An SDE approximation for stochastic differential delay equations with colored state-dependent noise

Austin McDaniel\textsuperscript{1}, Ozer Duman\textsuperscript{2}, Giovanni Volpe\textsuperscript{2,3}, and Jan Wehr\textsuperscript{1}

\textsuperscript{1}Department of Mathematics, University of Arizona, Tucson, Arizona 85721 USA
\textsuperscript{2}Soft Matter Lab, Department of Physics, Bilkent University, Cankaya, Ankara 06800, Turkey
\textsuperscript{3}UNAM - National Nanotechnology Research Center, Bilkent University, Cankaya, Ankara 06800, Turkey

Abstract

We consider a general multidimensional stochastic differential delay equation (SDDE) with colored state-dependent noises. We approximate it by a stochastic differential equation (SDE) system and calculate its limit as the time delays and the correlation times of the noises go to zero. The main result is proven using a theorem of convergence of stochastic integrals developed by Kurtz and Protter. The result formalizes and extends a method that has been used in the analysis of a noisy electrical circuit with delayed state-dependent noise, and may be further used as a working SDE approximation of an SDDE system modeling a real system, where noises are correlated in time and whose response to stimuli is delayed.

Introduction

Stochastic differential equations (SDEs) are widely employed to describe the time evolution of systems encountered in physics, biology, and economics among others \cite{1,2,3}. It is often natural to introduce a delay into the equations in order to account for the fact that the system's response to changes in its environment is not instantaneous. We are, therefore, led to consider stochastic differential delay equations (SDDEs). While there exists a general theory of SDDEs (see
Ref. [1] for a survey), it is much less developed and explicit than the theory of SDEs [1, 2, 3, 4]. It is thus useful to develop working approximations of SDDEs by SDEs. For example, such an approximation was applied in Ref. [5] to a physical system with one dynamical degree of freedom (the output voltage of a noisy electrical circuit), showing that the experimental system shifts from obeying Stratonovich calculus to obeying Itô calculus as the ratio between the driving noise correlation time and the feedback delay time changes (see [6] for related work). In this article, we employ the systematic and rigorous method developed in Ref. [7] to obtain much more general results which are applicable to systems with an arbitrary number of degrees of freedom, driven by several colored noises, and involving several time delays. More precisely, we derive an approximation of SDDEs driven by colored noise (or noises) in which the correlation time of the noise is of the same order as the response delay (or delays).

Mathematical Model

We consider the multidimensional SDDE system

\[
\frac{dx_t}{dt} = f(x_t)dt + g(x_{t-\delta})\eta_t dt
\]

(1)

where \( x_t = (x^1_t, \ldots, x^m_t)^T \) is the state vector (the superscript T denotes transpose), \( f(x_t) = (f^1(x_t), \ldots, f^m(x_t))^T \) where \( f \) is a vector-valued function describing the deterministic part of the dynamical system,

\[
g(x_{t-\delta}) = \begin{bmatrix}
g^{11}(x_{t-\delta}) & \ldots & g^{1j}(x_{t-\delta}) & \ldots & g^{1n}(x_{t-\delta}) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
g^{i1}(x_{t-\delta}) & \ldots & g^{ij}(x_{t-\delta}) & \ldots & g^{in}(x_{t-\delta}) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
g^{m1}(x_{t-\delta}) & \ldots & g^{mj}(x_{t-\delta}) & \ldots & g^{mn}(x_{t-\delta})
\end{bmatrix}
\]

where \( g \) is a matrix-valued function, \( x_{t-\delta} = (x^1_{t-\delta}, \ldots, x^m_{t-\delta})^T \) is the delayed state vector (note that each component is delayed by an independent amount \( \delta_i > 0 \)), and \( \eta_t = (\eta^1_t, \ldots, \eta^n_t)^T \) is a vector of independent noises \( \eta^j \), where \( \eta^j \) are colored (harmonic) noises with characteristic correlation times \( \tau_j \). These stochastic processes (defined precisely in equation (5)) have continuously differentiable realizations which makes the realizations of the solution process \( x_t \) twice continuously differentiable under natural assumptions on \( f \) and \( g \), made precise in the statement of Theorem 1.

Equation (1) is written componentwise as

\[
\frac{dx^i(t)}{dt} = f^i(x^1(t), \ldots, x^m(t)) + \sum_{j=1}^{n} g^{ij}(x^1(t - \delta_1), \ldots, x^m(t - \delta_m))\eta^j(t)
\]

(2)

We define the process \( y^i(t) = x^i(t - \delta_i) \). In terms of the \( y \) variables, equation (2)
becomes
\[
\frac{dy^i(t + \delta_i)}{dt} = f^i(y^1(t + \delta_1), \ldots, y^m(t + \delta_m)) + \sum_{j=1}^{n} g^{ij}(y^1(t), \ldots, y^m(t))\eta^j(t)
\]  \hspace{1cm} (3)

Expanding to first order in \(\delta_i\), we have \(\dot{y}^i(t + \delta_i) \approx \dot{y}^i(t) + \delta_i\ddot{y}^i(t)\) and
\[
f^i(y^1(t + \delta_1), \ldots, y^m(t + \delta_m)) \approx f^i(y^1(t), \ldots, y^m(t)) + \sum_{k=1}^{m} \delta_k \frac{\partial f^i(y^1(t), \ldots, y^m(t))}{\partial y_k} dy_k(t) + \sum_{j=1}^{n} g^{ij}(y^1(t), \ldots, y^m(t)) \eta^j(t)
\]

Substituting these approximations into equation (3), we obtain a new (approximate) system
\[
\frac{dy^i(t)}{dt} + \delta_i \frac{d^2y^i(t)}{dt^2} = f^i(y(t)) + \sum_{k=1}^{m} \delta_k \frac{\partial f^i(y(t))}{\partial y_k} dy_k(t) + \sum_{j=1}^{n} g^{ij}(y(t))\eta^j(t)
\]
where \(y(t) = (y^1(t), \ldots, y^m(t))^T\). We write these equations as the first order system
\[
\begin{align*}
\frac{dy^i_t}{dt} &= v^i_t dt \\
\frac{dv^i_t}{dt} &= \left[-\frac{1}{\delta_i} v^i_t + \frac{1}{\delta_i} f^i(y_t) + \frac{1}{\delta_i} \sum_{k=1}^{m} \delta_k \frac{\partial f^i(y_t)}{\partial y_k} v^k_t + \frac{1}{\delta_i} \sum_{j=1}^{n} g^{ij}(y_t)\eta^j_t\right] dt
\end{align*}
\]  \hspace{1cm} (4)

Supplemented by the equations defining the noise processes \(\eta^j\) (see equation (5)), these equations become the SDE system we study in this article.

**Derivation of Limiting Equation**

We study the limit of the system (4) as the time delays \(\delta_i\) and the correlation times of the colored noises go to zero. We take each \(\eta^j\) to be a harmonic noise process [8] defined as the stationary solution of the SDE
\[
\begin{align*}
\frac{d\eta^j_t}{dt} &= \frac{1}{\tau_j} \frac{\Gamma}{\Omega^2} \sigma^j_t dt \\
\frac{d\sigma^j_t}{dt} &= -\frac{1}{\tau_j} \frac{\Gamma^2}{\Omega^4} \sigma^j_t dt - \frac{1}{\tau_j} \Gamma \eta^j_t dt + \frac{1}{\tau_j} \Gamma dW^j_t
\end{align*}
\]  \hspace{1cm} (5)

where \(\Gamma > 0\) and \(\Omega\) are constants, \(W_t = (W^1_t, ..., W^j_t, ..., W^n_t)^T\) is an \(n\)-dimensional Wiener process, and \(\tau_j\) is the correlation time of the Ornstein-Uhlenbeck process obtained by taking the limit \(\Gamma, \Omega^2 \to \infty\) while keeping \(\frac{\Gamma}{\Omega^2}\) constant (see Appendix for details). As \(\tau_j \to 0\), the component \(\eta^j_t\) of the solution of equation (5) converges to a white noise.
We assume that the delay times \( \delta_i \) and the noise correlation times \( \tau_j \) are proportional to a single characteristic time \( \epsilon \), i.e., \( \delta_i = c_i \epsilon \) and \( \tau_j = k_j \epsilon \) where \( c_i, k_j > 0 \) remain constant in the limit \( \delta_i, \tau_j, \epsilon \to 0 \).

We consider the solution to equations (1) and (5) on a bounded time interval \( 0 \leq t \leq T \). Throughout this article, for an arbitrary vector \( a \in \mathbb{R}^d \), \( \|a\| \) will denote the Euclidean norm, and for a matrix \( A \in \mathbb{R}^{d \times d} \), \( \|A\| \) will denote the matrix norm induced by the Euclidean norm on \( \mathbb{R}^d \).

**Theorem 1.** Suppose that the \( f^i \) are bounded functions with bounded continuous first derivatives and bounded second derivatives and that the \( g^{ij} \) are bounded functions with bounded continuous first derivatives. Let \( (y_i^\epsilon, v_i^\epsilon, \eta_i^\epsilon, z_i^\epsilon) \) solve equations (1) and (5) (which depend on \( \epsilon \) through \( \delta_i, \tau_j \)) on \( 0 \leq t \leq T \) with initial conditions \( (y_0, v_0, \eta_0, z_0) \) the same for every \( \epsilon \), where \( (\eta_0, z_0) \) is distributed according to the stationary distribution corresponding to equation (5). Let \( y_i \) solve

\[
\frac{dy_i^\epsilon}{dt} = f^i(y_i^\epsilon) dt + \sum_{p,j} g^{pj}(y_i^\epsilon) \frac{\partial g^{ij}(y_i^\epsilon)}{\partial y_p} \left[ \frac{k_j(c_p \Gamma^2 + k_j \Omega^2 - c_p \Omega^2)}{2(c_p \Gamma^2 + c_p k_j \Gamma^2 + k_j^2 \Omega^2)} \right] dt \\
+ \sum_{j} g^{ij}(y_i^\epsilon) dW_j^\epsilon
\]

on \( 0 \leq t \leq T \) with the same initial condition \( y_0 \), and suppose strong uniqueness holds on \( 0 \leq t \leq T \) for (7) with the initial condition \( y_0 \). Then

\[
\lim_{\epsilon \to 0} P \left[ \sup_{0 \leq t \leq T} \|y_i^\epsilon - y_i\| > a \right] = 0
\]

for every \( a > 0 \).

