

Research article

Open Access

Amplitude distribution of stochastic oscillations in biochemical networks due to intrinsic noise

Moritz Lang, Steffen Waldherr and Frank Allgöwer

Address: Institute for Systems Theory and Automatic Control, Universität Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany

E-mail: Moritz Lang - moritz.lang@ist.uni-stuttgart.de; Steffen Waldherr* - waldherr@ist.uni-stuttgart.de;

Frank Allgöwer - allgower@ist.uni-stuttgart.de

* Corresponding author

Published: 17 November 2009

Received: 6 March 2009

PMC Biophysics 2009, 2:10 doi: 10.1186/1757-5036-2-10

Accepted: 17 November 2009

This article is available from: <http://www.physmathcentral.com/1757-5036/2/10>

© 2009 Lang et al

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Intrinsic noise is a common phenomenon in biochemical reaction networks and may affect the occurrence and amplitude of sustained oscillations in the states of the network. To evaluate properties of such oscillations in the time domain, it is usually required to conduct long-term stochastic simulations, using for example the Gillespie algorithm. In this paper, we present a new method to compute the amplitude distribution of the oscillations without the need for long-term stochastic simulations. By the derivation of the method, we also gain insight into the structural features underlying the stochastic oscillations. The method is applicable to a wide class of non-linear stochastic differential equations that exhibit stochastic oscillations. The application is exemplified for the MAPK cascade, a fundamental element of several biochemical signalling pathways. This example shows that the proposed method can accurately predict the amplitude distribution for the stochastic oscillations even when using further computational approximations.

PACS Codes: 87.10.Mn, 87.18.Tt, 87.18.Vf

MSC Codes: 92B05, 60G10, 65C30

1. Introduction

Oscillations are a widely occurring phenomenon in the dynamics of biological systems. On the intracellular level, oscillations occur for example in the activity of various genes or signalling proteins. To gain more insight into the processes related to these oscillations, mathematical models for the underlying biochemical and genetic networks are commonly constructed. Such models have proven helpful in connecting the underlying biochemistry with the temporal characteristics of emerging oscillations and the associated biological function. Examples for this type of results include the intracellular circadian clock [1] or the developmental process of somitogenesis [2].

In dynamical models for biological systems, oscillations are due to limit cycles or more complex attractors in deterministic frameworks, or may emerge from various stochastic effects in stochastic frameworks. Concerning the biological function, one would not expect that it makes a significant difference whether oscillations are due to a deterministic limit cycle or arise from stochastic effects. For the biological function, rather the temporal characteristics of oscillations are relevant, such as their frequency or amplitude. For deterministic limit cycle oscillations, these characteristics are easily computed by a numerical simulation of the model. For stochastic models, computing the temporal characteristics of oscillations is much more involved. Two approaches are common for solving stochastic systems. In the first approach, the so called chemical master equation (CME) is used [3]. The solution to the CME (or approximations thereof) yields the probabilities for each of the possible configurations of the system to be attained at a given point in time. However, since the temporal correlations of these probabilities are usually not obtained from the CME, the temporal characteristics of oscillations are not directly accessible from such a solution. In the second approach, a large number of realisations of the stochastic process are computed. These realisations can then be used to compute various temporal characteristics of the system, in particular oscillation amplitude and frequency distributions. Yet, the computation of a sufficiently large number of realisations typically entails a large computational effort.

In this paper, we develop a new method to compute the amplitude distribution for systems exhibiting stochastic oscillations. We thereby focus on systems where the deterministic part has a weakly stable equilibrium point (EP), typically with damped oscillations, and the stochastic effects induce sustained oscillations around this EP. In such systems, we can distinguish two mechanisms by which stochastic oscillations occur. For the first mechanism, the system needs to have the property that a certain small perturbation away from the equilibrium point leads to a large excursion in the state space before the system returns to the proximity of the EP [4]. If the noise is strong enough to reliably generate such a perturbation, but does not disturb the system during the round trip too much, we recognize regular oscillations with a well defined amplitude and frequency. This effect is called Coherence Resonance [5] or Stochastic Resonance [6, p. 149 ff.]. For the second mechanism, the deterministic part of the system already gives rise to damped oscillations, and the noise just serves to sustain the oscillations [7,8]. In this case, the density distribution of realizations in the state space will typically be similar to a Gaussian-like distribution with a maximum at the EP, while individual trajectories show pronounced oscillations with a frequency close to the frequency of the deterministic part. The amplitudes of this type of oscillations increase with the noise power, in contrast to the other type of systems, which mainly exhibit oscillations of a fixed amplitude independent of the noise power.

For the second mechanism, it is of interest to compute the amplitude distribution of the stochastic oscillations. Several methods to solve this problems were developed previously. In [9] the average change of a stochastic Lyapunov function was determined. By setting this average change to zero one may identify an orbit around which a stochastic realization

fluctuates, and thereby estimate a typical amplitude. In the context of this article the results of Kuske et al. [7,8] are very interesting. They apply a multi-scale ansatz to separate the properties of the oscillations mainly determined by the deterministic part of the analyzed system (the frequency) from those determined to a high percentage by the stochastic part (the amplitude). Therewith it is possible to calculate both the amplitude distribution and the mean frequency of the oscillations. Nevertheless this method has the disadvantage of requiring the discussed systems to be almost linear and the frequency of the oscillations not to be noteworthy disturbed by the stochastic effects.

In this article we focus on biochemical systems that can be modeled as a set of interconnected, possibly nonlinear, stochastic differential equations (SDEs). With the help of the Fokker-Planck equation (see e.g. [3, p. 193 ff.] or [10, p. 120 ff.]), we calculate the stationary density distribution in the state space. We provide a theorem allowing to calculate the amplitude distribution with only the knowledge of the density distribution and the corresponding SDE. The theorem allows not only to consider linear, but also a wide class of non-linear systems and therefore makes it possible to analyze not only systems with oscillations being of low amplitude compared to average concentrations but also of intermediate and high amplitudes. Although stochastic oscillations are a mainly two dimensional effect, we show that they also occur in higher order systems and give an example on how to analyze them then. As far as we are aware of, this paper introduces for the first time a method to analytically analyse higher order possibly non-linear systems being affected with intermediate to high amounts of noise, and thereby showing stochastic oscillations.

