J. Math. Biol. (2011) 63:433–457
DOI 10.1007/s00285-010-0376-2

Mathematical Biology

Sustained oscillations for density dependent Markov processes

Peter H. Baxendale · Priscilla E. Greenwood

Received: 4 June 2010 / Revised: 30 September 2010 / Published online: 13 November 2010

Abstract Simulations of models of epidemics, biochemical systems, and other bio-systems show that when deterministic models yield damped oscillations, stochastic counterparts show sustained oscillations at an amplitude well above the expected noise level. A characterization of damped oscillations in terms of the local linear structure of the associated dynamics is well known, but in general there remains the problem of identifying the stochastic process which is observed in stochastic simulations. Here we show that in a general limiting sense the stochastic path describes a circular motion modulated by a slowly varying Ornstein–Uhlenbeck process. Numerical examples are shown for the Volterra predator–prey model, Sel'kov's model for glycolysis, and a damped linear oscillator.

Keywords Sustained oscillations · Density dependent Markov processes · Ornstein–Uhlenbeck process · Stochastic averaging · Martingale problem

Mathematics Subject Classification (2000) 92C45 · 80A30 · 60J25 · 60H10

P. Baxendale supported in part by NSF Grant DMS-05-04853. P. Greenwood supported by the Statistical and Applied Mathematical Sciences Institute, Research Triangle Park, N.C., and the Mathematical, Computational and Modeling Sciences Center at Arizona State University.

P. H. Baxendale (✉)
Department of Mathematics, University of Southern California,
Los Angeles, CA, USA
e-mail: baxendal@usc.edu

P. E. Greenwood
Department of Mathematics, University of British Columbia,
Vancouver, BC, Canada
e-mail: Priscilla.Greenwood@asu.edu

1 Introduction

McKane et al. (2007) considered stochastic models for gene regulation and for ATP and ADP concentrations during phosphorylation. They translated equations of chemical reaction type into ordinary differential equations for the densities of molecules of four types in the system. Exact simulation of the associated stochastic model, obtained by interpreting jump rates as stochastic shows sustained random oscillations around a locally stable limit point of the deterministic system. In such examples the sets of deterministic equations would not be acceptable models for systems showing phenomena like biochemical oscillations. There are no limit cycles and no periodic input. Sustained oscillations, illustrated in Figure 7(a,b) of Kuske et al. (2007), result from the combination of random effects with dynamics in which a deterministic model has a locally stable fixed point to which the system converges via damped oscillations. It is intuitively clear that the combination of such dynamics with stochastic fluctuations will produce sample paths which do not converge, but which cycle around the locally stable point in some irregular fashion as in our Figs. 3 and 5. In Aparicio and Solari (2001) and Natiello and Solari (2007) an argument explaining sustained oscillations in terms of a Liapunov function of a deterministic model is given. This phenomenon appears also in spatial models, for example in a lattice based model (Morita and Tainaka 2006), and in an epidemic on a small world network (Telo da Gama and Nunes 2006). The semi-regular oscillations observed in sample-path plots of simulations are one dimensional time plots of irregular cycling in a two dimensional model.

Our object here is to identify this irregular cycling as, approximately, a specific stochastic process. We show that, in a limiting sense, the two dimensional stochastic path describes, up to an identified matrix multiple, a circular motion modulated by a slowly varying Ornstein–Uhlenbeck process.

Kuske et al. (2007) identified the sustained oscillations in an SIR epidemic model as approximating this same process, by a multi-scale analysis argument which has also been used in a number of similar contexts (Kłosek and Kuske 2005; Yu et al. 2006).

Most familiar biological system models are in the class we call density dependent, following Kurtz (1978). In Sect. 2, with the predator–prey model as an example, we sketch how, for large systems, such a model is associated with a stochastic diffusion, either by a stochastic argument as in Kurtz (1978) or by an expansion of the Kolmogorov equation due to van Kampen. The likeness of the results of these very different methods allows their alternative use, in particular when our attention is focussed, as here, on a stable fixed point of our system.

The Kurtz approach (Kurtz 1978) in which the original Markov chain model, normalized by system size N, is approximated for large system size by a diffusion model represented by a system of stochastic differential equations has the advantage that Kurtz has given us a pathwise bound on the error of approximation of order $\log N/N$. The overall standard deviation of the stochastic system is of order $N^{-1/2}$, and the limit of this system, indeed of the normalized Markov chain model, is deterministic and represented by a system of ordinary differential equations.

Having begun with a biosystem such as predator–prey, we are ready at the outset of Sect. 3 with an appropriate linear stochastic diffusion equation describing the local behavior of our model near a fixed point. Three changes of variables transform this

equation to one for a diffusion which has an interesting but simple stochastic limit, approached as the ratio of the slow decay of deterministic oscillations to the comparatively fast frequency goes to zero. We argue that when this ratio is small, as it is in the several examples of stochastically sustained oscillations, (e.g. Kuske et al. 2007; McKane and Newman 2005; McKane et al. 2007), the stochastic process observed in each biological example should be close to the one obtained from our limit description after reversing the changes of variable.

Evidence that our approximation is close to the solution of the local stochastic model near the fixed point is given by comparison, in a number of examples in Sect. 4, of computed power spectral densities. We give numerical results, with chosen parameter values, for the linear oscillator, for the predator–prey model of McKane and Newman (2005), and for the simple model of the phosphorylation step in glycolysis, due to Sel'kov, as studied by McKane et al. (2007).

2 Density dependent processes

The states in our Markov process model will be vectors X in $\mathbb{Z}^d$, where X represents the numbers of individuals in certain classes within a population. For example in an SIR epidemic model, X is the vector consisting of the numbers of susceptible, infected and recovered individuals in a population. In a chemical reaction model X is the vector consisting of the numbers of molecules of each of the chemicals involved in the reactions. For a specific example we will consider the predator–prey model described in McKane and Newman (2005). The stochastic system can be obtained from a corresponding deterministic system by taking some or all of the deterministic rates to be, instead, rates at which random interactions occur. The state of the system then changes due to random interactions between individuals. In non-closed systems the state can also change by individuals entering or leaving the system. Such events we will call interactions also.

Each type of interaction causes the vector X to jump by a vector $\mathbf{r} \in \mathbb{Z}^d$. The stochastic rates $W_N(X;\mathbf{r})$ at which these jumps $X \to X + \mathbf{r}$ take place are the jump rates in the Markov model. Here N is a large parameter which represents in some way the overall size of the system. The (finite) collection of possible jumps $\mathbf{r}$ and the corresponding jump rates $W_N(X;\mathbf{r})$ characterize the Markov process model. In particular, they provide exactly the information needed to simulate the process.

The examples we consider are Markov processes with jump rates which scale according to

$$W_N(X;\mathbf{r}) = N\Phi\left(\frac{1}{N}X;\mathbf{r}\right). \tag{1}$$

This is a special case of Van Kampen (1992, eqn (X.2.3)). Since N represents the size of the system, the vector $\frac{1}{N}X$ appearing in the function Φ represents the densities of the different classes within the population. Markov process with jump rates of the form (1) are called density dependent processes. Sometimes the term 'density-dependent' is used in a broader sense.

2.1 A predator–prey example

In the predator–prey model of McKane and Newman (2005), at a given time the state of the individual level model consists of n predators and m prey. The transition rates $T(n', m'|n, m)$ from state (n, m) to the state (n', m') are given in McKane and Newman (2005, eqn (1)):

$$\begin{aligned} T(n-1, m|n, m) &= d_1 n, \\ T(n, m+1|n, m) &= 2b\frac{m}{N}(N-n-m), \\ T(n, m-1|n, m) &= 2p_2\frac{nm}{N} + d_2 m, \\ T(n+1, m-1|n, m) &= 2p_1\frac{nm}{N}. \end{aligned}$$

This is a density dependent process with possible jumps $\mathbf{r} = (-1, 0)$,$(0, 1)$, $(0, -1)$, $(1, -1)$ and rates $N\Phi(\frac{1}{N}X; \mathbf{r})$ where

$$\Phi((x, y); \mathbf{r}) = \begin{cases} d_1 x & \text{if } \mathbf{r} = (-1, 0) \\ 2by(1-x-y) & \text{if } \mathbf{r} = (0, 1) \\ 2p_2 xy + d_2 y & \text{if } \mathbf{r} = (0, -1) \\ 2p_1 xy & \text{if } \mathbf{r} = (1, -1) \end{cases} \tag{2}$$

where $x = n/N$ and $y = m/N$.