**Remark 1.** Taking the limit \( \Gamma, \Omega^2 \to \infty \) in equation (1) while keeping \( \frac{\Gamma}{\Omega^2} \) constant, we get the simpler limiting equation

\[
\frac{dy_i^\epsilon}{dt} = f^i(y_i^\epsilon) dt + \sum_{p,j} g^{pj}(y_i^\epsilon) \frac{\partial g^{ij}(y_i^\epsilon)}{\partial y_p} \frac{1}{2(1 + \frac{\delta_p}{\tau_j})^{-1}} dt + \sum_{j} g^{ij}(y_i^\epsilon) dW_j^\epsilon
\]

**Preparation of the proof of Theorem 1.** In order to prove Theorem 1, it will be convenient to write equations (1) and (5) together in matrix form. To do this, we introduce the vector process

\[
X_i^\epsilon = (y_i^\epsilon, \xi_i^\epsilon, \zeta_i^\epsilon),
\]

where, as in the statement of the theorem, \( (y_i^\epsilon, v_i^\epsilon, \eta_i^\epsilon, z_i^\epsilon) \) solves equations (1) and (5). \( \xi_i^\epsilon = ((\xi_1^\epsilon), \ldots, (\xi_n^\epsilon)) \) where \( (\xi_j^\epsilon) = \frac{1}{\tau_j} \int_0^t (\zeta_j^\epsilon) ds \), and \( \zeta_i^\epsilon = ((\zeta_1^\epsilon), \ldots, (\zeta_n^\epsilon)) \) where \( (\zeta_j^\epsilon) = \int_0^t (\zeta_j^\epsilon) ds = \tau_j \frac{\Omega^2}{4} \left[ (\eta_j^\epsilon) - (\eta_0^\epsilon) \right] \). We let \( V_i^\epsilon = X_i^\epsilon \), so that
\( V_t^e = (v_t^e, \eta_t^e, z_t^e) \). Equations (4) and (5) can be written in terms of the processes \( X_t^e \) and \( V_t^e \) as

\[
\begin{align*}
\frac{dX_t^e}{\epsilon} &= V_t^e dt \\
\frac{dV_t^e}{\epsilon} &= \left[ \frac{F(X_t^e)}{\epsilon} - \frac{\gamma(X_t^e)}{\epsilon} V_t^e + \kappa(X_t^e) V_t^e \right] dt + \frac{\sigma}{\epsilon} dW_t
\end{align*}
\]

where \( F(X_t^e) \) is the vector of length \( m + 2n \) that is given, in block form, by

\[
F(X_t^e) = \begin{bmatrix} \hat{f}(y_t^e) \\ 0 \\ 0 \end{bmatrix}
\]

where \( \hat{f}(y_t^e) = \left( \frac{f_1(y_t^1)}{c_1}, \ldots, \frac{f_m(y_t^m)}{c_m} \right)^T \); \( \gamma(X_t^e) \) is the \((m+2n) \times (m+2n)\) matrix that is given, in block form, by

\[
\gamma(X_t^e) = \begin{bmatrix} D^1 & -\hat{g}(y_t^e) & 0 \\ 0 & 0 & -\frac{1}{\beta^2} D^2 \\ 0 & \Gamma D^2 & \frac{1}{\beta^2} D^2 \end{bmatrix}
\]

where

\[
(\hat{g}(y_t^e))_{ij} = \frac{g^{ij}(y_t^e)}{c_i},
\]

\[
D^1 = \begin{bmatrix} \frac{1}{c_1} & 0 & \ldots & 0 \\ 0 & \frac{1}{c_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{1}{c_m} \end{bmatrix},
\]

and

\[
D^2 = \begin{bmatrix} \frac{1}{k_1} & 0 & \ldots & 0 \\ 0 & \frac{1}{k_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{1}{k_m} \end{bmatrix};
\]

\( \kappa(X_t^e) \) is the \((m+2n) \times (m+2n)\) matrix that is given, in block form, by

\[
\kappa(X_t^e) = \begin{bmatrix} \hat{J}_f(y_t^e) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\]

where

\[
\hat{J}_f(y_t^e) = \begin{bmatrix} \frac{c_1}{c_1} \frac{\partial f_1(y_t^e)}{\partial y_1} & \frac{c_2}{c_1} \frac{\partial f_1(y_t^e)}{\partial y_2} & \ldots & \frac{c_m}{c_1} \frac{\partial f_1(y_t^e)}{\partial y_m} \\ \frac{c_1}{c_2} \frac{\partial f_2(y_t^e)}{\partial y_1} & \frac{c_2}{c_2} \frac{\partial f_2(y_t^e)}{\partial y_2} & \ldots & \frac{c_m}{c_2} \frac{\partial f_2(y_t^e)}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{c_1}{c_m} \frac{\partial f_m(y_t^e)}{\partial y_1} & \frac{c_2}{c_m} \frac{\partial f_m(y_t^e)}{\partial y_2} & \ldots & \frac{c_m}{c_m} \frac{\partial f_m(y_t^e)}{\partial y_m} \end{bmatrix};
\]
\( \sigma \) is the \((m + 2n) \times n\) matrix that is given, in block form, by

\[
\sigma = \begin{bmatrix}
0 \\
0 \\
\Gamma D^2
\end{bmatrix};
\]

and \( W_t \) is the \( n \)-dimensional Wiener process in equation (5). Using the introduced notation, we obtain the desired matrix form of equations (4) and (5). The equation for \( V_t^\epsilon \) becomes

\[
[\gamma(X_t^\epsilon) - \epsilon \kappa(X_t^\epsilon)] V_t^\epsilon dt = F(X_t^\epsilon) dt + \sigma dW_t - \epsilon dV_t^\epsilon.
\]

We claim that for \( \epsilon \) sufficiently small, \( \gamma(X) - \epsilon \kappa(X) \) is invertible for all \( X \in \mathbb{R}^{m+2n} \). To see this, we first note that the eigenvalues of \( \gamma(X) \) do not depend on \( X \) and are nonzero (see (19)). With this in mind, the claim follows from the boundedness of \( \kappa \), the continuity of the function that maps a matrix to the vector of its eigenvalues (repeated according to their multiplicity), and the fact that, for fixed \( \epsilon_0 > 0 \), the closure of \( \{\gamma(X) - \epsilon \kappa(X) : X \in \mathbb{R}^{m+2n}, 0 \leq \epsilon \leq \epsilon_0\} \) is compact since \( \gamma \) and \( \kappa \) are bounded. Thus, for \( \epsilon \) sufficiently small, we can solve for \( V_t^\epsilon dt \), rewriting the equation for \( X_t^\epsilon \) as

\[
dX_t^\epsilon = V_t^\epsilon dt = [\gamma(X_t^\epsilon) - \epsilon \kappa(X_t^\epsilon)]^{-1} [F(X_t^\epsilon) dt + \sigma dW_t - \epsilon dV_t^\epsilon].
\]

In integral form, this equation is

\[
X_t^\epsilon = X_0 + \int_0^t (\gamma(X_s^\epsilon) - \epsilon \kappa(X_s^\epsilon))^{-1} F(X_s^\epsilon) ds \\
+ \int_0^t (\gamma(X_s^\epsilon) - \epsilon \kappa(X_s^\epsilon))^{-1} \sigma dW_s \\
- \int_0^t \epsilon (\gamma(X_s^\epsilon) - \epsilon \kappa(X_s^\epsilon))^{-1} dV_s^\epsilon
\]

(11)

where \( X_0 = (y_0, 0, 0) \) is independent of \( \epsilon \) due to the fact that \( y_0 \) is the same for all \( \epsilon \).

To find the limit of equation (11) as \( \epsilon \to 0 \), we use the method of Hottovy et al. [7]. In particular, we use a theorem of Kurtz and Protter [9] which, for greater clarity, we state here in a less general but sufficient form. We consider a family of pairs of processes \((U^\epsilon, H^\epsilon)\) with paths in \( C([0, T], \mathbb{R}^{(m+2n) \times d}) \) (i.e. the space of continuous functions from \([0, T]\) to \( \mathbb{R}^{(m+2n) \times d} \)) where \( H^\epsilon \) is a semimartingale. Let \( H^\epsilon_t = M^\epsilon_t + \Lambda^\epsilon_t \) be the Doob-Meyer decomposition of \( H^\epsilon_t \) so that \( M^\epsilon_t \) is a local martingale and \( \Lambda^\epsilon_t \) is a process of locally bounded variation [19]. Let \( h^\epsilon : \mathbb{R}^{m+2n} \to \mathbb{R}^{(m+2n) \times d} \) be a family of matrix-valued functions and also let \( h : \mathbb{R}^{m+2n} \to \mathbb{R}^{(m+2n) \times d} \) be a matrix-valued function. Suppose that the process \( Y^\epsilon \), with paths in \( C([0, T], \mathbb{R}^{m+2n}) \), satisfies the stochastic integral equation

\[
Y_t^\epsilon = Y_0 + U_t^\epsilon + \int_0^t h^\epsilon(Y_s^\epsilon) dH_s^\epsilon
\]

(12)
with \( Y_0 \) independent of \( \epsilon \). Let \( H_t \) with paths in \( C([0, T], \mathbb{R}^d) \) be a semimartingale and let \( Y \) with paths in \( C([0, T], \mathbb{R}^{m+2n}) \) satisfy the stochastic integral equation

\[
Y_t = Y_0 + \int_0^t h(Y_s) dH_s \tag{13}
\]

Lemma 1 ([9, Theorem 5.4 and Corollary 5.6]). Suppose \((U^\epsilon, H^\epsilon) \to (0, H)\) in probability with respect to \( C([0, T], \mathbb{R}^{(m+2n) \times d}) \), i.e. for all \( a > 0 \),

\[
P \left[ \sup_{0 \leq s \leq T} \|U^\epsilon_s\| + \|H^\epsilon_s - H_s\| > a \right] \to 0 \tag{14}
\]
as \( \epsilon \to 0 \), and the following conditions are satisfied:

**Condition 1.** The total variations, \( \{V_t(A^\epsilon)\} \), are stochastically bounded for each \( t > 0 \), i.e. \( P[|V_t(A^\epsilon)| > L] \to 0 \) as \( L \to \infty \), uniformly in \( \epsilon \).

