

A Small Delay and Correlation Time Limit of Stochastic Differential Delay Equations with State-Dependent Colored Noise

Scott Hottovy¹, Austin McDaniel^{2,3}, and Jan Wehr²

¹Department of Mathematics, United States Naval Academy,
Annapolis, MD, 21401, USA

²Department of Mathematics, University of Arizona, Tucson, AZ
85721, USA

³School of Mathematical and Statistical Sciences, Arizona State
University, Tempe, AZ 85287, USA

Abstract

We consider a general stochastic differential delay equation (SDDE) with state-dependent colored noises and derive its limit as the time delays and the correlation times of the noises go to zero. The work is motivated by an experiment involving an electrical circuit with noisy, delayed feedback. An Ornstein-Uhlenbeck process is used to model the colored noise. The main methods used in the proof are a theorem about convergence of solutions of stochastic differential equations by Kurtz and Protter and a maximal inequality for sums of a stationary sequence of random variables by Peligrad and Utev.

1 Introduction

Stochastic differential equations (SDEs) are frequently used to describe the dynamics of physical and biological systems [23]. However, in situations where a system's response to stimuli is delayed, stochastic differential delay equations (SDDEs) provide more accurate models. This gain in accuracy comes at a price of greater mathematical difficulty because the theory of SDDEs is much less developed than the theory of SDEs. This is due in a large part to the fact that SDDEs are by nature non-Markovian systems, and so tools that follow from the Markov property, such as the generator and the Kolmogorov equations, do not apply to SDDEs. Thus, it is useful to develop approximations of SDDEs that are easier to work with than the original equations but still account for the effects of the delay(s). Such approximations have been used recently to show a

phase transition in the collective behavior of robots with sensorial delays [20] and a crossover from the Itô to the Stratonovich equation in the dynamics of an electrical circuit [27] (see [32] for a review).

Stochastic differential delay equations have been used to model, for example, population dynamics, neural networks [17], and genetic regulatory networks [30]. An approximation of SDDEs that is commonly made when the time delays are small is that obtained by Taylor expanding to first order in the time delays (see, e.g., [5, 16]). Another important application where SDDE models are used is semiconductor lasers with optical feedback. Light that is backscattered by different sources in front of a laser may reenter the laser cavity and produce a significant amount of feedback. This optical feedback is delayed due to the time that passes between when the reflected light exits the laser cavity and when it reenters. For semiconductor lasers, this delay time can be as large or larger than other important time scales of the system, and the delayed optical feedback can have dramatic effects on the laser dynamics [6, 14, 21, 31]. SDDE models then arise when the intrinsic noise in semiconductor lasers is taken into account (see, e.g., [18] and references therein).

In this article we consider a general SDDE system, driven by colored (i.e., temporally correlated) noise, which was motivated by an experiment involving an electrical circuit with a delayed feedback mechanism [27]. In the experiment the voltage is driven by a colored noise process with a rapidly decaying correlation function. An SDE approximation of a general SDDE that extends the model of this circuit was derived by first performing a Taylor expansion to first order in the time delays and then taking the limit in which the time delays and the correlation times of the noises go to zero at the same rate [27]. This limiting SDE contains *noise-induced drift* terms which depend on the ratios of the time delays to the noise correlation times. Convergence of the solution of the equation obtained by Taylor expansion to the solution of the limiting equation was proven later in [19].

The present article improves upon the approximation derived in [19, 27]. We study the same limit of the SDDE system but without first performing a Taylor approximation, and thereby get a more accurate result. Because we do not perform a Taylor expansion, we are able to use a simpler (less smooth) model of the colored noise process than the one used in [19, 27]; here we model the colored noise as a stationary Ornstein-Uhlenbeck process [4]. A key difference between this article and the previous works [19, 27] is that the latter study the Markovian system obtained by Taylor expansion whereas here we prove a theorem concerning the convergence of the solution of the non-Markovian SDDE system to the Markov process that satisfies the limiting SDE (the SDE system studied in [19, 27] can be cast as a Markovian system by treating the colored noise process as a state variable). We believe that the theorem we prove in this article is somewhat unique in the sense that it provides both a rigorous analytical result for a non-Markovian system as well as an effective approximation that can be directly used in applications. Our improvement on the results of [19, 27] can be used in practice as an SDE approximation of a general class of systems with delay.

We consider the multidimensional SDDE system

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t)dt + \boldsymbol{\sigma}(\mathbf{x}_{t-\boldsymbol{\delta}})\boldsymbol{\xi}_t dt \quad (1)$$

where $\mathbf{x}_t \in \mathbb{R}^m$ is the state vector, $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is a vector-valued function describing the deterministic part of the dynamical system, $\boldsymbol{\xi}$ is a vector, taking values in \mathbb{R}^n , of mean-zero, stationary, independent noise processes, $\boldsymbol{\sigma} : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times n}$ is a matrix-valued function, and $\mathbf{x}_{t-\boldsymbol{\delta}} = ((x_{t-\delta_1})_1, \dots, (x_{t-\delta_i})_i, \dots, (x_{t-\delta_m})_m)^\top$ (where \top denotes transpose) is the delayed state vector. Note that each component is delayed by a possibly different amount $\delta_i \geq 0$. We take each of the n independent noises ξ_j to be colored so that each has a correlation time $\tau_j > 0$. That is, $E[(\xi_t)_j(\xi_s)_j] = g(|t-s|\tau_j^{-1})$ where g is a function that decays quickly as its argument increases (for $i \neq j$, $E[(\xi_t)_i(\xi_s)_j] = 0$ by independence).

In the main theorem of this article (Theorem 1.1), we model the noise by taking the components of $\boldsymbol{\xi}$ to be independent stationary Ornstein-Uhlenbeck processes with correlation times τ_j . That is, we define $\xi_j = y_j$ where y_j is the stationary solution of

$$d(y_t)_j = -\frac{1}{\tau_j}(y_t)_j dt + \frac{1}{\tau_j}d(W_t)_j \quad (2)$$

where $\tau_j > 0$ and \mathbf{W} is an n -dimensional Wiener process. Equation (2) has a unique stationary measure and the solution of (2) with the initial condition distributed according to this stationary measure defines the stationary Ornstein-Uhlenbeck process that we use to model the colored noise in the SDDE system (1). We note that as $\tau_j \rightarrow 0$ for all $j = 1, \dots, n$, \mathbf{y} converges to an n -dimensional white noise (i.e., its correlation function converges to a delta function).

We study the limit of the system consisting of equations (1) and (2), with $\boldsymbol{\xi}_t = \mathbf{y}_t$, assuming that all δ_i and τ_j stay proportional to a single characteristic time $\epsilon > 0$ which goes to 0. Thus, we let $\delta_i = c_i \epsilon \rightarrow 0$ and $\tau_j = k_j \epsilon \rightarrow 0$ where $c_i, k_j > 0$ remain constant for all i, j as $\epsilon \rightarrow 0$.

We consider the solution to equations (1) and (2) on a bounded time interval $0 \leq t \leq T$. We let (Ω, \mathcal{F}, P) denote the underlying probability space. We will use the notation $\delta^* = \max_{1 \leq i \leq m} \delta_i$ and $c^* = \max_{1 \leq i \leq m} c_i = \delta^*/\epsilon$. In order to formulate a well-posed problem, because of the delays in (1), one needs to specify not only an initial condition but also the values of the process \mathbf{x} at all past times $t \in [-\delta^*, 0]$. Therefore, we assume that there is a $t_- < 0$ such that the values of \mathbf{x} are initially specified for $t \in [t_-, 0]$ and we only consider delays δ_i such that $\delta_i < |t_-|$ for all i . Let $\mathbf{x}^- : \Omega \times [t_-, 0] \rightarrow \mathbb{R}^m$ denote this *past condition* associated with (1). We assume that \mathbf{x}^- is independent of \mathbf{W}_t for all $t \geq 0$. We further assume that \mathbf{x}^- is defined so that there exists a unique solution to (1) with the past condition \mathbf{x}^- .

Theorem 1.1. *Suppose that \mathbf{f} is continuous and bounded and that $\boldsymbol{\sigma}$ is bounded with bounded, Lipschitz continuous first derivatives. Let $(\mathbf{x}^\epsilon, \mathbf{y}^\epsilon)$ solve equations*

(1) and (2) (which depend on ϵ through δ_i, τ_j) on $0 \leq t \leq T$ with the past condition \mathbf{x}^- the same for every ϵ and the initial condition \mathbf{y}_0^ϵ distributed according to the stationary distribution corresponding to (2). Let \mathbf{x} solve

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t)dt + \mathbf{S}(\mathbf{x}_t) dt + \boldsymbol{\sigma}(\mathbf{x}_t)d\mathbf{W}_t \quad (3)$$

on $0 \leq t \leq T$ with the same initial condition $\mathbf{x}_0 = \mathbf{x}_0^-$, where \mathbf{S} is defined componentwise as

$$S_i(\mathbf{x}) = \sum_{p,j} \frac{1}{2} e^{-\frac{\delta_p}{\tau_j}} \sigma_{pj}(\mathbf{x}) \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}) \quad (4)$$

and suppose strong uniqueness holds on $0 \leq t \leq T$ for (3) with the initial condition \mathbf{x}_0 . Then

$$\lim_{\epsilon \rightarrow 0} P \left[\sup_{0 \leq t \leq T} \|\mathbf{x}_t^\epsilon - \mathbf{x}_t\| > a \right] = 0 \quad (5)$$

for every $a > 0$.

An outline of the paper is as follows. In Section 2 we discuss the important aspects of modeling colored noise. In Section 3 we prove the main theorem of the paper. In Section 4 we discuss how the theorem relates to the previous works [19, 27]. We give conclusions in Section 5.

2 Colored noise process

In this section we discuss the model of colored noise that we use in this paper, i.e., the Ornstein-Uhlenbeck process. The noise process $\boldsymbol{\xi}$ driving the system (1) is colored, not white. The terms ‘‘colored’’ and ‘‘white’’ come from the Wiener-Khinchin theorem (see [4, Section 1.5.2]). This theorem states that the expected value of the modulus of the Fourier transform of the time series of a stationary noise process is equal to the Fourier transform of its time correlation function. Thus, the correlation function of a white noise process has constant power across all frequencies in the Fourier domain (much like white light contains all colors of the light spectrum in equal proportions). A colored noise is any (usually mean-zero) process ξ whose correlation function has a nonconstant spectral density. Any physical noise has a nonzero correlation time, but this correlation time is often small compared to other characteristic time scales in the physical system. Because of this it is often desirable to use as a model a stationary colored noise process ξ which has a correlation function of the form

$$E[\xi_t \xi_s] = \frac{1}{\epsilon} g\left(\frac{|t-s|}{\epsilon}\right)$$

where g is a function that decays rapidly as its argument increases. As $\epsilon \rightarrow 0$ this correlation function converges to the correlation function of white noise, i.e.,

$$E[\xi_t \xi_s] \rightarrow \delta(t-s).$$

Furthermore, it is typical to use as a model for colored noise a process ξ such that as $\epsilon \rightarrow 0$,

$$\int_0^t \xi_s ds \rightarrow W_t$$

in some probabilistic sense, where W is a Wiener process.

Along with the correlation function $g(r)$, $r \geq 0$, the smoothness of the noise process is another important consideration when modeling physical noise. For a mean-zero Gaussian process, the covariance function determines the entire law of the process, and hence also the smoothness of its realizations. Different processes with varying degrees of smoothness have been used for modeling colored noise in the literature. An infinitely differentiable process was used in [3] by convoluting a smooth function $h \in C^\infty([0, \infty))$ with a Wiener process. In [13] a piecewise differentiable approximation of a Wiener process was constructed and then differentiated to obtain a colored noise. In [19, 27] a differentiable harmonic noise process was used in order to make the solution twice differentiable and thus allow to use its Taylor expansion.