2.2 From Markov jump process to diffusion

There are two roughly equivalent paths which take us, for large N, from a Markov jump process defined by jump rates as in (1), exemplified here by (2), to a diffusion process. Whichever path one takes, the move to a diffusion will be part of our total argument which identifies sustained oscillations in the models we consider. One way is to follow Kurtz (1978) who focuses on the Markov process itself. The other is to follow Van Kampen (1992, Chap. X) who focuses instead on the forward Kolmogorov (or Fokker–Planck, or "master") equation which describes how the probability distribution of the Markov process evolves. We give a brief parallel account of these two approaches to clarify their relationship and their role in the present paper. Additional discussion can be found in the books of Gardiner (1990, Sect. 7.2) and Ethier and Kurtz (1986, Chap. 11). Some literature on our topic of sustained oscillations (Kuske et al. 2007) has involved the Kurtz development, while some other papers (Alonso et al. 2007; Di Patti and Fanelli 2009; McKane and Newman 2005; McKane et al. 2007) take the van Kampen route.

2.2.1 Method of Kurtz

The starting point for Kurtz (1978) is to represent the $\mathbb{Z}^d$-valued Markov jump process determined by (1) as

$$X(t) = X(0) + \sum_{\mathbf{r}} \mathbf{r} N^{(\mathbf{r})} \left(\int_0^t N\Phi\left(\frac{1}{N}X(s); \mathbf{r}\right) ds \right)$$

Here the sum is over the finite set of possible jumps $\mathbf{r}$, and $\{N^{(\mathbf{r})}(t) : t \geq 0\}$ is a collection of independent rate 1 scalar Poisson processes. The density dependent rate $N\Phi(\frac{1}{N}X(t); \mathbf{r})$ at which a jump $\mathbf{r}$ occurs appears as a time change within the Poisson process $N^{(\mathbf{r})}$. Dividing by N gives

$$X^N(t) = X^N(0) + \sum_{\mathbf{r}} \mathbf{r} \frac{1}{N} N^{(\mathbf{r})} \left(\int_0^t N\Phi\left(X^N(s); \mathbf{r}\right) ds \right) \tag{3}$$

where $X^N(t) \equiv \frac{1}{N}X(t)$. Replacing each compensated Poisson process $N^{(\mathbf{r})}(t) - t$ in (3) with a scalar Brownian motion $\widetilde{W}^{(\mathbf{r})}(t)$ gives the equation

$$\begin{aligned} \widetilde{X}^N(t) = \widetilde{X}^N(0) &+ \int_0^t \left(\sum_{\mathbf{r}} \mathbf{r}\Phi\left(\widetilde{X}^N(s); \mathbf{r}\right) \right) ds \\ &+ \sum_{\mathbf{r}} \mathbf{r} \frac{1}{N} \widetilde{W}^{(\mathbf{r})} \left(\int_0^t N\Phi\left(\widetilde{X}^N(s); \mathbf{r}\right) ds \right) \end{aligned}$$

and so the process $\widetilde{X}^N(t)$ satisfies a stochastic differential equation of the form

$$d\widetilde{X}^N(t) = \sum_{\mathbf{r}} \mathbf{r}\Phi\left(\widetilde{X}^N(t); \mathbf{r}\right) dt + \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} \mathbf{r} \sqrt{\Phi\left(\widetilde{X}^N(t); \mathbf{r}\right)} dW^{(\mathbf{r})}(t). \tag{4}$$

(The scalar Brownian motions $\widetilde{W}^{(\mathbf{r})}(t)$ and $W^{(\mathbf{r})}(t)$ are related through a random time change with rate $N\Phi(\widetilde{X}^N(t), \mathbf{r})$.) Kurtz uses a coupling of the compensated Poisson processes with the Brownian motions to show that the error introduced on bounded intervals of time by replacing $X^N(t)$ by $\widetilde{X}^N(t)$ is of order $(\log N)/N$ as $N \to \infty$.

For $z \in \mathbb{R}^d$ define

$$F_i(z) = \sum_{\mathbf{r}} \Phi(z; \mathbf{r})\mathbf{r}_i \quad \text{and} \quad B_{i,j}(z) = \sum_{\mathbf{r}} \Phi(z; \mathbf{r})\mathbf{r}_i\mathbf{r}_j. \tag{5}$$

Then $F(z) = (F_1(z), \ldots, F_d(z))$ is the vector field of means of the first term of the right side of (4) and $B(z) = \{B_{i,j}(z)\}_{1 \leq i,j \leq d}$ is the covariance function arising in the

second term. They are multidimensional versions of van Kampen's first and second jump moment functions $\alpha_{1,0}$ and $\alpha_{2,0}$, see (Van Kampen, 1992, eqn. (X.**2**.13)). In terms of this notation the defining Eq. (4) for $\widetilde{X}^N(t)$ can be written as

$$d\widetilde{X}^N(t) = F(\widetilde{X}^N(t))dt + \frac{1}{\sqrt{N}}C(\widetilde{X}^N(t))dW(t) \tag{6}$$

where $W(t)$ is a d-dimensional Brownian motion and the $d \times d$ matrix function $C(z)$ is chosen so that $C(z)C(z)^* = B(z)$. From Eq. (6) we can identify the differential equation

$$\frac{d\phi(t)}{dt} = F(\phi(t)) \tag{7}$$

for the deterministic limit $\phi(t)$, say, of $\frac{1}{N}X(t)$ as $N \to \infty$. The ODE (7) is the one obtained in the purely deterministic model for the macroscopic density using the reaction rates $\Phi(x; \mathbf{r})$. However it is the stochastic interpretation using the functions $\Phi(x; \mathbf{r})$ to determine the rates (1) in the jump Markov process $X(t)$ which gives rise to the second term on the right of (6). This term describes the order $1/\sqrt{N}$ fluctuations of $\frac{1}{N}x(t)$ away from $\phi(t)$, and is the term which can generate sustained oscillations in certain parameter regimes.

2.2.2 Method of van Kampen

Van Kampen starts with the Ansatz that the density process $\frac{1}{N}X(t)$ should have a deterministic limit as $N \to \infty$ and random fluctuations of order $1/\sqrt{N}$ away from this limit. Thus he writes

$$\frac{1}{N}X(t) = \phi(t) + \frac{1}{\sqrt{N}}\xi(t)$$

where $\phi(t)$ is deterministic and the stochasticity in the model is contained in the process $\xi(t)$. The information in (1) about the jump rates for $X(t)$ gives rise to the following Kolmogorov (or master) equation for the probability density function $\Pi(z, t)$ of the process $\xi(t)$.

$$\begin{aligned}
&\frac{\partial \Pi}{\partial t}(z,t) - \sqrt{N}\sum_{i=1}^{d}\frac{\partial \Pi}{\partial z_i}(z,t)\frac{d\phi_i}{dt}(t) \\
&\quad = N\sum_{\mathbf{r}}\left[\Phi\left(\phi(t) + \frac{z}{\sqrt{N}} - \frac{\mathbf{r}}{N}; \mathbf{r}\right)\Pi\left(z - \frac{\mathbf{r}}{\sqrt{N}}, t\right)\right. \\
&\qquad \left. -\Phi\left(\phi(t) + \frac{z}{\sqrt{N}}; \mathbf{r}\right)\Pi(z,t)\right].
\end{aligned}$$

Van Kampen's expansion in powers of $N^{-1/2}$, which involves the assumption that effects of a possibly discrete initialization of the process have been lost, yields

$$\frac{\partial \Pi}{\partial t}(z,t) = N^{1/2} \sum_{i=1}^{d} \frac{\partial \Pi}{\partial z_i}(z,t) \left(\frac{d\phi_i}{dt}(t) - F_i(\phi(t)) \right) - \sum_{i,j=1}^{d} \frac{\partial F_i}{\partial z_j}(\phi(t)) \frac{\partial}{\partial z_i} \left(z_j \Pi(z,t) \right) + \frac{1}{2} \sum_{i,j=1}^{d} B_{ij}(\phi(t)) \frac{\partial^2 \Pi}{\partial z_i z_j}(z,t) + O\left(N^{-1/2} \right).$$

The requirement that the term of order $N^{1/2}$ vanishes gives the Eq. (7) above. Thus van Kampen's deterministic limit agrees with Kurtz' deterministic limit, although the methods used are rather different. Continuing with van Kampen's expansion, the terms of order N^0 match up to yield the equation

$$\frac{\partial \Pi}{\partial t}(z,t) = - \sum_{i,j=1}^{d} \frac{\partial F_i}{\partial z_j}(\phi(t)) \frac{\partial}{\partial z_i} \left(z_j \Pi(z,t) \right) + \frac{1}{2} \sum_{i,j=1}^{d} B_{ij}(\phi(t)) \frac{\partial^2 \Pi}{\partial z_i z_j}(z,t). \quad (8)$$

This is the multidimensional version of Van Kampen (1992, eqn X.4.1). Expression (8) is the Kolmogorov equation for the process $\xi(t)$ given by

$$d\xi(t) = DF(\phi(t))\xi(t)dt + C(\phi(t))dW(t). \quad (9)$$

Terminating the van Kampen expansion as in (8) gives the approximation

$$\frac{1}{N} X(t) \approx \phi(t) + \frac{1}{\sqrt{N}} \xi(t) \quad (10)$$

in distribution as $N \to \infty$, where $\phi(t)$ is the deterministic solution of the ordinary differential equation (7) and $\xi(t)$ is given by the linear stochastic differential equation (9).