**Condition 2.**
1. \( \sup_{\theta \in \mathbb{R}^{m+2n}} \|h^\epsilon(\theta) - h(\theta)\| \to 0 \) as \( \epsilon \to 0 \)
2. \( h \) is continuous (see [9, Example 5.3])

Suppose that there exists a strongly unique global solution to equation (13). Then, as \( \epsilon \to 0 \), \( Y^\epsilon \to Y \) in probability with respect to \( C([0, T], \mathbb{R}^{m+2n}) \), i.e. for all \( a > 0 \),

\[
P \left[ \sup_{0 \leq s \leq T} \|Y^\epsilon_s - Y_s\| > a \right] \to 0 \text{ as } \epsilon \to 0
\]

**Proof of Theorem 1.** We cannot apply Lemma 1 directly to equation (11) because \( \epsilon V^\epsilon_t \) does not satisfy Condition 1. Instead, we integrate by parts the \( i \)th component of the last integral in equation (11). We then have

\[
\int_0^t \sum_j \epsilon \left( (\gamma(X^\epsilon_s) - \epsilon \kappa(X^\epsilon_s))^{-1} \right)_{ij} d(V^\epsilon_s)_j = \\
\sum_j \left( (\gamma(X^\epsilon_t) - \epsilon \kappa(X^\epsilon_t))^{-1} \right)_{ij} \epsilon(V^\epsilon_t)_j - \sum_j \left( (\gamma(X_0) - \epsilon \kappa(X_0))^{-1} \right)_{ij} \epsilon(V_0)_j \\
- \int_0^t \sum_{i,j} \frac{\partial}{\partial X^\epsilon_t} \left[ ((\gamma(X^\epsilon_s) - \epsilon \kappa(X^\epsilon_s))^{-1})_{ij} \right] \epsilon(V^\epsilon_t)_j d(X^\epsilon_t)_l \tag{15}
\]

where \( V_0 = (v_0, \eta_0, z_0) \). Note that

\[
d\left[ (\gamma(X^\epsilon_s) - \epsilon \kappa(X^\epsilon_s))^{-1} \right]_{ij} = \sum_l \frac{\partial}{\partial X^\epsilon_t} \left[ ((\gamma(X^\epsilon_s) - \epsilon \kappa(X^\epsilon_s))^{-1})_{ij} \right] d(X^\epsilon_t)_l
\]

because \( X^\epsilon_s \) is continuously differentiable. The Itô term in the integration by parts formula is zero for a similar reason.


Since \(d(X^*_s)_t = (V^*_s)ds\), we can write the last integral in equation (15) as
\[
\int_0^t \sum_{i,j} \frac{\partial}{\partial X_i} \left[ ((\gamma(X^*_s) - \epsilon \kappa(X^*_s))^{-1})_{i,j} \right] \epsilon(V^*_s)_j(V^*_s)_i ds
\]
The product \(\epsilon(V^*_s)_j(V^*_s)_i\) that appears in the above integral is the \((j, l)\) entry of the outer product matrix \(\epsilon V^*_s(V^*_s)^T\). Our next step is to express this matrix as the solution of a certain equation. We start by using the Itô product formula to calculate
\[
d[\epsilon V^*_s(\epsilon V^*_s)^T] = \epsilon(d(V^*_s))((\epsilon V^*_s)^T + \epsilon V^*_s(\epsilon d(V^*_s)^T) + d(\epsilon V^*_s)d(\epsilon V^*_s)^T,
\]
so that, using equation (9),
\[
d[\epsilon V^*_s(\epsilon V^*_s)^T] = [\epsilon F(X^*_s)(V^*_s)^T - \epsilon \gamma(X^*_s)V^*_s(V^*_s)^T + \epsilon^2 \kappa(X^*_s)V^*_s(V^*_s)^T]ds
+ \epsilon \sigma dW_s(V^*_s)^T
+ [\epsilon V^*_s(F(X^*_s))^T - \epsilon V^*_s(V^*_s)^T(\gamma(X^*_s))^T + \epsilon^2 V^*_s(V^*_s)^T(\kappa(X^*_s))^T]ds
+ \epsilon V^*_s(\sigma dW_s)^T + \sigma \sigma^T ds
\]
We will show later that the terms that include \(\epsilon V^*_s\) converge to zero (see Lemma 4). Defining
\[
\tilde{U}_t^\epsilon = \int_0^t [\epsilon V^*_s(F(X^*_s))^T + \epsilon V^*_s(\epsilon V^*_s)^T(\kappa(X^*_s))^T]ds + \int_0^t \epsilon V^*_s(\sigma dW_s)^T
\]
and combining the last two equations, we have
\[
- \epsilon V^*_t(\epsilon V^*_t)^T(\gamma(X^*_t))^T dt - \epsilon \gamma(X^*_t)V^*_t(V^*_t)^T dt
= d[\epsilon V^*_t(\epsilon V^*_t)^T] - \sigma \sigma^T dt - d\tilde{U}_t^\epsilon - d(\tilde{U}_t^\epsilon)^T
\]
Our goal is to write the differential \(\epsilon V_t^\epsilon(\epsilon V_t^\epsilon)^T dt\) in another form and substitute it back into equation (15). With \(A = -\gamma(X_t^\epsilon), B = \epsilon V_t^\epsilon(\epsilon V_t^\epsilon)^T dt,\) and \(C = d[\epsilon V_t^\epsilon(\epsilon V_t^\epsilon)^T] - \sigma \sigma^T dt - d\tilde{U}_t^\epsilon - d(\tilde{U}_t^\epsilon)^T,\) equation (18) becomes
\[
AB + BA^T = C
\]
An equation of this form (to be solved for \(B\)) is called Lyapunov’s equation [11 [12]. By Ref. [12 Theorem 6.4.2], if the real parts of all eigenvalues of \(A\) are negative, it has a unique solution
\[
B = -\int_0^\infty e^{Ay}Ce^{A^Ty}dy
\]
for any \(C\). The eigenvalues of \(\gamma(X_t^\epsilon)\) are
\[
\frac{1}{c_i}, i = 1, \ldots, m, \quad \text{and} \quad \frac{T^2}{2k^j\Omega^2} \left[ 1 \pm \sqrt{1 - 4\frac{\Omega^2}{T^2}} \right], j = 1, \ldots, n \quad (19)
\]
so that the eigenvalues of $\gamma(X_t^\varepsilon)$ do not depend on $X_t^\varepsilon$ and have positive real parts (since $c_i > 0$ and $k_j > 0$ for $i = 1, ..., m$, $j = 1, ..., n$). Thus, all eigenvalues of $A = -\gamma(X_t^\varepsilon)$ have negative real parts, so we have
\[
\epsilon V_t^\varepsilon (V_t^\varepsilon)^T dt = - \int_0^\infty e^{-\gamma(X_t^\varepsilon)y} \left( d[eV_t^\varepsilon(eV_t^\varepsilon)^T] - \sigma \sigma^T dt - d(U_t^\varepsilon) - d(\tilde{U}_t^\varepsilon) \right) e^{-(\gamma(X_t^\varepsilon))^y} dy
\]
\[
= - \int_0^\infty e^{-\gamma(X_t^\varepsilon)y} d[eV_t^\varepsilon(eV_t^\varepsilon)^T] e^{-(\gamma(X_t^\varepsilon))^y} dy
\]
\[
+ \int_0^\infty e^{-\gamma(X_t^\varepsilon)y} (\sigma \sigma^T dt) e^{-(\gamma(X_t^\varepsilon))^y} dy
dC_t^1
\]
\[
+ \int_0^\infty e^{-\gamma(X_t^\varepsilon)y} (dU_t^\varepsilon + d(\tilde{U}_t^\varepsilon))^T e^{-(\gamma(X_t^\varepsilon))^y} dy.
dC_t^2
\]

After substituting the above expression into equation (15), a part of the term containing $dC_t^1$ will be included in the function $h^\varepsilon$ (in the notation of Lemma 1) and the other part will be included in the differential of the $H_t^\varepsilon$ process. Neither part will contribute to the limiting equation (8). The term containing $dC_t^2$ will contribute a noise-induced drift term to the limiting equation. Finally, the term containing $dC_t^3$ will become a part of $U_t^\varepsilon$, which will be shown to converge to zero, and so this term will not contribute to the limiting equation. First, we have
\[
(dC_t^1)_{jl} = \sum_{k_1,k_2} \int_0^\infty (e^{-\gamma(X_t^\varepsilon)y})_j k_1 \left. d[(eV_t^\varepsilon)_{k_1}(eV_t^\varepsilon)^T] (e^{-(\gamma(X_t^\varepsilon))^y})_{k_2} \right. dy
\]
\[
= \sum_{k_1,k_2} d[(eV_t^\varepsilon)_{k_1}(eV_t^\varepsilon)^T] \int_0^\infty (e^{-\gamma(X_t^\varepsilon)y})_j k_1 (e^{-(\gamma(X_t^\varepsilon))^y})_{k_2} dy.
\]

