Multi-scale analysis of stochastic delay differential equations

M. M. Klosek* and R. Kuske †

Abstract

We apply multi-scale analysis to stochastic delay-differential equations, deriving approximate stochastic equations for the amplitudes of oscillatory solutions near critical delays of deterministic systems. Such models are particularly sensitive to noise when the system is near a critical point, which marks a transition to sustained oscillatory behavior in the deterministic system. In particular, we are interested in the case when the combined effects of the noise and the proximity to criticality amplify oscillations which would otherwise decay in the deterministic system. The derivation of reduced equations for the envelope of the oscillations provides an efficient analysis of the dynamics by separating the influence of the noise from the intrinsic oscillations over long time scales. We focus on two well-known problems: the linear SDDE, and the logistic equation with delay. In addition to the envelope equations, the analysis identifies scaling relationships between small noise and the proximity of the bifurcation due to the delay which enhances the resonance of the noise with the intrinsic oscillations of the systems.

1 Introduction

In this paper we give a multi-scale analysis to study the effect of noise near critical delays for stochastic delay-differential equations. We focus on the case where the noise is small, but the dynamics are sensitive to the noise through a resonance. For example, in the presence of noise, oscillations can be sustained in the subcritical region, where oscillations otherwise decay over time in the absence of noise.

To illustrate this sensitivity, numerical simulations of the linear delay equation $dx = [-\alpha x(t) + \beta x(t - \tau)]dt + \delta dw$ are shown in Figure 1. In the absence of noise, a standard linear analysis shows that oscillations grow for $\tau < \tau_c$ and decay for $\tau > \tau_c$. The simulations suggest that when $1 \gg \tau_c - \tau > 0$, the evidence of the periodic behavior of the deterministic model is amplified by the combination of noise and delay. This effect is clear even when $\delta$ is small compared with $\tau - \tau_c$ (see Figure 1). As we show in this paper, the interaction of noise and delay support these oscillations, even over long time intervals. In the absence of external periodic forcing this phenomenon has been called autonomous stochastic resonance [1], where the noise excites the oscillations intrinsic to the deterministic dynamics. It has been observed in many systems with delays, including models of neurons, lasers, and a variety of oscillators with delayed feedback (see [2]-[6], and references therein). It has also been observed in systems without delays, where

---

*Laboratory of Computational and Experimental Biology, NCI, NIH, 12 South Drive, Bethesda, MD 20892-5677, email: klosek@nih.gov

†Department of Mathematics, University of British Columbia, Vancouver, BC Canada V6T1Z2, email: rachel@math.ubc.ca
Figure 1: The numerical simulation of (2.1) for $\alpha = 1$, $\beta = -\sqrt{2}$, $\tau = 3\pi/4 - \epsilon^2$, and $\delta = .01$. In the top figure $\epsilon^2 = 1.5$. In the bottom figure the solid line is for $\epsilon^2 = .5$ and the dash-dotted line corresponds to $\epsilon^2 = .1$. As $\epsilon$ decreases, the solution has an oscillatory behavior, even though $\tau < \tau_c$.

it has been studied in the context of stochastic bifurcations [7]-[12]. The interest in the influence of noise in stochastic differential delay equations has led to a large number of numerical studies, but there have been few quantitative analytical results which describe their dynamics.

In this paper we consider both linear and logistic stochastic delay equations. In the linear case we consider both the additive and multiplicative noise cases,

$$dx = \left[-\alpha x(t) + \beta x(t-\tau)\right]dt + \delta \left\{ \frac{dw}{x(t)dw} \right\}, \quad (1.1)$$

and we consider additive noise for the logistic case,

$$dx = r x(t)(1 - x(t - \tau))dt + \delta dw. \quad (1.2)$$

Using multi-scale analysis, we derive envelope equations for the stochastic amplitude of the oscillations. This approach is similar to the one used for the stochastic van der Pol-Duffing equation in [12], which yields results in agreement with other methods [8]-[11]. In [12] the multi-scale approach provides a new viewpoint of the dynamics in which the oscillations have the deterministic oscillation frequency associated with the Hopf bifurcation and are modulated by a slowly varying stochastic amplitude. It also has an insightful physical interpretation which is directly related to resonance with Fourier-type components.

In our analysis we consider the case where $\tau$ is near $\tau_c$, where $\tau_c$ is a critical point for the deterministic system. That is, for $\delta = 0$ and $\tau < \tau_c$, oscillations decay over time and for $\tau > \tau_c$ there is sustained oscillatory behavior: in the linear case the amplitude of the oscillations grows over time, while in the logistic case there is a stable oscillatory solution [13]. We explore the noise-sensitivity of the system for small values of $\tau - \tau_c$, restricting our analysis to $\delta \ll 1$ in order to understand where small noise can play a significant role in the dynamics, even though
it does not dominate the overall behavior of the system. We obtain relevant parameter scalings related to this noise-sensitivity.

For oscillator models, it is well known that in the absence of noise ($\delta = 0$), one can give an asymptotic approximation to the solution when the parameters correspond to close proximity to a critical point such as a Hopf bifurcation [14]. The method of multiple scales has been used in a few deterministic models with delay [15]-[17] to give reduced systems on a long time scale. We briefly review the multi-scale methods for determining the behavior of the envelope in the deterministic case.

In the setting of a Hopf bifurcation, a multi-scale approximation explicitly employs the natural frequency, say $\omega$, of the oscillation associated with the bifurcation [14]. The form of the solution is then

$$
\dot{x} \sim A(T) \cos \omega t + B(T) \sin \omega t, \quad T = \epsilon^2 t,
$$

(1.3)

where $\epsilon^2$ is the parameter measuring the proximity to the bifurcation. Here $A(T)$ and $B(T)$ are functions of a slow time $T$, which are treated as constants with respect to the fast oscillations with frequency $\omega$ on the $t$ time scale. By deriving equations on the slow $T$ scale for the amplitude, or envelope, described by $A(T)$ and $B(T)$, one gets an asymptotic approximation for the process near the bifurcation. The method takes $x$ as a function of both times $t$ and $T$, which are treated as independent. Then a perturbation expansion $x \sim x_0 + \epsilon x_1 + \ldots$ is used, with $x_0$ given by (1.3) and derivative $x_t$ replaced by $x_t + \epsilon^2 x_T$. Proceeding with the perturbation expansion, the equation for the higher order contributions $x_j$ for $j > 0$, are subject to solvability conditions, which give envelope equations for $A(T)$ and $B(T)$. These solvability conditions are often in the form of conditions of orthogonality to the oscillatory modes $\cos \omega t$ and $\sin \omega t$. The benefits of analyzing the envelope equations are that they are often relatively simple compared to the original model and that they allow an analysis or computation on the long time scale.

In the remainder of this paper, we derive such envelope or amplitude equations in the stochastic case $\delta \neq 0$. We show that under certain parameter restrictions, the envelope dynamics can be described by the usual deterministic process on the long time scale plus a diffusion process which includes resonances with the oscillations. The form of this diffusion process is model dependent, and we use a multi-scale point of view in order to take advantage of the resonances with the natural modes in the system. This approach is described further below and demonstrated in detail in the following sections.

1.1 Main results

We summarize the form of the envelope equations as found in the following sections. For the linear model with additive noise we use the approximation (1.3) and obtain

$$
\begin{align*}
\,dA &= \psi_A dT + \frac{\delta}{\epsilon} dw_1 \\
\,dB &= \psi_B dT + \frac{\delta}{\epsilon} dw_2,
\end{align*}
$$

(1.4)

where $\psi_A$ and $\psi_B$ are given by (2.18), and $w_{1,2}$ are independent Brownian motions.
For additive noise in the logistic model with delay, we look for small amplitude oscillations about unity, \( x \sim 1 + \epsilon A(T) \cos \omega t + \epsilon B(t) \sin \omega t + \ldots \), which yields

\[
\begin{align*}
\frac{dA}{dT} &= \psi_A \frac{dT}{T} + \frac{\delta}{\epsilon^2} dW_1, \\
\frac{dB}{dT} &= \psi_B \frac{dT}{T} + \frac{\delta}{\epsilon^2} dW_2.
\end{align*}
\]

The additional factor of \( \epsilon^{-1} \) in (1.5) is a result of seeking small amplitude oscillations. The different scalings of the noise are discussed further in Section 4. For the subcritical case \( \tau < \tau_c \), \( \omega = r \) and \( \psi_A \) and \( \psi_B \) are given in (3.5). For the supercritical case, \( \omega = r \sigma \) for \( \sigma \) in (3.8), and \( \psi_A \) and \( \psi_B \) are given in (3.12). As in (1.4), \( w_{1,2} \) are independent Brownian motions.