The paper is structured as follows. To give an easily accessible introduction to the material, we first present our results on the calculation of the amplitude distribution of stochastic oscillations and afterwards state the underlying assumptions and theorems. Then we explain the application of our theory by calculating the amplitude distribution of a simple example, the damped harmonic oscillator in the presence of additive noise, to get a first insight into the reasons for stochastic oscillations. We want to remark that this example simplifies, due to its structure, the necessary calculations significantly. As a more realistic example, we discuss oscillations in the MAP kinase cascade with an incorporated negative feedback with limited amounts of entities of each molecular species. For this system, the stationary probability distribution can be estimated by a linear approximation, or it can be computed numerically. We compare the amplitude distributions predicted by our method, based on these two approaches, to an "experimental" amplitude distribution obtained from a long-term stochastic simulation.

2. Results and Discussion

2.1 Amplitude Distribution of the Stochastic Oscillations

In this article we develop a method to derive the amplitude distribution of stochastic oscillations from the knowledge of the stationary density distribution of a stochastic

differential equation (SDE). Here we first give a short outline of our results for which we present the corresponding theorems and proofs in the next sections.

Biochemical networks with stochastic dynamics can be modelled by the Langevin equation [11]. For the derivation of the theoretical results, we consider a two-dimensional Langevin equation having an equilibrium point at zero:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \Sigma(\mathbf{x})\Gamma \tag{1}$$

with the state vector $\mathbf{x} \in \mathcal{R}^2$, the system's dynamics $\mathbf{f} = (f_1, f_2)^T$, where f_1 and f_2 are smooth functions ($\in C^\infty$), Σ a 2-by-2 matrix of smooth functions and $\Gamma = (\Gamma_1, \Gamma_2)^T$, where Γ_1 and Γ_2 are uncorrelated, statistical independent Gaussian white noise with zero mean and variance of one.

Let us denote the stationary density distribution of the system (1) by $P(\mathbf{x})$, and the amplitude distribution for the oscillations by $P_A(\chi)$, where χ denotes the oscillation amplitude. Under certain assumptions (see following sections) the amplitude distribution $P_A(\chi)$ then satisfies

$$P_A(\chi) \propto v(\mathbf{x}(\chi)) \cdot P(\mathbf{x}(\chi)) \tag{2}$$

with $v(\mathbf{x}) = \|\mathbf{f}(\mathbf{x})\|$ the average speed at state \mathbf{x} and $\hat{\mathbf{x}}(\chi)$ given by

$$\hat{\mathbf{x}}(\chi) = \arg \max_{\mathbf{x} \in A} P(\mathbf{x}). \tag{3}$$

Hereby A is defined as $A = \{\mathbf{x} \in \mathcal{R}^2 \mid \langle \mathbf{x} \mid \bar{\mathbf{e}} \rangle = \chi\}$ and $\bar{\mathbf{e}}$ is the unit vector in the direction where the amplitude is measured.

The product on the right hand side of (2) represents the steady state flux of reactions through a certain state $\hat{\mathbf{x}}$. Intuitively speaking, $\hat{\mathbf{x}}$ is chosen so that most of the realizations going through $\hat{\mathbf{x}}$ will also reach their maximal value in a small neighborhood of $\hat{\mathbf{x}}$, so that under certain assumptions (see following sections) it is justifiable to approximate the amplitude of such a realization by the value reached at $\hat{\mathbf{x}}$.

The apparent complicated definition of $\hat{\mathbf{x}}$ is a result of the freedom of choice in which direction $\bar{\mathbf{e}}$ the amplitude is measured. In a biological system it is obviously important to distinguish whether one measures the oscillations of the concentration of species A or of species B . Sometimes it is experimentally not possible to distinguish between two species, e.g. if B represents the phosphorylated version of a protein A , so that one can only measure the amplitude of $[A] + [B]$. Additionally, for nonlinear networks, the amplitude distribution may be different depending on whether one measures the amplitude of the oscillations in the positive direction from the steady state or in the negative direction. If one e.g. measures the

amplitude distribution of the concentration of the first species x_1 in the positive direction, (3), i.e. $\vec{e} = (1, 0)^T$, can be simplified to $\hat{\mathbf{x}}(\chi) = (\chi, \arg \max_{x_2 \in \mathbb{R}} P(x_1 = \chi, x_2))^T$.

To normalize the probability distribution P_A of the amplitude, we have to divide equation (2) through the integral of the probability of all amplitudes:

$$\tilde{P}_A(\chi) = \left(\int_{\xi=0}^{\infty} P_A(\xi) d\xi \right)^{-1} P_A(\chi). \tag{4}$$

Simple models may have additional properties allowing us to simplify (2). For linear systems the speed $v(\hat{\mathbf{x}}(\chi))$ is directly proportional to $\|\hat{\mathbf{x}}(\chi)\|$ simplifying (2) to

$$P_A(\chi) \propto \|\mathbf{x}(\chi)\| \cdot P(\mathbf{x}(\chi)), \tag{5}$$

with $\hat{\mathbf{x}}$ given as above.

We want to emphasize here that (2) allows to analytically or numerically calculate the amplitude distribution only from the knowledge of certain properties of the SDE describing the biological system without the necessity to run numerical simulations. The formula yields a good approximation for a wide class of nonlinear problems and can be used to calculate the amplitude distribution even for certain systems having more than two dimensions as we demonstrate in the second example of this paper.