Next, we have $dC_t^2 = J(X_t^\varepsilon) dt$ where $J$ is the unique solution of the Lyapunov equation
\[
J - \gamma J + \gamma J = \sigma \sigma^T.
\]
Finally, using equation (17) for $\tilde{U}^\varepsilon$ we see that
\[
(dC_t^3)_{jl} = \sum_{k_1,k_2} \left[ \int_0^\infty (e^{-\gamma(X_t^\varepsilon)y})_j k_1 (e^{-(\gamma(X_t^\varepsilon))^y})_{k_2} dy \left( eV_t^\varepsilon (F(X_t^\varepsilon))^T \right)_{k_1,k_2} dt \right.
\]
\[
+ [eV_t^\varepsilon (V_t^\varepsilon)^T (F(X_t^\varepsilon))^T]_{k_1,k_2} dt + [\epsilon V_t^\varepsilon (\sigma dW_t)^T]_{k_1,k_2}
\]
\[
+ [F(X_t^\varepsilon) (eV_t^\varepsilon)^T]_{k_1,k_2} dt + [\kappa (X_t^\varepsilon) eV_t^\varepsilon (eV_t^\varepsilon)^T]_{k_1,k_2} dt
\]
\[
+ [\sigma dW_t (eV_t^\varepsilon)^T]_{k_1,k_2} \right].
\]
We are now ready to rewrite equation (11) and apply Lemma 1. Substituting the expression for $\epsilon V_s^t (V_s^t)^T dt$ into equation (15), equation (11) becomes

\[
(X_s^t)_i = (X_0)_i + (U_s^t)_i + \int_0^t \left( (\gamma(X_s^i) - \epsilon \kappa(X_s^i))^{-1} F(X_s^i) \right)_i ds \\
+ \left( \int_0^t \left( (\gamma(X_s^i) - \epsilon \kappa(X_s^i))^{-1} \sigma d W_s \right)_i \right) \\
+ \sum_{l,j} \int_0^t \frac{\partial}{\partial X_l} \left[ \left( (\gamma(X_s^i) - \epsilon \kappa(X_s^i))^{-1} \right)_{ij} \right] J_{jl} (X_s^i) ds \\
+ \sum_{l,j} \left[ \int_0^t \frac{\partial}{\partial X_l} \left[ \left( (\gamma(X_s^i) - \epsilon \kappa(X_s^i))^{-1} \right)_{ij} \right] \times \\
\sum_{k_1,k_2} \left( - \int_0^\infty (e^{-\gamma(X_s^i)^y})_{k_1} (e^{-(\gamma(X_s^i))} y)_{k_2} ds \right) e V_s^t k_1 (e V_s^t)^T k_2 \right] \\
+ \sum_{k_1,k_2} \left[ (e V_s^t (F(X_s^i))^{T})_{k_1,k_2} ds + [e V_s^t (e V_s^t)^T (\kappa(X_s^i))^{T}]_{k_1,k_2} ds \\
+ (e V_s^t (\sigma d W_s)^T k_1,k_2 + [F(X_s^i)(e V_s^t)^T]_{k_1,k_2} ds \\
+ [\kappa(X_s^i) e V_s^t (e V_s^t)^T]_{k_1,k_2} ds + [\sigma d W_s (e V_s^t)^T]_{k_1,k_2} \right] \right]
\] (21)

where the components of $U_s^t$ are

\[
(U_s^t)_i = - \sum_j (\gamma(X_s^i) - \epsilon \kappa(X_s^i))^{-1} \epsilon (V_s^t)_{ij} + \sum_j (\gamma(X_0) - \epsilon \kappa(X_s^i))^{-1} \epsilon (V_s^t)_{ij} \\
+ \sum_{l,j} \left[ \int_0^t \frac{\partial}{\partial X_l} \left[ \left( (\gamma(X_s^i) - \epsilon \kappa(X_s^i))^{-1} \right)_{ij} \right] \times \\
\sum_{k_1,k_2} \left( - \int_0^\infty (e^{-\gamma(X_s^i)^y})_{k_1} (e^{-(\gamma(X_s^i))} y)_{k_2} ds \right) e V_s^t k_1 (e V_s^t)^T k_2 \right] \\
+ (e V_s^t (F(X_s^i))^{T})_{k_1,k_2} ds + [e V_s^t (e V_s^t)^T (\kappa(X_s^i))^{T}]_{k_1,k_2} ds \\
+ (e V_s^t (\sigma d W_s)^T k_1,k_2 + [F(X_s^i)(e V_s^t)^T]_{k_1,k_2} ds \\
+ [\kappa(X_s^i) e V_s^t (e V_s^t)^T]_{k_1,k_2} ds + [\sigma d W_s (e V_s^t)^T]_{k_1,k_2} \right] \right]
\] (22)

We can now write equation (21) in the form of Lemma 1

\[
Y_s^t = Y_0 + U_s^t + \int_0^t h^s(Y_s^t) dH_s^t
\]

by letting $h^s : \mathbb{R}^{(m+2n)} \rightarrow \mathbb{R}^{(m+2n) \times (1+n+1+(m+2n)^2)}$ be the matrix-valued function given by

\[
h^s(Y) = \left( (\gamma(Y)-\epsilon \kappa(Y))^{-1} F(Y), (\gamma(Y)-\epsilon \kappa(Y))^{-1} \sigma, S^s(Y), \Lambda^1(Y), \ldots, \Lambda^{m+2n}(Y) \right)
\] (23)

where $S^s : \mathbb{R}^{(m+2n)} \rightarrow \mathbb{R}^{(m+2n)}$ is the vector-valued function defined componentwise as

\[
S^s(Y) = \sum_{l,j} \frac{\partial}{\partial Y_l} \left[ \left( (\gamma(Y) - \epsilon \kappa(Y))^{-1} \right)_{ij} \right] J_{jl}(Y)
\]

\] 10
with $J$ the solution to equation (20), and $\Lambda^{k_2} : \mathbb{R}^{(m+2)n} \to \mathbb{R}^{(m+2n) \times (m+2n)}$ defined componentwise as

$$
\Lambda^{k_2}_{ik_1}(Y) = \sum_{l,j} \frac{\partial}{\partial Y_l} \left[ \left( (\gamma(Y)^{-1} \cdot e \cdot \kappa(Y))^{-1} \right)_{ij} \right] \left[ \int_0^\infty \left( e^{-\gamma(Y)y} \right)_{jk_1} (e^-{(\gamma(Y)^T)y})_{k_2} dy \right].
$$

and by letting $H'_t$ be the process with paths in $C([0, T], \mathbb{R}^{1+n+1+(m+2n)^2})$ given by

$$
H'_t = \begin{bmatrix}
    t \\
    W_t \\
    (eV^*_t) \cdot eV^*_t - (eV_0) \cdot eV_0 \\
    \vdots \\
    (eV^*_t)_{(m+2n)} \cdot eV^*_t - (eV_0)_{(m+2n)} \cdot eV_0
\end{bmatrix}.
$$

We now define

$$
h(Y) = \left( (\gamma(Y)^{-1} F(Y), (\gamma(Y)^{-1} \sigma, S(Y), \Psi^1(Y), \ldots, \Psi^{m+2n}(Y) \right)
$$

where $S$ is defined componentwise as

$$
S_t(Y) = \sum_{l,j} \frac{\partial}{\partial Y_l} \left[ \left( (\gamma(Y)^{-1}) \right)_{ij} \right] J_{jl}(Y)
$$

and $\Psi^{k_2}$ is defined componentwise as

$$
\Psi^{k_2}_{ik_1}(Y) = \sum_{l,j} \frac{\partial}{\partial Y_l} \left[ \left( (\gamma(Y)^{-1}) \right)_{ij} \right] \left[ \int_0^\infty \left( e^{-\gamma(Y)y} \right)_{jk_1} (e^-{(\gamma(Y)^T)y})_{k_2} dy \right].
$$

Letting

$$
H_t = \begin{bmatrix}
    t \\
    W_t \\
    t \\
    0 \\
    \vdots \\
    0
\end{bmatrix},
$$

we show in the next section that $U', h', H', h$, and $H$ satisfy the assumptions of Lemma 1. It follows that, as $\epsilon \to 0$, $X'_t$ converges to the solution of the equation

$$
dX_t = \left[ (\gamma(X_t))^{-1} F(X_t) + S(X_t) \right] dt + (\gamma(X_t))^{-1} \sigma dW_t.
$$

Letting $X_t = (y_t, \xi_t, \zeta_t)$ (i.e., analogously to $X'_t$, we let $y_t$ stand for the vector of the first $m$ components of $X_t$, $\xi_t$ stand for the vector of the next $n$ components, and $\zeta_t$ stand for the vector of the last $n$ components), we have

$$
(\gamma(X_t))^{-1} = \begin{bmatrix}
    (D^1)^{-1} \\
    0 \\
    \tilde{\gamma}(y_t) \cdot (D^2)^{-1} \\
    0 \\
    \frac{1}{\tilde{\gamma}(y_t)} \cdot (D^2)^{-1} \\
    \tilde{\gamma}(y_t) \cdot (D^2)^{-1} \\
    0
\end{bmatrix}
$$

(28)
where \((g(y_t))_{ij} = k_j g^{ij}(y_t)\). Thus, from (27), we obtain the following limiting equation for \(y_t\)

\[
dy_t^i = f^i(y_t)dt + \sum_{p,j} g^{pj}(y_t) \frac{\partial g^{ij}(y_t)}{\partial y_p} \left[ \frac{k_j (c_p \Gamma^2 + k_j \Omega^2 - c_p \Omega^2)}{2(c_p \Gamma^2 + c_p k_j \Gamma^2 + k_j \Omega^2)} \right] dt \tag{29}
+ \sum_j g^{ij}(y_t)dW_t^j
\]

Taking the limit \(\Gamma, \Omega^2 \to \infty\) while keeping \(\frac{\Gamma}{\Omega^2}\) constant, this becomes

\[
dy_t^i = f^i(y_t)dt + \sum_{p,j} g^{pj}(y_t) \frac{\partial g^{ij}(y_t)}{\partial y_p} \frac{1}{2} \left( 1 + \frac{\delta_p}{\tau_j} \right)^{-1} dt + \sum_j g^{ij}(y_t)dW_t^j \tag{30}
\]

Q.E.D.