In this paper we use a stationary Ornstein-Uhlenbeck (OU) process to model the colored noise. Its key properties of being a stationary, Gaussian process with a rapidly decaying correlation function make this process a good model of physical noise in many situations. Physical noise is often Gaussian (to a very good approximation) because of the central limit theorem. The OU process is a widely used model (see, e.g., [7, 12, 24, 25]) that also has the advantage of being a continuous process that can be expressed as the solution to a linear SDE. That is, a stationary OU process y with paths in $C([0, T], \mathbb{R})$ is defined as the solution to the SDE

$$dy_t = -\alpha y_t dt + \sigma dW_t, \quad y_0 = w \quad (6)$$

where $\alpha, \sigma > 0$ and w is distributed according to the stationary measure. This process y is a mean-zero Gaussian process with covariance function [4, Section 4.5.4]

$$E[y_t y_s] = \frac{\sigma^2}{2\alpha} e^{-\alpha|t-s|}. \quad (7)$$

In Section 1, we defined the colored noises in (1) to be n independent, stationary OU processes with correlation times $\tau_j = k_j \epsilon$. In other words, $(\xi_t)_j = (y_t^\epsilon)_j$ is the solution of (6) with $\alpha = \sigma = \tau_j^{-1}$, $W = W_j$, and $(y_0^\epsilon)_j = w_j$, where w_j is a random variable, independent of the process W_j , having the stationary distribution. That is, w_j is a Gaussian random variable with mean zero and variance $1/(2k_j \epsilon)$. Another reason the OU process is an advantageous model is because equation (6) is exactly solvable and every moment of \mathbf{y}_t^ϵ can be calculated. Furthermore, the covariance moments $E[(y_t^\epsilon)_j (y_s^\epsilon)_j]^n$ can be calculated for all $n \in \mathbb{N}$ by using Wick's theorem [9, Theorem 1.3.8].

In Figure 1, OU noise is compared to white noise. White noise (black) is generated by taking the differential of a standard Wiener process. Ornstein-Uhlenbeck noise (dark gray) is generated by using the SDE (2) with $\tau = 5$ (arbitrary time units). In panel a) the average power as a function of frequency

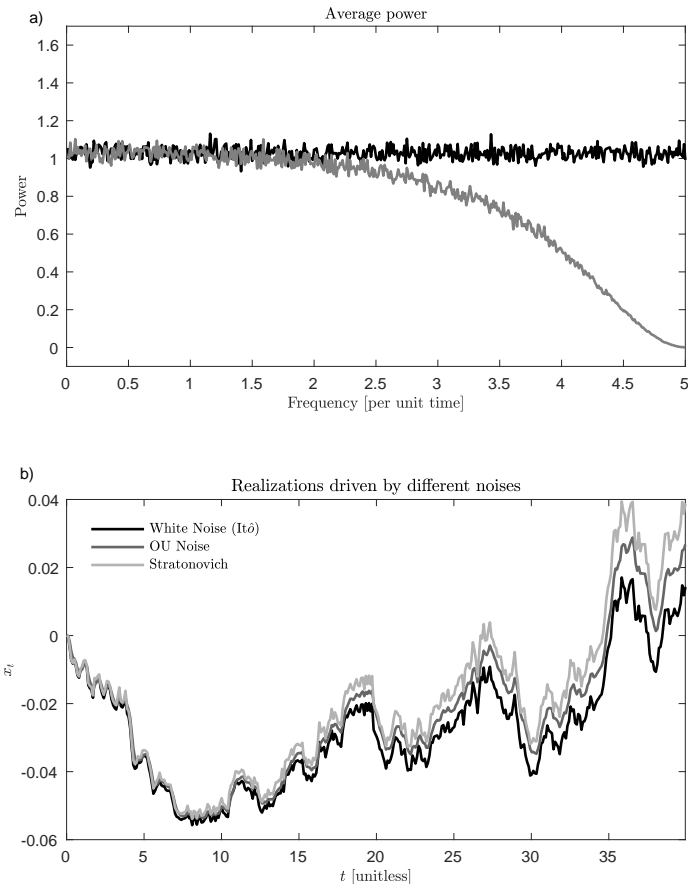


Figure 1: Comparison of white and Ornstein-Uhlenbeck noises. Plot a) is the average power, or modulus of the Fourier time series, of each noise. Plot b) is a realization of the solution to system (1) with $\delta = 0$, $f(x) = ax$, $\sigma(x) = (bx + c)$, and the initial condition $x_0 = 0$, where ξ is each type of noise. The simulation is carried out using the Euler-Maruyama method [11] with step size $\Delta t = 0.001$. Here, the same realization of the same Wiener process is used to generate both the white noise and the Ornstein-Uhlenbeck noise.

is plotted for each process. This is the Fourier transform pair of the correlation function. For the white noise process the average power is constant as a function of frequency, while the average power of the OU process decays rapidly after frequency 2. In panel b), a realization is shown of a one-dimensional example of system (1), with no delay, driven by the noises. Notice that the process driven by the OU noise and that driven by the white noise diverge from each other after time $t = 20$. This is a result of the Stratonovich correction that appears in (3) when one sets the delay equal to zero.

Theorem 1.1, as stated, only applies to the case where the noise is modeled by a stationary Ornstein-Uhlenbeck process. This is a Markov process that satisfies a simple stochastic differential equation. The theorem can be modified to handle any noise process that is stationary and that solves a linear SDE with additive (white) noise (e.g. the harmonic noise process used in [19, 27]). In addition, a wide class of Gaussian noises can be represented as linear transformations of multidimensional Ornstein-Uhlenbeck processes. More precisely, every stationary, mean-zero Gaussian process with a covariance function which is a Bohl function (i.e., a linear combination of products of exponentials, polynomials, and trigonometric functions) can be obtained by applying a constant linear functional to the stationary solution of an equation of the form

$$d\boldsymbol{\eta}_t = \mathbf{A}\boldsymbol{\eta}_t dt + \mathbf{B}d\mathbf{W}_t$$

where $\boldsymbol{\eta}$ is a process taking values in \mathbb{R}^n , \mathbf{W} is an m -dimensional Wiener process, and \mathbf{A} and \mathbf{B} are constant matrices of appropriate dimensions. This construction, known in stochastic control theory as *stochastic realization* [10], can be thought of as exhibiting a hidden Markov nature of a class of Gaussian processes. It has been used recently to obtain small mass limits of SDE systems driven by a class of colored noises [15] and it is very likely that by applying it, one would be able to generalize the results obtained here to cases where the noise is modeled by a broad class of non-Markov Gaussian processes. We conclude this discussion with the important remark that for different noises, the coefficients of the additional drift in the analogous limiting equation depend on the covariance function of the noise.

3 Proof of Theorem 1.1

In this section we prove the main theorem of the paper, Theorem 1.1. The main tool that we use is a theory of convergence of solutions of SDEs by Kurtz and Protter [13]. A similar technique is used in [8] and in [19]. The structure of the section is as follows. In Section 3.1 we introduce the theory of convergence of solutions of SDEs that we will use to prove the theorem. In Section 3.2 we use integration by parts and substitution to write system (1) in the form necessary for applying the Kurtz-Protter theorem. In Section 3.3 we complete the proof of Theorem 1.1 by verifying that the conditions of the Kurtz-Protter theorem are satisfied.

We begin with the theory of convergence of solutions of SDEs, where we state (in a less general but sufficient form) a theorem of Kurtz and Protter [13].

3.1 Convergence of solutions of SDEs

We fix a probability space (Ω, \mathcal{F}, P) and an n -dimensional Wiener process \mathbf{W} on it. Let \mathcal{F}_t will be (the usual augmentation of) $\sigma(\{\mathbf{W}_s : s \leq t\})$, the filtration generated by W up to time t [29].

Suppose \mathbf{H} is an $\{\mathcal{F}_t\}$ -adapted semimartingale with paths in $C([0, T], \mathbb{R}^k)$, whose Doob-Meyer decomposition is $\mathbf{H}_t = \mathbf{M}_t + \mathbf{A}_t$ so that \mathbf{M} is an \mathcal{F}_t -local martingale and \mathbf{A} is a process of locally bounded variation, such that $\mathbf{A}_0 = 0$ [29]. For a continuous $\{\mathcal{F}_t\}$ -adapted process \mathbf{Y} with paths in $C([0, T], \mathbb{R}^{d \times k})$ and for $t \leq T$ consider the Itô integral

$$\int_0^t \mathbf{Y}_s d\mathbf{H}_s = \lim \sum_i \mathbf{Y}_{t_i} (\mathbf{H}_{t_{i+1}} - \mathbf{H}_{t_i}) , \quad (8)$$

where $\{t_i\}$ is a partition of $[0, t]$ and the limit is taken as the maximum of $t_{i+1} - t_i$ goes to zero. For a continuous processes \mathbf{Y} such that

$$P \left[\int_0^T \|\mathbf{Y}_s\|^2 d\langle \mathbf{M} \rangle_s + \int_0^T \|\mathbf{Y}_s\| dV_s(\mathbf{A}) < \infty \right] = 1 ,$$

where $\langle \mathbf{M} \rangle_s$ is the quadratic variation of \mathbf{M} and $V_s(\mathbf{A})$ is the total variation of \mathbf{A} , the limit in equation (8) exists in the sense that

$$\sup_{0 \leq t \leq T} \left\| \int_0^t \mathbf{Y}_s d\mathbf{H}_s - \sum_i \mathbf{Y}_{t_i} (\mathbf{H}_{t_{i+1}} - \mathbf{H}_{t_i}) \right\| \rightarrow 0$$

in probability. This and related convergence modes will be used throughout the paper [28].

Consider $(\mathbf{U}^\epsilon, \mathbf{H}^\epsilon)$ with paths in $C([0, T], \mathbb{R}^m \times \mathbb{R}^k)$ adapted to $\{\mathcal{F}_t\}$ where \mathbf{H}^ϵ is a semimartingale with respect to \mathcal{F}_t . Let $\mathbf{H}_t^\epsilon = \mathbf{M}_t^\epsilon + \mathbf{A}_t^\epsilon$ be its Doob-Meyer decomposition. Let $\mathbf{h} : \mathbb{R}^k \rightarrow \mathbb{R}^{k \times n}$ be a matrix-valued function and let \mathbf{X}^ϵ , with paths in $C([0, T], \mathbb{R}^m)$, satisfy the SDE

$$\mathbf{X}_t^\epsilon = \mathbf{X}_0 + \mathbf{U}_t^\epsilon + \int_0^t \mathbf{h}(\mathbf{X}_s^\epsilon) d\mathbf{H}_s^\epsilon , \quad (9)$$

where $\mathbf{X}_0^\epsilon = \mathbf{X}_0 \in \mathbb{R}^m$ is the same initial condition for all ϵ . Define \mathbf{X} , with paths in $C([0, T], \mathbb{R}^m)$, to be the solution of

$$\mathbf{X}_t = \mathbf{X}_0 + \int_0^t \mathbf{h}(\mathbf{X}_s) d\mathbf{H}_s . \quad (10)$$

Note that (9) implies $\mathbf{U}_0^\epsilon = \mathbf{0}$ for all ϵ .