2.2 Derivation of the Results

To derive the results presented in the proceeding section we analyze nonlinear SDE systems of order two as given by (1), with an asymptotically stable equilibrium point and damped oscillations in the deterministic part. We restrict our analysis to two dimensional systems, because many higher dimensional systems can be reduced to two dimensions for the purpose of analysing stochastic oscillations. In the derivation we assume w.l.o.g. that we want to compute the amplitude of the oscillations in the positive direction of x_1 , i.e. $\vec{e} = (1, 0)^T$, and that the equilibrium point of the deterministic formulation ($\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$) of (1) is at the origin. Both assumptions can easily be satisfied for an arbitrary system by an appropriate coordinate transformation.

We establish an angular phase relationship of the state vector $\mathbf{x}(t)$ with respect to a fixed reference vector in the state space [12]. An oscillation period of the system (1) is defined as the time during which the angle between the state vector and the fixed reference vector changes by 2π . Further, the amplitude of an oscillation is defined as the maximal value of x_1 during the corresponding oscillation period.

We make the following assumptions on the system (1):

1. There exists a stationary density distribution $P: \mathcal{R}^2 \rightarrow \mathcal{R}^+$: $\mathbf{x} \mapsto P(\mathbf{x})$ for (1) which is sufficiently smooth in \mathbf{x} . We demand the curvature of the level curves of this distribution to

exist and not to change its sign. For the computation of the amplitude distribution $P_A(\chi)$, $P(\mathbf{x})$ may be computed analytically, numerically, or may be obtained by the long term limit of a measurement.

2. The deterministic formulation ($\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$) of (1), which is at the state $\tilde{\mathbf{x}} := \mathbf{x}(t_0)$ at time $t = t_0$ with the probability of the state $\tilde{\mathbf{x}}$ being $P(\tilde{\mathbf{x}}) =: \tilde{p}$, will approximately evolve tangential to the level curve of the probability distribution, thus for small Δt the probability of the state at time $t + \Delta t$ won't have changed significantly ($P(\mathbf{x}(t_0 + \Delta t)) \approx \tilde{p}$) This assumption can be easily checked by calculating the Lie derivative of P with respect to \mathbf{f} :

$$L_{\mathbf{f}}P = \frac{\partial P}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}), \tag{6}$$

which has to be small for almost all $\mathbf{x} \in \mathbb{R}^2$ having a non-negligible probability $P(\mathbf{x})$.

3. In addition we require that the average speed $v = \|\mathbf{f}(\mathbf{x})\|$ of (1) does not vanish for almost all $\mathbf{x} \in \mathbb{R}^2$ having a non-negligible probability $P(\mathbf{x})$.

Assumption 3 is obviously not satisfied in a small area around the EP where for most systems v becomes very small. However when the system is very close to the EP, the motion of most realizations can be approximated by a random walk. Thus during the time when the system is in the small neighborhood of the EP it is not justifiable to speak about oscillations anymore until the system leaves the neighborhood again. Away from the EP, $\|\mathbf{f}(\mathbf{x})\|$ is significantly bigger than zero for most biological relevant systems so that Assumption 3 is fulfilled.

With the help of these assumptions we are able to formulate a theorem for the calculation of the amplitude distribution of stochastic oscillations. For the formulation of the theorem let us define

$$\hat{X} = \left\{ \hat{\mathbf{x}} \in \mathbb{R}^2 : \frac{\partial P(\mathbf{x})}{\partial x_2} \Big|_{\mathbf{x}=\hat{\mathbf{x}}} = 0, \hat{x}_1 > 0 \right\}, \tag{7}$$

i.e. the unit vector \bar{e}_{x_1} in the direction of x_1 is orthogonal to the tangent on the level curve of the probability distribution in $\hat{\mathbf{x}}$ (see Figure 1). Because of Assumption 1 there exists exactly one state $\hat{\mathbf{x}} \in \hat{X}$ for every amplitude χ satisfying $\hat{x}_1 = \chi$. Furthermore we define the following variables, which can be calculated with the knowledge of the system (1) and the stationary probability distribution P :

$$a(\mathbf{x}) = \frac{\partial P}{\partial x_1} \Big|_{\mathbf{x}=\mathbf{x}} \tag{8a}$$

$$\kappa(\mathbf{x}) = \frac{\partial^2 P}{\partial x_2^2} \Big|_{\mathbf{x}=\mathbf{x}} \tag{8b}$$

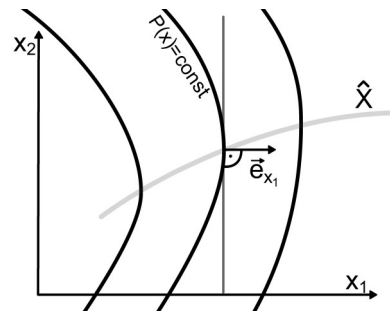


Figure 1
Definition of \hat{x} . The solid black lines are the level curves of $P(\mathbf{x})$. In $\hat{x} \in \hat{X}$ the unit vector \vec{e}_{x_1} in the direction of x_1 is orthogonal to the tangent on the level curve trough \hat{x} .

$$L(\mathbf{x}) = -\sqrt{\text{var}(L_{f+\Sigma\Gamma}P(\mathbf{x}))} \Big|_{\mathbf{x}=\mathbf{x}} \tag{8c}$$

$$\approx -\frac{\partial P}{\partial x_1} \sqrt{\Sigma_{11}^2(\mathbf{x}) + \Sigma_{12}^2(\mathbf{x})} \Big|_{\mathbf{x}=\mathbf{x}} \tag{8d}$$

$$v(\mathbf{x}) = \|\mathbf{f}(\mathbf{x})\| \Big|_{\mathbf{x}=\mathbf{x}} \tag{8e}$$

$$\approx \|f_2(\mathbf{x})\| \Big|_{\mathbf{x}=\mathbf{x}}.$$

The variable v can be thought of as the deterministic speed, L is the standard deviation of the Lie-derivative of P with respect to the right hand side of (1), a the derivative of P in the direction of x_1 , and κ the curvature of P with respect to x_2 , each evaluated at the states $\hat{x} \in \hat{X}$. The variables are illustrated in Figure 2.