**Verification of Conditions**

In this section we verify that the conditions of Lemma 1 are satisfied. We first prove four lemmas.

**Lemma 2.** Let the functions \(f^i\) and \(g^{ij}\) satisfy the assumptions of Theorem 1, and let \(\epsilon_0 > 0\) be such that \(\gamma(X) - \epsilon \kappa(X)\) is invertible for \(0 \leq \epsilon \leq \epsilon_0\), \(X \in \mathbb{R}^{m+2n}\) (we have previously shown that such an \(\epsilon_0\) exists). Then there exists \(C > 0\) such that for \(0 \leq \epsilon \leq \epsilon_0\) and \(X \in \mathbb{R}^{m+2n}\), and \(1 \leq l \leq m + 2n\),

\[
\left\| \frac{\partial}{\partial X_l} [(\gamma(X) - \epsilon \kappa(X))^{-1}] \right\| < C.
\]

**Proof.** By differentiating the identity \((\gamma(X) - \epsilon \kappa(X))^{-1}(\gamma(X) - \epsilon \kappa(X)) = I\) we have

\[
\frac{\partial}{\partial X_l} [(\gamma(X) - \epsilon \kappa(X))^{-1}] = \tag{31}
-(\gamma(X) - \epsilon \kappa(X))^{-1} \left[ \frac{\partial}{\partial X_l} [\gamma(X) - \epsilon \kappa(X)] \right] (\gamma(X) - \epsilon \kappa(X))^{-1}
\]

From the assumption that the derivatives of the \(g^{ij}\) and the second derivatives of the \(f^i\) are bounded, it follows that \(\frac{\partial \gamma}{\partial X_l}\) and \(\frac{\partial \kappa}{\partial X_l}\) are bounded functions of \(X\). Also, there exists \(C_1 > 0\) such that for \(0 \leq \epsilon \leq \epsilon_0\) and \(X \in \mathbb{R}^{m+2n}\), \(\| (\gamma(X) - \epsilon \kappa(X))^{-1} \| < C_1\). This follows from the fact that the map that takes a matrix to its inverse is a continuous function on the space of invertible matrices and the fact that the closure of \(\{\gamma(X) - \epsilon \kappa(X) : X \in \mathbb{R}^{m+2n}, 0 \leq \epsilon \leq \epsilon_0\}\) is compact since \(\gamma\) and \(\kappa\) are bounded. □
Lemma 3. For each $\epsilon > 0$, let $X^i$ be any process with paths in $C([0, T], \mathbb{R}^{m+2n})$ and let the functions $f^i$ and $g^j$ satisfy the assumptions of Theorem 1. Let $p(t)$ with paths in $C([0, T], \mathbb{R}^{m+2n})$ solve the equation

$$\frac{d}{dt} p(t) = -\frac{1}{\epsilon} (\gamma(X^i) - \epsilon \kappa(X^i)) p(t)$$

Then there exist $C > 0$ and $C_d > 0$ such that for $0 \leq t_0 \leq t \leq T$,

$$\|p(t)\| \leq C \|p(t_0)\| \exp \left\{-\frac{C_d(t - t_0)}{\epsilon} \right\}.$$  \hfill (32)

Proof. Let $\beta(t)$ stand for the vector of the first $m$ components of $p(t)$, let $\mu(t)$ stand for the vector of the next $n$ components of $p(t)$, and let $\nu(t)$ stand for the vector of the last $n$ components of $p(t)$, so that $p(t) = (\beta(t), \mu(t), \nu(t))$. Then $(\mu(t), \nu(t))$ solves the constant coefficient system

$$\frac{d}{dt} \begin{bmatrix} \mu(t) \\ \nu(t) \end{bmatrix} = -\frac{1}{\epsilon} A \begin{bmatrix} \mu(t) \\ \nu(t) \end{bmatrix}$$

where

$$A = \begin{bmatrix} 0 & -\frac{\Gamma D^2}{\Omega^2} \\ \Gamma^2 D^2 & \frac{\Gamma^2}{\Omega^2} D^2 \end{bmatrix}.$$  \hfill (32)

The eigenvalues $\lambda_1, \lambda_2, ..., \lambda_{2n}$ of $A$ are equal to

$$\frac{\Gamma^2}{2 k_j \Omega^2} \left[ 1 \pm \sqrt{1 - \frac{4 \Omega^2}{\Gamma^2}} \right], \quad j = 1, ..., n$$

and $A$ is diagonalizable if $\Gamma^2 \neq 4 \Omega^2$ (if $\Gamma^2 = 4 \Omega^2$, an argument similar to the one below follows using the Jordan form of $A$). So, writing $A = P \Lambda P^{-1}$, where $\Lambda$ is a diagonal matrix consisting of $\lambda_1, \lambda_2, ..., \lambda_{2n}$, gives

$$\begin{bmatrix} \mu(t) \\ \nu(t) \end{bmatrix} = P \begin{bmatrix} e^{-\frac{\Gamma D^2}{\Omega^2} (t-t_0)} & 0 & ... & 0 \\ 0 & e^{-\frac{\Gamma D^2}{\Omega^2} (t-t_0)} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & e^{-\frac{\Gamma D^2}{\Omega^2} (t-t_0) \lambda_{2n}} \end{bmatrix} P^{-1} \begin{bmatrix} \mu(t_0) \\ \nu(t_0) \end{bmatrix}.$$

Let $c_\lambda = \min_{1 \leq j \leq 2n} Re(\lambda_j) > 0$. Then we have

$$\| (\mu(t), \nu(t)) \| \leq C_1 \| (\mu(t_0), \nu(t_0)) \| e^{-\frac{c_\lambda (t-t_0)}{\epsilon}}.$$  \hfill (33)

where $C_1$ is a constant.

Next, $\beta(t)$ solves

$$\frac{d}{dt} \beta(t) = \left( -\frac{1}{\epsilon} D^1 + \tilde{J}_f(y^i) \right) \beta(t) + \frac{1}{\epsilon} \tilde{\gamma}(y^i) \mu(t),$$
so that
\[
\beta(t) = \psi_{t_0}(t)\beta(t_0) + \int_{t_0}^{t} \psi_{t_0}(t)(\psi_{t_0}(s))^{-1} \frac{1}{c} \hat{g}(y_s^i) \mu(s) ds
\]
where \(\psi_{t_0}(t)\) is the particular fundamental solution matrix of the equation
\[
\frac{d}{dt}\psi(t) = \left(-\frac{1}{c} D^1 + \hat{J}_f(y_t^i)\right) \psi(t)
\]
satisfying \(\psi_{t_0}(t_0) = I\).

We derive an upper bound on the norm of \(\psi_{t_0}(t)\). Let \(u(t)\) with paths in \(C([0, T], \mathbb{R}^m)\) solve the equation
\[
\frac{d}{dt}u(t) = \left(-\frac{1}{c} D^1 + \hat{J}_f(y_t^i)\right) u(t).
\]
Then
\[
\frac{d}{dt}\left(\|u(t)\|^2\right) = \frac{d}{dt}(u(t)^T u(t))
\]
\[
= 2\left(\left(-\frac{1}{c} D^1 + \hat{J}_f(y_t^i)\right) u(t)\right)^T u(t)
\]
\[
\leq -\frac{2}{c} \|u(t)\|^2 + 2\|\hat{J}_f(y_t^i) u(t)\| \|u(t)\|
\]
where \(c = \max_{1 \leq i \leq m} c_i > 0\) (recall that \(D^1\) is the diagonal matrix with entries \(\frac{1}{c_i}\)),
\[
\leq 2 \left(-\frac{1}{c} + C_2\right) \|u(t)\|^2
\]
where \(C_2\) is a constant that bounds \(\|\hat{J}_f(y_t^i)\|\) (such a bound exists by the assumption that the first derivatives of the \(f_i^t\) are bounded).