2.2.3 Diffusion near a stable fixed point

So far we have made no assumptions about the behavior of the deterministic process $\phi(t)$ as $t \to \infty$. Suppose now that the ordinary differential equation (7) has a single fixed point z_{eq}, say, and that $\phi(t) \to z_{\text{eq}}$ as $t \to \infty$ for all starting positions $\phi(0)$. We can study the fluctuations of $\frac{1}{N} X(t)$ away from z_{eq} using either of the methods above. Using the method of van Kampen, we can replace $\phi(t)$ with its limiting value z_{eq} in the Eq. (9) to obtain, for large t,

$$d\xi(t) = A_0 \xi(t)dt + C_0 dW(t) \quad (11)$$

where $A_0 = DF(z_{\text{eq}})$ and $C_0 = C(z_{\text{eq}})$, so that $C_0 C_0^* = B(z_{\text{eq}}) = B_0$, say. Alternatively, using the method of Kurtz we can linearize the Eq. (6) at z_{eq} and rescale by a factor $1/\sqrt{N}$ to obtain, for large t,

$$\frac{1}{N} X(t) \approx \widetilde{X}^N(t) \approx z_{\text{eq}} + \frac{1}{\sqrt{N}} \widetilde{\xi}(t)$$

where

$$d\widetilde{\xi}(t) = DF(z_{\text{eq}})\widetilde{\xi}(t)dt + C(z_{\text{eq}})dW(t).$$

In this setting the methods of Kurtz and van Kampen give the same approximation

$$\frac{1}{N} X(t) \approx z_{\text{eq}} + \frac{1}{\sqrt{N}} \xi(t) \tag{12}$$

where $\xi(t)$ is given by (11)

Suppose now $d = 2$ and that A_0 has a pair of complex eigenvalues $-\lambda \pm i\omega$. Since A_0 is the linearization of the vector field F at z_{eq}, the deterministic trajectories $\phi(t)$ spiral in towards the stable fixed point z_{eq}. In the large N limit this is all we will see. However if N is large but still finite then we see also the effect of the process $\xi(t)$, modulated by the factor $1/\sqrt{N}$. The equation for the process $\xi(t)$ has the same linear term $A_0\xi(t)dt$ as the linearized equation for $\phi(t)$ near z_{eq}, but it also has the noise term $C_0 dW(t)$ which creates non-trivial stationary behavior. It is the interplay between the two terms which we study in the next section.

3 A simple approximation of a linear diffusion

In this section we will obtain a simple approximate description of the solution $\xi(t)$ of the diffusion equation

$$d\xi(t) = A_0\xi(t)dt + C_0 dW(t) \tag{13}$$

in dimension 2 in the case when the matrix A_0 has complex eigenvalues $-\lambda \pm i\omega$ with $0 < \lambda \ll \omega$. The distribution of the noise is determined by the covariance matrix $B_0 = C_0 C_0^*$.

3.1 Preparation

First we diagonalize the matrix A_0 in order to see more clearly the separate effects of the (slow) dilation due to λ and the (fast) rotation due to ω. Let Q be a 2×2 matrix such that

$$Q^{-1} A_0 Q = \begin{bmatrix} -\lambda & \omega \\ -\omega & -\lambda \end{bmatrix} \equiv A,$$

and write $Y(t) = Q^{-1}\xi(t)$. Then

$$dY(t) = AY(t)\,dt + C\,dW(t) \tag{14}$$

where $C = Q^{-1}C_0$. The noise in Eq. (14) has covariance matrix

$$B = CC^* = Q^{-1}C_0C_0^*(Q^{-1})^* = Q^{-1}B_0(Q^{-1})^*.$$

Define

$$\sigma^2 = \frac{1}{2}\mathrm{tr}(B) = \frac{1}{2}(B_{11} + B_{22}) = \frac{1}{2}\sum_{i,j=1}^{2} C_{ij}^2.$$

Next, in order to study the motion of $Y(t)$ relative to the fast rotation due to ω, we write

$$Y(t) = R_{-\omega t}Z(t)$$

where the matrix

$$R_s = \begin{bmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{bmatrix}$$

corresponds to rotation through the angle s. From the SDE (14) for $Y(t)$ we obtain (using Itô's formula)

$$dZ(t) = -\lambda Z(t)dt + R_{\omega t}C\,dW(t).$$

Finally, in order to compare the process $Z(t)$ with a standardized two-dimensional Ornstein–Uhlenbeck process, we rescale in time and space by writing

$$U(t) = \frac{\sqrt{\lambda}}{\sigma}Z(t/\lambda).$$

Using the substitution $s = u/\lambda$ we have

$$\begin{aligned} U(t) - U(0) &= \frac{\sqrt{\lambda}}{\sigma}\left(-\lambda\int_0^{t/\lambda} Z(s)\,ds + \int_0^{t/\lambda} R_{\omega s}C\,dW(s)\right) \\ &= \frac{\sqrt{\lambda}}{\sigma}\left(-\int_0^{t} Z(u/\lambda)\,du + \frac{1}{\sqrt{\lambda}}\int_0^{t} R_{\omega u/\lambda}C\,d\widetilde{W}(u)\right) \\ &= -\int_0^{t} U(u)\,du + \int_0^{t} R_{\omega u/\lambda}D\,d\widetilde{W}(u) \end{aligned}$$

where $D = (1/\sigma)C$, so that $\mathrm{tr}(DD^*) = 2$, and $\widetilde{W}(t) = \sqrt{\lambda}W(t/\lambda)$ is a standard Brownian motion. Therefore

$$dU(t) = -U(t)dt + R_{\omega t/\lambda} D\, d\widetilde{W}(t). \tag{15}$$

3.2 Averaging

In the SDE (15) suppose that the parameter ω is a function of λ, and that the covariance matrix DD^* depends in some way on λ subject only to the condition that $\mathrm{tr}(DD^*) = 2$. We write $U(t) = U^\lambda(t)$ to denote the dependence on λ. The process $U^\lambda(t)$ defined by (15) has generator

$$(\mathcal{L}^\lambda f)(x,t) = -\sum_{i=1}^{2} x_i \frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{2} a_{ij}(t)\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

where

$$a_{ij}(t) = \left[R_{\omega t/\lambda} DD^* R_{-\omega t/\lambda}\right]_{ij}.$$

Each entry of $a_{ij}(t)$ is a linear combination of fast varying terms of the form $\cos^2 \omega t/\lambda$ or $\sin^2 \omega t/\lambda$ or $(\sin \omega t/\lambda)(\cos \omega t/\lambda)$. If we replace each of these terms by its time average we obtain the constant matrix $\overline{a} = I$ (for details of this calculation see later), and if we replace the diffusion matrix $a_{ij}(t)$ in the generator $\mathcal{L}^\lambda$ by its time average $\overline{a} = I$ we obtain the generator

$$\overline{\mathcal{L}} f(x) = -\sum_{i=1}^{2} x_i \frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i=1}^{2} \frac{\partial^2 f}{\partial x_i^2}(x).$$

of the standard 2-dimensional Ornstein–Uhlenbeck process. A rigorous result based on these observations follows.

Theorem 1 *For each fixed T and $x \in \mathbb{R}^2$ the distribution of $\{U^\lambda(t) : 0 \le t \le T\}$ given by (15) with $U^\lambda(0) = x$ converges as $\lambda/\omega \to 0$ to the distribution of the standard 2-dimensional Ornstein–Uhlenbeck process $\{S(t) : 0 \le t \le T\}$ generated by the SDE*

$$dS(t) = -S(t)\, dt + dW(t)$$

with $S(0) = x$.

To prove the weak convergence of $U^\lambda(\cdot)$ with generator $\mathcal{L}^\lambda$ to $S(\cdot)$ with generator $\overline{\mathcal{L}}$ we will use a martingale problem convergence argument. First we introduce some terminology and background.