With the help of these definitions we can formulate the following two lemmas and a theorem, which characterizes the amplitude of the stochastic oscillations.

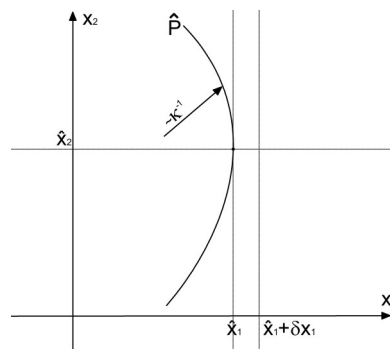


Figure 2
Visualization of the variables. Visualization of the variables needed to calculate the amplitude distribution of stochastic oscillations.

Lemma 1 Assume the Assumptions 1-3 are satisfied. Let Ψ_f be a realization of (1) being at the state $\hat{\mathbf{x}} \in \hat{X}$ at time t_0 . Then the amplitude of Ψ_f during the current oscillation will lie with a probability of 70.8% in the set $[\hat{x}_1, \hat{x}_1 + \delta x_1]$ and with a probability of 95.5% in the set $[\hat{x}_1, \hat{x}_1 + 2\sqrt[3]{2}\delta x_1]$, with δx_1 defined by $\delta x_1 = \frac{3}{4} \frac{L}{a} \sqrt[3]{\frac{L}{2v^2\kappa}}$.

Lemma 2 Assume the Assumptions 1-3 are satisfied. Then the net flux density $\bar{\varphi}(\mathbf{x})$ of realizations at the state $\mathbf{x} \in \mathcal{R}^2$ has an absolute value proportional to the product of the probability $P(\mathbf{x})$ and the average speed $v(\mathbf{x})$ of (1) and is tangential to the level curve of the density distribution P at this point.

Lemma 2 is derived by considering the flux of realizations through an infinitesimally small region around the state x , as illustrated in Figure 3.

Theorem 1 (Distribution of amplitudes) Under the Assumptions 1-3 and if δx_1 as defined in Lemma 1 is small, the probability $P_A(\chi)$ of an amplitude χ of the stochastic oscillations of (1) in the direction x_1 is proportional to the absolute value of the net flux density $\|\bar{\varphi}(\mathbf{x})\|$ at the state $\hat{x}_1 \in \hat{X}$ satisfying $\hat{x}_1 = \chi$.

Our results for the calculation of the amplitude distribution $P_A(\chi)$ presented in the previous section (2) directly follow from the application of Theorem 1. The proofs of the respective lemmas and theorem can be found in the appendix.

2.3 Remarks

- To check if the assumptions necessary for the application of Theorem 1 are fulfilled, one can calculate δx_1 according to the formula in Lemma 1. If δx_1 is significantly smaller than \hat{x}_1 for almost all $\hat{\mathbf{x}} \in \hat{X}$ having a non-negligible probability $P(\hat{\mathbf{x}})$, Theorem 1 may be applied. For a general stochastic differential equation, it is hard to exactly indicate how much smaller δx_1 must be compared to \hat{x}_1 . This highly depends on the structure of the system and also on the accuracy being required for the solution. Nevertheless, as a rule of thumb, the theorem

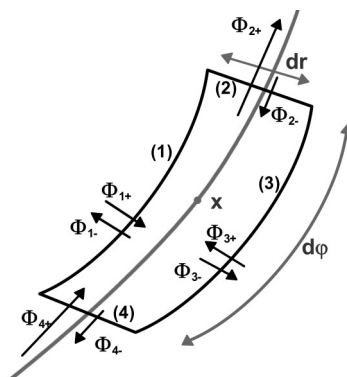


Figure 3
Flux through unit cell. The fluxes through the edges of an infinitesimal small unit cell around a state $\mathbf{x} \in \mathbb{R}^2$.

gives good approximations if $\frac{\delta x_1}{x_1} < 0.2$. It is not necessary that this is true for every $\hat{\mathbf{x}} \in \hat{X}$, but the calculated density of the amplitude may differ from the real one for the amplitudes $\chi \in \mathbb{R}$ for which the corresponding δx_1 is not small enough. Nevertheless the accuracy of the amplitude distribution for other amplitudes is not affected (see proof).

- It may be possible to relax the assumption that the curvature of the level curves does not change its sign globally. Intuitively it seems to be sufficient for most systems that the curvature does not change its sign only locally around the states $\hat{\mathbf{x}} \in \hat{X}$, and that the amplitude in one oscillation period reaches its maximum in the area around $\hat{\mathbf{x}}$. However this has to be shown for the system of interest explicitly.
- If the Lie derivative $L_f P$ of the density distribution P is not small enough, the net flux density $\bar{\varphi}$ cannot be approximated to be tangential to the level curves of P anymore. Depending on the size of $L_f P$ this may lead to a significant bias in the calculation of the amplitude distribution for certain systems. In this case, the derivation of the amplitude distribution would be more involved, and also depend on the term $\frac{\partial \hat{x}_2}{\partial \chi}$, starting from an appropriate modification of (66) in the appendix.

2.4 The Damped Harmonic Oscillator in the Presence of Additive Gaussian White Noise

2.4.1 Equations and Properties of the System

In this section we give a motivational example for the application of Theorem 1 by discussing the behavior of the damped harmonic oscillator in the presence of additive Gaussian white noise. The oscillator satisfies the following Langevin equation:

$$\dot{x}_1 = x_2 - kx_1 + \sigma\Gamma_1(t) \tag{9a}$$

$$\dot{x}_2 = -x_1 - kx_2 + \sigma\Gamma_2(t), \tag{9b}$$

with Γ_i , $i \in \{1, 2\}$ uncorrelated, statistical independent Gaussian white noise with zero mean and variance of one, $\sigma \in \mathbb{R}^+$ and $k \in \mathbb{R}$.