Thus, by Gronwall’s inequality, we have
\[
\|u(t)\|^2 \leq \|u(t_0)\|^2 e^{\int_{t_0}^{t} \left(-\frac{1}{c} + C_2\right) ds}
\]

Now, let \((\psi_{t_0}(t))_{j}\) denote the \(j^{th}\) column of \(\psi_{t_0}(t)\). Then, by the chain of inequalities
\[
\|\psi_{t_0}(t)\| \leq C_3 \|\psi_{t_0}(t)\|_1 = C_3 \max_j \|(\psi_{t_0}(t))_{j}\|_1 \leq C_4 \max_j \|(\psi_{t_0}(t))_{j}\|, \quad (34)
\]
where \(\|\cdot\|_1\) denotes the induced matrix \(l^1\) norm or the vector \(l^1\) norm depending on its argument, and \(C_3\) and \(C_4\) are constants, we have
\[
\|\psi_{t_0}(t)\| \leq C_4 e^{-\frac{(t-t_0)}{c}}
\]
for $0 \leq t_0 \leq t \leq T$, where $C'$ depends on $T$. Thus, since $\psi_{t_0}(t) = \psi_s(t)\psi_{t_0}(s)$ and since $\hat{g}(y^*)$ is bounded by assumption, we have, for $0 \leq t_0 \leq t \leq T$,  
\[ \|\beta(t)\| \leq \|\beta(t_0)\|C'e^{-\frac{(t-t_0)}{\epsilon}} + C_5\|\langle \mu(t_0), \nu(t_0)\rangle\| \int_{t_0}^{t} \frac{1}{\epsilon} C'e^{-\frac{(t-s)}{\epsilon}} C_1 e^{-\frac{\gamma(s-t_0)}{\epsilon}} ds \]
where $C_5$ is a constant that bounds $\|\hat{g}(y^*)\|$,  
\[ \leq C_6\|p(t_0)\| \left( e^{-\frac{(t-t_0)}{\epsilon}} + \frac{t-t_0}{\epsilon} e^{-\frac{\epsilon'(t-t_0)}{2\epsilon}} \right) \]
where $\epsilon' = \min\{c_\lambda, \frac{1}{\epsilon}\}$ and $C_6$ is a constant,  
\[ = C_6\|p(t_0)\| \left( e^{-\frac{(t-t_0)}{\epsilon}} + \frac{t-t_0}{\epsilon} e^{-\frac{\epsilon'(t-t_0)}{2\epsilon}} \right) \leq C_6\|p(t_0)\| \left( e^{-\frac{(t-t_0)}{\epsilon}} + C_7 e^{-\frac{\epsilon'(t-t_0)}{2\epsilon}} \right) \]
where $C_7$ is a constant that bounds $xe^{-\frac{\epsilon'}{2\epsilon}}$ for all $x \geq 0$, so that we have  
\[ \|\beta(t)\| \leq C''\|p(t_0)\|e^{-\frac{\epsilon'(t-t_0)}{2\epsilon}} \]  
for $0 \leq t_0 \leq t \leq T$, where $C''$ is a constant. The bound (32) then follows from (30) and (31).

\[ \text{(35)} \]

**Lemma 4.** For each $\epsilon > 0$, let $X_t^\epsilon$ be any process with paths in $C([0,T], \mathbb{R}^{m+2n})$ and define $V_t^\epsilon$ as the solution to the SDE given by the second equation in (9), where the functions $f^i$ and $g^{ij}$ satisfy the assumptions of Theorem 1. Then $\epsilon V_t^\epsilon \to 0$ as $\epsilon \to 0$ in $L^2$, and therefore in probability, with respect to $C([0,T], \mathbb{R}^{m+2n})$, i.e.  
\[ \lim_{\epsilon \to 0} E \left[ \left( \sup_{0 \leq t \leq T} \|\epsilon V_t^\epsilon\| \right)^2 \right] = 0 \]
and so, for all $a > 0$,  
\[ \lim_{\epsilon \to 0} P \left( \sup_{0 \leq t \leq T} \|\epsilon V_t^\epsilon\| > a \right) = 0 \]

**Proof.** We solve the second equation in (9). This equation is a linear SDE with variable coefficients so its solution is [2]  
\[ V_t^\epsilon = \Phi(t)V_0 + \frac{1}{\epsilon} \int_0^t \Phi(t)(\Phi(s))^{-1} F(X^\epsilon_s)ds + \frac{1}{\epsilon} \int_0^t \Phi(t)(\Phi(s))^{-1} \sigma dW_s \]
where $\Phi(t)$ is the fundamental solution matrix of the equation  
\[ \frac{d}{dt} \Phi(t) = -\frac{1}{\epsilon} (\gamma(X^\epsilon_t) - \epsilon \kappa(X^\epsilon_t)) \Phi(t) \]  
1

13

15
satisfying $\Phi(0) = I$, so $\Phi(t)$ denotes $\Phi_0(t)$ in the notation introduced in Lemma 3. By Lemma 3, we have

$$
\|(\Phi_{t_0}(t))_j\| \leq C \exp\left\{- \frac{C_d(t - t_0)}{\epsilon}\right\}
$$

for $0 \leq t_0 \leq t \leq T$, where $(\Phi_{t_0}(t))_j$ denotes the $j^{th}$ column of $\Phi_{t_0}(t)$, so by the same inequalities as in (33),

$$
\|\Phi_{t_0}(t)\| \leq C_1 \exp\left\{- \frac{C_d(t - t_0)}{\epsilon}\right\}
$$

for $0 \leq t_0 \leq t \leq T$, where $C_1$ is a constant. Then, using $\Phi_{t_0}(t) = \Phi_s(t)\Phi_{t_0}(s)$, we have, for $f : [0, \infty) \rightarrow \mathbb{R}^{m+2n}$ and $t \leq T$,

$$
\int_0^t \|\Phi(t)(\Phi(s))^{-1}f(s)\|ds \leq \int_0^t C_1 \exp\left\{- \frac{C_d(t - s)}{\epsilon}\right\} \|f(s)\|ds.
$$

Now, $F$ is bounded since, by assumption, the $f^i$ are bounded, so there exists a constant $C_2$ such that $\|F(X)\| \leq C_2$ for all $X \in \mathbb{R}^{m+2n}$. Thus, using the time substitution $\tilde{s} = s/\epsilon$, we have

$$
\sup_{0 \leq t \leq T} \left\|\int_0^t \Phi(t)(\Phi(s))^{-1}F(X_s)ds\right\| \leq \sup_{0 \leq t \leq T} \int_0^t \|\Phi(t)(\Phi(s))^{-1}F(X_s)\|ds
$$

$$
\leq \sup_{0 \leq t \leq T} C_2 \int_0^t C_1 \exp\left\{- \frac{C_d(t - s)}{\epsilon}\right\} ds
$$

$$
\leq C_1 C_2 \int_0^T \exp\left\{- \frac{C_d(T - s)}{\epsilon}\right\} ds
$$

$$
= C_1 C_2 \epsilon \int_0^T \exp\left\{- \frac{C_d(s - \tilde{s})}{\epsilon}\right\} d\tilde{s}
$$

$$
= \frac{C_1 C_2 \epsilon}{C_d} \left(1 - e^{-\frac{C_d T}{\epsilon}}\right) \leq C_3 \epsilon
$$

where $C_3 = \frac{C_1 C_2}{C_d}$.

For the stochastic integral, using the Itô isometry (see [2, Theorem (4.4.14)]), we have

$$
E\left[\left\|\int_0^T \Phi(T)(\Phi(s))^{-1}\sigma dW_s\right\|^2\right] = \int_0^T E\left[\|\Phi(T)(\Phi(s))^{-1}\sigma\|_{HS}^2\right]ds
$$

$$
\leq D_1 \int_0^T E\left[\|\Phi(T)(\Phi(s))^{-1}\sigma\|^2\right]ds
$$

where $D_1$ is a constant and, for a matrix $A$, $\|A\|_{HS} = \sqrt{\sum_{i,j} A_{ij}^2}$ denotes the Hilbert-Schmidt norm of $A$. Using similar bounds as above and the same time
substitution $\tilde{s} = s/\epsilon$, we have

$$E \left[ \left\| \int_0^T \Phi(T)(\Phi(s))^{-1} \sigma dW_s \right\|^2 \right] \leq \epsilon D_2 \int_0^T \exp \left\{ -2C_d \left( \frac{T - \tilde{s}}{\epsilon} \right) \right\} d\tilde{s} \leq D_3 \epsilon$$

for constants $D_2$ and $D_3$. Thus,

$$E \left[ \sup_{0 \leq t \leq T} \| \epsilon V_t^\epsilon \| \right] \leq 3E \left[ \sup_{0 \leq t \leq T} \| \epsilon \Phi(t)V_0 \| \right]$$

$$+ 3E \left[ \sup_{0 \leq t \leq T} \left\| \int_0^t \Phi(t)(\Phi(s))^{-1} F(X_t^\epsilon) ds \right\| \right]$$

$$+ 3E \left[ \sup_{0 \leq t \leq T} \left\| \int_0^t \Phi(t)(\Phi(s))^{-1} \sigma dW_s \right\| \right] \leq C \epsilon$$ (36)

where $C$ is a constant, and where we have used the Cauchy-Schwarz inequality,

$$\left( \sum_{i=1}^N |a_i| \right)^2 \leq N \sum_{i=1}^N |a_i|^2,$$

and Doob’s maximal inequality $\mathbb{E}$ for the Itô integral, which is a martingale. \qed

**Lemma 5.** Let $X_t^\epsilon \in \mathbb{R}^{m+2n}$ and let $g : \mathbb{R}^{m+2n} \to \mathbb{R}$ be bounded. Then

$$\lim_{\epsilon \to 0} E \left[ \sup_{0 \leq t \leq T} \left| \int_0^t g(X_s^\epsilon) \epsilon(V_s^\epsilon)_i ds \right| \right] = 0$$ (37)

and

$$\lim_{\epsilon \to 0} E \left[ \left( \sup_{0 \leq t \leq T} \left| \int_0^t g(X_s^\epsilon) \epsilon(V_s^\epsilon)_i d(W_s)_j \right| \right)^2 \right] = 0$$ (38)

for all $i = 1, ..., m + 2n$ and $j = 1, ..., n$.