For the linear case with multiplicative noise, we again seek an approximation of the form (1.3), obtaining

\[
\begin{align*}
\frac{dA}{dB} = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} dT + \Sigma_0 \begin{pmatrix} A \\ B \end{pmatrix} d\xi_0(T) + \Sigma_1 \begin{pmatrix} A \\ B \end{pmatrix} d\xi_1(T) + \Sigma_2 \begin{pmatrix} A \\ B \end{pmatrix} d\xi_2(T),
\end{align*}
\]

where \( \Sigma_j \) are constant matrices given in (2.27), the \( \xi_j \) are independent Brownian motions, and \( \psi_A \) and \( \psi_B \) are the same as in (1.4). Note that the drift terms \( \psi_A, B \) in (1.4)-(1.6) include terms with small delays on the \( T \) time scale. While this may complicate the analysis of the amplitude equations, the fact that the amplitude equations are in terms of the slow time makes simulation over long time scales very efficient, as compared with the original equation.

The equations (1.4)-(1.6) are obtained by considering the noise in terms of a Fourier series-type representation, viewing the noise as fast deterministic oscillations which correspond to resonances with the deterministic frequencies, and coefficients which are Brownian motions on the slow time scale \( T \).

\[
K_0 dW_0(T) + \sum_{j=1}^{\infty} K_{j,1} \cos j \omega t dW_{j,1}(T) + K_{j,2} \sin j \omega t dW_{j,2}(T).
\]

The coefficients \( K_0 \) and \( K_{j,m} \), \( m = 1, 2 \) may be functions of \( x \), depending on whether the noise is additive or multiplicative, and the \( W_0 \), \( W_{j,1} \), \( W_{j,2} \) are independent. Note that the series (1.7) differs from a Fourier-Wiener series (A.14), in which the stochastic coefficients are (time-independent) standard normal random variables. Our goal is to determine evolution equations for the stochastic envelope, so we use a series in which the coefficients are diffusion processes on the slow time \( T \). The resulting amplitude equations include only a finite number of noise terms, since the projection onto the primary modes of the deterministic system (\( \cos \omega t \) and \( \sin \omega t \) in this case) eliminate all but a few of these terms. For example, in (1.6) there are three such terms, as follows from the form of multiplicative noise \( x dW \) in (1.1). In Appendix A we compare moments of the stochastic process obtained via a Fourier-Wiener series with the multi-scale approximation.

The paper is organized as follows. In Section 2 we give a detailed description of the multi-scale analysis, applying it to the linear problem with both additive and multiplicative noise. In Section 3 we show how the method can be used in a nonlinear example, the logistic model with noise, considering both sub- and supercritical values of the delay. In Section 4 we summarize our results and discuss the method in a broader context.
2 Linear case: differential-delay equation with noise

We begin with the linear model, considering both additive and multiplicative noise. We use this problem as a relatively simple setting in which to demonstrate the approach; in later sections we use similar methods, but omit more of the details. There have been a number of previous analyses of linear stochastic delay differential equations with additive noise, primarily concerned with the corresponding Fokker-Planck equation and the invariant density [18]-[21]. For small delays, the approach of [20] and [21] has been extended to nonlinear systems and multiplicative noise. Note that the multi-scale analysis of this paper is not limited to small delays, and is applicable to systems with multiplicative noise and nonlinearities as shown in the following sections.

2.1 Additive noise

We first consider a linear SDDE with small noise.

\[ dx = (\alpha x(t) + \beta x(t - \tau))dt + \delta dw. \]  

(2.1)

Here \( w \) is Brownian motion and we consider the case when \( \delta \ll 1 \), in order to examine sensitivity to small noise. When \( \delta = 0 \), it is well-known that the solution can be written as

\[ x(t) = e^{\lambda t}, \quad \text{Re} \lambda < \langle > 0 \quad \text{for} \quad \tau < \langle > \tau_c. \]  

(2.2)

We assume the parameter values to be such that the system is just below the threshold for growth or decay of oscillatory solutions such as \( x = \cos bt \). Specifically,

\[ b = \sqrt{\beta^2 - \alpha^2}, \quad \beta \cos b\tau_c = \alpha, \quad b = -\beta \sin b\tau_c \]

\[ \tau = \tau_c + \epsilon^2 \tau_2. \]  

(2.3)

Without loss of generality we have defined the threshold as \( \tau_c \), keeping the other parameters fixed. The parameter \( \epsilon \) will be used as a measurement of proximity to this threshold, and we assume that \( 1 \gg \epsilon > 0 \).

In order to capture the influence of the noise over a long time, we look for a periodic solution which has an amplitude that varies stochastically on a slow time scale \( T = \epsilon^2 t \).

\[ \dot{x}(t) = A(T) \cos bt + B(T) \sin bt, \quad T = \epsilon^2 t. \]  

(2.4)

The choice of slow time is not unexpected, since solving (2.1) with (2.3) gives an eigenvalue \( \lambda \) with \( O(\epsilon^2) \) real part.

Now we derive equations for \( A(T) \) and \( B(T) \), assuming the following forms:

\[ dA = \psi_A dT + \sigma_A d\xi_1(T) \]

\[ dB = \psi_B dT + \sigma_B d\xi_2(T). \]  

(2.5)

The coefficients \( \psi_A, \sigma_A, \psi_B, \sigma_B \) are unknown. We determine these coefficients by deriving two equations for \( dx \), first using Ito’s formula, and then using the equation (2.1).
First, using Ito’s formula we have

\[
dx = \frac{\partial x}{\partial t} dt + \frac{\partial x}{\partial A} dA + \frac{\partial x}{\partial B} dB + \frac{\sigma_A^2}{2} \frac{\partial^2 x}{\partial A^2} dT + \frac{\sigma_B^2}{2} \frac{\partial^2 x}{\partial B^2} dT. \tag{2.6}\]

Using the expressions from (2.5), and the fact that the second derivatives of \(x\) with respect to \(A\) and \(B\) vanish, we get

\[
dx = (-bA(T) \sin bt + bB(T) \cos bt) dt + (\psi_A \cos bt + \psi_B \sin bt) dT + \sigma_A \cos bt d\xi_1 + \sigma_B \sin bt d\xi_2. \tag{2.7}\]

Note that in the spirit of multi-scale analysis, we treat the slow time \(T = \epsilon^2 t\) as independent of \(t\). Next we give the expression for \(dx\) by substitution of (2.4) into (2.1)

\[
dx = \left[ -\alpha (A(T) \cos bt + B(T) \sin bt) + \beta \left\{ A(T - \epsilon^2 \tau) (\cos bt \cos b\tau + \sin bt \sin b\tau) + B(T - \epsilon^2 \tau) (\sin bt \cos b\tau - \cos bt \sin b\tau) \right\} \right] dt + \delta dw(t). \tag{2.8}\]

Now we wish to equate these two expressions (2.7) and (2.8) for \(dx\), and thus determine the coefficients \(\psi_A, \psi_B, \sigma_A,\) and \(\sigma_B\). We will restrict our analysis to the case that \(\delta \ll 1\) and \(\epsilon \ll 1\), and use an asymptotic expansion. The validity of this analysis is discussed in Section 4. First we write

\[
\cos b\tau = \cos (b\tau_c + \epsilon^2 \tau_2) \sim \cos b\tau_c - b\epsilon^2 \tau_2 \sin b\tau_c + O(\epsilon^4);
\]

\[
\sin b\tau = \sin (b\tau_c + \epsilon^2 \tau_2) \sim \sin b\tau_c + b\epsilon^2 \tau_2 \cos b\tau_c + O(\epsilon^4). \tag{2.9}\]