Lemma 3.1. [13, Theorem 5.4 and Corollary 5.6] Suppose $(\mathbf{U}^\epsilon, \mathbf{H}^\epsilon) \rightarrow (\mathbf{0}, \mathbf{H})$ in probability with respect to $C([0, T], \mathbb{R}^m \times \mathbb{R}^k)$, i.e., for all $a > 0$,

$$P \left[\sup_{0 \leq s \leq T} \left(\|\mathbf{U}_s^\epsilon\| + \|\mathbf{H}_s^\epsilon - \mathbf{H}_s\| \right) > a \right] \rightarrow 0 \quad (11)$$

as $\epsilon \rightarrow 0$, and the following conditions are satisfied:

Condition 1. [Tightness condition] The family of total variations evaluated at T , $\{V_T(\mathbf{A}^\epsilon)\}$, is stochastically bounded, i.e., $P[V_T(\mathbf{A}^\epsilon) > L] \rightarrow 0$ as $L \rightarrow \infty$, uniformly in ϵ .

Condition 2. \mathbf{h} is continuous (see [13, Example 5.3]).

Suppose that there exists a strongly unique global solution to equation (10). Then, as $\epsilon \rightarrow 0$, \mathbf{X}^ϵ converges to \mathbf{X} , the solution of equation (10), in probability with respect to $C([0, T], \mathbb{R}^m)$.

3.2 Derivation of the limiting equation

Proof of Theorem 1.1. To write system (1) (with $\boldsymbol{\xi}_t = \mathbf{y}_t^\epsilon$) in the form of (9), we will use integration by parts and substitution. To this end, we write equation (2) in matrix form. Recalling that $\tau_j = k_j \epsilon$ and defining

$$\mathbf{D} = \begin{bmatrix} \frac{1}{k_1} & 0 & \dots & 0 \\ 0 & \frac{1}{k_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{k_n} \end{bmatrix},$$

equation (2) becomes

$$d\mathbf{y}_t^\epsilon = \frac{\mathbf{D}}{\epsilon} \left(-\mathbf{y}_t^\epsilon dt + d\mathbf{W}_t \right). \quad (12)$$

We solve equation (12) for $\mathbf{y}_t^\epsilon dt$ and substitute it into equation (1) to obtain

$$d\mathbf{x}_t^\epsilon = \mathbf{f}(\mathbf{x}_t^\epsilon) dt + \boldsymbol{\sigma}(\mathbf{x}_{t-c\epsilon}^\epsilon) d\mathbf{W}_t - \boldsymbol{\sigma}(\mathbf{x}_{t-c\epsilon}^\epsilon) \epsilon \mathbf{D}^{-1} d\mathbf{y}_t^\epsilon$$

where $\mathbf{x}_{t-c\epsilon}^\epsilon = ((x_{t-c_1\epsilon}^\epsilon)_1, \dots, (x_{t-c_i\epsilon}^\epsilon)_i, \dots, (x_{t-c_m\epsilon}^\epsilon)_m)^T$. In integral form, this equation becomes

$$\mathbf{x}_t^\epsilon = \mathbf{x}_0 + \int_0^t \mathbf{f}(\mathbf{x}_s^\epsilon) ds + \int_0^t \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) d\mathbf{W}_s - \int_0^t \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) \epsilon \mathbf{D}^{-1} d\mathbf{y}_s^\epsilon. \quad (13)$$

In the limit as $\epsilon \rightarrow 0$, we expect the second and the third terms on the right-hand side of equation (13) to converge to the analogous terms of the limiting equation (3). In addition, $\epsilon \mathbf{y}^\epsilon$ goes to zero as $\epsilon \rightarrow 0$, as we will show later (see Lemma 3.2). Because of this, one might expect the last term of the right-hand side to converge to zero as well. This is the case when $\boldsymbol{\sigma}(\mathbf{x}) = \boldsymbol{\sigma}$ is a constant function. For non-constant $\boldsymbol{\sigma}(\mathbf{x})$, in order to be able to apply Lemma 3.1 directly the process $\epsilon \mathbf{y}^\epsilon$ would have to satisfy Condition 1. This is not the case, nor is it true for any family of colored noise processes which converge to white noise (see [13]).

To resolve this issue, we first split the last integral into a part that involves values of the past condition and a part that does not. We then integrate by parts the i^{th} component of the latter integral to obtain

$$\begin{aligned}
(x_t^\epsilon)_i &= (x_0)_i + \int_0^t f_i(\mathbf{x}_s^\epsilon) ds + \left(\int_0^t \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) d\mathbf{W}_s \right)_i \\
&\quad - \left(\int_0^{c^*\epsilon} \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) \epsilon \mathbf{D}^{-1} d\mathbf{y}_s^\epsilon \right)_i - \sum_j \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon) k_j \epsilon (y_t^\epsilon)_j \\
&\quad + \sum_j \sigma_{ij}(\mathbf{x}_{c^*\epsilon-c\epsilon}^\epsilon) k_j \epsilon (y_{c^*\epsilon}^\epsilon)_j + \int_{c^*\epsilon}^t \sum_{p,j} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j d(x_{s-c_p\epsilon}^\epsilon)_p.
\end{aligned}$$

Recall that $c^*\epsilon$ denotes the largest delay. We note that the last term is a Lebesgue integral since $d(x_{s-c_p\epsilon}^\epsilon)_p$ is the differential of a continuous process with bounded variation. We substitute equation (1) into the term containing $d(x_{s-c_p\epsilon}^\epsilon)_p$ to obtain

$$\begin{aligned}
(x_t^\epsilon)_i &= (x_0)_i + \int_0^t f_i(\mathbf{x}_s^\epsilon) ds + \left(\int_0^t \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) d\mathbf{W}_s \right)_i \\
&\quad + \int_{c^*\epsilon}^t \sum_{p,j} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds \\
&\quad + \int_{c^*\epsilon}^t \sum_{p,j,\ell} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) (y_{s-c_p\epsilon}^\epsilon)_\ell ds \\
&\quad - \left(\int_0^{c^*\epsilon} \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) \epsilon \mathbf{D}^{-1} d\mathbf{y}_s^\epsilon \right)_i \\
&\quad - \sum_j \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon) k_j \epsilon (y_t^\epsilon)_j + \sum_j \sigma_{ij}(\mathbf{x}_{c^*\epsilon-c\epsilon}^\epsilon) k_j \epsilon (y_{c^*\epsilon}^\epsilon)_j \tag{14}
\end{aligned}$$

where $\mathbf{x}_{s-c_p\epsilon}^\epsilon = ((x_{s-c_p\epsilon}^\epsilon)_1, \dots, (x_{s-c_p\epsilon}^\epsilon)_i, \dots, (x_{s-c_p\epsilon}^\epsilon)_m)^\top$ and $\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon = ((x_{s-c_p\epsilon-c\epsilon}^\epsilon)_1, \dots, (x_{s-c_p\epsilon-c\epsilon}^\epsilon)_i, \dots, (x_{s-c_p\epsilon-c\epsilon}^\epsilon)_m)^\top$.

In the above equation, we will show that the boundary terms and the second Lebesgue integral (the one whose integrand contains the factor $f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon)$) go to zero because $\epsilon \mathbf{y}^\epsilon$ goes to zero. To prove convergence, we add and subtract like terms without delays on the right-hand side of the above equation. In addition, we add and subtract the integral of $S_i(\mathbf{x})$, where $S_i(\mathbf{x})$ is defined in equation (4). In the resulting equation for \mathbf{x}^ϵ , we keep with the notation of Lemma 3.1 by collecting the terms which we will show directly go to zero into a new process called \mathbf{U}^ϵ . At this point, it is convenient to extend the process \mathbf{y}^ϵ so that \mathbf{y}_s^ϵ is defined for $s \geq t_-$, where $t_- < 0$ was introduced before the statement of Theorem 1.1. We do this in such a way so that \mathbf{y}^ϵ considered on the interval $s \geq t_-$ is a stationary process.

Let \mathbf{U}^ϵ be defined componentwise as

$$\begin{aligned}
(\mathbf{U}_t^\epsilon)_i &= - \sum_j \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon) k_j \epsilon (y_t^\epsilon)_j + \sum_j \sigma_{ij}(\mathbf{x}_{c^*\epsilon-c\epsilon}^\epsilon) k_j \epsilon (y_{c^*\epsilon}^\epsilon)_j \\
&\quad - \left(\int_0^{c^*\epsilon} \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) \epsilon \mathbf{D}^{-1} d\mathbf{y}_s^\epsilon \right)_i \\
&\quad + \left(\int_0^t (\boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \boldsymbol{\sigma}(\mathbf{x}_s^\epsilon)) d\mathbf{W}_s \right)_i \\
&\quad + \int_{c^*\epsilon}^t \sum_{p,j} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds \\
&\quad + \int_{c^*\epsilon}^t \sum_{p,j,\ell} \left[\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) \right. \\
&\quad \quad \quad \left. - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds \\
&\quad - \int_0^{c^*\epsilon} \sum_{p,j,\ell} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds.
\end{aligned} \tag{15}$$

Then \mathbf{x}_t^ϵ can be written componentwise as

$$\begin{aligned}
(x_t^\epsilon)_i &= (x_0)_i + (\mathbf{U}_t^\epsilon)_i + \int_0^t f_i(\mathbf{x}_s^\epsilon) ds + \left(\int_0^t \boldsymbol{\sigma}(\mathbf{x}_s^\epsilon) d\mathbf{W}_s \right)_i \\
&\quad + \int_0^t \sum_{p,j} \frac{1}{2} e^{-\frac{c_p}{k_j}} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{pj}(\mathbf{x}_s^\epsilon) ds \\
&\quad + \int_0^t \sum_{p,j} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{pj}(\mathbf{x}_s^\epsilon) \left(k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_j - \frac{1}{2} e^{-\frac{c_p}{k_j}} \right) ds \\
&\quad + \int_0^t \sum_{\substack{p,j,\ell \\ j \neq \ell}} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds.
\end{aligned} \tag{16}$$

We now write equation (16) in the form (9) of Lemma 3.1 by letting $\mathbf{h} : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times (1+n+1+mn^2)}$ be the matrix-valued function given by

$$\begin{aligned}
\mathbf{h}(\mathbf{X}) &= \left(\mathbf{f}(\mathbf{X}), \boldsymbol{\sigma}(\mathbf{X}), \mathbf{S}(\mathbf{X}), \boldsymbol{\Lambda}^{11}(\mathbf{X}), \dots, \boldsymbol{\Lambda}^{1n}(\mathbf{X}), \right. \\
&\quad \left. \boldsymbol{\Lambda}^{21}(\mathbf{X}), \dots, \boldsymbol{\Lambda}^{2n}(\mathbf{X}), \boldsymbol{\Lambda}^{n1}(\mathbf{X}), \dots, \boldsymbol{\Lambda}^{nn}(\mathbf{X}) \right)
\end{aligned} \tag{17}$$

where $\mathbf{S} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is the vector-valued function defined in equation (4) and $\boldsymbol{\Lambda}^{j\ell} : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times m}$ is defined componentwise as

$$\Lambda_{ip}^{j\ell}(\mathbf{X}) = \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{X}) \sigma_{p\ell}(\mathbf{X})$$

and by letting \mathbf{H}^ϵ be the process, with paths in $C([0, T], \mathbb{R}^{(1+n+1+mn^2)})$, given by

$$\mathbf{H}_t^\epsilon = \begin{bmatrix} t \\ \mathbf{W}_t \\ t \\ \mathbf{G}_t^{11} \\ \vdots \\ \mathbf{G}_t^{nn} \end{bmatrix}, \quad (18)$$

where $\mathbf{G}^{j\ell}$ is the process, with paths in $C([0, T], \mathbb{R}^m)$, given by

$$(\mathbf{G}_t^{j\ell})_p = \int_0^t \left(k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p \epsilon}^\epsilon)_\ell - E \left[k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p \epsilon}^\epsilon)_\ell \right] \right) ds. \quad (19)$$

Note that the expectation in the integrand in (19) is equal to $\frac{1}{2} e^{-c_p/k_j}$ for $j = \ell$ and equal to zero for $j \neq \ell$.