Let $\mathcal{A}$ be the differential generator of a d-dimensional diffusion process $X(t)$:

$$\mathcal{A}f(x,t) = \sum_{i=1}^{d} b_i(x,t)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x,t)\frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

We say that a process $X(\cdot)$ solves the **martingale problem** for the operator $\mathcal{A}$ if

$$f(X(t)) - f(X(0)) - \int_0^t \mathcal{A}f(X(s),s)\,ds \tag{16}$$

is a martingale for each f in the space $C_0^\infty(\mathbb{R}^d)$ of smooth functions on $\mathbb{R}^d$ with compact support, see Stroock and Varadhan (1979). Itô's formula shows that a solution of

$$dX(t) = b(X(t),t)dt + \sigma(X(t),t)dW(t)$$

satisfies the martingale problem for $\mathcal{A}$, where $a(x,t) = \sigma(x,t)(\sigma(x,t))^*$. For our purposes here, it is important to note that the martingale problem for the Ornstein–Uhlenbeck operator $\overline{\mathcal{L}}$ defined above has a solution which is unique in law, see Stroock and Varadhan (1979, Thm 10.2.2).

If we wish to prove for a family of processes X^λ with corresponding generators $\mathcal{A}^\lambda$ that any limit process is the solution of the martingale problem for $\mathcal{A}$, one might expect a proof that shows the martingales

$$f(X^\lambda(t)) - f(X^\lambda(0)) - \int_0^t \mathcal{A}^\lambda f(X^\lambda(s),s)\,ds$$

converging to the martingale (16). However in many cases, including the present one, this argument does not work directly. Happily, it was found (Papanicolaou et al. 1977; Kushner 1984) that a sufficient condition for weak convergence can involve a suitable perturbation of each test function f. A simple version of (Kushner, 1984, Theorem 2, p. 44) which we will use is

Proposition 1 *Suppose the martingale problem for the diffusion generator $\mathcal{A}$ has a unique solution for each initial condition. Fix $T > 0$. Let $\mathcal{A}^\lambda$ be a family of differential operators for diffusion processes X^λ. Suppose for each f in a dense subset of $C_0^2(\mathbb{R}^d)$ there exists a corresponding family $F^\lambda(x,t)$ in the domain $D(\mathcal{A}^\lambda)$ of $\mathcal{A}^\lambda$ such that as $\lambda \to 0$,*

$$\left|F^\lambda(x,t) - f(x)\right| \to 0 \quad \textit{uniformly on } 0 \le t \le T, \tag{17}$$

and

$$\left|\left(\frac{\partial}{\partial t}+\mathcal{A}^{\lambda}\right)F^{\lambda}(x,t)-\mathcal{A}f(x,t)\right|\to 0 \quad \textit{uniformly on } 0\le t\le T. \tag{18}$$

Suppose that the laws of the processes X^{λ} are tight on $[0,T]$, and $X^{\lambda}(0)\to x_0$. Then the processes $X^{\lambda}(\cdot)$ converge weakly to $X(\cdot)$ on $[0,T]$, where X is the unique solution to the martingale problem for $\mathcal{A}$ with initial condition $X(0)=x_0$.

Proof Note that since the test function f is a function of x only, there is no need to include a $\frac{\partial f}{\partial t}$ term in (18). An application of Itô's formula shows that

$$F^{\lambda}(X^{\lambda}(t),t)-F^{\lambda}(X^{\lambda}(0),0)-\int_0^t\left(\frac{\partial}{\partial s}+\mathcal{A}^{\lambda}\right)F^{\lambda}(X^{\lambda}(s),s)\,ds$$

is a martingale. This martingale property can be expressed as: for any finite set of times $t_1,t_2,\ldots,t_k\le t<u$ and smooth bounded functions $h:\mathbb{R}^{kd}\to\mathbb{R}$,

$$E\left(h(X^{\lambda}(t_i),i\le k)\left[F^{\lambda}(X^{\lambda}(u),u)-F^{\lambda}(X^{\lambda}(t),t)-\int_t^u\left(\frac{\partial}{\partial s}+\mathcal{A}^{\lambda}\right)F^{\lambda}(X^{\lambda}(s),s)\,ds\right]\right)=0$$

Our hypotheses allow us to replace F^{λ} by f and $(\frac{\partial}{\partial t}+\mathcal{A}^{\lambda})F^{\lambda}$ by $\mathcal{A}f$ to obtain

$$\mathbb{E}\left(h(X^{\lambda}(t_i),i\le k)\left[f(X^{\lambda}(u),u)-f(X^{\lambda}(t),t)-\int_t^u\mathcal{A}f(X^{\lambda}(s),s)\,ds\right]\right)=\delta(\lambda)$$

where $\delta(\lambda)\to 0$ as $\lambda\to 0$. Consider a weakly convergent subsequence of $X^{\lambda}(\cdot)$, converging to $X(\cdot)$, say. Then

$$\mathbb{E}\left(h(X(t_i),i\le k)\left[f(X(u),u)-f(X(t),t)-\int_t^u\mathcal{A}f(X(s),s)\,ds\right]\right)=0.$$

Therefore $X(\cdot)$ solves the martingale problem for $\mathcal{A}$, and hence is the unique process with generator $\mathcal{A}$ and initial value $X(0)=x_0$. □

Proof of Theorem 1 The coefficients in the SDE (15) are all globally Lipschitz continuous, uniformly in λ, so tightness is standard. It remains to check the uniqueness of any subsequential limit of the U^{λ}, and we will do this using Proposition 1. The only difference between the generators $\mathcal{L}^{\lambda}$ and $\overline{\mathcal{L}}$ is in the diffusion coefficient.

For brevity of notation, write $\omega t/\lambda = u$ and $DD^* = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$ where $\alpha + \gamma = \mathrm{tr}(DD^*) = 2$ and $\beta^2 \le \alpha\gamma$. The matrix $a(t) = a_{ij}(t)$ of diffusion coefficients in $\mathcal{L}^\lambda$ is

$$\begin{bmatrix} \alpha\cos^2 u - 2\beta \sin u \cos u + \gamma \sin^2 u & (\alpha-\gamma)\sin u\cos u + \beta(\cos^2 u - \sin^2 u) \\ (\alpha-\gamma)\sin u\cos u + \beta(\cos^2 u - \sin^2 u) & \alpha \sin^2 u + 2\beta \sin u \cos u + \gamma \cos^2 u \end{bmatrix}.$$

The time average of this matrix is

$$\overline{a} = \frac{1}{2}\begin{bmatrix} \alpha+\gamma & 0 \\ 0 & \alpha+\gamma \end{bmatrix} = I,$$

and this is the matrix of diffusion coefficients in $\overline{\mathcal{L}}$. The difference $a_{ij}(t) - \overline{a}_{ij}$ has time average 0 and so it may be written as the time derivative of some periodic function $\widehat{a}_{ij}(t)$. In fact

$$a_{ij}(t) = \overline{a}_{ij} + \frac{d\widehat{a}_{ij}}{dt}(t)$$

where

$$\widehat{a}(t) = \frac{\lambda}{4\omega}\begin{bmatrix} (\alpha-\gamma)\sin 2u + 2\beta\cos 2u & (\gamma-\alpha)\cos 2u + 2\beta \sin 2u \\ (\gamma-\alpha)\cos 2u + 2\beta \sin 2u & (\gamma-\alpha)\sin 2u - 2\beta\cos 2u \end{bmatrix}.$$

Notice that each entry in this matrix has absolute value at most

$$\frac{\lambda}{4\omega}\sqrt{(\alpha-\gamma)^2 + 4\beta^2} \le \frac{\lambda}{4\omega}\sqrt{(\alpha-\gamma)^2 + 4\alpha\gamma} = \frac{\lambda}{2\omega}. \tag{19}$$

For any $f \in C_0^4(\mathbb{R}^2)$ define

$$F^\lambda(x,t) = f(x) - \frac{1}{2}\sum_{i,j=1}^{2} \widehat{a}_{ij}(t)\frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

The estimate (19) gives the first condition (17) in Proposition 1. Moreover

$$\begin{aligned}\left(\frac{\partial}{\partial t} + \mathcal{L}^\lambda\right) F^\lambda(x,t) &= -\frac{1}{2}\sum_{i,j=1}^{2} \frac{d\widehat{a}_{ij}}{dt}(t)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \mathcal{L}^\lambda f(x) \\ &\quad -\frac{1}{2}\sum_{i,j=1}^{2} \widehat{a}_{ij}(t)\mathcal{L}^\lambda\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right) \\ &= \overline{\mathcal{L}} f(x) - \frac{1}{2}\sum_{i,j=1}^{2} \widehat{a}_{ij}(t)\mathcal{L}^\lambda\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right),\end{aligned}$$

so that the estimate (19) also gives the second condition (18) in Proposition 1. Thus we may apply Proposition 1, and the theorem is proved. □

Remark 1 Using the fact that $\mathrm{tr}(DD^*) = 2$, the factor $\sqrt{(\alpha-\gamma)^2+4\beta^2}$ in (19) can be rewritten in terms of the Hilbert-Schmidt norm of $DD^* - I$:

$$\sqrt{(\alpha-\gamma)^2+4\beta^2} = \sqrt{2}\|DD^* - I\|_{\mathrm{HS}} \le 2.$$

Thus the conclusion of Theorem 1 remains true under the weaker assumption $(\lambda/\omega)\|DD^* - I\|_{\mathrm{HS}} \to 0$.