Then we substitute (2.9) into the right hand side of (2.8), equate this with the right hand side of (2.7), and collect the coefficients of like powers of \(\epsilon\). Then the \(O(1)\) terms cancel, and we get

\[
(\psi_A \cos bt + \psi_B \sin bt) dT + \sigma_A \cos bt d\xi_1(T) + \sigma_B \sin bt d\xi_2(T) \tag{2.10}
\]

\[
\begin{align*}
&= \left\{ \cos bt \left[ -\beta A(T) \epsilon^2 \tau_2 b \sin b\tau_c - \beta B(T) \epsilon^2 \tau_2 b \cos b\tau_c 
- \frac{\epsilon^2 \beta}{\epsilon^2} \frac{B(T - \epsilon^2 \tau) - B(T)}{e^2} \sin b\tau_c + \frac{\epsilon^2 \beta}{\epsilon^2} \frac{A(T - \epsilon^2 \tau) - A(T)}{e^2} \cos b\tau_c \right] 
+ \sin bt \left[ \beta A(T) \epsilon^2 \tau_2 b \cos b\tau_c - \beta B(T) \epsilon^2 \tau_2 b \sin b\tau_c 
+ \frac{\epsilon^2 \beta}{\epsilon^2} \frac{A(T - \epsilon^2 \tau) - A(T)}{e^2} \sin b\tau_c + \frac{\epsilon^2 \beta}{\epsilon^2} \frac{B(T - \epsilon^2 \tau) - B(T)}{e^2} \cos b\tau_c \right] \right\} dt + O(\epsilon^4) + \delta dw(t) \\
&= \left\{ \left[ b\tau_2 (-\alpha B(T) + bA(T)) + b \frac{B(T - \epsilon^2 \tau) - B(T)}{e^2} + \alpha \frac{A(T - \epsilon^2 \tau) - A(T)}{e^2} \right] \cos bt 
+ \left[ b\tau_2 (\alpha A(T) + bB(T)) - b \frac{A(T - \epsilon^2 \tau) - A(T)}{e^2} + \alpha \frac{B(T - \epsilon^2 \tau) - B(T)}{e^2} \right] \sin bt \right\} dT \\
&+ O(\epsilon^4) + \delta dw(t). \tag{2.11}
\end{align*}
\]
Here we have written $\epsilon^2 dt = dT$ and
\[ A(T - \epsilon^2 \tau) = A(T) + \epsilon^2 \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2}, \]
\[ B(T - \epsilon^2 \tau) = B(T) + \epsilon^2 \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2}, \] (2.12)
and have treated $\epsilon^{-2}[A(T - \epsilon^2 \tau) - A(T)]$ and $\epsilon^{-2}[B(T - \epsilon^2 \tau) - B(T)]$ as $O(1)$.

Then we neglect the $O(\epsilon^4)$ terms and obtain the drift and diffusion coefficients $\psi_A, \psi_B, \sigma_A$ and $\sigma_B$, in the equations for $A$ and $B$, using the method of multi-scale analysis. That is, we project the equations onto $\cos bt$ and $\sin bt$ while treating functions of $T$ as independent of $t$. In order to determine the effect of noise in the slow time variable $T$, we rewrite the noise in terms of two Brownian motions $w_j(T), j = 1, 2$, using standard identities
\[ dw(t) = \sin bt dw_1(t) + \cos bt dw_2(t) \] (2.13)
\[ \delta dw_j(t) = \delta \epsilon^{-1} dw_j(T). \]

We apply the orthogonality of $\cos bt$ and $\sin bt$ to obtain a condition on the noise terms in (2.10)-(2.11), using (2.14), which yields
\[ \int_0^{2\pi/b} \begin{cases} \cos bt \\ \sin bt \end{cases} [\sigma_A \cos bt d\xi_1(T) + \sigma_B \sin bt d\xi_2(T)] dt = \]
\[ \int_0^{2\pi/b} \begin{cases} \cos bt \\ \sin bt \end{cases} \frac{\delta}{\epsilon} [\sin bt dw_1(T) + \cos bt dw_2(T)] dt, \] (2.14)
which implies
\[ \sigma_A = \sigma_B = \frac{\delta}{\epsilon}. \] (2.15)

In the spirit of multi-scale analysis, we have treated the functions of slow time $T$ as independent of $t$, and we equate the independent Brownian motions, $\xi_1(T) = w_1(T), \xi_2(T) = w_2(T)$. Similarly, we compare the drift terms in (2.10)-(2.11), using $\epsilon^2 dt = dT$,
\[ (\psi_A \cos bt + \psi_B \sin bt) dT = \]
\[ \left\{ b\tau_2 (-\alpha B(T) + bA(T)) + b \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2} + \alpha \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} \right\} \cos bt \]
\[ + \left\{ b\tau_2 (\alpha A(T) + bB(T)) - b \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} + \alpha \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2} \right\} \sin bt \] (2.17)
using the orthogonality as in (2.14) to obtain $\psi_A, \psi_B$, in the equations (2.5) for $A$ and $B$
\[ \psi_A = b\tau_2 (-\alpha B(T) + bA(T)) + b \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2} + \alpha \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} \] \[ \psi_B = b\tau_2 (\alpha A(T) + bB(T)) - b \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} + \alpha \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2}. \] (2.18)

The drifts $\psi_A$ and $\psi_B$ correspond to the long time dynamics obtained using a multi-scale analysis for the deterministic problem $\delta = 0$. This correspondence also holds for the non-linear logistic example considered in Section 3.
Given these expressions for $A$ and $B$, the system (2.5) can be written as

\[
\begin{bmatrix}
dA(T) \\
 dB(T)
\end{bmatrix} = \left\{ \mathbf{P} \begin{bmatrix} A(T) \\
 B(T) \end{bmatrix} + \mathbf{Q} \begin{bmatrix} A(T - \tau^2) \\
 B(T - \tau^2) \end{bmatrix} \right\}dT + \frac{\delta}{\epsilon} \begin{bmatrix}
d\xi_1 \\
 d\xi_2
\end{bmatrix}
\]  

(2.19)

We can again view this as a system of linear delay-differential equations, so that the deterministic problem would have a solution of the form $C_0 e^{\Lambda T}$. The eigenvalues of $\mathbf{P} + \mathbf{Q} e^{-\Lambda T} - \Lambda \mathbf{I}$ have negative real parts for $\tau_2 < 0$, and then the process (2.19) can be stationary. Starting with arbitrary initial conditions, the process approaches its stationary behavior at a rate, on the $T$ scale, determined by the real part of these eigenvalues. If $\tau_2 > 0$, then the process (2.19) is not stationary, and the multi-scale approximation is valid only for short times; then the original model (2.1) is dominated by exponential growth.

In Figure 2 we compare the numerical approximation to the density of $x$ using $\hat{x}$ (2.4), and the original problem (2.1) for $\tau_2 < 0$. In order to obtain an approximation for the invariant density $p(x)$ as shown, the simulations must be run for large $t$ (typically $t > 100$). Then the density is constructed from a large number (5000) of such realizations of the process. From this result it is evident that the variance of the process increases with the ratio of $\delta$ to $\epsilon$. In terms of the dynamics, the invariant density indicates the range of amplitudes of the oscillations that are observed as $t \to \infty$. While the amplitude slowly varies over time, the variance of the process indicates the amplitudes that are likely to be observed for all time. As either $\delta$ increases or $\epsilon$ decreases, larger amplitude oscillations are observed more frequently. The analysis suggests that the equations for $A$ and $B$ are valid only for $\delta/\epsilon \ll 1$, since this assumption is underlying the multi-scale ansatz (2.5) and the asymptotic expansion. Additional simulations demonstrate that the agreement is reasonable up to $\delta \sim \epsilon$. This restriction is discussed further in Section 4.

There are clear computational advantages to using the multi-scale approximation (2.4), since the approximation requires the simulation of the equations for $A$ and $B$ (2.5) on the $T$ time scale, rather than the original scale $t$. Then the simulation of the process is much faster using (2.5) rather than the original equation (2.1).

Note that there are delays in the drift terms of the amplitude equations, but they are small delays on the $T$ time scale, so that an analytical approximation to the density is possible. We give some analytical results for small delays in Appendix A. Additional approaches are demonstrated in [20]-[21].

### 2.2 Multiplicative noise

In this section we consider the linear example again, but this time with multiplicative noise,

\[
dx = (-\alpha x(t) + \beta x(t - \tau))dt + \delta x(t)dw. \tag{2.20}
\]

As before, for $\tau$ near $\tau_c$ we look for a multi-scale approximation of the form $\hat{x}(t) = A(T) \cos bt + B(T) \sin bt$ with $T = \epsilon^2 t$, $\tau = \epsilon^2 \tau_2$, $\tau_2 < 0$, and $b$ given in (2.3).