We will show that the processes $\mathbf{G}^{j\ell}$ converge to zero in L^2 with respect to $C([0, T], \mathbb{R}^m)$ for all $j, \ell = 1, \dots, n$ (Lemma 3.5). Thus, the limiting process \mathbf{H} is given by

$$\mathbf{H}_t = \begin{bmatrix} t \\ \mathbf{W}_t \\ t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}. \quad (20)$$

We show in the next subsection that $\mathbf{U}^\epsilon, \mathbf{H}^\epsilon, \mathbf{H}$, and \mathbf{h} satisfy the assumptions of Lemma 3.1. Thus, given the definitions of \mathbf{h} and \mathbf{H} above, by computing the right-hand side of

$$\mathbf{x}_t = \mathbf{x}_0 + \int_0^t \mathbf{h}(\mathbf{x}_s) d\mathbf{H}_s \quad (21)$$

we get the limiting SDE (3).

3.3 Verifying Conditions of Lemma 3.1

In this subsection, we complete the proof of Theorem 1.1 by showing that the conditions of Lemma 3.1 are satisfied. That is, we show $\mathbf{U}^\epsilon \rightarrow \mathbf{0}$ in probability with respect to $C([0, T], \mathbb{R}^m)$, $\mathbf{H}^\epsilon \rightarrow \mathbf{H}$ in probability with respect to $C([0, T], \mathbb{R}^{(1+n+1+mn^2)})$, and Conditions 1 and 2 are satisfied. We begin with lemmas that we will need. First we show that $\epsilon \mathbf{y}^\epsilon \rightarrow \mathbf{0}$ in L^2 with respect to $C([0, T], \mathbb{R}^n)$. Nelson showed a similar result, namely that $\sup_{0 \leq t \leq T} \|\epsilon \mathbf{y}_t^\epsilon\| \rightarrow 0$ as $\epsilon \rightarrow 0$ with probability one [22].

Lemma 3.2. *For each $\epsilon > 0$, let \mathbf{y}^ϵ be the solution to equation (2) with \mathbf{y}_0^ϵ distributed according to the stationary distribution corresponding to (2). There*

exists $C' > 0$ independent of ϵ such that

$$E \left[\left(\sup_{0 \leq t \leq T} \left| \int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} d(W_s)_j \right| \right)^2 \right] \leq C' \epsilon^{1/2}, \quad (22)$$

and this implies that there exists $C > 0$ independent of ϵ such that

$$E \left[\sup_{0 \leq t \leq T} \|\epsilon \mathbf{y}_t^\epsilon\|^2 \right] \leq C \epsilon^{1/2}. \quad (23)$$

Thus, as $\epsilon \rightarrow 0$, $\epsilon \mathbf{y}^\epsilon \rightarrow \mathbf{0}$ in L^2 , and therefore in probability, with respect to $C([0, T], \mathbb{R}^n)$.

Proof. We begin by proving (22) by following the first part of the argument of [24, Lemma 3.7]. We fix $\alpha \in (0, \frac{1}{2})$ and use the factorization method from [2] (see also [1, Section 5.3]) to rewrite

$$\begin{aligned} I(t) &= \int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} d(W_s)_j \\ &= \frac{\sin(\pi\alpha)}{\pi} \int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} (t-s)^{\alpha-1} Y(s) ds, \end{aligned}$$

where

$$Y(s) = \int_0^s e^{-\frac{(s-u)}{k_j \epsilon}} (s-u)^{-\alpha} d(W_u)_j$$

and we used the identity

$$\int_u^t (t-s)^{\alpha-1} (s-u)^{-\alpha} ds = \frac{\pi}{\sin(\pi\alpha)}, \quad 0 < \alpha < 1.$$

We fix $m > \frac{1}{2\alpha}$ and use the Hölder inequality:

$$|I(t)|^{2m} \leq C_1 \left(\int_0^t \left| e^{-\frac{(t-s)}{k_j \epsilon}} (t-s)^{\alpha-1} \right|^{\frac{2m}{2m-1}} ds \right)^{2m-1} \int_0^t |Y(s)|^{2m} ds.$$

Using the change of variables $z = \frac{2m}{2m-1} \frac{(t-s)}{k_j \epsilon}$, we have

$$\begin{aligned} & \int_0^t e^{-\frac{2m}{2m-1} \frac{(t-s)}{k_j \epsilon}} (t-s)^{\frac{2m}{2m-1}(\alpha-1)} ds \\ &= \left(\frac{k_j(2m-1)}{2m} \right)^{\frac{2m\alpha-1}{2m-1}} \epsilon^{\frac{2m\alpha-1}{2m-1}} \int_0^{\frac{t}{k_j \epsilon} \frac{2m}{2m-1}} e^{-z} z^{\frac{2m}{2m-1}(\alpha-1)} dz \\ &\leq C_2 \epsilon^{\frac{2m\alpha-1}{2m-1}} \int_0^\infty e^{-z} z^{\frac{2m}{2m-1}(\alpha-1)} dz \\ &\leq C_3 \epsilon^{\frac{2m\alpha-1}{2m-1}} \end{aligned}$$

where in the above we have used the fact that $e^{-z} z^{\frac{2m}{2m-1}(\alpha-1)} \in L^1(\mathbb{R}^+)$ since $m > \frac{1}{2\alpha}$. Therefore, we have

$$E \left[\sup_{0 \leq t \leq T} |I(t)|^{2m} \right] \leq C_4 \epsilon^{2m\alpha-1} E \left[\int_0^T |Y(s)|^{2m} ds \right].$$

Let $\frac{1}{4} < \alpha < \frac{1}{2}$ and $m = 2$. Then, using first the fact that $Y(t)$ is a mean-zero Gaussian, and then the Itô isometry, we have

$$\begin{aligned} E[|Y(t)|^4] &= 3 \left(E \left[(Y(t))^2 \right] \right)^2 \\ &= 3 \left(\int_0^t e^{-\frac{2(t-u)}{k_j \epsilon}} (t-u)^{-2\alpha} du \right)^2 \\ &= 3 \left(\frac{k_j \epsilon}{2} \right)^{2(1-2\alpha)} \left(\int_0^{\frac{2t}{k_j \epsilon}} e^{-s} s^{-2\alpha} ds \right)^2 \\ &\leq C_5 \epsilon^{2(1-2\alpha)} \end{aligned}$$

using the fact that $e^{-s} s^{-2\alpha} \in L^1(\mathbb{R}^+)$ since $\alpha < \frac{1}{2}$. Thus,

$$E \left[\sup_{0 \leq t \leq T} |I(t)|^4 \right] \leq C_6 \epsilon,$$

where C_6 is a constant that depends on T , so we get (22) by the Cauchy-Schwarz inequality:

$$E \left[\sup_{0 \leq t \leq T} |I(t)|^2 \right] \leq \left(E \left[\sup_{0 \leq t \leq T} |I(t)|^4 \right] \right)^{1/2} \leq C' \epsilon^{1/2}.$$

Now we prove (23). The solution of (2) is

$$(y_t^\epsilon)_j = e^{-\frac{t}{k_j \epsilon}} (y_0^\epsilon)_j + \frac{1}{k_j \epsilon} \int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} d(W_s)_j.$$

Thus, we have

$$\begin{aligned} E \left[\sup_{0 \leq t \leq T} |(\epsilon y_t^\epsilon)_j|^2 \right] &\leq 2E \left[\sup_{0 \leq t \leq T} \left| e^{-\frac{t}{k_j \epsilon}} (\epsilon y_0^\epsilon)_j \right|^2 \right] \\ &\quad + \frac{2}{k_j^2} E \left[\sup_{0 \leq t \leq T} \left(\int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} d(W_s)_j \right)^2 \right]. \end{aligned}$$

For the first term we use that $(y_0^\epsilon)_j$ is distributed according to the stationary distribution corresponding to (2) and thus has mean zero and variance

$$E[|(y_0^\epsilon)_j|^2] = (2k_j\epsilon)^{-1}.$$

$$E \left[\sup_{0 \leq t \leq T} \left| e^{-\frac{t}{k_j\epsilon}} (\epsilon y_0^\epsilon)_j \right|^2 \right] \leq E \left[|(\epsilon y_0^\epsilon)_j|^2 \right] = \frac{\epsilon}{2k_j}.$$

The bound (23) then follows from this bound together with (22). \square

The next lemma is elementary but we include its proof for completeness.

Lemma 3.3. *For each $\epsilon > 0$, let \mathbf{y}^ϵ be defined on $t \geq 0$ as the solution to equation (2) with \mathbf{y}_0^ϵ distributed according to the stationary distribution corresponding to (2), and let \mathbf{y}^ϵ be defined on $t_- \leq t < 0$ in such a way so that \mathbf{y}^ϵ considered on the interval $t \geq t_-$ is a stationary process. Then there exists $C > 0$ such that for all $\epsilon > 0$,*

$$E \left[\left| \epsilon (y_t^\epsilon)_j (y_s^\epsilon)_\ell \right|^2 \right] < C$$

for all $1 \leq j, \ell \leq n$ and $t, s \geq t_-$.

Proof. For all $t \geq t_-$, $(y_t^\epsilon)_j$ is a mean-zero Gaussian random variable with variance

$$E \left[|(y_t^\epsilon)_j|^2 \right] = \frac{1}{2k_j\epsilon}, \quad (24)$$

and so

$$E \left[|(y_t^\epsilon)_j|^4 \right] = \frac{3}{4k_j^2\epsilon^2}.$$

Thus, by the Cauchy-Schwarz inequality

$$E \left[\left| \epsilon (y_t^\epsilon)_j (y_s^\epsilon)_\ell \right|^2 \right] \leq \left(E \left[(\sqrt{\epsilon} y_t^\epsilon)_j^4 \right] \right)^{1/2} \left(E \left[(\sqrt{\epsilon} y_s^\epsilon)_\ell^4 \right] \right)^{1/2} = \frac{3}{4k_j k_\ell}.$$

\square

Lemma 3.4. *For each $\epsilon > 0$, let \mathbf{x}^ϵ be defined as in the statement of Theorem 1.1 where \mathbf{f} is bounded and $\boldsymbol{\sigma}$ is bounded with bounded derivatives. Also, for $\mathbf{t} = (t_1, \dots, t_m) \in [0, T]^m$ and $\mathbf{h} = (h_1, \dots, h_m) \in [0, 1]^m$, let $\mathbf{x}_\mathbf{t}^\epsilon = ((x_{t_1}^\epsilon)_1, \dots, (x_{t_m}^\epsilon)_m)$, let $\mathbf{x}_{\mathbf{t}+\mathbf{h}}^\epsilon = ((x_{t_1+h_1}^\epsilon)_1, \dots, (x_{t_m+h_m}^\epsilon)_m)$, and let $h_M = \max_{1 \leq i \leq m} h_i$. Then there exists C independent of ϵ such that for all $\mathbf{t} \in [0, T]^m$, $\mathbf{h} \in [0, 1]^m$, and $\epsilon > 0$,*

$$E \left[\|\mathbf{x}_{\mathbf{t}+\mathbf{h}}^\epsilon - \mathbf{x}_\mathbf{t}^\epsilon\|^2 \right] \leq C \left(h_M + \epsilon^{1/2} \right).$$