It is easy to show that the deterministic system ($\sigma = 0$) is asymptotically stable for $k > 0$. The eigenvalues of the system matrix

$$A = \begin{bmatrix} -k & 1 \\ -1 & -k \end{bmatrix} \tag{10}$$

are at $\lambda_{1,2} = -k \pm i$, therefore the deterministic system will show damped oscillations upon perturbation from its steady state.

2.4.2 Calculation of the Density Distribution

For $\sigma > 0$ we can calculate with the help of the multivariate Fokker-Planck equation the density distribution of x_i , $i \in \{1, 2\}$, in the state space. The general multivariate Fokker-Planck equation (see, for example, [3, p. 193 ff.] or [13, p. 96 ff.]) is given by

$$\frac{\partial P(\mathbf{x},t)}{\partial t} = -\sum_{i=1}^n \frac{\partial}{\partial x_i} (A_i(\mathbf{x})P) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (B_{i,j}(\mathbf{x})P) \tag{11}$$

with $\hat{A}_i(\mathbf{x})$ and $B_{i,j}(\mathbf{x}) > 0$ real functions describing the influence of the system dynamics and the noise on the distribution. The diffusion term \mathbf{B} may be calculated to

$$\mathbf{B}(\mathbf{x}) = \mathbf{\Sigma}(\mathbf{x}) \cdot \mathbf{\Sigma}^T(\mathbf{x}), \tag{12}$$

with $\mathbf{\Sigma}$ the matrix of the noise terms of (1). We call (11) linear if $\forall_i \hat{A}_i(\mathbf{x})$ is a linear function in \mathbf{x} and $\forall_i, j B_{i,j}(\mathbf{x})$ is constant. Then (11) can be simplified to

$$\frac{\partial P(\mathbf{x},t)}{\partial t} = -\sum_{i,j=1}^n A_{i,j} \frac{\partial}{\partial x_i} (x_j P) + \frac{1}{2} \sum_{i,j=1}^n B_{ij} \frac{\partial^2}{\partial x_i \partial x_j} P. \tag{13}$$

For the system (9), $A_{i,j}$ are given by the system matrix A from (10), and $B_{i,j}$ are the elements of the matrix $\mathbf{B} = \sigma^2 I_2$, where I_2 is the second order identity matrix. To get the stationary density distribution, we set $\frac{\partial P(\mathbf{x})}{\partial t} = 0$. Following [3], the stationary density distribution for the linear Fokker-Planck equation (13) is given by

$$P(\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \langle \mathbf{x} \rangle)^T \mathbf{\Xi}^{-1}(\mathbf{x} - \langle \mathbf{x} \rangle)\right), \tag{14}$$

with $\langle \mathbf{x} \rangle$ the mean value of \mathbf{x} and $\mathbf{\Xi}$ the matrix of the second moments of P , which is the solution of the equation $\mathbf{A}\mathbf{\Xi} + \mathbf{\Xi}\mathbf{A}^T + \mathbf{B} = 0$. In system (9), $\langle \mathbf{x} \rangle = \mathbf{0}$ and $\mathbf{\Xi}$ may be calculated to $\mathbf{\Xi} = \frac{\sigma^2}{2k} I_2$. The state of the system is therefore Gaussian distributed with the maximum of the density being at the EP.

2.4.3 Determination of the Amplitude Distribution

To get the amplitude distribution $P_A(\chi)$ of the oscillations of the harmonic oscillator (9), we apply Theorem 1 after checking the assumptions made in the theorem. Notice that due to the easy structure of this first example and its symmetry, we are always measuring the oscillations in a direction tangential to a principal axis of the distribution. Therefore $\hat{\mathbf{x}}(\chi) = (\chi, 0)^T$ for all χ . Because of this it is easy to calculate the necessary variables with the help of the definitions (8a)-(8e) to

$$a = -\frac{2k\hat{x}_1}{\sigma^2} P(\hat{\mathbf{x}}) \tag{15a}$$

$$\kappa = -\frac{2k}{\sigma^2} P(\mathbf{x}) \tag{15b}$$

$$L \approx -\frac{2k\hat{x}_1}{\sigma} P(\hat{\mathbf{x}}) \tag{15c}$$

$$v \approx \|\hat{x}_1\|. \tag{15d}$$

After some calculations we obtain

$$\delta x_1 \leq \frac{3}{4} \sigma^3 \sqrt{\frac{\sigma}{2\hat{x}_1}}. \tag{16}$$

For $s = 10$ and $k = 0.01$, δx_1 is getting small compared to \hat{x}_1 for $\hat{x}_1 \geq 20$, so our approximation should at least hold for every amplitude greater or equal to 20. The results will nevertheless show that we even get good estimations for amplitudes much smaller than 20.

Due to the linearity of the oscillator (9) and the symmetry of the harmonic oscillator, we can determine the amplitude distribution by utilizing formula (5) as

$$P_A(\chi) \propto \chi \exp\left(-\frac{\chi^2 k}{\sigma^2}\right). \tag{17}$$

For $s = 10$ and $k = 0.01$ the predicted amplitude distribution P_A is plotted over the amplitude χ in Figure 4. For comparison we also displayed the amplitude distribution of the stochastic oscillations as measured in a long term simulation run.