We assume the form of equations for $A$ and $B$

\[
\begin{aligned}
\frac{dA}{dB} &= \left( \begin{array}{c} \psi_A \\ \psi_B \end{array} \right) dT + \Sigma_0 \begin{bmatrix} A \\
 B \end{bmatrix} d\xi_0(T) + \Sigma_1 \begin{bmatrix} A \\
 B \end{bmatrix} d\xi_1(T) + \Sigma_2 \begin{bmatrix} A \\
 B \end{bmatrix} d\xi_2(T), \tag{2.21}
\end{aligned}
\]

where the $\Sigma_j$’s are matrices. In the following we show that the drift coefficients $\psi_A$ and $\psi_B$ are the same as in the additive noise case, but the noise terms are different. The form of the
Figure 2: Comparison of the invariant density $p(x)$ obtained from (2.4) ($\ast$’s) and (2.1) for $\epsilon^2 = .1$ and $\epsilon^2 = .01$. The approximation improves as $\epsilon$ decreases, keeping $1 \gg \epsilon > \delta > 0$. Here $\delta = .01$.

noise terms is not unexpected, since the noise in the original problem is of the form $x \, dw$, so we would expect multiplicative noise of the form $A \, dw_1$ and $B \, dw_2$ in the evolution equations. The appearance of three noise terms follows from the resonances, as shown below. The form of equations is also motivated by the results of Baxendale [9], who found averaged equations of this form for describing the two-point motion and the associated Lyapunov exponent for the stochastic van der Pol-Duffing equation with multiplicative noise. In [12] it was shown that a multi-scale approximation for the stochastic van der Pol-Duffing gives the same result as in [9]. That is, equations are obtained for the slowly varying amplitudes of oscillatory modes corresponding to the Hopf bifurcation. The coefficients in the noise terms are determined from a consistency comparison of the infinitesimal generators for the original and averaged processes, which we discuss below in the context of (2.20).

The matrices $\Sigma_j$ can be obtained using the multi-scale approach as follows. First, we write the noise as a Fourier series-type representation as in (1.7), which is again a Brownian motion [22]. In Appendix B we show that only three terms contribute to the amplitude equations

$$\delta x(t) dw = \delta K_0 x \, dw_0 + \delta K_1 x \cos 2bt \, dw_1 + \delta K_2 x \sin 2bt \, dw_2, \quad (2.22)$$

where $w_j$ are independent Brownian motions. Note that the full Fourier series-type representation includes other terms in addition to those in (2.22), such as $Kx \sin kbt \, dw_k$ for $k > 2$, and we could include such terms here. However, under the projection (2.14) onto $\cos bt$ and $\sin bt$ combined with the the multi-scale approach which treats $w_k(t)$ as $\epsilon^{-1} w_k(T)$, these terms will not contribute to the stochastic amplitude equations. That is, the terms in (2.22) correspond to resonances when using the Fourier-series type representation, as in (1.7).

In Appendix B we give the details of the derivation of the amplitude equations,

$$\begin{pmatrix} dA \\ dB \end{pmatrix} = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} dT + \frac{\delta}{\epsilon} \begin{pmatrix} K_0 & K_1 \\ -K_0 & K_2 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} d\xi_0(T) + \frac{K_1}{2} \begin{pmatrix} A \\ -B \end{pmatrix} d\xi_1(T) + \frac{K_2}{2} \begin{pmatrix} B \\ A \end{pmatrix} d\xi_2(T), \quad (2.23)$$

9
We obtain the same drift terms (2.18) as in the linear case with additive noise, but the noise terms are different. While we have the form of these terms, the coefficients $K_j$ have not yet been determined. In order to find these coefficients we look for consistency in the generator for the process (2.23) with the averaged generator for the process (2.20). In particular, the diffusion operator in the generator corresponding to the noise in (2.23) is given by summing the entries of $\frac{\delta^2}{2\epsilon^2}M$ where

$$M \equiv K_0^2 \left( \begin{array}{cc} A^2 \frac{\partial^2}{\partial x^2} & AB \frac{\partial^2}{\partial x \partial A} \\ AB \frac{\partial^2}{\partial A \partial x} & B^2 \frac{\partial^2}{\partial x^2} \end{array} \right) + \frac{K_2^2}{4} \left( \begin{array}{cc} A^2 \frac{\partial^2}{\partial x^2} & -AB \frac{\partial^2}{\partial A \partial B} \\ -AB \frac{\partial^2}{\partial B \partial A} & B^2 \frac{\partial^2}{\partial x^2} \end{array} \right) \quad (2.24)$$

$$+ \frac{K_2^2}{4} \left( \begin{array}{cc} B^2 \frac{\partial^2}{\partial x^2} & AB \frac{\partial^2}{\partial B \partial A} \\ AB \frac{\partial^2}{\partial A \partial B} & A^2 \frac{\partial^2}{\partial x^2} \end{array} \right).$$

Then we compare this with the diffusion operator corresponding to the noise process $\delta x dw(t) = \delta t^{-1} x dw(T)$ in (2.20), written in terms of $A$ and $B$

$$\frac{\delta^2}{2\epsilon^2} \frac{\partial^2}{\partial x^2} = \frac{\delta^2}{2\epsilon^2} (A \cos bt + B \sin bt)^2 \left( \cos \frac{bt}{\epsilon} \frac{\partial}{\partial A} + \sin \frac{bt}{\epsilon} \frac{\partial}{\partial B} \right)^2. \quad (2.25)$$

Averaging this over one period of the fast oscillation while treating $A$ and $B$ as constants, yields the diffusion operator given by summing the entries of

$$\frac{\delta^2}{2\epsilon^2} \left( \begin{array}{c} \frac{3}{8} A^2 + \frac{1}{8} B^2 \frac{\partial^2}{\partial A^2} \\ \frac{1}{4} AB \frac{\partial^2}{\partial A \partial B} \\ \frac{1}{4} \frac{\partial^2}{\partial B^2} \frac{3}{8} B^2 + \frac{1}{8} A^2 \end{array} \right). \quad (2.26)$$

Comparing with (2.24) we find that $K_0 = 1/2$, $K_1 = K_2 = 1/\sqrt{2}$ so that the matrices $\Sigma_j$ in (2.20) are given by

$$\Sigma_0 = \frac{\delta}{\epsilon^2} \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right), \quad \Sigma_1 = \frac{\delta}{\epsilon \sqrt{2}} \left( \begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right), \quad \Sigma_2 = \frac{\delta}{\epsilon \sqrt{2}} \left( \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right),$$

and $\xi_j(T) = w_j(T)$.

Hence we see that we keep three terms in the Fourier series representation of $x(t) dw$, corresponding to resonances with the intrinsic oscillation related to the bifurcation, in order to describe the stochastic modulations. This result is similar to [9] and [12] which study multiplicative noise in the stochastic van der Pol-Duffing equation. The combination of the condition of keeping all terms which correspond to a resonance with the primary mode and the condition of consistency with the averaged generator specifies this representation of the diffusion in terms of three independent Brownian motions. For one-point motion in van der Pol-Duffing, studied in [11], consistency under the averaging was obtained using only two terms and a different approach.

We compare the simulation of the multi-scale approximation based on (2.21) with the solution of the original problem (2.20) in Figure 3. Approximations for the probability density $p(x, t)$ are constructed using 5000 realizations.
Figure 3: Comparison of the densities obtained from the multi-scale approximation of $x$ given by $A(T) \cos bt + B(T) \sin bt$ from (2.21) (solid line) and numerical simulation of (2.20) (dash-dotted lines) for $t = 1000$, $\beta = -\sqrt{2}$, $\alpha = 1$, $\tau_2 = -1$, $\epsilon = .1$, and $\delta = .03$ and $\delta = .1$. Here the initial condition on $[-\tau, 0]$ is $x = .25 \cos bt$.

3 Logistic equation with delay and noise

Next we consider the logistic differential delay equation with noise

$$dx = (rx(s)(1 - x(s - \tau)))ds + \delta dw. \quad (3.1)$$

We first review some results for the deterministic case $\delta = 0$. We consider values of $r$ for $r > 0$, where the solution $x = 0$ is unstable. Then the oscillations about $x = 1$ decay for subcritical delays ($\tau < \tau_c$) and are sustained for supercritical values ($\tau > \tau_c$), where $r\tau_c = \pi/2$. In the supercritical case, there is a stable limit cycle which is discussed in Section 3.2. In order to look for the reduced model for the noisy case on the long time scale, we first write the problem for $y(t) = x(s) - 1$, with $s = t/\sigma$,

$$\sigma dy = -ry(t - \sigma t)(1 + y(t)) dt + \delta dW(t), \quad (3.2)$$

where $W(t) = w(t/\sigma)$ and $\delta = \hat{\delta} \sqrt{\sigma}$. The factor $\sigma$ is a phase factor which is unity for the subcritical case, but in the supercritical case it is of the form $\sigma = 1 + \epsilon^2 \sigma_2$.