Proof. From equation (14) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
& E \left[\left| (x_{t_i+h_i}^\epsilon)_i - (x_{t_i}^\epsilon)_i \right|^2 \right] \\
& \leq 6E \left[\left(\int_{t_i}^{t_i+h_i} f_i(\mathbf{x}_s^\epsilon) ds \right)^2 \right] + 6E \left[\left(\left(\int_{t_i}^{t_i+h_i} \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) d\mathbf{W}_s \right)_i \right)^2 \right] \\
& \quad + 6E \left[\left(\int_{t_i}^{t_i+h_i} \sum_{p,j} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds \right)^2 \right] \\
& \quad + 6E \left[\left(\int_{t_i}^{t_i+h_i} \sum_{p,j,\ell} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) (y_{s-c_p\epsilon}^\epsilon)_\ell ds \right)^2 \right] \\
& \quad + 6E \left[\left(\sum_j \sigma_{ij}(\mathbf{x}_{t_i+h_i-c\epsilon}^\epsilon) k_j \epsilon (y_{t_i+h_i}^\epsilon)_j \right)^2 \right] \\
& \quad + 6E \left[\left(\sum_j \sigma_{ij}(\mathbf{x}_{t_i-c\epsilon}^\epsilon) k_j \epsilon (y_{t_i}^\epsilon)_j \right)^2 \right].
\end{aligned}$$

Using the boundedness of f_i , the Itô isometry, and the Cauchy-Schwarz inequality then gives

$$\begin{aligned}
& E \left[\left| (x_{t_i+h_i}^\epsilon)_i - (x_{t_i}^\epsilon)_i \right|^2 \right] \\
& \leq C_1 h_i^2 + 6C_2 \int_{t_i}^{t_i+h_i} E \left[\|\boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon)\|^2 \right] ds \\
& \quad + 6h_i E \left[\int_{t_i}^{t_i+h_i} \left(\sum_{p,j} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) \right)^2 ds \right] \\
& \quad + 6h_i E \left[\int_{t_i}^{t_i+h_i} \left(\sum_{p,j,\ell} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) (y_{s-c_p\epsilon}^\epsilon)_\ell \right)^2 ds \right] \\
& \quad + 6nE \left[\sum_j \left(\sigma_{ij}(\mathbf{x}_{t_i+h_i-c\epsilon}^\epsilon) k_j \epsilon (y_{t_i+h_i}^\epsilon)_j \right)^2 \right] + 6nE \left[\sum_j \left(\sigma_{ij}(\mathbf{x}_{t_i-c\epsilon}^\epsilon) k_j \epsilon (y_{t_i}^\epsilon)_j \right)^2 \right].
\end{aligned}$$

Using the boundedness of \mathbf{f} and the boundedness of $\boldsymbol{\sigma}$ and its derivatives, we

get

$$\begin{aligned}
& E \left[\left| (x_{t_i+h_i}^\epsilon)_i - (x_{t_i}^\epsilon)_i \right|^2 \right] \\
& \leq C_1 h_i^2 + C_3 h_i + 6C_4 h_i \int_{t_i}^{t_i+h_i} E \left[\sum_j \left(\epsilon(y_s^\epsilon)_j \right)^2 \right] ds \\
& \quad + 6C_5 h_i \int_{t_i}^{t_i+h_i} E \left[\sum_{j,\ell} \left(\epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell \right)^2 \right] ds \\
& \quad + 6nC_6 E \left[\sum_j \left(\epsilon(y_{t_i+h_i}^\epsilon)_j \right)^2 \right] + 6nC_6 E \left[\sum_j \left(\epsilon(y_{t_i}^\epsilon)_j \right)^2 \right].
\end{aligned}$$

Thus, using Lemma 3.2 and Lemma 3.3,

$$E \left[\left| (x_{t_i+h_i}^\epsilon)_i - (x_{t_i}^\epsilon)_i \right|^2 \right] \leq C_1 h_i^2 + C_3 h_i + 6C_7 \epsilon^{1/2} h_i^2 + 6C_8 h_i^2 + C_9 \epsilon^{1/2}$$

from which the statement follows. \square

3.3.1 U^ϵ converges to zero

Now we are ready to prove that U^ϵ goes to zero in probability with respect to $C([0, T], \mathbb{R}^m)$ as $\epsilon \rightarrow 0$. To do this, we split $(U^\epsilon)_i$ into three types of terms.

Recall that $(U^\epsilon)_i$ is defined in equation (15) as

$$\begin{aligned}
(U_t^\epsilon)_i &= - \underbrace{\int_0^{c^*\epsilon} \sum_{p,j,\ell} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds}_{I} \\
&\quad - \underbrace{\left(\int_0^{c^*\epsilon} \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) \epsilon D^{-1} d\mathbf{y}_s^\epsilon \right)}_I \\
&\quad + \underbrace{\left(\int_0^t (\boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \boldsymbol{\sigma}(\mathbf{x}_s^\epsilon)) d\mathbf{W}_s \right)}_I \\
&\quad + \int_{c^*\epsilon}^t \sum_{p,j,\ell} \left[\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) \right. \\
&\quad \quad \quad \left. - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] \underbrace{k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds}_{II} \\
&\quad - \underbrace{\sum_j \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon) k_j \epsilon (y_t^\epsilon)_j + \sum_j \sigma_{ij}(\mathbf{x}_{c^*\epsilon-c\epsilon}^\epsilon) k_j \epsilon (y_{c^*\epsilon}^\epsilon)_j}_{III} \\
&\quad + \underbrace{\int_{c^*\epsilon}^t \sum_{p,j} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds}_{III} .
\end{aligned}$$

We will prove that these terms all converge to zero as $\epsilon \rightarrow 0$. The type *I* terms converge to zero because the length of the interval of integration goes to zero. Convergence of the type *II* terms to zero will be a consequence of Lemma 3.4 and the Lipschitz continuity of $\boldsymbol{\sigma}$ and its first derivatives. The proof of convergence of type *III* terms to zero will follow from $\epsilon \mathbf{y}^\epsilon \rightarrow \mathbf{0}$ as $\epsilon \rightarrow 0$. Recall that for $p \geq 1$, convergence in L^p with respect to $C([0, T], \mathbb{R})$ implies convergence in probability with respect to $C([0, T], \mathbb{R})$. Thus, to show that these terms converge to zero in probability with respect to $C([0, T], \mathbb{R})$, we show that the fourth term goes to zero in L^1 with respect to $C([0, T], \mathbb{R})$ and the other terms go to zero in L^2 with respect to $C([0, T], \mathbb{R})$. We note that it is possible to show that the fourth term goes to zero in L^2 with respect to $C([0, T], \mathbb{R})$, but it would take slightly more work.

We start with the first two terms of $(U_t^\epsilon)_i$, i.e., the type *I* terms. For the first term, by two separate applications of the Cauchy-Schwarz inequality, the

boundedness of σ and its first derivatives, and Lemma 3.3, we have

$$\begin{aligned} E \left[\left(\sup_{0 \leq t \leq T} \left| \int_0^{c^* \epsilon} \sum_{p,j,\ell} \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p \epsilon}^\epsilon) \ell ds \right| \right)^2 \right] \\ \leq C_1 \epsilon \sum_{p,j,\ell} \int_0^{c^* \epsilon} E \left[\left| \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p \epsilon}^\epsilon) \ell \right|^2 \right] ds \\ \leq C_2 \epsilon^2 . \end{aligned}$$

For the second term, using (12) and two separate applications of the Cauchy-Schwarz inequality, we have

$$\begin{aligned} E \left[\left(\sup_{0 \leq t \leq T} \left| \left(\int_0^{c^* \epsilon} \sigma(\mathbf{x}_{s-c\epsilon}^\epsilon) \epsilon \mathbf{D}^{-1} d\mathbf{y}_s^\epsilon \right)_i \right| \right)^2 \right] \\ = E \left[\left| \sum_j \int_0^{c^* \epsilon} \sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) \left(- (y_s^\epsilon)_j ds + d(W_s)_j \right) \right|^2 \right] \\ \leq C_1 \sum_j E \left[\left| \int_0^{c^* \epsilon} \sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) (y_s^\epsilon)_j ds \right|^2 \right] \\ + C_1 \sum_j E \left[\left| \int_0^{c^* \epsilon} \sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) d(W_s)_j \right|^2 \right] . \end{aligned}$$

Using the Cauchy-Schwarz inequality, the Itô isometry, and the boundedness of σ then gives

$$\begin{aligned} E \left[\left(\sup_{0 \leq t \leq T} \left| \left(\int_0^{c^* \epsilon} \sigma(\mathbf{x}_{s-c\epsilon}^\epsilon) \epsilon \mathbf{D}^{-1} d\mathbf{y}_s^\epsilon \right)_i \right| \right)^2 \right] \\ \leq C_2 \epsilon \sum_j \int_0^{c^* \epsilon} E \left[\left| \sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) (y_s^\epsilon)_j \right|^2 \right] ds + C_1 \sum_j E \left[\int_0^{c^* \epsilon} \sigma_{ij}^2(\mathbf{x}_{s-c\epsilon}^\epsilon) ds \right] \\ \leq C_3 \epsilon . \end{aligned}$$

We now turn to the type *II* terms, for which it suffices to show that for every i, j, ℓ , and p ,

$$\lim_{\epsilon \rightarrow 0} E \left[\left(\sup_{0 \leq t \leq T} \left| \int_0^t \left(\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon) \right) d(W_s)_j \right| \right)^2 \right] = 0 \quad (25)$$

and

$$\lim_{\epsilon \rightarrow 0} E \left[\sup_{c^* \epsilon \leq t \leq T} \left| \int_{c^* \epsilon}^t \left[\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds \right| \right] = 0. \quad (26)$$

We note that in (26) we take the supremum over $c^* \epsilon \leq t \leq T$ because in the preceding we have implicitly defined this term, as well as the type *III* terms, to be zero for $0 \leq t < c^* \epsilon$. First we show (25). We will use the Lipschitz continuity of σ_{ij} for all i, j , which follows from the assumption of Theorem 1.1 that $\boldsymbol{\sigma}$ has bounded first derivatives. Also, we note that the Itô integral in (25) is a martingale because the integrand is still nonanticipating in the presence of the delays. Thus, by Doob's maximal inequality and the Itô isometry, we have

$$\begin{aligned} E \left[\left(\sup_{0 \leq t \leq T} \left| \int_0^t (\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon)) d(W_s)_j \right| \right)^2 \right] \\ \leq 2E \left[\left(\int_0^T (\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon)) d(W_s)_j \right)^2 \right] \\ = 2 \int_0^T E \left[(\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon))^2 \right] ds. \end{aligned}$$