3.3 Consequence

The sequence of conversions $\xi(\cdot)$ to $Y(\cdot)$ to $Z(\cdot)$ to $U(\cdot)$ to $S(\cdot)$ gives

$$\xi(t) = QY(t) = QR_{-\omega t}Z(t) = \frac{\sigma}{\sqrt{\lambda}}QR_{-\omega t}U(\lambda t) \approx \frac{\sigma}{\sqrt{\lambda}}QR_{-\omega t}S(\lambda t), \quad (20)$$

where $S(t)$ is a standard 2-dimensional Ornstein–Uhlenbeck process, for $\lambda \ll \omega$. From the approximation in (20) we see

(i) oscillations of $\xi(t)$ with frequency ω,
(ii) slow variation given by the OU process $S(\lambda t)$,
(iii) the stationary OU process has $E\|S(\lambda t)\|^2 = 1$,
(iv) typical amplitude of $\xi(t)$ is $(\sigma/\sqrt{\lambda})\|Q\|$.

For future reference write the approximations

$$Y^{\mathrm{app}}(t) = \frac{\sigma}{\sqrt{\lambda}}R_{-\omega t}S(\lambda t)$$

and

$$\xi^{\mathrm{app}}(t) = QY^{\mathrm{app}}(t) = \frac{\sigma}{\sqrt{\lambda}}QR_{-\omega t}S(\lambda t).$$

The two components $Y_1^{\mathrm{app}}(t)$ and $Y_2^{\mathrm{app}}(t)$ have the same magnitude and are out of phase by $\pi/2$. The matrix Q contains information about the relative magnitudes and the phase difference of the two components $\xi_1^{\mathrm{app}}(t)$ and $\xi_2^{\mathrm{app}}(t)$. For more discussion of this issue see Example 4.2.

There is an explicit formula for the power spectral density for the stationary solution $\xi(t)$ of any equation of the form (13), see e.g. Gardiner (1990, Sect 4.4.4). This has been used by some authors, see McKane and Newman (2005, Fig. 2), McKane et al. (2007, Fig. 6), to compare the (theoretical) power spectral density of the diffusion model described in Subsect. 2.2.3 with the results obtained by simulating the original jump Markov process $X(t)$ with rates (1). In the examples in Sect. 4 we will compare instead the (theoretical) power spectral density for $\xi(t)$ with the (theoretical) power

spectral density for our approximation $\xi^{\text{app}}(t)$. In order to calculate the power spectral density for $\xi^{\text{app}}(t)$, we observe that the transition from $U(t)$ to $S(t)$ in Theorem 1 can be achieved by replacing DD^* with I. It follows that $Y^{\text{app}}(t)$ satisfies the equation

$$dY^{\text{app}}(t) = AY^{\text{app}}(t)\,dt + \sigma\,dW(t)$$

obtained by replacing C by σI in (14), and hence

$$d\xi^{\text{app}}(t) = A_0\xi^{\text{app}}(t)\,dt + \sigma Q\,dW(t) \tag{21}$$

The power spectral density for $\xi^{\text{app}}(t)$ can now be calculated using the coefficients in the Eq. (21).

4 Examples

4.1 Linear oscillator

It is well known, see e.g. Landa and McClintock (2000), that a weakly damped harmonic oscillator with white noise forcing has a narrow bandwidth output. Putting $y = \dot{x}/\omega_0$ in

$$\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = \sigma_0\dot{W}_t \tag{22}$$

gives the 2-dimensional SDE (driven by 1-dimensional noise)

$$d\xi(t) = \begin{bmatrix} 0 & \omega_0 \\ -\omega_0 & -2\lambda \end{bmatrix}\xi(t)\,dt + \begin{bmatrix} 0 \\ \sigma_0/\omega_0 \end{bmatrix}dW(t).$$

Put $\xi(t) = QY(t)$ where $Q = \begin{bmatrix} 1 & 0 \\ -\lambda/\omega_0 & \omega/\omega_0 \end{bmatrix}$ and $\omega = \sqrt{\omega_0^2 - \beta^2/4}$, then

$$dY(t) = \begin{bmatrix} -\lambda & \omega \\ -\omega & -\lambda \end{bmatrix}Y(t)\,dt + \begin{bmatrix} 0 \\ \sigma_0/\omega \end{bmatrix}dW(t).$$

Assuming $0 < \lambda \ll \omega$ we can apply Theorem 1 to obtain

$$Y(t) \approx Y^{\text{app}}(t) = \frac{\sigma_0}{\omega\sqrt{2\lambda}}R_{-\omega t}S(\lambda t)$$

and then

$$\xi(t) \approx \xi^{\text{app}}(t) = QY^{\text{app}}(t) = \frac{\sigma_0}{\omega\sqrt{2\lambda}}QR_{-\omega t}S(\lambda t).$$

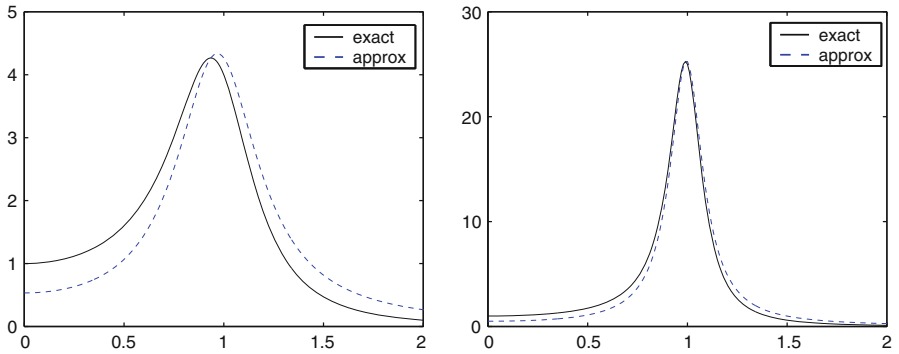

Fig. 1 Power spectral density as a function of frequency for the exact output $x(t)$ and approximate output $x^{\text{app}}(t)$ of the linear oscillator (22) with parameter values $\omega_0 = \sigma = 1$ and two different values of λ. On *left* $\lambda = 0.25$. On *right* $\lambda = 0.1$

In particular the output of the oscillator is

$$x(t) \approx x^{\text{app}}(t) = \xi_1^{\text{app}}(t) = \frac{\sigma_0}{\omega_0\sqrt{2\lambda}}\left(S_1(\lambda t)\cos\omega t + S_2(\lambda t)\sin\omega t\right).$$

In this abstract setting we can compare the power spectral densities for the exact solution $x(t)$ and our approximation $x^{\text{app}}(t)$ for different values of the ratio λ/ω_0, see Fig. 1.

4.2 McKane and Newman (2005)

We return again to the predator–prey model of McKane and Newman (2005). Using the parameter values $b = 0.1$, $d_1 = 0.1$, $d_2 = 0$, $p_1 = 0.25$ and $p_2 = 0.05$ given in McKane and Newman (2005, Figure 1) we get $x_{\text{eq}} = y_{\text{eq}} = 0.2$ and then

$$A_0 = \begin{bmatrix} 0 & 0.1 \\ -0.16 & -0.04 \end{bmatrix} \quad \text{and} \quad B_0 = \begin{bmatrix} 0.04 & -0.02 \\ -0.02 & 0.048 \end{bmatrix}.$$

Take

$$Q = \begin{bmatrix} 1 & 0 \\ -0.2 & 1.25 \end{bmatrix}$$

then

$$A = Q^{-1}A_0Q = \begin{bmatrix} -0.02 & 0.125 \\ -0.125 & -0.02 \end{bmatrix}$$

and

$$B = Q^{-1} B_0 (Q^{-1})^* = \begin{bmatrix} 0.04 & -0.00961 \\ -0.00961 & 0.0267 \end{bmatrix}.$$

We have $\lambda = 0.02$, $\omega = 0.125$, so that $\lambda/\omega = 0.16$ which is fairly small. Also $\sigma^2 = (1/2)\mathrm{tr}(B) = 0.0333$, and so

$$\frac{\sigma}{\sqrt{\lambda}} = \sqrt{\frac{\sigma^2}{\lambda}} = \sqrt{\frac{0.0333}{0.02}} = 1.29.$$

We get

$$Y(t) \approx Y^{\mathrm{app}}(t) = 1.29 R_{-\omega t} S(\lambda t)$$

where S is the standard 2-dimensional Ornstein–Uhlenbeck process, and then

$$\xi(t) \approx \xi^{\mathrm{app}}(t) = Q Y^{\mathrm{app}}(t) = 1.29 Q R_{-\omega t} S(\lambda t) = \begin{bmatrix} 1.29 & 0 \\ -0.258 & 1.61 \end{bmatrix} R_{-\omega t} S(\lambda t).$$

Figure 2 shows the power spectral densities of the predator and prey components of the exact diffusion model $\xi(t)$ as compared with the corresponding components for $\xi^{\mathrm{arr}}(t)$.