3.1 Subcritical case $\tau < \tau_c$

We consider the case when $\tau < \tau_c$, so that, in the absence of noise, $x \to 1$ as $t \to \infty$. Then we set

$$y = \epsilon[A(T) \cos rt + B(T) \sin rt] + \epsilon^2[C(T) \cos 2rt + D(T) \sin 2rt] + O(\epsilon^3) \quad (3.3)$$

$$\tau = \tau_c + \epsilon^2 \tau_2, \quad \sigma = 1, \quad \tau_2 < 0. \quad (3.4)$$
The functions $C(T)$ and $D(T)$ are quadratic functions of $A(T)$ and $B(T)$ (see (C.1)). They are the coefficients of the higher order harmonics which are slaved to the dominant mode $\cos rt$, $\sin rt$ through the nonlinearities. As shown in Appendix C, they are chosen to eliminate the $O(\epsilon^2)$ terms in the multi-scale analysis, leaving the evolution equations at $O(\epsilon^3)$. In the following, we omit the argument $T$ from $A(T)$, $B(T)$, $C(T)$, $D(T)$, except when the argument is at the delayed time $T - \epsilon^2 \tau$.

Following the same approach as in the linear additive noise case in Section 2.1, we look for equations of the form (2.5) for $A$ and $B$. Using Ito’s formula we get an equation for $dy$ replacing $x$ with $y$ in (2.6). The resulting equation (C.2) is given in Appendix C. Note that in this case $y_{AA} \neq 0$ and $y_{BB} \neq 0$.

We get a second equation (C.3) for $dy$ by substitution of (3.3) into (3.2). In Appendix C we show that after subtracting (C.3) from (C.2), there is cancellation up to $O(\epsilon^3)$ in the drift terms. Comparing the remaining drift terms and noise terms as in the linear case, using the orthogonality of $\cos rt$ and $\sin rt$ (2.14), and substituting $\epsilon^2 dt = dT$ and treating functions of $T$ as independent of $t$ under the multi-scale assumption, we get

$$\psi_A = r^2 \tau_2 A + r \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2} - \frac{r}{2} \left[ AD - BC \right] + \frac{r}{2} \left[ AC + BD \right]$$

$$\psi_B = r^2 \tau_2 B - r \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} + \frac{r}{2} \left[ AC + BD \right] - \frac{r}{2} \left[ - AD + BC \right].$$

Furthermore, by rewriting the noise as in (2.13)-(2.14), we get

$$\epsilon \sigma_A d\xi_1(T) = \frac{\delta}{\epsilon} dw_1(T) \quad \epsilon \sigma_B d\xi_2(T) = \frac{\delta}{\epsilon} dw_2(T).$$

In Figure 4 we compare the simulation of the original model (3.1) with the result for (3.3) given by the amplitude equation. We have run the simulations for sufficiently large $t$, so that we approximate the invariant density $p(x)$.

### 3.2 Supercritical case $\tau > \tau_c$

In this section we consider the case $\tau > \tau_c$, for $r\tau_c = \frac{\pi}{2}$ and $\tau = \tau_c + \epsilon^2 \tau_2$ for $\tau_2 > 0$. Then the stable oscillatory behavior for $\delta = 0$ can be approximated by

$$x \sim 1 + \epsilon \cos(\sigma \tau)$$

$$\sigma \sim 1 + \epsilon^2 \sigma_2, \quad \sigma_2 = -\frac{3}{20}, \quad \epsilon = \sqrt{\frac{\tau - \tau_c}{\tau_2}}$$

$$\tau_2 = -\frac{1}{20r} - \sigma_2 \tau_c.$$
Figure 4: Comparison of the invariant density \( p(x) \) in the subcritical case \( \tau < \tau_c \ (\tau_2 = -1) \), constructed from 5000 realizations. The multi-scale approximation of \( x \) is \( 1 + A(T) \cos rt + B(T) \sin rt \), with \( A \) and \( B \) given by (2.5) with (3.5) and (3.6) (‘s and o’s). It is compared with the numerical simulation of (3.1) (solid lines) for \( r = 1 \). On the left, \( \epsilon = .2 \), and \( \delta = .005 \) (‘s) and \( \delta = .002 \) (o’s). On the right \( \delta = .005 \) and \( \epsilon = .2 \) (‘s) and \( \epsilon = .07 \) (o’s).

3.2.1 Nonlinear evolution equation

We write

\[
\begin{align*}
    y &= \epsilon [A(T) \cos rt + B(T) \sin rt] + \epsilon^2 [C(T) \cos 2rt + D(T) \sin 2rt] + O(\epsilon^3) \\
    \tau &= \tau_c + \epsilon^2 \tau_2, \quad \tau_2 > 0.
\end{align*}
\]

(3.10) (3.11)

The parameters \( \epsilon \) and \( \sigma_2 \) are related to \( \tau_2 \) as defined above. Then \( A^2(0) + B^2(0) = 1 \), as follows from the choice of \( \sigma \) and \( \epsilon \). Given the expression of the oscillatory solution in (3.8), we have \( A(0) = 1 \) and \( B(0) = 0 \) for convenience. One could also write the oscillation as a combination of \( \sin rt \) and \( \cos rt \), so that both \( A(0) \) and \( B(0) \) are non-zero, which is equivalent to introducing a phase shift in the form of the deterministic oscillation.

Following the same procedure as in the previous cases, we find that

\[
\begin{align*}
    \psi_A &= -\sigma_2 r B + r^2 (\tau_2 + \sigma_2 \tau_c) A + r \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2} \\
    &\quad - \frac{r}{2} \left[ AD - BC \right] + \frac{r}{2} \left[ AC + BD \right] \\
    \psi_B &= \sigma_2 r A + r^2 (\tau_2 + \sigma_2 \tau_c) B - r \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} \\
    &\quad + \frac{r}{2} \left[ AC + BD \right] - \frac{r}{2} \left[ -AD + BC \right] \\
    \sigma_A &= \sigma_B = \frac{\delta}{\epsilon^2}.
\end{align*}
\]

(3.12) (3.13)

Where the arguments are omitted for \( A, B, C, \) and \( D, \) the argument is simply \( T \). The details of the derivation are shown in Appendix C.

In the absence of noise \( \delta = 0 \), there is a steady state solution to the equations \( A'(T) = \psi_A \) and \( B'(T) = \psi_B \), given by \( A = 1, B = 0 \) and \( \sigma \) given by (3.8). That is, the drift gives attracting
dynamics to this steady state for $A$ and $B$, and the noise gives fluctuations about this steady state. As we show in the next section, one can derive linear evolution equations describing fluctuations about this steady state. The range of parameters for which this approximation is valid is discussed further in Section 4.

The comparison for the simulation of the original model with the amplitude equation is given in Figure 5. The simulations were run for $t$ large, so that we can approximate the invariant density $p(x)$ for the model.

![Comparison of the simulation of the original model with the amplitude equation](image)

**Figure 5**: Comparison of the approximation of the invariant density $p(x)$ obtained from 5000 realizations for large $t$ of (3.10) ($*$’s and ‘o’$’s) and (3.1) (solid lines) for $r = 1$. On the left, $\epsilon = .2$ and $\delta = .005$. On the right $\epsilon = .2$ and $\delta = .002$. Note that the asymptotic form (3.10) captures the bimodal shape of the density, and reflects that it becomes less pronounced as $\delta/\epsilon^2$ increases.

### 3.2.2 Linear evolution equation

In the previous section we noted that the stochastic amplitudes behave as fluctuations about the deterministic steady state for these amplitudes. This leads to an alternative derivation of amplitude equations for the deviation from $x = 1 + \epsilon \cos rt$, writing $x = 1 + y$ with

$$
y = \epsilon \left[ (1 + A(T)) \cos rt + B(T) \sin rt \right] +$$

$$\epsilon^2 \left[ (C_0 + C(T)) \cos 2rt + (D_0 + D(T)) \sin 2rt \right].$$

(3.14)

The constants $C_0 = 1/5$ and $D_0 = 1/10$ follow from the previous analysis, setting $A = 1$ and $B = 0$ in (C.1). We assume that $A$ and $B$ are stochastic quantities, again described by equations of the form (2.5). In this case we take

$$C(T) = \frac{2}{5} A(T) - \frac{1}{5} B(T) + O(\epsilon^2)$$

$$D(T) = \frac{1}{5} A(T) + \frac{2}{5} B(T) + O(\epsilon^2).$$

(3.15)
We follow the procedure outlined for the previous cases to get

\[
\psi_A = -\sigma_2 r B + r^2 (\tau_2 + \sigma_2 \tau_C) A + r \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2} + \frac{3r}{20} (A - B) \quad (3.16)
\]

\[
\psi_B = \sigma_2 r A + r^2 (\tau_2 + \sigma_2 \tau_C) B - r \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} + \frac{r}{20} (9A + B) .
\]

These equations are simply (3.12) with

\[
A(T) = 1 + A(T) \quad B(T) = B(T) \quad (3.17)
\]

\[
C(T) = C_0 + C(T) \quad D(T) = D_0 + D(T) , \quad (3.18)
\]

linearized about \( A = B = 0 \). In this case we keep only linear terms, since this is sufficient to describe the variation about the deterministic solution \( x \sim 1 + \epsilon \cos rt \). The diffusion coefficients are the same as in (3.13). The results compare well with the nonlinear evolution equation of the previous subsection.