**Proof.** We have, using the Cauchy-Schwarz inequality,

$$E \left[ \left( \sup_{0 \leq t \leq T} \left| \int_0^t g(X_s^\epsilon) \epsilon(V_s^\epsilon)_i \right| ds \right)^2 \right] \leq E \left[ \left( \int_0^T |g(X_s^\epsilon)\epsilon(V_s^\epsilon)_i| ds \right)^2 \right]$$

$$\leq T \int_0^T E \left[ (g(X_s^\epsilon)\epsilon(V_s^\epsilon)_i)^2 \right] ds$$

$$\leq D^2 T \int_0^T E \left[ (\epsilon(V_s^\epsilon)_i)^2 \right] ds$$

where $D$ is a constant that bounds $g$. Taking the limit as $\epsilon \to 0$ of both sides, and using the Lebesgue dominated convergence theorem and Lemma 4, we get
decomposition of $H_\epsilon$ columns of the matrix of expression for the integral of the right side of equation (16): are made up of the column of the Lebesgue integrals that are present in the decomposition $X$ function of the stochastic integrals are local martingales, the last of other terms we observe that there exist $\kappa$ and have positive real parts. With these facts in mind, equation (39) follows from sections (22), (24), and (26) respectively. The fact that $H^c \to H$ in probability with respect to $C([0,T], \mathbb{R}^{m+2n})$ is an immediate consequence of Lemma 4. To show $U^c$ converges to zero as $\epsilon \to 0$ in probability with respect to $C([0,T], \mathbb{R}^{m+2n})$, it suffices to show that

$$\lim_{\epsilon \to 0} E \left[ \left( \sup_{0 \leq t \leq T} \|U^c_t\| \right)^2 \right] = 0$$

In considering the first two terms of each component of $U^c_t$, we again observe that there exist $\epsilon_0 > 0$ and $C > 0$ such that for $0 \leq \epsilon \leq \epsilon_0$ and $X \in \mathbb{R}^{m+2n}$, $\|\gamma(X) - \epsilon \kappa(X)\| < C$ (this is shown in the proof of Lemma 2). In considering the other terms we observe that $\int_0^T \epsilon \left( e^{-\gamma(X)^2} \right)_{j\ell} (e^{-\left(\gamma(X)^2\right)^{j\ell}})_{k\ell} dy$ is a bounded function of $X^c$ since the eigenvalues of $\gamma(X)^2$ are independent of the value of $X^c$ and have positive real parts. With these facts in mind, equation (39) follows from Lemmas 2, 4, and 5.

We now check Condition 1 of Lemma 1. To do this, we find the Doob-Meyer decomposition of $H^c$, i.e. the decomposition $H^c = M^c + A^c$ where $M^c$ is a local martingale and $A^c$ is a process of locally bounded variation. First, note that the columns of the matrix $\epsilon V^c T - \epsilon V_0 T$ make up the last $(m + 2n)^2$ rows of $H^c$: $(\epsilon V^c T - \epsilon V_0 T)_{11}$ is the first column of $\epsilon V^c T - \epsilon V_0 T$, $(\epsilon V^c T - \epsilon V_0 T)_{22}$ is the second column of $\epsilon V^c T - \epsilon V_0 T$, and so on. Consider the expression for $d(V^c T)$ given by equation (16). Because the stochastic integrals are local martingales, the last $(m + 2n)^2$ rows of $A^c$ are made up of the column of the Lebesgue integrals that are present in the expression for the integral of the right side of equation (16):
\[ ((A_1^t)^1, (A_1^t)^2, \ldots, (A_1^t)^{m+2n}) = \]
\[ \int_0^t \epsilon V_s^\epsilon (F(X_s^\epsilon))^T ds + \int_0^t F(X_s^\epsilon)(\epsilon V_s^\epsilon)^T ds \]
\[ - \int_0^t \epsilon V_s^\epsilon (V_s^\epsilon)^T (\gamma(X_s^\epsilon))^T ds - \int_0^t \gamma(X_s^\epsilon)V_s^\epsilon \epsilon (V_s^\epsilon)^T ds \]
\[ + \int_0^t \epsilon^2 V_s^\epsilon (V_s^\epsilon)^T (\kappa(X_s^\epsilon))^T ds + \int_0^t \epsilon \kappa(X_s^\epsilon)V_s^\epsilon (V_s^\epsilon)^T ds + \int_0^t \sigma \sigma^T ds \]

Thus, to show that Condition 1 holds, it suffices to show (since \( \int_0^t \sigma \sigma^T ds \) is just a constant) that the family (indexed by \( \epsilon \))
\[ \int_0^t \| \epsilon V_s^\epsilon (F(X_s^\epsilon))^T \| ds + \int_0^t \| F(X_s^\epsilon)(\epsilon V_s^\epsilon)^T \| ds \]
\[ + \int_0^t \| \epsilon V_s^\epsilon (V_s^\epsilon)^T (\gamma(X_s^\epsilon))^T \| ds + \int_0^t \| \gamma(X_s^\epsilon)V_s^\epsilon \epsilon (V_s^\epsilon)^T \| ds \]
\[ + \int_0^t \| \epsilon^2 V_s^\epsilon (V_s^\epsilon)^T (\kappa(X_s^\epsilon))^T \| ds + \int_0^t \| \epsilon \kappa(X_s^\epsilon)V_s^\epsilon (V_s^\epsilon)^T \| ds \]

is stochastically bounded (see the statement of Lemma 1 for the definition of a stochastically bounded family). The first two and last two terms go to zero in \( L^2 \) as \( \epsilon \to 0 \) by Lemma 4 and the fact that \( \kappa \) and \( F \) are bounded (by the assumptions of Theorem 1). Thus, these terms go to zero in probability, and so it suffices to show that the third and fourth terms are stochastically bounded.

Since \( \gamma \) is bounded (by the assumptions of Theorem 1), it suffices to show that \( E[\sup_{0 \leq s \leq t} \| \epsilon V_s^\epsilon (V_s^\epsilon)^T \|] \) is bounded uniformly in \( \epsilon \). This follows from (30) and the fact that for a vector \( v \) and outer product \( vv^T \), \( \| vv^T \| = \| v \|^2 \):
\[ E \left[ \sup_{0 \leq s \leq t} \| \epsilon V_s^\epsilon (V_s^\epsilon)^T \| \right] = E \left[ \sup_{0 \leq s \leq t} \epsilon \| V_s^\epsilon \|^2 \right] \leq C \]

We now check Condition 2 of Lemma 1, where \( h^\epsilon \) and \( h \) are defined in equations (23) and (25) respectively. We first note that \( J \) is continuous and bounded given the assumption that the \( g^{ij} \) are continuous and bounded (we have explicitly computed \( J \) in order to arrive at equation (20)). Part 1 of Condition 2 then follows from the boundedness of \( F, \kappa, \gamma, \frac{\partial \kappa}{\partial X}, \) and \( \frac{\partial \gamma}{\partial X} \), equation (21), the fact that taking the matrix inverse is a continuous function, and the fact that, for fixed \( \epsilon_0 > 0 \), the closure of \( \{ \gamma(X) - \epsilon \kappa(X) : X \in \mathbb{R}^{m+2n}, 0 \leq \epsilon \leq \epsilon_0 \} \) is compact since \( \gamma \) and \( \kappa \) are bounded. Part 2 of Condition 2 is immediate given equation (23) and the assumptions that the \( f^i \) are continuous and the \( g^{ij} \) have continuous derivatives.
Discussion

The main result of this article reduces the system of stochastic differential delay equations (1) to a simpler system (equations 6 and 8). First we use Taylor expansion to obtain the (approximate) system of SDEs (4) and then we further simplify it by taking the limit as the time delays and correlation times of the noises go to zero. This is useful for applications as the final equations are easier to analyze than the original ones [5].

Figure 1: Dependence of the coefficients $\alpha_{jp}$ of the noise-induced drift on the ratio between the corresponding delay time $\delta_p$ and noise correlation time $\tau_j$ (see equation 41). For $\delta_p/\tau_j \to 0$, the solution converges to the Stratonovich integral of equation 42, while, for $\delta_p/\tau_j \to \infty$, the solution converges to its Itô integral.

As a result of dependence of the noise coefficients on the state of the system (multiplicative noise), a noise-induced drift appears in equation 6. It has a form analogous to that of the Stratonovich correction to the Itô equation with the noise term $\sum_j g^{ij}(y_t)dW^j_t$. Each drift is a linear combination of the terms $g^{pj}(y_t)\frac{\partial g^{ij}(y_t)}{\partial y_p}$, but, while in the Stratonovich correction they all enter with coefficients equal to $\frac{1}{2}$, their coefficients in the additional drift of the limiting equation 6 are

$$ k_j(c_p\Gamma^2 + k_j\Omega^2 - c_p\Omega^2) 2(c_p^2\Gamma^2 + c_pk_j\Gamma^2 + k_j^2\Omega^2)^{-1}. $$

As explained in Remark 1, these coefficients approach their limiting value

$$ \alpha_{jp} = \frac{1}{2} \left(1 + \frac{\delta_p}{\tau_j}\right)^{-1}, $$

(41)
as the harmonic noise approaches the Ornstein-Uhlenbeck process, i.e. taking the limit $\Gamma, \Omega^2 \rightarrow \infty$ while keeping $\frac{\Gamma}{\Omega}$ constant (see Fig. 1). One can interpret the terms of the noise-induced drift as representing different stochastic integration conventions, a point that is further explained in Ref. [5]. For example, if all $\frac{\delta_p}{\tau_j} \rightarrow 0$, the solution converges to the Stratonovich integral of

$$dy_t = f^i(y_t)dt + \sum_j g^{ij}(y_t)dW_t^j,$$

which is equation (41) without the noise-induced drift terms; if all $\frac{\delta_p}{\tau_j} \rightarrow \infty$, the solution converges to the Itô integral of the above equation.

