion control. *Proc. IFAC World Congress*, Barcelona, Spain.
- [39] Vinnicombe, G. (2002). On the stability of end-to-end congestion control for the Internet. *Univ. Cambridge Tech. Rep. CUED/F-INFENG/ TR.398 [Online]*. Available at <http://www.eng.cam.ac.uk/~gv>
- [40] Williams, R.J. (1998). An invariance principle for semimartingale reflecting Brownian motions in an orthant. *Queueing Systems Theory Appl.* **30** 5–25.
- [41] Ying, L., Dullerud, G., and Srikant, R. (2006). Global stability of Internet congestion controllers with heterogeneous delays. *IEEE/ACM Transactions on Networking* **14** 579–591.