4 Discussion

We have adapted multi-scale analysis methods to study stochastic envelopes of oscillations in stochastic delay equations. The multi-scale analysis allows a separation of stochastic and deterministic effects by looking for long-time behavior of the stochastic envelopes of deterministic oscillations. The results allow us to study the long time behavior of the stochastic process with delay, and determine critical scalings for sustained oscillations which appear deterministic but are due purely to the presence of noise. For example, from Figure 2 we see that the invariant density has a variance which increases as the noise increases, or as the proximity to the critical delay \( \tau_c \) decreases, even when this delay is \( \tau < \tau_c \). In the deterministic case these oscillations would decay exponentially with time. However, for additive noise the behavior of the invariant density with non-trivial variance indicates that the oscillations do not decay for long time, but rather they have an envelope that varies over a significant range. The resonance of the noise with the natural modes of the system causes oscillations to be sustained for delays which are subcritical, but near critical. This is true even when the noise is very small, since the magnitude of the oscillations also depends on the size of \( \tau_c - \tau \). In the case of multiplicative noise, where the oscillations in the presence of noise do decay over time, the noise can slow their decay so that the density has a significant variance for large but finite time.

There are a number of challenges in studying the behavior of stochastic dynamics with delays. The delay causes the system to be infinite dimensional, which complicates the analysis. The multi-scale method uses the projection onto the basis for the deterministic process, as used in dynamical studies of delay systems without noise [23], which allows an approach which appears two-dimensional. Indeed, as we see from the results for the stochastic evolution equations, it is not possible to remove the delays from these equations. However, under the multi-scale method the delays are small, so that approximations are possible and simulation is efficient on the long time scale. The projection used in the multi-scale method has been used in systems without delays [12], where it was shown to be related to other averaging or projection methods which have been used for stochastic systems. However, these other methods have not been used to derive averaged equations in the context of delays.
We have shown that the multi-scale method can be applied for problems with either linear and non-linear drifts. This is not surprising, since the method has been used for deterministic problems of this type without delays, and does not depend on the form of the drift. It is also valuable for both additive and multiplicative noise. The form of the stochastic envelope equation differs depending on the type of non-linearity and the type of noise, but the multi-scale approach is the same for each case.

We also show that the multi-scale method can be used in both the sub- and super-critical cases, as in Section 3. In the super-critical case, we show that the envelope equations describe stochastic fluctuations about a non-trivial steady state amplitude. Then the density has a bimodal form due to the underlying deterministic oscillations. Again the variance and overall shape of the bimodal distribution is sensitive to small noise for delays near critical, as shown in Section 3.2. In this paper we restrict our attention to small amplitude oscillations for the non-linear case. We extend the multi-scale method to large amplitude oscillations in [24].

The drift terms for these envelope equations correspond to those obtained by applying a multi-scale approach to the deterministic problem. To determine the diffusion terms for the envelope equations, one writes the noise in terms of resonant modes with slowly varying stochastic coefficients. The reduced system retains a finite number of these terms, as it focuses on resonances on the fast time scale. In the case of additive noise, the form is given using two independent Brownian motions. In the case of multiplicative noise, the system requires the use of three independent Brownian motions, due to additional resonances in the multiplicative noise.

Another aspect of the multi-scale analysis is the rescaling of the noise in the long time equation. For multiplicative noise the scaling is given by $\delta/\epsilon$. This is the same scaling as found for the stochastic van der Pol-Duffing case [9]-[12]. For additive noise in the linear case, the scaling is also given by $\delta/\epsilon$.

In the nonlinear case of Section 3, the scaling is given by $\delta/\epsilon^2$, which also corresponds to the scaling used for additive noise in the stochastic van der Pol-Duffing model [9]. The scaling $\delta/\epsilon^2$, as compared to $\delta/\epsilon$ in the linear case, follows from the fact that we seek small amplitude solutions $x \sim \epsilon A \cos t + \epsilon B \sin t + \ldots$ in the non-linear case of Section 3. We recover the $\delta/\epsilon$ scaling by denoting $\tilde{A} = \epsilon A$ and $\tilde{B} = \epsilon B$ in the linear evolution equations (3.16), which become,

$$\begin{bmatrix} d\tilde{A} \\ d\tilde{B} \end{bmatrix} = \begin{bmatrix} \psi_{\tilde{A}} \\ \psi_{\tilde{B}} \end{bmatrix} dT + \frac{\delta}{\epsilon} \begin{bmatrix} d\xi_1 \\ d\xi_2 \end{bmatrix} \quad (4.1)$$

with $\psi_{\tilde{A},\tilde{B}}$ given by replacing $A$ and $B$ with $\tilde{A}$ and $\tilde{B}$, respectively, in (3.16). Under a similar rescaling $\tilde{A} = \epsilon A$, $\tilde{B} = \epsilon B$ for the nonlinear evolution equations (3.12), the coefficient of the noise is again $\delta/\epsilon$, and the nonlinear terms like $AD$ and $BC$ become $\epsilon^{-2}\tilde{A}\tilde{D}$ and $\epsilon^{-2}\tilde{B}\tilde{C}$. While these terms may appear large, in fact the dynamics described by the drift are dissipative about the steady state, so these terms do not result in large fluctuations. The dissipative nature of the amplitudes is evidenced by the good agreement between the nonlinear and linear evolution equations given in Section 3. There we showed that the linear evolution equations correspond to a linearization of the nonlinear evolution equations about the attracting steady state amplitudes of the deterministic dynamics.

The diffusion coefficients $\delta/\epsilon$ and $\delta/\epsilon^2$ in the envelope equations indicate the amplification of the noise over the long time dynamics. This scaling also indicates the validity of the multi-scale approximation. Implicit in the multi-scale approach is the assumption that there is a
competition between the deterministic dynamics, described by the drift, and the stochastic
dynamics. Under this assumption, one looks for solutions which are described on the fast scale
by deterministic oscillations, and on the slow scale by the stochastic amplitudes. However, if the
noise dominates the dynamics in the original equation, then the multi-scale assumption is not
valid; the fluctuations on the fast scale will dominate deterministic oscillations on the fast scale,
and the separation in terms of two time scales will not be possible. This is also true if the noise
dominates the dynamics on the slow scale. The stochastic effects do not dominate the dynamics
of the envelope on the slow time scale when the noise is relatively small, e.g. \( \delta/\epsilon \ll 1 \) for the
linear cases, and \( \delta/\epsilon^2 \ll 1 \) for the nonlinear case. For values of the noise \( \delta \gg \epsilon \), the stochastic
effects dominate, and the multi-scale approximation is no longer appropriate. In practice, from
the simulations we find that the multi-scale approximation gives a good approximation of the
dynamics as long as the diffusion coefficients are not large: that is, for the linear case, for \( \epsilon \)
and for the nonlinear case, for \( \delta \sim \epsilon^2 \), the multi-scale approximation still yields reasonable
results.

A Additional comments on the linear case with additive noise

We revisit the linear example (2.5) of the previous case. First we give the analytical results for
the density for \( A \) and \( B \) in the case of stationarity, as considered in Section 2. Then we compare
the covariance for the multi-scale approximation with the original system.