We bound the above integral by using the Lipschitz continuity of σ_{ij} and Lemma 3.4. Since Lemma 3.4 applies only to differences of values of the process \mathbf{x}^ϵ at nonnegative times, we split the above integral into two terms. The first term involves values of \mathbf{x}^ϵ at negative times, i.e., values of the past condition \mathbf{x}^- , and the second term only involves values of \mathbf{x}^ϵ at nonnegative times. Recalling that $c^* = \max_{1 \leq i \leq m} c_i$ and letting M be the Lipschitz constant for σ_{ij} , we have

$$\begin{aligned} E \left[\left(\sup_{0 \leq t \leq T} \left| \int_0^t (\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon)) d(W_s)_j \right| \right)^2 \right] \\ \leq 2 \int_0^{c^* \epsilon} E \left[(\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon))^2 \right] ds + 2 \int_{c^* \epsilon}^T E \left[(M \|\mathbf{x}_{s-c\epsilon}^\epsilon - \mathbf{x}_s^\epsilon\|)^2 \right] ds \\ \leq C_1 \epsilon + C_2 T (\epsilon + \epsilon^{1/2}) \end{aligned}$$

using the boundedness of σ_{ij} and Lemma 3.4, from which (25) follows. To prove

(26):

$$\begin{aligned}
& E \left[\sup_{c^*\epsilon \leq t \leq T} \left| \int_{c^*\epsilon}^t \left[\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) \right. \right. \right. \\
& \qquad \qquad \qquad \left. \left. \left. - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_{j\epsilon}(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds \right| \right] \\
& \leq \int_{c^*\epsilon}^T E \left[\left| \left[\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) \right. \right. \right. \\
& \qquad \qquad \qquad \left. \left. \left. - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_{j\epsilon}(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell \right| \right] ds \\
& \leq \int_{c^*\epsilon}^T \left(E \left[\left(\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right] \right)^{1/2} \\
& \qquad \qquad \qquad \times \left(E \left[\left(k_{j\epsilon}(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell \right)^2 \right] \right)^{1/2} ds \\
& \leq C_1 \int_{c^*\epsilon}^T \left(E \left[\left(\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right] \right)^{1/2} ds
\end{aligned}$$

by Lemma 3.3. Again, in order to apply Lemma 3.4, we split the above integral into a part that involves values of the past condition \mathbf{x}^- and a part that only

involves values of \mathbf{x}^ϵ at nonnegative times. We have

$$\begin{aligned}
& E \left[\sup_{c^*\epsilon \leq t \leq T} \left| \int_{c^*\epsilon}^t \left[\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) \right. \right. \right. \\
& \qquad \qquad \qquad \left. \left. \left. - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds \right| \right] \\
& \leq C_1 \int_{c^*\epsilon}^{2c^*\epsilon} \left(E \left[\left(\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) \right. \right. \right. \\
& \qquad \qquad \qquad \left. \left. \left. - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right] \right)^{1/2} ds \\
& + C_1 \int_{2c^*\epsilon}^T \left(E \left[2 \left(\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right. \right. \\
& \qquad \qquad \qquad \left. \left. + 2 \left(\frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) - \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_s^\epsilon) \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right] \right)^{1/2} ds \\
& \leq C_2 \epsilon + C_3 \int_{2c^*\epsilon}^T \left(E \left[\|\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon - \mathbf{x}_s^\epsilon\|^2 + \|\mathbf{x}_{s-c\epsilon}^\epsilon - \mathbf{x}_s^\epsilon\|^2 \right] \right)^{1/2} ds \\
& \quad \text{(using the boundedness and Lipschitz continuity of } \boldsymbol{\sigma} \text{ and its first derivatives)} \\
& \leq C_2 \epsilon + C_4 T \left(\epsilon + \epsilon^{1/2} \right)^{1/2}
\end{aligned}$$

by Lemma 3.4, which gives (26).

For the type *III* terms, it suffices to show that for every i, j , and p ,

$$\lim_{\epsilon \rightarrow 0} E \left[\sup_{0 \leq t \leq T} \left| \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon) k_j \epsilon (y_t^\epsilon)_j \right|^2 \right] = 0 \tag{27}$$

$$\lim_{\epsilon \rightarrow 0} E \left[\sup_{0 \leq t \leq T} \left| \sigma_{ij}(\mathbf{x}_{c^*\epsilon-c\epsilon}^\epsilon) k_j \epsilon (y_{c^*\epsilon}^\epsilon)_j \right|^2 \right] = 0 \tag{28}$$

and

$$\lim_{\epsilon \rightarrow 0} E \left[\sup_{c^*\epsilon \leq t \leq T} \left| \int_{c^*\epsilon}^t \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds \right|^2 \right] = 0. \tag{29}$$

Equations (27) and (28) follow immediately from the boundedness of σ_{ij} and Lemma 3.2. To prove equation (29), we first use the boundedness of $\frac{\partial \sigma_{ij}}{\partial x_p}$ and f_p and then apply the Cauchy-Schwarz inequality:

$$\begin{aligned} E \left[\sup_{c^* \epsilon \leq t \leq T} \left| \int_{c^* \epsilon}^t \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_{s-c\epsilon}^\epsilon) k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p \epsilon}^\epsilon) ds \right|^2 \right] \\ \leq CE \left[\left(\int_{c^* \epsilon}^T |\epsilon (y_s^\epsilon)_j| ds \right)^2 \right] \\ \leq CT \int_0^T E \left[|\epsilon (y_s^\epsilon)_j|^2 \right] ds \end{aligned}$$

which goes to zero by Lemma 3.2.

Therefore $\mathbf{U}^\epsilon \rightarrow \mathbf{0}$ as $\epsilon \rightarrow 0$ in probability with respect to $C([0, T], \mathbb{R}^m)$, as claimed.

3.3.2 $\mathbf{H}^\epsilon \rightarrow \mathbf{H}$

Here we show that $\mathbf{H}^\epsilon \rightarrow \mathbf{H}$ in L^2 , and thus in probability, with respect to $C([0, T], \mathbb{R}^{(1+n+1+mn^2)})$. Note that the first three components of \mathbf{H}_t^ϵ , defined in (18), are independent of ϵ . Thus, it suffices to show that $(G^{j\ell})_p \rightarrow 0$ in L^2 with respect to $C([0, T], \mathbb{R})$ for all $j, \ell = 1, \dots, n$ and $p = 1, \dots, m$.

A heuristic argument that provides some insight into why $(G^{j\ell})_p$ converges to zero is as follows. Recall that $(G_t^{j\ell})_p$ is defined as

$$(G_t^{j\ell})_p = \int_0^t \left(k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p \epsilon}^\epsilon)_\ell - E \left[k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p \epsilon}^\epsilon)_\ell \right] \right) ds. \quad (30)$$

For each j , define the new process $(z)_j$ by $(z_r)_j = \sqrt{\epsilon} (y_{\epsilon r}^\epsilon)_j$. Then $(z)_j$ solves the ϵ -independent equation

$$d(z_r)_j = -\frac{1}{k_j} (z_r)_j dr + \frac{1}{k_j} d(\tilde{W}_r)_j$$

with the Wiener process $\tilde{W}_r = \epsilon^{-1/2} W_{\epsilon r}$. In terms of the process $(z)_j$, $(G_t^{j\ell})_p$ can be written as

$$(G_t^{j\ell})_p = \epsilon \int_0^{t/\epsilon} \left(k_j (z_u)_j (z_{u-c_p})_\ell - E \left[k_j (z_u)_j (z_{u-c_p})_\ell \right] \right) du.$$

The above integral can be thought of as the sum of $m = O(1/\epsilon)$ identically distributed random variables with zero mean. Furthermore, these random variables are weakly correlated, since the covariance function for $(z)_j$ decays exponentially with an exponential decay constant of order 1. Thus, we expect the L^2 -norm of this sum to grow about as fast as $O(1/\sqrt{\epsilon})$. Since $(G_t^{j\ell})_p$ is equal

to this integral multiplied by ϵ , we expect $(G_t^{j\ell})_p$ to converge to zero as $\epsilon \rightarrow 0$ for all $t \in [0, T]$. For fixed $t \in [0, T]$, convergence of $(G_t^{j\ell})_p$ to zero in L^2 can be shown by expressing the square of the integral as a double integral and then using Wick's theorem to compute the expectation of the terms in the integrand. However, to control the supremum norm, more work is required, and this is the content of the next lemma.

Lemma 3.5. *For each $\epsilon > 0$, let \mathbf{y}^ϵ be defined on $t \geq 0$ as the solution to equation (2) with \mathbf{y}_0^ϵ distributed according to the stationary distribution corresponding to (2), and let \mathbf{y}^ϵ be defined on $t_- \leq t < 0$ in such a way so that \mathbf{y}^ϵ considered on the interval $t \geq t_-$ is a stationary process. Then for all j, ℓ , and p ,*

$$\lim_{\epsilon \rightarrow 0} E \left[\left(\sup_{0 \leq t \leq T} \left| \int_0^t \left(k_j \epsilon (y_u^\epsilon)_j (y_{u-c_p \epsilon})_j - \frac{1}{2} e^{-\frac{c_p}{k_j}} \right) du \right| \right)^2 \right] = 0 \quad (31)$$

and

$$\lim_{\epsilon \rightarrow 0} E \left[\left(\sup_{0 \leq t \leq T} \left| \int_0^t k_j \epsilon (y_u^\epsilon)_j (y_{u-c_p \epsilon})_\ell du \right| \right)^2 \right] = 0 \quad \text{for } j \neq \ell. \quad (32)$$

To show this, a Riemann sum approximation is used in order to apply the following maximal inequality for sums of a stationary sequence of random variables.

Lemma 3.6. *[26, Proposition 2.3] Let $\{X_i : i \in \mathbb{N}\}$ be a stationary sequence of random variables and let $S_n = \sum_{i=1}^n X_i$. Let \mathcal{F}_0 be the σ -algebra generated by X_1 . Let $n \in \mathbb{N}$ and let $r = \lceil \log_2(n) \rceil$ (i.e., r is the smallest integer greater than or equal to $\log_2(n)$). Then*

$$E \left[\max_{1 \leq i \leq n} S_i^2 \right] \leq n \left(2\sqrt{E[X_1^2]} + (1 + \sqrt{2}) \Delta_r \right)^2 \quad (33)$$

where

$$\Delta_r = \sum_{j=0}^{r-1} \frac{\sqrt{E[E[S_{2^j} | \mathcal{F}_0]^2]}}{2^{j/2}}.$$

Remark 3.7. *By the conditional Jensen's inequality, $E[S_{2^j} | \mathcal{F}_0]^2 \leq E[S_{2^j}^2 | \mathcal{F}_0]$. Thus, (33) implies*

$$E \left[\max_{1 \leq i \leq n} S_i^2 \right] \leq n \left(2\sqrt{E[X_1^2]} + (1 + \sqrt{2}) \tilde{\Delta}_r \right)^2 \quad (34)$$

where

$$\tilde{\Delta}_r = \sum_{j=0}^{r-1} \frac{\sqrt{E[S_{2^j}^2]}}{2^{j/2}}.$$

Proof of Lemma 3.5. Let

$$\Psi_{j\ell p} = E \left[k_j \epsilon (y_u^\epsilon)_j (y_{u-c_p\epsilon}^\epsilon)_\ell \right] = \frac{\delta_{j\ell}}{2} e^{-\frac{c_p}{k_j}}.$$

Then showing both (31) and (32) is equivalent to showing that for all j, ℓ , and p ,

$$\lim_{\epsilon \rightarrow 0} E \left[\left(\sup_{0 \leq t \leq T} \left| \int_0^t \left(k_j \epsilon (y_u^\epsilon)_j (y_{u-c_p\epsilon}^\epsilon)_\ell - \Psi_{j\ell p} \right) du \right| \right)^2 \right] = 0. \quad (35)$$

Define the new process $\mathbf{z} = ((z)_1, \dots, (z)_n)$ by defining each component $(z)_j$ by $(z_r)_j = \sqrt{\epsilon} (y_{\epsilon r}^\epsilon)_j$. Then $(z)_j$ solves the ϵ -independent equation

$$d(z_r)_j = -\frac{1}{k_j} (z_r)_j dr + \frac{1}{k_j} d(\tilde{W}_r)_j, \quad (z_0)_j = \sqrt{\epsilon} (y_0^\epsilon)_j$$

with the Wiener process $\tilde{W}_r = \epsilon^{-1/2} W_{\epsilon r}$. Letting $r = u/\epsilon$ and $N = 1/\epsilon$ gives

$$\begin{aligned} \int_0^t \left(k_j \epsilon (y_u^\epsilon)_j (y_{u-c_p\epsilon}^\epsilon)_\ell - \Psi_{j\ell p} \right) du &= \epsilon \int_0^{t/\epsilon} \left(k_j \epsilon (y_{\epsilon r}^\epsilon)_j (y_{\epsilon r-c_p\epsilon}^\epsilon)_\ell - \Psi_{j\ell p} \right) dr \\ &= \frac{1}{N} \int_0^{Nt} \left(k_j (z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr. \end{aligned}$$