We can write

$$\begin{bmatrix} -0.258 & 1.61 \end{bmatrix} = 1.63 \begin{bmatrix} \cos\phi & \sin\phi \end{bmatrix}$$

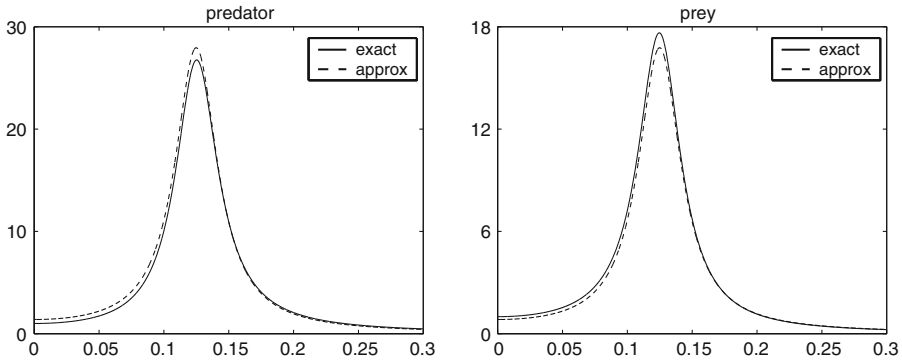

Fig. 2 Power spectral densities for the predator–prey model (McKane and Newman 2005), comparing each of the components of the exact diffusion model $\xi(t)$ with the corresponding component of our approximation $\xi^{\mathrm{app}}(t)$. For each component the power spectral density $P(\omega)$ for the exact diffusion model has been normalized to 1 at $\omega = 0$. On *left* is predator, on *right* is prey

where $\phi = 1.73$. Write also

$$S(t) = \|S(t\| \begin{bmatrix} \cos\theta(t) \\ \sin\theta(t) \end{bmatrix}$$

Then

$$\xi_1(t) \approx \xi_1^{\text{app}}(t) = 1.29\|S(\lambda t)\| \cos(\omega t - \theta(\lambda t)) \tag{23}$$

and

$$\xi_2(t) \approx \xi_2^{\text{app}}(t) = 1.63\|S(\lambda t)\| \cos(\omega t + 1.73 - \theta(\lambda t)). \tag{24}$$

The diffusion model for the predator density n/N and prey density m/N is

$$\left(\frac{n}{N}, \frac{m}{N}\right) = (x_{\text{eq}}, y_{\text{eq}}) + \frac{1}{\sqrt{N}}\xi(t) \tag{25}$$

The variations of the predator density are of magnitude roughly $1.29/\sqrt{3200} = 0.023$, whereas the variations of the prey density are somewhat larger with magnitude roughly $1.63/\sqrt{3200} = 0.029$. These values are consistent with the simulation shown in McKane and Newman (2005, Fig 1). The approximations (23, 24) also show a phase difference of 1.73 radians between the two components. Thus peak values of the predator density typically occur about $1.73/\omega$ later than the corresponding peak value of the prey density. This lag can also be seen in McKane and Newman (2005, Fig 1).

The significance of the scaling factor $\sigma/\sqrt{\lambda}$ and the diagonalizing matrix Q can be seen in Fig. 3, which shows a simulated trajectory of the right side of (25) where $\xi(t)$ is given by the SDE (11) with the matrices A_0 and $B_0 = C_0 C_0^*$ given above. The simulation was carried out using a first order Euler method with time step 0.01, and the results were plotted every 0.25 units of time. After the early transient behavior, the trajectory has roughly elliptical paths with slowly varying amplitude. The superimposed white ellipse is obtained by replacing $\xi(t)$ by $\xi^{\text{app}}(t)$ and freezing the Ornstein–Uhlenbeck process S with amplitude $\|S\|$ equal to its root mean square value 1. Equivalently, it may be obtained by taking the image of the unit circle under the matrix mapping $N^{-1/2}(\sigma/\sqrt{\lambda})Q$ and centering it at the point $(x_{\text{eq}}, y_{\text{eq}})$.

4.3 McKane et al. (2007), Sel'kov model

Our approximation can also be useful in higher dimensional models. The Sel'kov model studied by McKane et al. (2007) captures the essence of the reaction which is the key step in glycolysis, in which the enzyme PFK1 catalyzes the phosphorylation of F6P. Because there are four reactants ATP, ADP, PFK1/ADP and PFK1/ADP/ATP, the model is 4 dimensional (m_1, m_2, n_1, n_2) with two large parameters N_1 and N_2. Since N_1 is a measure of size for the first two components and N_2 is a measure of size for the last two components it is natural to look for a macroscopic limit and diffusion approximation of the form

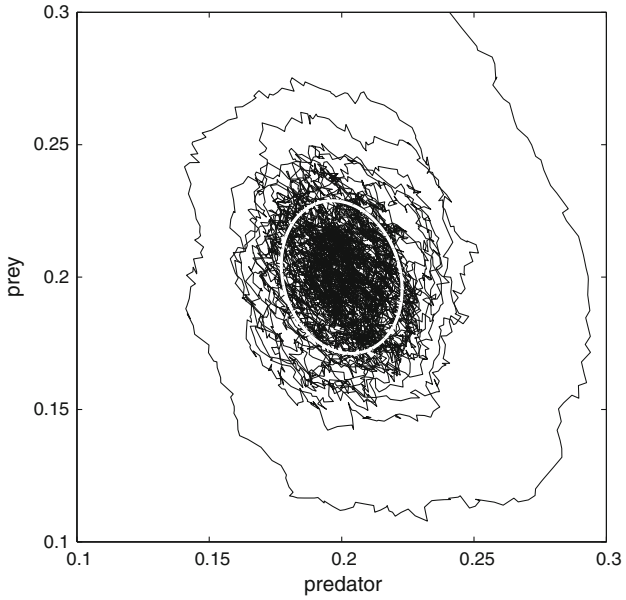

Fig. 3 Simulated trajectory of the diffusion model (25) for Example 4.2. See text for explanation of the *white ellipse*

$$
\begin{aligned}
s_1 &= \frac{m_1}{N_1} = \phi_1(t) + \frac{1}{\sqrt{N_1}}\xi_1(t) \\
s_2 &= \frac{m_2}{N_1} = \phi_2(t) + \frac{1}{\sqrt{N_1}}\xi_2(t) \\
a &= \frac{n_1}{N_2} = \phi_3(t) + \frac{1}{\sqrt{N_2}}\xi_3(t) \\
b &= \frac{n_2}{N_2} = \phi_4(t) + \frac{1}{\sqrt{N_2}}\xi_4(t)
\end{aligned}
$$

Assuming the ratio $N_1/N_2 = \sigma$ is fixed as N_1 and N_2 tend to infinity, this model fits in the setting described in Sect. 2. The only difference is that the factor σ or $\sqrt{\sigma}$ appears at appropriate places in the formulas (5) for the vector field F and the covariance function B, see McKane et al. (2007, eqn (31)–(34) and (A.2)).