The system (2.5) is a special case of the system of linear delay stochastic differential equations

\[
\frac{d}{dt} \mathbf{R}(t) = \mathbf{P} \mathbf{R}(t) dt + \mathbf{Q} \mathbf{R}(t - \tau) dt + \mathbf{D} \, dw(t),
\]

where

\[
\mathbf{P} = \begin{bmatrix}
\beta^2 \tau_2 - \alpha/\epsilon^2 & -\alpha \beta \tau_2 - b/\epsilon^2 \\
\alpha \beta \tau_2 + b/\epsilon^2 & \beta^2 \tau_2 - \alpha/\epsilon^2
\end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix}
\alpha/\epsilon^2 & b/\epsilon^2 \\
-b/\epsilon^2 & \alpha/\epsilon^2
\end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix}
\delta/\epsilon & 0 \\
0 & \delta/\epsilon
\end{bmatrix},
\]

and \( \tau = \epsilon^2 \tau_2 \). First we state general results on the covariance of the process \( \mathbf{R}(t) \) in (A.1) and
then we specify them to the system (2.5).

We define by \( \Phi(t) \) the fundamental solution to the deterministic system

\[
\frac{d}{dt} \Phi(t) = \mathbf{P} \, \Phi(t) + \mathbf{Q} \, \Phi(t - \tau)
\]

\( \Phi(0) = \mathbf{I}, \quad \Phi(t) = \mathbf{0} \) for \( t < 0 \). (A.3)

Below we use the fact that under conditions for stationarity, the covariance of the process \( \mathbf{R}(t) \)
also satisfies (A.3). The covariance \( \mathbf{K}(t_1, t_2) \) of the process \( \mathbf{R}(t) \) is given by

\[
\mathbf{K}(t_1, t_2) = \int_0^{\min(t_1, t_2)} \Phi(t_1 - s) \mathbf{D} \mathbf{D}^T \Phi(t_2 - s) \, ds + \Phi(t_1) \text{cov}(\mathbf{R}(0), \mathbf{R}(0)) \Phi^T(t_2). \quad (A.5)
\]

If all roots to the corresponding characteristic equation of (A.3) have negative real parts
then the process \( \mathbf{R}(t) \) is stationary and \( \mathbf{K}(t_1, t_2) = \mathbf{K}(t_1 - t_2) = \mathbf{K}(u) \) (if an initial condition
\( \mathbf{R}(0) \) is suitably chosen or as \( t \to \infty \)). In this case we conclude from (A.5) that

\[
\mathbf{D} \mathbf{D}^T + \mathbf{P} \mathbf{K}(0) + \mathbf{Q} \mathbf{K}(\tau) + \mathbf{K}(\tau) \mathbf{P}^T + \mathbf{K}(\tau) \mathbf{Q}^T = 0.
\]
We note that when $Q \equiv 0$, the expressions (A.3), (A.5) and (A.6) reduce to the well known results for linear stochastic systems [25].

Since $K(u)$ satisfies (A.3) then (A.3) and (A.6) imply that

$$\frac{d}{du} K(0) = -\frac{1}{2} DD^T. \quad (A.7)$$

Also we find

$$\frac{d^2}{du^2} K(u) = PK(u)P^T - Q K(u)Q^T + P \frac{d}{du} K(u) - \frac{d}{du} K(u)P^T \quad \text{for } u \in [0, \bar{\tau}]. \quad (A.8)$$

Then (A.6), (A.7), and (A.8) generalize results of [18] to systems of linear SDDE's.

Now we calculate the covariance of the process $R(T) \equiv [A(T), B(T)]$ defined in (2.5) with $P$, $Q$ and $D$ given by (A.2). We solve (A.3) for $K(u)$ with the initial condition (A.7). For $\tau_2 < 0$, $\beta > 0$ and $\alpha > 0$ the roots of the characteristic equation are complex conjugates $\lambda = \mu \pm i\nu$, and the corresponding eigenvectors of the matrix $P + Q \exp(-\lambda \tau)$ are $\mathbf{v}_1 = \mathbf{v} \pm i\mathbf{w}$.

Then the solution to (A.3) with the initial condition (A.7) is given by

$$K(u) = \frac{1}{2} e^{\mu u} \left[ \begin{array}{c} \mathbf{v} \cos(\nu u) - \mathbf{w} \sin(\nu u) \\ \mathbf{v} \sin(\nu u) + \mathbf{w} \cos(\nu u) \end{array} \right] \left[ \begin{array}{c} \mu \mathbf{v} - \nu \mathbf{w} \\ \mu \mathbf{w} - \nu \mathbf{v} \end{array} \right]^{-1}. \quad (A.9)$$

In the asymptotic limit of $\epsilon \ll 1$ the characteristic equation for (A.3) yields

$$\lambda_{1,2} = \frac{b^2 \tau_2}{1 + b^2 \tau^2 + (1 + \alpha \tau)^2} \pm \frac{b \tau_2 (\alpha + \tau (\alpha^2 + \beta^2))}{b^2 \tau_2^2 + (1 + \alpha \tau)^2} + O(\epsilon^2) \equiv \mu_0 \pm i\nu_0 + O(\epsilon^2). \quad (A.10)$$

We write

$$K(u) = -\frac{1}{2} \frac{\delta^2}{\epsilon^2 \tau \beta^2} \left( 1 + O(\epsilon^3) \right) e^{(\mu_0 + O(\epsilon^2))u} \left[ \begin{array}{cc} \cos(\nu_0 + O(\epsilon^2))u & -\sin(\nu_0 + O(\epsilon^2))u \\ \sin(\nu_0 + O(\epsilon^2))u & \cos(\nu_0 + O(\epsilon^2))u \end{array} \right]. \quad (A.11)$$

In particular, we note that to leading order

$$\text{var} (A(T) \cos(bt) + B(T) \sin(bt)) = -\frac{\delta^2}{\epsilon^2 \tau \beta^2}, \quad (A.12)$$

where $b = \sqrt{\beta^2 - \alpha^2}$.

We also calculate the exact variance of the process $x(t)$ in (2.1) using (A.3) specified for $t = 0$ and (A.7) as initial conditions for (A.8)

$$\text{var} x(t) = \frac{\delta^2 b - \beta \sin[b(\tau + \epsilon^2 \tau_2)]]}{2b \alpha - \beta \cos[b(\tau + \epsilon^2 \tau_2)]} = -\frac{\delta^2}{\epsilon^2 \tau \beta^2} + \frac{\delta^2}{12 \tau \epsilon^2} + O(\epsilon^6). \quad (A.13)$$

We note that approximating the delay terms in (2.18) by first order derivatives yields a 2-dim stochastic process whose covariance matches to leading order the covariance in (A.9). Then this approximation gives $A$ and $B$ as coupled Ornstein-Uhlenbeck processes. In the case of the nonlinear models, as in Section 3, the approximation by first order derivatives is not appropriate [26].
The multi-scale approximation together with the stationarity of $A$ and $B$ suggests that similar results could be obtained using a Fourier-Wiener representation for $w(t)$ on an appropriate finite time interval $t = [0, T]$ 

$$dw(t) = \sqrt{\frac{2}{T}} \sum_{n=1}^{\infty} f_n \sin b_n t + g_n \cos b_n t \quad \text{for } t \in [0, T], \tag{A.14}$$

where $f_n$ and $g_n$ are independent identically distributed normal random variables [27], and $T = O(\epsilon^{-2})$. Here $b_n = n b$ for $b$ defined in (2.3). Note that the coefficients in (A.14) are not time dependent. We will look for the long time behavior of $x$, using a representation with a finite number of terms from this series. Then we assume a Fourier series form for $x(t)$,

$$x(t) = \sum c_n \cos b_n t + d_n \sin b_n t, \tag{A.15}$$

where the coefficients $c_n$ and $d_n$ are time-independent random variables. In order to compare with the long-time behavior of the multi-scale approximation we substitute (A.15) and (A.14) into the linear equation (2.1), taking the case $\alpha = 0$ and $\beta = -1$, for simplicity. This yields

$$-\epsilon^2 b^2 \tau_2 (c_1 \cos bt + d_1 \sin bt) + \sum_{n=2}^{\infty} \hat{c}_n \cos b_n t + \hat{d}_n \sin b_n t = \delta \sqrt{\frac{2}{T}} \left[ f_1 \cos bt + g_1 \sin bt + \sum_{n=2}^{\infty} f_n \cos b_n t + g_n \sin b_n t \right], \tag{A.16}$$

where $\hat{c}_n$ and $\hat{d}_n$ are $O(1)$ quantities. Comparing coefficients in (A.16) yields

$$c_1 = -\sqrt{\frac{2}{T}} \frac{\delta}{\epsilon^2 b^2 \tau_2} f_1 + O(\epsilon^3)$$

$$d_1 = -\sqrt{\frac{2}{T}} \frac{\delta}{\epsilon^2 b^2 \tau_2} g_1 + O(\epsilon^3) \quad \text{for } \tau_2 < 0$$

$$c_n = O(\delta), \quad d_n = O(\delta), \quad (n \neq 1). \tag{A.17}$$

Then we see that for the characteristic time scale $T = 2/(\epsilon^2 b^2 \tau_2)$, we recover the leading order result (A.12) using a finite number of terms in the Fourier series (A.14).