We approximate the above integral by a Riemann sum. Let $\{t_i\}_{i=1, \dots, m}$ be the partition of $[0, NT]$ into m equal parts of size $\Delta t = t_{i+1} - t_i$, where $t_1 = 0$ and $t_{i+1} - t_i = NTm^{-1}$ for all i . Near the end of the proof, we will choose m in terms of $N = 1/\epsilon$. Let $b(t) = \max\{i : iNTm^{-1} \leq Nt\}$. We add and subtract the corresponding Riemann sum and use the Cauchy-Schwarz inequality to obtain

$$\begin{aligned} &E \left[\sup_{0 \leq t \leq T} \left| \int_0^t \left(k_j \epsilon (y_u^\epsilon)_j (y_{u-c_p\epsilon}^\epsilon)_\ell - \Psi_{j\ell p} \right) du \right|^2 \right] \\ &\leq 2E \left[\sup_{0 \leq t \leq T} \left| \frac{1}{N} \left(\int_0^{Nt} \left(k_j (z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \right. \right. \right. \\ &\quad \left. \left. \left. - \frac{NT}{m} \sum_{i=1}^{b(t)} \left(k_j (z_{t_i})_j (z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right) \right|^2 \right] \\ &+ \frac{2}{N^2} E \left[\sup_{0 \leq t \leq T} \left| \frac{NT}{m} \sum_{i=1}^{b(t)} \left(k_j (z_{t_i})_j (z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right|^2 \right]. \quad (36) \end{aligned}$$

Each of the two terms on the right-hand side converges to zero for a different reason. The first term goes to zero because as m increases, the Riemann sum

better approximates the integral. The second term goes to zero because the sum grows like \sqrt{m} since it is a sum of on the order of m weakly correlated, mean-zero random variables.

We will start with the first term. First, writing

$$\frac{NT}{m} \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) = \int_{t_i}^{t_{i+1}} \left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr$$

we have

$$\begin{aligned} & E \left[\sup_{0 \leq t \leq T} \left| \frac{1}{N} \int_0^{Nt} \left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \right. \right. \\ & \qquad \qquad \qquad \left. \left. - \frac{T}{m} \sum_{i=1}^{b(t)} \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right|^2 \right] \\ &= E \left[\sup_{0 \leq t \leq T} \left| \frac{1}{N} \sum_{i=1}^{b(t)} \int_{t_i}^{t_{i+1}} \left[\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right] dr \right. \right. \\ & \qquad \qquad \qquad \left. \left. + \frac{1}{N} \int_{t_{b(t)}}^{Nt} \left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \right|^2 \right] \\ &\leq 2E \left[\sup_{0 \leq t \leq T} \left| \frac{1}{N} \sum_{i=1}^{b(t)} \int_{t_i}^{t_{i+1}} \left[\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right] dr \right|^2 \right] \\ & \quad + 2E \left[\sup_{0 \leq t \leq T} \left| \frac{1}{N} \int_{t_{b(t)}}^{Nt} \left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \right|^2 \right]. \tag{37} \end{aligned}$$

We first bound the first term on the right-hand side of (37). By the Cauchy–

Schwarz inequality,

$$\begin{aligned}
& E \left[\sup_{0 \leq t \leq T} \left| \frac{1}{N} \sum_{i=1}^{b(t)} \int_{t_i}^{t_{i+1}} \left[\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) \right. \right. \right. \\
& \qquad \qquad \qquad \left. \left. \left. - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right] dr \right|^2 \right] \\
& \leq E \left[\left(\frac{1}{N} \sum_{i=1}^{b(T)} \int_{t_i}^{t_{i+1}} \left| \left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) \right. \right. \right. \\
& \qquad \qquad \qquad \left. \left. \left. - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right| dr \right)^2 \right] \\
& \leq \frac{1}{N^2} b(T) \Delta t \sum_{i=1}^{b(T)} \int_{t_i}^{t_{i+1}} E \left[\left(\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) \right. \right. \\
& \qquad \qquad \qquad \left. \left. - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] dr .
\end{aligned}$$

Note that $b(T) = m$. We compute the expectation that makes up the integrand. First we expand the square:

$$\begin{aligned}
& E \left[\left(\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\
& = E \left[\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right)^2 \right] + E \left[\left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right)^2 \right] \\
& \quad - 2E \left[\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right] .
\end{aligned}$$

By using Wick's Theorem [9] to compute each expectation, we obtain

$$\begin{aligned}
& E \left[\left(\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\
& = \frac{k_j}{2k_\ell} \left(1 - e^{-\frac{|r-t_i|}{k_j}} e^{-\frac{|r-t_i|}{k_\ell}} \right) + \frac{\delta_{j\ell}}{2} \left(e^{-2\frac{c_p}{k_j}} - e^{-\frac{|r-t_i+c_p|}{k_j}} e^{-\frac{|r-t_i-c_p|}{k_j}} \right) .
\end{aligned}$$

Note that $r \geq t_i$ and $r - t_i \leq \Delta t$, so that for $\Delta t < \min_p c_p$,

$$-\frac{|r-t_i+c_p|}{k_j} - \frac{|r-t_i-c_p|}{k_j} = -\frac{r-t_i+c_p}{k_j} + \frac{r-t_i-c_p}{k_j} = -2\frac{c_p}{k_j} .$$

Therefore, letting $C_1 = \max_{j,\ell} \frac{k_j}{2k_\ell}$ and $C_2 = \max_{j,\ell} \frac{k_j+k_\ell}{k_j k_\ell}$, we have, for $\Delta t = \frac{NT}{m}$ sufficiently small,

$$E \left[\left(\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \leq C_1 \left(1 - e^{-C_2 \frac{NT}{m}} \right).$$

Thus, for N/m sufficiently small, we have the following bound for the first term in (37):

$$\begin{aligned} & 2E \left[\sup_{0 \leq t \leq T} \left| \frac{1}{N} \sum_{i=1}^{b(t)} \int_{t_i}^{t_{i+1}} \left[\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right] dr \right|^2 \right] \\ & \leq \frac{2}{N^2} m \Delta t \sum_{i=1}^m \int_{t_i}^{t_{i+1}} E \left[\left(\left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] dr \\ & \leq \frac{2}{N^2} m^2 (\Delta t)^2 C_1 \left(1 - e^{-C_2 \frac{NT}{m}} \right) \\ & = 2C_1 T^2 \left(1 - e^{-C_2 \frac{NT}{m}} \right). \end{aligned} \tag{38}$$

We now bound the second term in (37):

$$\begin{aligned} & 2E \left[\sup_{0 \leq t \leq T} \left| \frac{1}{N} \int_{t_b(t)}^{Nt} \left(k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \right|^2 \right] \\ & \leq \frac{2}{N^2} E \left[\max_{1 \leq i \leq m} \left(\int_{t_i}^{t_{i+1}} \left| k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right| dr \right)^2 \right] \\ & \leq \frac{2}{N^2} E \left[\sum_{i=1}^m \left(\int_{t_i}^{t_{i+1}} \left| k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right| dr \right)^2 \right] \\ & \leq \frac{2}{N^2} C_3 m (\Delta t)^2 \\ & = \frac{2C_3 T^2}{m}. \end{aligned} \tag{39}$$

Together, (38) and (39) give a bound for the first term in (36). We now turn our attention to the second term in (36). To bound this term we will use Lemma 3.6. First, we note

$$\begin{aligned} & \frac{1}{N^2} E \left[\sup_{0 \leq t \leq T} \left| \frac{NT}{m} \sum_{i=1}^{b(t)} \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right|^2 \right] \\ &= E \left[\max_{0 \leq n \leq m} \left| \frac{T}{m} \sum_{i=1}^n \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right|^2 \right]. \end{aligned}$$

Using the notation of Lemma 3.6, let $X_i = k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p}$. Note that $\{X_i\}$ is a stationary sequence since \mathbf{z} is a stationary process. We now estimate the quantity $\tilde{\Delta}_r$ in (34). Consider the Riemann sum

$$\frac{NT}{m} \sum_{i=1}^m \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right).$$

We have

$$\begin{aligned} & E \left[\left(\frac{NT}{m} \sum_{i=1}^m \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\ &= E \left[\sum_{i=1}^m \frac{NT}{m} \sum_{q=1}^m \frac{NT}{m} \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \left(k_j(z_{t_q})_j(z_{t_q-c_p})_\ell - \Psi_{j\ell p} \right) \right] \\ &= \sum_{i=1}^m \frac{NT}{m} \sum_{q=1}^m \frac{NT}{m} E \left[k_j^2(z_{t_i})_j(z_{t_i-c_p})_\ell(z_{t_q})_j(z_{t_q-c_p})_\ell - \Psi_{j\ell p} k_j(z_{t_i})_j(z_{t_i-c_p})_\ell \right. \\ & \quad \left. - \Psi_{j\ell p} k_j(z_{t_q})_j(z_{t_q-c_p})_\ell + (\Psi_{j\ell p})^2 \right] \\ &= \sum_{i=1}^m \frac{NT}{m} \sum_{q=1}^m \frac{NT}{m} \left[\frac{k_j}{4k_\ell} e^{-\frac{|t_i-t_q|}{k_j}} e^{-\frac{|t_i-t_q|}{k_\ell}} + \frac{\delta_{j\ell}}{4} e^{-\frac{|t_i-t_q+c_p|}{k_j}} e^{-\frac{|t_i-t_q-c_p|}{k_j}} \right] \\ &\leq C_4 NT \end{aligned}$$

where C_4 does not depend on N , m , or T , by comparison with the corresponding integral (recall that $\frac{NT}{m} = t_{i+1} - t_i$). Thus,

$$E \left[\left(\sum_{i=1}^m \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \leq C_4 \frac{m^2}{NT}.$$

Applying the last bound to partial sums $S_{2^j} = \sum_{i=1}^{2^j} X_i$, with NT replaced by $\frac{2^j}{m}NT$, we obtain

$$\tilde{\Delta}_r = \sum_{j=0}^{r-1} \frac{\sqrt{E[S_{2^j}^2]}}{2^{j/2}} \leq C_5 \sum_{j=0}^{r-1} \frac{2^j}{\sqrt{\frac{2^j}{m}NT}} 2^{-j/2} = C_5 r \sqrt{\frac{m}{NT}}.$$

Thus, by Lemma 3.6

$$\begin{aligned} E \left[\max_{1 \leq n \leq m} \left(\sum_{i=1}^n \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\ \leq m \left(2C_6 + (1 + \sqrt{2}) C_5 \sqrt{\frac{m}{NT}} \log_2(m) \right)^2 \end{aligned}$$

where $C_6 = \left(E \left[\left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right)^2 \right] \right)^{1/2}$. From this we obtain the following, which provides a bound for the second term in (36):

$$\begin{aligned} E \left[\max_{1 \leq n \leq m} \left(\frac{T}{m} \sum_{i=1}^n \left(k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\ \leq \frac{T^2}{m} \left(2C_6 + (1 + \sqrt{2}) C_5 \sqrt{\frac{m}{NT}} \log_2(m) \right)^2. \quad (40) \end{aligned}$$

Putting together (36), (37), (38), (39), and (40), we have, for N/m sufficiently small,

$$\begin{aligned} E \left[\sup_{0 \leq t \leq T} \left| \int_0^t \left(k_j \epsilon(y_u^\epsilon)_j(y_{u-c_p\epsilon})_\ell - \Psi_{j\ell p} \right) du \right|^2 \right] \\ \leq 2C_1 T^2 \left(1 - e^{-C_2 \frac{NT}{m}} \right) + \frac{2C_3 T^2}{m} \\ + \frac{2T^2}{m} \left(2C_6 + (1 + \sqrt{2}) C_5 \sqrt{\frac{m}{NT}} \log_2(m) \right)^2. \end{aligned}$$

Setting $m = N^2$ proves (35) and thus (31) and (32). \square

3.3.3 Conditions 1 and 2

Now we show that Conditions 1 and 2 are satisfied. First we show that \mathbf{H}^ϵ satisfies Condition 1. To do so, we must show that for every j, p , and $\ell \neq j$,

$$\int_0^T \left| k_j \epsilon(y_s^\epsilon)_j(y_{s-c_p\epsilon})_\ell - \frac{1}{2} e^{-\frac{c_p}{k_j}} \right| ds$$

and

$$\int_0^T \left| k_{j\epsilon}(y_s^\epsilon)_j (y_{s-c_p\epsilon})_\ell \right| ds$$

are stochastically bounded. This follows from Lemma 3.3 by the Chebyshev inequality, since the second moments of the integrands are bounded by a constant, independent of ϵ .