The simulation results shown in McKane et al. (2007, Fig 7) have $N_1 = N_2 = 4096$ and parameter values giving fixed points $s_1 = 0.342$, $s_2 = 0.0108$, $a = 0.0033$ and $b = 0.0025$. The diffusion approximation is given by $d\xi(t) = A_0\xi(t) + C_0 dW(t)$ where

$$
A_0 = \begin{bmatrix}
-0.00506 & -0.00175 & -0.342 & 0.324 \\
-0.00381 & -0.233 & 0.650 & 0.132 \\
-0.00175 & 0.200 & -0.991 & 0.451 \\
0.00506 & 0.00175 & 0.342 & -0.453
\end{bmatrix}
$$

and

$$C_0 C_0^* = B_0 = \begin{bmatrix} 0.00227 & 0 & 0.00194 & -0.00194 \\ 0 & 0.00494 & -0.00397 & -0.000324 \\ 0.00194 & -0.00397 & 0.00656 & -0.00227 \\ -0.00194 & -0.000324 & -0.00227 & 0.00227 \end{bmatrix}.$$

Write $\xi(t) = QY(t)$ where

$$Q = \begin{bmatrix} 0.290 & 0.325 & -0.125 & -0.976 \\ -0.442 & 0.806 & -0.154 & 0.0625 \\ 0.786 & -0.0967 & -0.0484 & 0.0152 \\ -0.322 & -0.485 & -0.0386 & 0.00121 \end{bmatrix},$$

then

$$dY(t) = \begin{bmatrix} -1.29 & 0 & 0 & 0 \\ 0 & -0.391 & 0 & 0 \\ 0 & 0 & -0.000680 & 0.00501 \\ 0 & 0 & -0.00501 & -0.000680 \end{bmatrix} Y(t)\,dt + C\,dW(t) \quad (26)$$

where

$$B = CC^* = \begin{bmatrix} 0.0104 & -0.000123 & -0.000575 & 0.000196 \\ -0.000123 & 0.00492 & 0.00190 & -0.000665 \\ -0.000575 & 0.00190 & 0.00890 & -0.00247 \\ 0.000196 & -0.000665 & -0.00247 & 0.00122 \end{bmatrix}. \quad (27)$$

From (26, 27) we see that $Y_1(t)$ satisfies an equation

$$dY_1(t) = -1.29 Y_1(t)\,dt + \sqrt{0.0104}\,dB_1(t)$$

where $B_1(t)$ is scalar Brownian motion, and so $Y_1(t)$ is a scalar Ornstein–Uhlenbeck process with stationary variance $0.0104/(2 \times 1.29) = (0.0635)^2$ and hence typical magnitude 0.0635. Similarly

$$dY_2(t) = -0.391 Y_2(t)\,dt + \sqrt{0.00492}\,dB_2(t)$$

where $B_2(t)$ is scalar Brownian motion, and so $Y_2(t)$ is a scalar Ornstein–Uhlenbeck process with stationary variance $0.00492/(2 \times 0.391) = (0.0793)^2$ and hence typical magnitude 0.0793. It remains to study the pair $\begin{bmatrix} Y_3(t) \\ Y_4(t) \end{bmatrix} \equiv \widetilde{Y}(t)$, say. We have

$$d\widetilde{Y}(t) = \begin{bmatrix} -0.000680 & 0.00501 \\ -0.00501 & -0.000680 \end{bmatrix} \widetilde{Y}\,dt + \widetilde{C}\,dW(t)$$

where

$$\widetilde{B} = \widetilde{C}\widetilde{C}^* = \begin{bmatrix} 0.00890 & -0.00247 \\ -0.00247 & 0.00122 \end{bmatrix}.$$

Here we have $\lambda = 0.000680$ and $\omega = 0.00501$, so that $\lambda/\omega = 0.136$. Also $\sigma^2 = \text{trace}(\widetilde{B})/2 = 0.00506$, so that $\sigma/\sqrt{\lambda} = 2.73$. By Theorem 1 we obtain

$$\widetilde{Y}(t) \approx \begin{bmatrix} Y_3^{\text{app}}(t) \\ Y_4^{\text{app}}(t) \end{bmatrix} = 2.73 R_{-\omega t} S(\lambda t).$$

Notice that $Y_1(t)$ and $Y_2(t)$ are small relative to the vector $\widetilde{Y}(t)$, so that

$$Y(t) \approx Y^{\text{app}}(t) = \begin{bmatrix} 0 \\ 0 \\ Y_3^{\text{app}}(t) \\ Y_4^{\text{app}}(t) \end{bmatrix}.$$

Therefore

$$\xi(t) \approx \xi^{\text{app}}(t) = QY^{\text{app}}(t) = \begin{bmatrix} -0.340 & -2.66 \\ -0.421 & 0.171 \\ -0.132 & 0.0414 \\ -0.105 & 0.00329 \end{bmatrix} R_{-\omega t} S(\lambda t). \tag{28}$$

The variations of the ATP concentration s_1 are given by $\xi_1(t)/\sqrt{N_1}$ and so are of magnitude roughly $\sqrt{(-0.340)^2 + (-2.66)^2}/\sqrt{4096} = 0.042$. Similarly the variations of the ADP concentration s_2 are of magnitude roughly $\sqrt{(-0.421)^2 + (0.171)^2}/\sqrt{4096} = 0.0071$. These values are consistent with the simulation results shown in McKane et al. (2007, Fig 7(a,b)). It can be seen in Fig. 4 that there is good agreement between the (normalized) power spectral densities for the first two components of the exact diffusion model $\xi(t)$ (shown in McKane et al. 2007, Fig 7(c,d)) and the power spectral densities for the corresponding components of our approximation $\xi^{\text{app}}(t)$.

The approximation (28) also gives multidimensional information about the diffusion process $\xi(t)$. It implies that the oscillations in the system take place in the two-dimensional subspace $H \subset \mathbb{R}^4$ spanned by the 3rd and 4th vectors of Q. A similar calculation to the one leading to (23, 24) yields

$$\xi_i(t) \approx \xi_i^{\text{app}}(t) = r_i \|S(\lambda t)\| \cos(\omega t + \phi_i - \theta(\lambda t)), \quad 1 \le i \le 4,$$

where

$$r = \begin{bmatrix} 2.68 \\ 0.454 \\ 0.139 \\ 0.105 \end{bmatrix} \quad \text{and} \quad \phi = \begin{bmatrix} 4.59 \\ 2.76 \\ 2.84 \\ 3.11 \end{bmatrix}$$

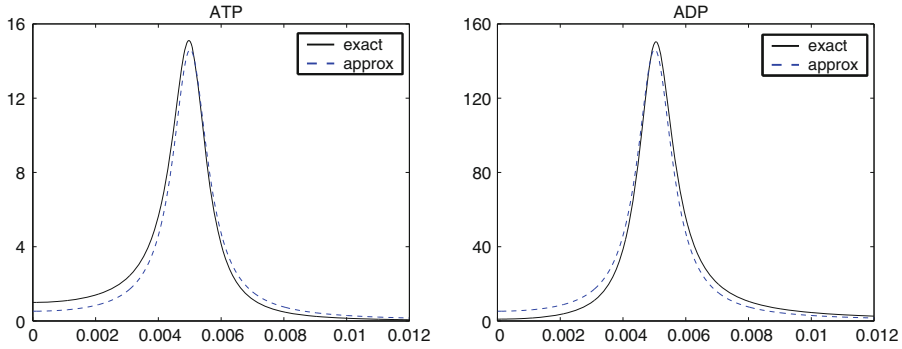

Fig. 4 Power spectral densities for the Sel'kov model (McKane et al. 2007), comparing the first two components of the exact diffusion model $\xi(t)$ with the corresponding component of our approximation $\xi^{\text{app}}(t)$. For each component the power spectral density $P(\omega)$ for the exact diffusion model has been normalized to 1 at $\omega = 0$. On *left* is ATP, on *right* is ADP

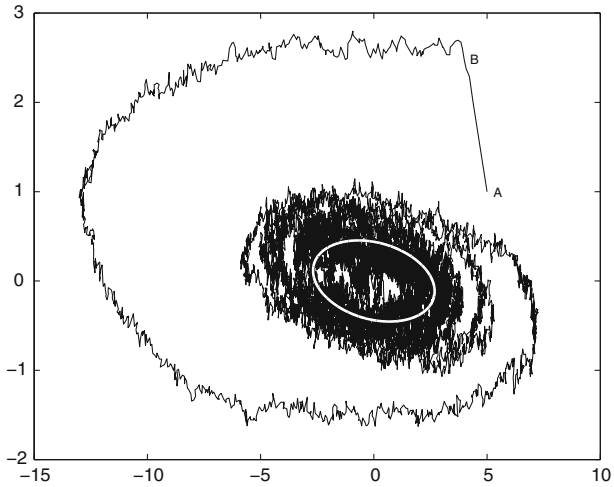

Fig. 5 First two components of simulated trajectory of the diffusion model $\xi(t)$ for Example 4.3. See text for explanation of the points A and B and the *white ellipse*

and the slowly varying function $\theta(\lambda t)$ denotes the direction of $S(\lambda t)$. The different values of r show the different typical magnitudes of the four component in the diffusion system. There is a phase difference of $\phi_1 - \phi_2 = 1.83$ between $\xi_1(t)$ and $\xi_2(t)$; this implies that the peak values of ADP density typically occur about $1.83/\omega$ later than the corresponding peak of ATP density. Moreover the relative closeness of ϕ_2, ϕ_3 and ϕ_4 implies that the components $\xi_2(t)$, $\xi_3(t)$ and $\xi_4(t)$ are oscillating almost exactly in phase.