This easy example was discussed to give an insight into the reasons for stochastic oscillations. In the next section, we show the practical applicability of our algorithm by

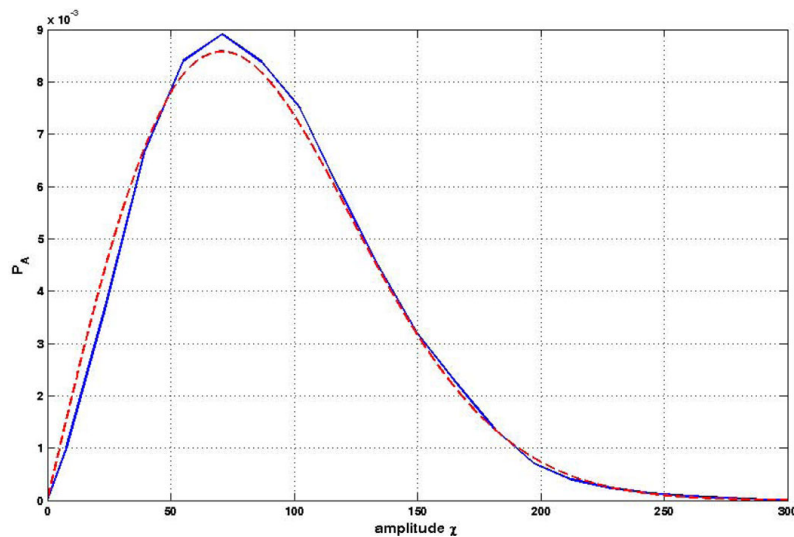


Figure 4
Amplitude distribution of the oscillations of the damped harmonic oscillator. Amplitude distribution of the oscillations of x_2 of the damped harmonic oscillator (9) in the presence of white noise with autocorrelation $\sigma = 10$ and $k = 0.01$. The solid curve corresponds to the data from simulation, the dashed curve is the prediction according to Theorem 1. The simulation was run over a time scale of 100000 s which corresponds to around 16000 oscillation periods.

predicting the amplitude distribution of a complex biochemical system and therewith give an example of a more biologically relevant application for the results of the paper.

2.5 Oscillations in the MAP Kinase Signaling Cascade

In the following section we apply the algorithm developed in this paper to a realistic example from biochemical signal transduction. A frequent module in many eucaryotic cells from yeast to mammals is the mitogen activated protein (MAP) kinase signaling cascade. MAPK cascades are typically activated by extracellular stimuli such as growth factors, and regulate the activity of various genes, thereby provoking a cellular response to the applied stimulus. Many important cellular functions such as differentiation, proliferation and death are controlled by MAPK cascades [14]. MAPK cascades consist of three layers of kinases, where each kinase phosphorylates and thereby activates the kinase on the next layer, as shown in Figure 5. The kinases are named MAPKKK, MAPKK, and MAPK, in the order of activation, and the phosphorylated, active forms are denoted with a star. The basic structure may be complemented by additional feedback interconnections, giving rise to deterministic limit cycle oscillations [15]. More recently, oscillations in the MAPK cascade have been determined experimentally in yeast cells, where the oscillatory activity of the MAPK controls periodic changes in cell shape during the mating process [16].

2.5.1 The Deterministic Model

For this example, we use a basic ODE model of the MAP kinase signaling cascade. The model contains a negative feedback interconnection from the last kinase MAPK to the activation of the first kinase MAPKKK [15]. We use a simpler model compared to the one in [15], and therefore sustained oscillations do not occur in the deterministic version of the model considered here.

Denote x_1 , x_2 and x_3 the concentrations of MAPKKK*, MAPKK* and MAPK*, respectively. Using three conservation relations

$$x_1 + [MAPKKK] = T_1 = 100 \text{ nM} \tag{18a}$$

$$x_2 + [MAPKK] = T_2 = 300 \text{ nM} \tag{18b}$$

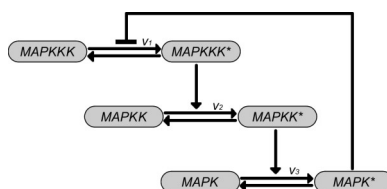


Figure 5
Schematic view on the MAP kinase signaling cascade. Schematic view on the MAP kinase signaling cascade. The species with the stars (*) are the phosphorylated versions of the species without stars. MAPKKK* catalyzes the phosphorylation of MAPKK, MAPKK* in turn catalyzes the phosphorylation of MAPK. A negative feedback is given by the repression of the activation of MAPKKK by MAPK*.

$$x_3 + [\text{MAPK}] = T_3 = 300 \text{ nM}, \tag{18c}$$

we get the deterministic description of the system using mass balancing with Michaelis-Menten reaction kinetics:

$$\dot{x}_1 = v_1 = v_{1,f} - v_{1,b} \tag{19a}$$

$$\dot{x}_2 = v_2 = v_{2,f} - v_{2,b} \tag{19b}$$

$$\dot{x}_3 = v_3 = v_{3,f} - v_{3,b}, \tag{19c}$$

with the reaction rates given by

$$v_{1,f} = \frac{k_{11}(T_1 - x_1)}{1 + k_{12}(T_1 - x_1) + k_{12}(T_1 - x_1)k_{13}x_3 + k_{13}x_3} \tag{20a}$$

$$v_{1,b} = \frac{p_{11}x_1}{1 + p_{12}x_1} \tag{20b}$$

$$v_{2,f} = \frac{k_{21}(T_2 - x_2)x_1}{1 + k_{22}(T_2 - x_2)} \tag{20c}$$

$$v_{2,b} = \frac{p_{21}x_2}{1 + p_{22}x_2} \tag{20d}$$

$$v_{3,f} = \frac{k_{31}(T_3 - x_3)x_2}{1 + k_{32}(T_3 - x_3)} \tag{20e}$$

$$v_{3,b} = \frac{p_{31}x_3}{1 + p_{32}x_3}. \tag{20f}$$

The parameter sets for the phosphorylation (k_{ij}) and for the dephosphorylation (p_{kl}) are given in Table 1.