Condition 2 is the requirement that \mathbf{h} is continuous. In view of the definition (17) of \mathbf{h} , this is true by the assumptions of Theorem 1.1 that \mathbf{f} is continuous and that $\boldsymbol{\sigma}$ has continuous first derivatives.

We now complete the proof of Theorem 1.1. From Section 3.3.1, $\mathbf{U}^\epsilon \rightarrow \mathbf{0}$ in probability with respect to $C([0, T], \mathbb{R}^m)$. From Lemma 3.5, $\mathbf{H}^\epsilon \rightarrow \mathbf{H}$ in L^2 with respect to $C([0, T], \mathbb{R}^{(1+n+1+mn^2)})$. Therefore, $(\mathbf{U}^\epsilon, \mathbf{H}^\epsilon) \rightarrow (\mathbf{0}, \mathbf{H})$ in probability with respect to $C([0, T], \mathbb{R}^m \times \mathbb{R}^{(1+n+1+mn^2)})$. Furthermore \mathbf{H}^ϵ satisfies Condition 1 by Lemma 3.3 and \mathbf{h} satisfies Condition 2. Thus, by Lemma 3.1, $\mathbf{x}^\epsilon \rightarrow \mathbf{x}$ in probability with respect to $C([0, T], \mathbb{R}^m)$. \square

4 Discussion

We have derived the limiting equation for a general stochastic differential delay equation with multiplicative colored noise as the time delays and correlation times of the noises go to zero at the same rate. As a result of the dependence of the noise coefficients on the state of the system (multiplicative noise), a *noise-induced drift* appears in the limiting equation. The limiting SDE that we derived is useful for applications as it provides a simpler model than the original equation while at the same time still accounting for the effects of the time delays through the coefficients of the noise-induced drift.

The noise-induced drift has a form analogous to that of the *Stratonovich correction* to the Itô equation with noise term $\boldsymbol{\sigma}(\mathbf{x}_t)d\mathbf{W}_t$. That is, the noise-induced drift is a linear combination of the terms $\sigma_{pj}(\mathbf{x}_t) \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}_t)$. The coefficients of this linear combination in the limiting equation (3) are

$$\alpha_{jp} = \frac{1}{2} e^{-\frac{\delta_p}{\tau_j}}, \quad (41)$$

whereas the coefficients of the Stratonovich correction would all be equal to $1/2$. Thus, as explained in [19, 27], one can interpret the terms of the noise-induced drift as representing different stochastic integration conventions. We note that the exponential factor in (41) comes from the form (i.e., exponentially decaying) of the covariance function of the particular noise process that we use, that is, the Ornstein-Uhlenbeck process.

The limiting equation that we have derived here is a more accurate approximation of the delayed system than the limiting equation derived in [19, 27]. In particular, in [19, 27], equation (1) was Taylor expanded to first order in the time delays and the limiting equation corresponding to the expanded equation

was derived. It was shown that this limiting equation contains a noise-induced drift $\tilde{\mathbf{S}}(\mathbf{x}_t)$ given componentwise by

$$\tilde{S}_i(\mathbf{x}) = \sum_{p,j} \frac{1}{2} \left(1 + \frac{\delta_p}{\tau_j}\right)^{-1} \sigma_{pj}(\mathbf{x}) \frac{\partial \sigma_{ij}}{\partial x_p}(\mathbf{x}). \quad (42)$$

The coefficient $\frac{1}{2} \left(1 + \frac{\delta_p}{\tau_j}\right)^{-1}$ is a first-order Taylor expansion in the parameter δ_p/τ_j of the coefficient $\frac{1}{2} e^{-\delta_p/\tau_j}$ in (4) in the sense that $\left(1 + \frac{\delta_p}{\tau_j}\right)^{-1}$ is obtained when one expands the denominator of $1/e^{\delta_p/\tau_j}$ to first order in δ_p/τ_j . Thus, while the two limiting equations are close when all the ratios δ_p/τ_j are small, the limiting equation derived here is overall a better approximation of the delayed system.

5 Conclusion

We have proven a result concerning convergence of the solution of a general SDDE driven by state-dependent colored noise to the solution of an SDE driven by white noise. The main theorem (Theorem 1.1) was proven using direct analysis of the SDDE without Taylor expansion and thus it improves upon a result from previous works. The resulting limiting equation (3) was compared to the previous results in [19, 27]. The noise-induced drift derived there was seen to be a first-order expansion in the ratios δ_p/τ_j of the one found here. The limiting equation derived here can be used as an approximation to study the dynamics of real systems modeled by the SDDE.

Acknowledgements

A.M. and J.W. were partially supported by the NSF grants DMS 1009508 and DMS 0623941.

References

- [1] Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions, *Encyclopedia Math. Appl.*, vol. 44. Cambridge University Press, Cambridge, UK (1992)
- [2] Da Prato, G., Kwapien, S., Zabczyk, J.: Regularity of solutions of linear stochastic equations in Hilbert spaces. *Stochastics* **23**(1), 1–23 (1988)
- [3] Freidlin, M.: Some remarks on the Smoluchowski-Kramers approximation. *J. Stat. Phys.* **117**, 617–634 (2004)

- [4] Gardiner, C.W.: Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, *Springer Series in Synergetics*, vol. 13, third edn. Springer-Verlag, Berlin (2004)
- [5] Guillouxic, S., L'Heureux, I., Longtin, A.: Small delay approximation of stochastic delay differential equations. *Phys. Rev. E* **59**(4), 3970–3982 (1999)
- [6] Heil, T., Fischer, I., Elsässer, W., Krauskopf, B., Green, K., Gavrielides, A.: Delay dynamics of semiconductor lasers with short external cavities: bifurcation scenarios and mechanisms. *Phys. Rev. E* **67**, 066214 (2003)
- [7] Hottovy, S., Volpe, G., Wehr, J.: Thermophoresis of Brownian particles driven by coloured noise. *EPL (Europhys. Lett.)* **99**, 60002 (2012)
- [8] Hottovy, S., McDaniel, A., Volpe, G., Wehr, J.: The Smoluchowski-Kramers limit of stochastic differential equations with arbitrary state-dependent friction. *Commun. Math. Phys.* **336**(3), 1259–1283 (2015)
- [9] Janson, S.: Gaussian Hilbert Spaces, *Cambridge Tracts in Mathematics*, vol. 129. Cambridge University Press, Cambridge (1997). DOI 10.1017/CBO9780511526169
- [10] Kalman, R.E.: A new approach to linear filtering and prediction problems. *J. Basic Eng.* **82**(1), 35–45 (1960)
- [11] Kloeden, P.E., Platen, E.: Numerical solution of stochastic differential equations, *Applications of Mathematics (New York)*, vol. 23. Springer-Verlag, Berlin (1992)
- [12] Kupferman, R., Pavliotis, G.A., Stuart, A.M.: Itô versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise. *Phys. Rev. E* **70**, 036120 (2004)
- [13] Kurtz, T.G., Protter, P.: Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* **19**(3), 1035–1070 (1991)
- [14] Lenstra, D., Verbeek, B.H., den Boef, A.J.: Coherence collapse in single-mode semiconductor lasers due to optical feedback. *IEEE J. Quantum Electron.* **QE-21**(6), 674–679 (1985)
- [15] Lim, S.H., Wehr, J.: Homogenization for a class of generalized Langevin equations with an application to thermophoresis. To appear in *J. Stat. Phys.*
- [16] Longtin, A.: Stochastic delay-differential equations. In: F.M. Atay (ed.) *Complex Time-Delay Systems*, pp. 177–195. Springer-Verlag, Berlin (2010)
- [17] Mao, X.: Stochastic differential equations and applications. Woodhead Publishing, Cambridge (2007)

- [18] McDaniel, A., Mahalov, A.: Stochastic differential equation model for spontaneous emission and carrier noise in semiconductor lasers. *IEEE J. Quantum Electron.* **54**(1), 2000206 (2018)
- [19] McDaniel, A., Duman, O., Volpe, G., Wehr, J.: An SDE approximation for stochastic differential delay equations with state-dependent colored noise. *Markov Process. Relat.* **22**(3), 595–628 (2016)
- [20] Mijalkov, M., McDaniel, A., Wehr, J., Volpe, G.: Engineering sensorial delay to control phototaxis and emergent collective behaviors. *Phys. Rev. X* **6**(1), 011008 (2016)
- [21] Mulet, J., Mirasso, C.R.: Numerical statistics of power dropouts based on the Lang-Kobayashi model. *Phys. Rev. E* **59**(5), 5400–5405 (1999)
- [22] Nelson, E.: *Dynamical Theories of Brownian Motion*. Princeton University Press, Princeton, N.J. (1967)
- [23] Øksendal, B.: *Stochastic Differential Equations*, Sixth edn. Universitext. Springer-Verlag, Berlin (2003). An Introduction with Applications
- [24] Pavliotis, G.A., Stuart, A.M.: Analysis of white noise limits for stochastic systems with two fast relaxation times. *Multiscale Model. Simul.* **4**(1), 1–35 (2005). DOI 10.1137/040610507
- [25] Pavliotis, G.A., Stuart, A.M.: *Multiscale Methods: Averaging and Homogenization*, *Texts in Applied Mathematics*, vol. 53. Springer, New York (2008)
- [26] Peligrad, M., Utev, S.: A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.* **33**(2), 798–815 (2005)
- [27] Pesce, G., McDaniel, A., Hottovy, S., Wehr, J., Volpe, G.: Stratonovich-to-Itô transition in noisy systems with multiplicative feedback. *Nat. Commun.* **4**, 2733 (2013). DOI 10.1038/ncomms3733
- [28] Protter, P.E.: *Stochastic Integration and Differential Equations*, *Stochastic Modelling and Applied Probability*, vol. 21, second edn. Springer-Verlag, Berlin (2005)
- [29] Revuz, D., Yor, M.: *Continuous Martingales and Brownian Motion*, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 293, third edn. Springer-Verlag, Berlin (1999)
- [30] Tian, T., Burrage, K., Burrage, P.M., Carletti, M.: Stochastic delay differential equations for genetic regulatory networks. *J. Comput. Appl. Math* **205**, 696–707 (2007)
- [31] Torcini, A., Barland, S., Giacomelli, G., Marin, F.: Low-frequency fluctuations in vertical cavity lasers: Experiments versus Lang-Kobayashi dynamics. *Phys. Rev. A* **74**(6), 063801 (2006)

- [32] Volpe, G., Wehr, J.: Effective drifts in dynamical systems with multiplicative noise: a review of recent progress. *Rep. Prog. Phys.* **79**(5), 053901 (2016)