Figure 5 shows the projection onto the first and second coordinates of a simulation of the 4-dimensional process $\xi(t)$. The simulation was carried out using a first order Euler method with time step 0.1, and the results were plotted every 1 unit of time. The transition A to B corresponds to the fast motion of $\xi(t)$ from its initial value towards

the subspace H. The trajectory from B onwards shows the combination of the fast rotation together with the more slowly changing Ornstein–Uhlenbeck process. The white ellipse denotes the trajectory of $(\xi_1^{\text{app}}, \xi_2^{\text{app}})$ given by (28) when the Ornstein–Uhlenbeck process is frozen with $\|S\| = 1$. Equivalently, it is the projection image of the unit circle under the matrix $\left[\begin{smallmatrix} -0.3404 & -2.663 \\ -0.4207 & 0.1706 \end{smallmatrix}\right]$ consisting of the first two rows of the 4×2 matrix appearing in (28).

5 Discussion

A biological system such as a predator–prey, epidemic, or biochemical system may show sustained oscillatory behavior about some fixed point, owing to stochastic effects, even though there is no periodic input or limit cycle. We have shown here how one may explicitly identify a simple diffusion process which approximates the behavior of such a large-population system, starting with a description of the system in terms of Markov transition rates. The calculation of the approximating diffusion is rather elementary, even though the justification involves a stochastic averaging result proved by the martingale problem method. The martingale problem argument is presented here in a self-contained way for convenience and completeness.

From the expression (20) for the approximate diffusion we see that the amplitude of the process of oscillations about the fixed point is of order $\sigma \|Q\|/\sqrt{\lambda}$, where σ, Q, λ come from algebraic calculations. The product $\sigma \|Q\|$ is a measure of the noise in the linear diffusion (13), while the factor $1/\sqrt{\lambda}$ can be seen as compensating for the rescaling of time which also results in the slowing of the Ornstein–Uhlenbeck process appearing in (20). None of these is directly related to the fixed point x_{eq} of the deterministic system.

Embedded in the matrix Q is the information of the relative phases of the components of the 2-dimensional process, as seen for the predator–prey example in (23, 24).

If our model is of dimension greater than two, the transformation matrix Q, as in Sect. 3.1, reveals the subspace inside which the oscillations take place. If the eigenvalues corresponding to the eigenvectors in directions orthogonal to this subspace are negative and large in magnitude relative to the damping rate of the oscillations, which we continue to call λ in the Sel'kov example in Sect. 4.3, then the processes in these other directions are relatively small and the higher dimensional oscillatory process can be approximated from our 2-dimensional result as illustrated by (28).

We saw in Subsect. 2.2.1 that the step from Markov jump process to the approximating SDE is appropriate when populations are large enough so that $\log N/N$ is small. It may be that corresponding results can be obtained for smaller populations by looking directly at the discrete Markov chain, as suggested by the discussion in Natiello and Solari (2007) but we do not address this question. We avoid the question of extinction, which usually arises in SDE approximations, by assuming the fixed point of our system is well away from a zero or absorbing state. In other words, our real starting point is the linear diffusion equation arrived at in e.g. McKane and Newman (2005) and McKane et al. (2007) just before they compute the power spectral density of their process. To see the value of our result (20) for biology, beyond what is

done in studies like (McKane and Newman 2005; McKane et al. 2007), we need to pose the question: what is the benefit of having a description of oscillations sustained by noise in terms of a rotation modulated by an Ornstein–Uhlenbeck process? What additional insight does this afford beyond knowing the power spectral density? Here are some examples of potential results and investigations which have now become possible.

Consider the SIR epidemic model (Kuske et al. 2007) or the SIRS model (Dushoff et al. 2004) where (20) also applies. A series of observed epidemics of, say a childhood disease, may come from a time interval where the OU process in our analysis is near 0 or from a time interval where the OU process is having a large excursion from 0. The expression (20) gives us the understanding that, given the limited time interval of the data, the overall variance, corresponding to epidemic amplitude, derived from epidemic parameters may have little correspondence with the observed sizes of epidemics, unless our data is a sufficiently long series. A similar point can be made about the oscillations in biochemical reactions.

Epidemic data often includes the effect of periodic forcing due, for example, to seasonal temperature or behavior patterns. In Dushoff et al. (2004) it was suggested that dynamical resonance can account for seasonality of influenza epidemics, and simulations dramatically validated this idea. Having, now, the technique to compute the approximation (20) to an epidemic process without forcing, we are in a position to incorporate seasonal forcing into an epidemic model and explore mathematically the phenomenon of re-enforcement.

Suppose we are interested in the distribution of the time a quasi-stationary process will spend in a potential well. If the process is one of sustained oscillations, as is evidently common in bio-systems, then this time is simply related to the boundary crossing time of an OU process. An analysis of the Morris Lecar neuron firing model (Ditlevsen and Greenwood 2010) is facilitated by this line of thought.

The authors of Aparicio and Solari (2001) and Natiello and Solari (2007) study the behavior of sample paths of Markov jump processes in the neighborhood of a stable fixed point of focus type (complex eigenvalues) of the deterministic process. Using the stochastic variation of a Liapunov function as a functional of the sample path as a tool, they show that 'the stochastic evolution of the angle variable $\phi = \arctan(y/x)$ is monotonic in the proximity of the equilibrium point.' (see Natiello and Solari 2007 for details). It is a challenging question to look for a direct connection between this result of Natiello and Solari (2007) and our result expressed by (20).

A prime object of study in connection with sustained oscillations has been the power spectral density of the stochastic system (McKane and Newman 2005; McKane et al. 2007; Risau-Gusman and Abramson 2007), which can be computed via the explicit formula in Gardiner (1990, Sec 4.4.4). In particular, Risau-Gusman and Abramson (2007) the formula is used to develop an attractive dimensionless, scale-invariant definition of 'quality', meaning relative peakedness, of a power spectral density peak. They found that 'good' stochastic oscillations are present only when their measure, ϵ, is small. Their ϵ, defined in terms of parameters of the deterministic model, is very nearly our λ/ω when ϵ is small, so we have consistency of these very different results.

References

Alonso D, McKane AJ, Pascual M (2007) Stochastic amplification in epidemics. J R Soc Interface 4:575–582

Aparicio J, Solari H (2001) Sustained oscillations in stochastic settings. Math Biosci 169:15–25

Di Patti F, Fanelli D (2009) Can a microscopic stochastic model explain the emergence of pain cycles in patients? J Stat Mech. doi:10.1088/1742-5468/2009/P01004

Ditlevsen S, Greenwood P (2010) The Morris-Lecar neuron model embeds the leaky integrate-and-fire model (in preparation)

Dushoff J, Plotkin J, Levin S, Earn D (2004) Dynamical resonance can account for seasonality of influenza epidemics. Proc Natl Acad Sci 101(48):16915–16916

Ethier S, Kurtz T (1986) Markov processes. Wiley, New York

Gardiner CW (1990) Handbook of stochastic methods for physics, chemistry and the natural sciences. Springer, New York

Kłosek M, Kuske R (2005) Multiscale analysis of stochastic delay differential equations. Multiscale Model Simul 3:706–729

Kurtz T (1978) Strong approximation theorems for density dependent Markov chains. Stoch Process Appl 6:223–240

Kushner HJ (1984) Approximation and weak convergence methods for random processes, with applications to stochastic systems theory. MIT Press, Cambridge

Kuske R, Gordillo L, Greenwood P (2007) Sustained oscillations via coherent resonance in SIR. J Theor Biol 245:459–469

Landa PS, McClintock PVE (2000) Changes in the dynamical behavior of nonlinear systems induced by noise. Phys Rep 323:1–80

McKane AJ, Newman TJ (2005) Predator-prey cycles from resonant amplification of demographic stochasticity. Phys Rev Lett 94:218102

McKane AJ, Nagy JD, Newman TJ, Stefanini MO (2007) Amplified biochemical oscillations in cellular systems. J Stat Phys 128:165–191

Morita S, Tainaka K (2006) Undamped oscillations in prey-predator models on a finite size lattice. Popul Ecol 48:99–105

Natiello M, Solari H (2007) Blowing-up of deterministic fixed points in stochastic population dynamics. Math Biosci 209:319–335

Papanicolaou GG, Stroock DW, Varadhan SRS (1977) Martingale approach to some limit theorems. In: 1976 Duke turbulence conference. Duke Univ Math Ser III, Chap VI, pp 1–120

Risau-Gusman S, Abramson G (2007) Bounding the quality of stochastic oscillations in population models. Eur Phys J B 60:515–520

Stroock DW, Varadhan SRS (1979) Multidimensional diffusion processes. Springer, Berlin

Telo da Gama M, Nunes A (2006) Epidemics in small world networks. Eur Phys J B 50:205–208

Van Kampen NG (1992) Stochastic processes in physics and chemistry, 2nd edn. North-Holland, Amsterdam

Yu N, Kuske R, Li Y (2006) Stochastic phase dynamics: multiscale behavior and coherence measures. Phys Rev E 73:056205