The biochemically relevant equilibrium point x_0 of the system is computed as

Table 1: Parameter Set for the MAP Kinase Signaling Cascade

Parameter	Value	Unit
k_{11}	0.256	l/s
k_{12}	0.0872	l/nM
k_{13}	0.1144	l/nM
k_{21}	$0.4229 \cdot 10^{-2}$	l/s
k_{22}	0.0347	l/nM
k_{31}	$0.2049 \cdot 10^{-2}$	l/s
k_{32}	0.0243	l/nM
p_{11}	0.1003	l/s
p_{12}	0.1428	l/nM
p_{21}	0.1218	l/s
p_{22}	0.0794	l/nM
p_{31}	0.1343	l/s
p_{32}	0.0167	l/nM

$$x_{0,1} = 11.056 \text{ nM} \tag{21a}$$

$$x_{0,2} = 46.968 \text{ nM} \tag{21b}$$

$$x_{0,3} = 44.122 \text{ nM.} \tag{21c}$$

We can determine the local stability of this EP by analyzing the Jacobian $A = \left. \frac{\partial y}{\partial x} \right|_{x=x_0}$ of the system's dynamics evaluated at the EP. The eigenvalues of A are obtained as

$$\lambda_1 = -0.06699 \frac{1}{\text{s}} \tag{22a}$$

$$\lambda_{2,3} = (-0.0004747 \pm 0.03208i) \frac{1}{\text{s}}, \tag{22b}$$

and thus the deterministic system is locally asymptotically stable.

2.5.2 The Stochastic Model

For validation purposes and to get a first insight into the system's dynamics, we did stochastic simulations of the biochemical network model of the MAPK cascade as described by (19). We used the stochastic simulation software Dizzy by the Institute for Systems Biology [17], which is able to do stochastic simulations of models not only with mass action kinetics, but also with arbitrary kinetics such as Michaelis-Menten kinetics, which are used in our model. For the simulations we assumed a cell volume of $V = 1 \text{ pl}$, which corresponds to the absolute number of molecules of each enzyme of

$$N_A \cdot V \cdot T_1 = 60000 \text{ molec} \tag{23a}$$

$$N_A \cdot V \cdot T_2 = 180000 \text{ molec} \tag{23b}$$

$$N_A \cdot V \cdot T_3 = 180000 \text{ molec,} \tag{23c}$$

where N_A is the Avogadro constant. We provide a typical plot of the oscillations of x_1 , x_2 and x_3 against the time in Figure 6. As can be seen from the plot, the oscillations don't vanish as the deterministic calculations suggested, but are sustained oscillations with a determined amplitude range. In the following we will show that these oscillations can be analyzed with the help of Theorem 1.

As discussed in [11] it is possible, under certain assumptions, to describe a system of biochemical reactions by a set of stochastic differential equations, giving rise to the so called Langevin approximation. These assumptions are, roughly speaking, fulfilled if the amount of entities of each species does not get too small, like in our system. The Langevin approximation of the MAPK cascade model (Figure 5) is given by

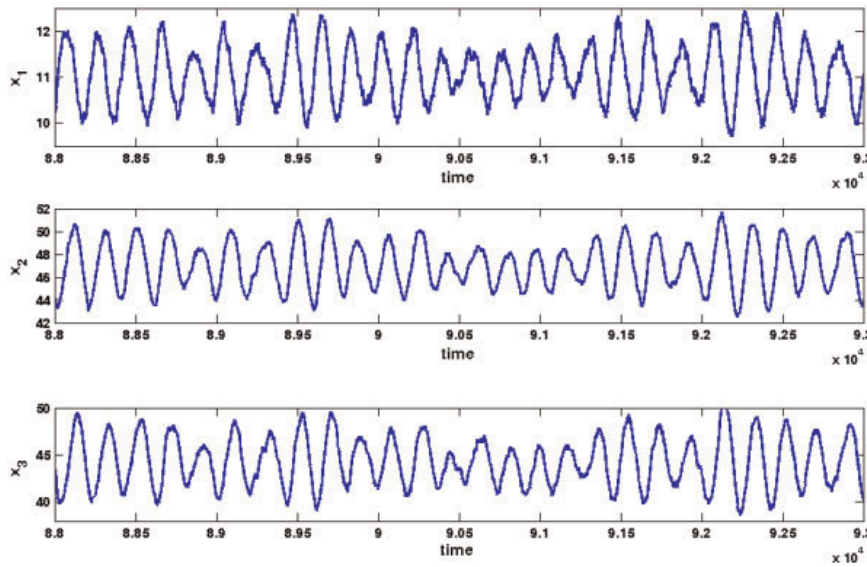


Figure 6
Oscillations of the MAP kinase signaling cascade. Plot of the oscillations of the MAP kinase signaling cascade. The plot is the result of a stochastic simulation via the Direct Gillespie approach with Dizzy.

$$\dot{x}_1 = v_{1,f} - v_{1,b} + \sqrt{\frac{v_{1,f}}{N_A V}} \Gamma_{1,f} + \sqrt{\frac{v_{1,b}}{N_A V}} \Gamma_{1,b} \tag{24a}$$

$$\dot{x}_2 = v_{2,f} - v_{2,b} + \sqrt{\frac{v_{2,f}}{N_A V}} \Gamma_{2,f} + \sqrt{\frac{v_{2,b}}{N_A V}} \Gamma_{2,b} \tag{24b}$$

$$\dot{x}_3 = v_{3,f} - v_{3,b} + \sqrt{\frac{v_{3,f}}{N_A V}} \Gamma_{3,f} + \sqrt{\frac{v_{3,b}}{N_A V}} \Gamma_{3,b} \tag{24c}$$