![Figure 2](image_url)

Figure 2: (a-d) Drift fields (arrows) estimated from a numerical solution of the SDDEs (43) with colored noises ($A = B = 0.1$ and $\sigma = 0.2$) for various values of the ratios $\delta_1/\tau_1$ and $\delta_2/\tau_2$. The circles represent the zero-drift points. (e) Modulus of the displacement of the zero-drift point from the equilibrium position corresponding to equations (43) without noise ($\sigma = 0$) as a function of $\delta_1/\tau_1$ and $\delta_2/\tau_2$. (f-i) Drift fields (arrows) of the solution of the limiting SDEs (8) corresponding to the SDDEs (43). $\alpha_{11}$ and $\alpha_{22}$ are given as functions of $\delta_1/\tau_1$ and $\delta_2/\tau_2$ by equation (41). The circles represent the zero-drift points. There is good agreement between (f-i) and (a-d). (j) Modulus of the displacement of the zero-drift point from the equilibrium position corresponding to equations (43) without noise ($\sigma = 0$) for the solution of the limiting SDEs (8) corresponding to the SDDEs (43) as a function of $\alpha_{11}$ and $\alpha_{22}$. Again, (j) and (e) are in good agreement.

While convergence of equations (4) to (8) is rigorously proven in this article, a specific system with non-zero values of $\delta_p$ and $\tau_j$ is more accurately described by (4) than by (8). In addition, equations (4) were obtained from the original
system \(1\) by an approximation (Taylor expansion). It is thus important to compare the behavior of the numerical solutions of \(4\) and \(8\) in a particular case. As an example, we consider the two-dimensional system

\[
\begin{align*}
\frac{dx_1}{dt} &= A x_1 (1 - x_1 - B x_2^2) dt + \sigma x_1 - \delta_1 \eta_1 t dt \\
\frac{dx_2}{dt} &= A x_2 (1 - x_2^2 - B x_1) dt + \sigma x_2 - \delta_2 \eta_2 t dt 
\end{align*}
\] (43)

where \(A, B,\) and \(\sigma\) are non-negative constants, \(\eta_1 t\) and \(\eta_2 t\) are colored noises with correlation times \(\tau_1\) and \(\tau_2\) respectively, and \(\delta_1\) and \(\delta_2\) are the delay times. These equations can describe, e.g., the dynamics of a noisy ecosystem where two populations are present whose sizes are proportional to the state variables \(x_1\) and \(x_2\). In the absence of noise (\(\sigma = 0\)) the system described by equations (43) is known as the competitive Lotka-Volterra model [14] and has only one stable fixed point at \(x_{eq}^1 = x_{eq}^2 = (1 + B)^{-1}\) for which \(x_{eq}^1, x_{eq}^2 \neq 0\). For a noisy system (with or without delay) there are no fixed points. One can still resort to an estimation of the system’s drift field, as done in Ref. [4, Methods], and identify the points in the state space where the drift is zero. For the system described by equations (43), the drift fields and the coordinates of the zero-drift point (for which \(x_1, x_2 \neq 0\)) depend on \(\delta_1/\tau_1\) and \(\delta_2/\tau_2\), as shown in Figs. 2(a-e) for \(A = B = 0.1\) and \(\sigma = 0.2\). We now calculate the drift fields and zero-drift points of the corresponding limiting SDEs [5]. The results, shown in Figs. 2(f-j), are in good agreement with the ones obtained by directly simulating equation (43).

Acknowledgments

We would like to thank the referee at Nature Communications who made the suggestion that we should consider the main equation in multiple dimensions and with different time delays; this suggestion led to a more general result (the one contained in this article) that more clearly reveals the interplay between the time delays and the correlation times of the noises. A.M. and J.W. were partially supported by the NSF grants DMS 1009508 and DMS 0623941. G.V. was partially supported by the Marie Curie Career Integration Grant (MC-CIG) No. PCIG11 GA-2012-321726.

Appendix

Here, we make some remarks about the harmonic noise process defined by [5]. The stationary harmonic noise process, defined as the stationary solution to [5], satisfies \(E[\eta^2] = 0\) and has covariance [8, 15]

\[
E[\eta_j^s \eta_{j+s}^t] = \frac{1}{2\tau_j} e^{-\frac{\Gamma^2}{2\Omega^2} s} \left[ \cos(\omega_1 s) + \frac{\Gamma^2}{2\Omega^2} \sin(\omega_1 s) \right], \ s \geq 0 \] (44)

where

\[
\omega_1 = \frac{\Gamma}{\Omega \tau_j} \sqrt{1 - \frac{\Gamma^2}{4\Omega^2}}
\]
We state a result concerning the convergence of the harmonic noise process to an Ornstein-Uhlenbeck process as \( \Gamma, \Omega^2 \to \infty \) while the ratio \( \frac{\Gamma}{\Omega^2} \) remains constant. Letting \( \tilde{\eta}_t^j = \tau_j \frac{\Omega^2}{\Gamma} \eta_t^j \), equation (5) becomes

\[
\begin{align*}
\frac{d\tilde{\eta}_t^j}{dt} &= z_t^j dt \\
\frac{dz_t^j}{dt} &= -\frac{1}{\tau_j} \Gamma z_t^j dt - \frac{1}{\tau_j} \Gamma \tilde{\eta}_t^j dt + \frac{1}{\tau_j} \Gamma dW_t^j
\end{align*}
\]

Note that this is a system of linear SDEs with constant coefficients, and so it can be solved explicitly. Thus, its limit can be studied directly, and we have the following result (this result can also be shown using the theorem of Hottovy et al. [7]). Let \( \tilde{\chi}_t^j \) be the solution to

\[
\begin{align*}
\frac{d\tilde{\chi}_t^j}{dt} &= -\frac{1}{\tau_j} \tilde{\chi}_t^j dt + \frac{\Omega^2}{\Gamma} dW_t^j
\end{align*}
\]

Then, as \( \Gamma, \Omega^2 \to \infty \) while the ratio \( \frac{\Gamma}{\Omega^2} \) remains constant, \( \tilde{\eta}_t^j \) converges to \( \tilde{\chi}_t^j \) in \( L^2 \) with respect to \( C([0, T], \mathbb{R}) \), that is,

\[
\lim_{\Gamma \to \infty \left( \frac{\Gamma}{\Omega^2} \right. \text{ constant})} E \left[ \left( \sup_{0 \leq t \leq T} |\tilde{\eta}_t^j - \tilde{\chi}_t^j| \right)^2 \right] = 0
\]

Thus, letting \( \chi_t^j \) be the solution to

\[
\begin{align*}
\frac{d\chi_t^j}{dt} &= -\frac{1}{\tau_j} \chi_t^j dt + \frac{1}{\tau_j} dW_t^j
\end{align*}
\]

so that \( \chi_t^j \) is an Ornstein-Uhlenbeck process with correlation time \( \tau_j \), we have that as \( \Gamma, \Omega^2 \to \infty \) while the ratio \( \frac{\Gamma}{\Omega^2} \) remains constant, \( \eta_t^j \) converges to \( \chi_t^j \) in \( L^2 \) (and therefore in probability) with respect to \( C([0, T], \mathbb{R}) \).

References

[1] Øksendal, B. (2007). *Stochastic Differential Equations*. Springer, Berlin.

[2] Arnold, L. (1974). *Stochastic Differential Equations: Theory and Applications*. Wiley-Interscience [John Wiley & Sons], New York. Translated from German.

[3] Karatzas, I. and Shreve, S. (1998). *Brownian Motion and Stochastic Calculus*. Springer, New York.

[4] Ivanov, A. F., Kazmerchuk, Y. I., and Swishchuk, A. V. (2003). Theory, stochastic stability and applications of stochastic delay differential equations: a survey of recent results. *Differential Equations and Dynamical Systems*. 11 55-115.
[5] Pesce, G. et al. (2013). Stratonovich-to-Itô transition in noisy systems with multiplicative feedback. Nat. Commun. 4:2733 doi: 10.1038/ncomms3733.

[6] Polettini, M. (2013). Generally covariant state-dependent diffusion. J. Stat. Mech. P07005.

[7] Hottovy, S., Wehr, J., and Volpe, G. The Smoluchowski-Kramers approximation for stochastic differential equations with arbitrary state dependent friction. Submitted for publication.

[8] Schimansky-Geier, L. and Zulicke, C. (1990). Harmonic noise: Effect on bistable systems. Z. Phys. B. 79 451-460.

[9] Kurtz, T. and Protter, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. Ann. Probab. 19(3) 1035-1070.

[10] Revuz, D. and Yor, M. (1999). Continuous Martingales and Brownian Motion, volume 293 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, third edition.

[11] Bellman, R. (1997). Introduction to Matrix Analysis, volume 19 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia. Reprint of the second (1970) edition.

[12] Ortega, J. (1987). Matrix Theory: A Second Course. The University Series in Mathematics. Plenum Press, New York.

[13] Hartman, P. (2002). Ordinary Differential Equations, volume 38 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia. Corrected reprint of the second (1982) edition.

[14] Murray, J. D. (2002). Mathematical Biology. Springer, Heidelberg.

[15] Wang, M. C. and Uhlenbeck, G. E. (1945). On the theory of the Brownian motion II. Rev. Mod. Phys. 17 323-342.

[16] Itô, K. (1994). Stochastic integral. Proc. Japan Acad. A 20 519-524.

[17] Stratonovich, R. L. (1966). A new representation for stochastic integrals and equations. SIAM J. Control 4 362-371.

[18] Ermak, D. L. and McCammon, J. A. (1978). Brownian dynamics with hydrodynamic interactions. J. Chem. Phys. 69 1352-1360.

[19] Volpe, G., Helden, L., Brettschneider, T., Wehr, J., and Bechinger, C. (2010). Influence of noise on force measurements. Phys. Rev. Lett. 104 170602.

[20] Kupferman, R., Pavliotis, G. A., and Stuart, A. M. (2004). Itô versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise. Phys. Rev. E 70 036120.
[21] Van Kampen, N. G. (1981). Itô versus Stratonovich. *J. Stat. Phys.* **24** 175-187.