**B  Details for the case of multiplicative noise**

As in the case of additive noise, we use Ito’s formula to get (2.6) and the substitution of $\dot{x} = A(T) \cos bt + B(T) \sin bt$ into (2.20), equating these two expressions for $dx$ to get equations for the drift and diffusions coefficients. The expansion of the drift terms follows the same steps as in the additive noise case. Then we use the projection (2.14) to solve for $A$ and $B$, which are the same as in the Section 2.1 (2.18).

For the noise terms, we start by rewriting them as in (2.22)

$$\delta x(t) dw = \delta K_0 x \, dw_0 + \delta K_1 x \cos 2bt \, dw_1 + \delta K_2 x \sin 2bt \, dw_2 + \ldots, \tag{B.1}$$

where the form of the terms given by $+ \ldots$ is given in (1.7). We again write $w_j(t) = w_j(T)/\epsilon$ in (2.22) and substitute $\dot{x} = A(T) \cos bt + B(T) \sin bt$ into (1.7). Then we apply the projection onto
\cos bt \text{ and } \sin bt \text{ as in (2.14) treating } w_j(T) \text{ as independent of } t \text{ under the multi-scale assumption. This yields two expressions, }

\begin{align*}
\frac{b}{\pi} \int_0^{2\pi/b} \cos bt \left[ \frac{\delta}{\epsilon} (K_0 x \, dw_0(T) + K_1 x \cos 2bt \, dw_1(T) + K_2 x \sin 2bt \, dw_2(T) + \ldots ) \right] dt &= \frac{\delta}{\epsilon} \left[ K_0 A + \frac{1}{2} K_1 A + \frac{1}{2} K_2 B \right] \\
\frac{b}{\pi} \int_0^{2\pi/b} \sin bt \left[ \frac{\delta}{\epsilon} (K_0 x \, dw_0(T) + K_1 x \cos 2bt \, dw_1(T) + K_2 x \sin 2bt \, dw_2(T) + \ldots ) \right] dt &= \frac{\delta}{\epsilon} \left[ K_0 B - \frac{1}{2} K_1 B + \frac{1}{2} K_2 A \right].
\end{align*}

(B.2)

Note that the additional terms + \ldots \text{ in (B.1) that are proportional to } \sin mbt \text{ and } \cos mbt \text{ for } m > 2, \text{ do not contribute to (B.2). Then we see that the form of the equations (2.21) is indeed consistent with (B.2), with the matrices } \Sigma_i \text{ of the form}

\begin{align*}
\Sigma_0 &= \frac{\delta}{\epsilon} K_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\
\Sigma_1 &= \frac{\delta}{\epsilon} K_1 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \\
\Sigma_2 &= \frac{\delta}{\epsilon} K_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.
\end{align*}

(B.3)

Next we compare the diffusion terms in the generators to determine the constants } \mathcal{K}_j. \text{ For the stochastic amplitudes } A \text{ and } B, \text{ these terms are given in (2.24). We compare these terms with the averaged diffusion terms for } x = A \cos bt + B \sin bt. \text{ In particular, we calculate the average of (2.25)}

\begin{align*}
\frac{\delta^2}{2\epsilon^2} \frac{b}{\pi} \int_0^{2\pi/b} \left( A^2 \cos^2 bt + 2AB \cos bt \sin bt + B^2 \sin^2 bt \right) \cdot \\
\left( \cos^2 bt \frac{\partial^2}{\partial A^2} + 2 \cos bt \sin bt \frac{\partial}{\partial A} + \sin^2 bt \frac{\partial^2}{\partial B^2} \right) dt &= \frac{\delta^2}{2\epsilon^2} \left\{ \left( \frac{3}{8} A^2 + \frac{1}{8} B^2 \right) \frac{\partial^2}{\partial A^2} + \frac{1}{2} AB \frac{\partial}{\partial A} \frac{\partial}{\partial B} + \left( \frac{3}{8} B^2 + \frac{1}{8} A^2 \right) \frac{\partial^2}{\partial B^2} \right\}.
\end{align*}

(B.4)

Then we compare this result with (2.24), which yields (2.26) and the matrices } \Sigma_i \text{ given in (2.27).

C \quad Details for the logistic equation

We start with the two equations for } dy \text{ obtained from Ito’s formula and substitution for the logistic case. We derive the nonlinear evolution equations for both the subcritical } \tau_2 < 0 \text{ and supercritical } \tau_2 > 0 \text{ case. In both cases we take}

\begin{align*}
C(T) &= \frac{1}{5} (A^2 - B^2 - AB) \\
D(T) &= \frac{1}{10} (A^2 - B^2 + 4AB).
\end{align*}

(C.1)

This choice leads to some cancellation, which we describe below. From Ito’s formula (2.6) we have,

\begin{align*}
dy &= \epsilon (-rA(T) \sin rt + rB(T) \cos rt - 2erC(T) \sin 2rt + 2erD(T) \cos 2rt) dt +
\end{align*}
\[
\begin{align*}
\epsilon \cos rt &+ \epsilon^2 \left( \frac{\partial C}{\partial A} \cos 2rt + \frac{\partial D}{\partial A} \sin 2rt \right) \left( \psi_A dT + \sigma_A d\xi_1(T) \right) \\
+ \epsilon \sin rt &+ \epsilon^2 \left( \frac{\partial C}{\partial B} \cos 2rt + \frac{\partial D}{\partial B} \sin 2rt \right) \left( \psi_B dT + \sigma_B d\xi_2(T) \right) \\
+ \epsilon^2 &\left( \frac{\sigma_A^2}{2} \left[ \frac{\partial^2 C}{\partial A^2} \cos 2rt + \frac{\partial^2 D}{\partial A^2} \sin 2rt \right] + \frac{\sigma_B^2}{2} \left[ \frac{\partial^2 C}{\partial B^2} \cos 2rt + \frac{\partial^2 D}{\partial B^2} \sin 2rt \right] \right) dT,
\end{align*}
\]
and from substitution of (3.3) into (3.2) we have

\[
\sigma dy = -\epsilon r [A \sin rt - B \cos rt] \\
- \epsilon^2 r \left[ -C \cos 2rt - D \sin 2rt - A^2 \sin rt \cos rt - B^2 \cos rt \sin rt + AB(\sin^2 rt - \cos^2 rt) \right] \\
- \epsilon^3 r \left[ r(\tau_2 + \sigma_2 \tau_c)(-A \cos rt - B \sin rt) \right] \\
+ \left( \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} \sin rt + \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2} \cos rt \right) \\
+ (A \sin rt - B \cos rt) (C \cos 2rt + D \sin 2rt) \\
+ (A \cos rt + B \sin rt) (-C \cos 2rt - D \sin 2rt) \right] + O(\epsilon^4).
\]

We have used the fact that \( r \tau_c = \frac{\pi}{2} \) and have written

\[
A(T - \epsilon^2 \tau) = A(T) + \epsilon^2 \left( \frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} \right)
\]

\[
B(T - \epsilon^2 \tau) = B(T) + \epsilon^2 \left( \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2} \right),
\]
assuming that

\[
\frac{A(T - \epsilon^2 \tau) - A(T)}{\epsilon^2} = O(1), \quad \frac{B(T - \epsilon^2 \tau) - B(T)}{\epsilon^2} = O(1).
\]

Then we substitute (C.4) and (C.1) into (C.3), equating (C.2) and (C.3). Then the \( O(\epsilon) \) and \( O(\epsilon^2) \) terms cancel, using (C.5) and the definition of \( C \) and \( D \) (C.1). The equations for \( \psi_A, \psi_B, \sigma_A \) and \( \sigma_B \) are obtained by projecting (C.2)-(C.3) onto the modes \( \sin rt \) and \( \cos rt \), as in (2.14). The diffusion terms are written as in (2.13), and we integrate with respect to the fast time scale, treating the functions of \( T \) as independent of \( t \). Then using the projections (2.14) we identify \( \xi_1(T) = w_1(\epsilon^2 t) \) and \( \xi_2(T) = w_2(\epsilon^2 t) \) which yields (3.6). From the drift terms we get (3.5) for the subcritical case and (3.12) for the supercritical case.

References