with $\Gamma_{i,f}$ and $\Gamma_{j,b}$, $i, j \in \{1, 2, 3\}$ uncorrelated, statistical independent Gaussian white noise with zero mean and a variance of one, N_A the Avogadro constant and V the cell volume. This system corresponds to the deterministic one with an additional noise part for every reaction added, with a standard deviation corresponding to the square root of the magnitude of the reaction rate divided by the cell volume. The factors of the white noises, $\sqrt{\frac{v_{i,f}}{N_A V}}$ and $\sqrt{\frac{v_{i,b}}{N_A V}}$ with $i, j \in \{1, 2, 3\}$, are monotone in x and don't change significantly in an area around the EP, so that we can approximate them with their steady state values. Afterwards we can combine each function's noise terms by adding their variances $\sigma_i^2 = \left(\frac{v_{i,f} + v_{i,b}}{N_A V} \right) \Big|_{x=x_0}$, leading to the following equations:

$$\dot{x}_1 = v_{1,f} - v_{1,b} + \sigma_1 \Gamma_1 \tag{25a}$$

$$\dot{x}_2 = v_{2,f} - v_{2,b} + \sigma_2 \Gamma_2 \tag{25b}$$

$$\dot{x}_3 = v_{3,f} - v_{3,b} + \sigma_3 \Gamma_3 \tag{25c}$$

with $\sigma_1 = 3.786 \cdot 10^{-2} \frac{nM}{s}$, $\sigma_2 = 6.350 \cdot 10^{-2} \frac{nM}{s}$ and $\sigma_3 = 10.664 \cdot 10^{-2} \frac{nM}{s}$ for a cell volume of $V = 1$ pl.

2.5.3 Transformation and Model Reduction

The theoretical results in this paper have been developed for two-dimensional systems only. In higher order systems stochastic oscillations may appear on a two dimensional manifold in the state space. Such a system can be reduced to an order two system by computing the slow manifold, making use of a time scale separation [18]. Following this approach, we first transform the system (25) to

$$\dot{z}_1 = A_{fast} z_1 + \psi_{fast}(z_1, z_{slow}) + \Sigma_{fast}(z_1, z_{slow}) \Gamma \tag{26a}$$

$$\dot{z}_{slow} = A_{slow} z_{slow} + \psi_{slow}(z_1, z_{slow}) + \Sigma_{slow}(z_1, z_{slow}) \Gamma, \tag{26b}$$

with $z_1 \in \mathbb{R}$ the coordinate of the fast and $z_{slow} = (z_2, z_3)^T$ the coordinates of the slow manifold, ψ_{fast} and $\psi_{slow} = (\psi_{slow, 1}, \psi_{slow, 2})^T$ vectors of polynomials of order two and higher, ϕ_{fast} and ϕ_{slow} matrices of the noise strength and $\Gamma = (\Gamma_1, \Gamma_2, \Gamma_3)^T$ the noise vector of system (25).

The transformation needs to be done in such a way that the absolute values of the real parts of the eigenvalues of A_{fast} are much larger than these of A_{slow} . To this end, we compute the eigenvectors corresponding to the eigenvalues (22a)-(22b) as

$$e_1 = (0.1506, -0.2699, 0.9510)^T \tag{27a}$$

$$e_{2,3} = (-0.07633 \pm 0.1615i, 0.4923 \pm 0.3441i, 0.7793)^T. \tag{27b}$$

The two eigenvectors with non-zero imaginary parts, e_2 and e_3 , are both eigenvectors corresponding to eigenvalues having real parts with small absolute values. Therefore it is straightforward to define the desired transformation as

$$z = T^{-1}(x - x_0), \tag{28}$$

with the transformation matrix $T = (e_1, e_2 + e_3, (e_2 - e_3) \cdot i)$. We can get the reaction rates of the transformed system by calculating

$$v_z(z) = T^{-1} v(x) \Big|_{x=Tz+x_0}. \tag{29}$$

Because all of the eigenvalues (22a)-(22b) have non-zero real parts, the Hartman-Grobman theorem states that there exists a local transformation $z = H(\zeta)$ such that the system in ζ -coordinates obeys the differential equation

Table 2: Parameter Set of the Slow Manifold of the MAP Kinase Signaling Cascade

Parameter	Value	Parameter	Value	Parameter	Value
$a_{0,0}$	0	$a_{0,2}$	0.01800	$a_{4,0}$	$7.327 \cdot 10^{-8}$
$a_{1,0}$	0	$a_{3,0}$	$-8.476 \cdot 10^{-6}$	$a_{3,1}$	$-8.191 \cdot 10^{-7}$
$a_{0,1}$	0	$a_{2,1}$	$3.625 \cdot 10^{-5}$	$a_{2,2}$	$-8.162 \cdot 10^{-7}$
$a_{2,0}$	0.003604	$a_{1,2}$	$2.425 \cdot 10^{-4}$	$a_{1,3}$	$3.951 \cdot 10^{-6}$
$a_{1,1}$	0.01102	$a_{0,3}$	$5.750 \cdot 10^{-4}$	$a_{0,4}$	$1.785 \cdot 10^{-5}$

$$\dot{\zeta} = \begin{bmatrix} A_{fast} & \mathbf{0} \\ \mathbf{0} & A_{slow} \end{bmatrix} \zeta. \tag{30}$$

Because in the ζ -system there exists a slow manifold spanned by the eigenvectors corresponding to the two eigenvalues with real parts having low absolute values, in the z -system there exists a slow manifold, too (see [19]). We now search the function $z_1 = m(z_2, z_3)$ which describes the dependence of the states on the manifold. We approximate this function with a truncated Taylor series expansion:

$$z_1 = \sum_{j=0}^N \sum_{k=0}^j a_{j,k} z_2^{j-k} z_3^k + O(\|z_{2,3}\|^{N+1}). \tag{31}$$

We substitute (31) in $\dot{z}_1 = v_{z_1}$ and can therewith calculate the coefficients $a_{j,k}$ (see Table 2). The resulting manifold in the original x -coordinates is shown in Figure 7. The figure also includes an example trajectory. As can be seen, this trajectory first converges exponentially fast to an ϵ -neighborhood of the slow manifold and afterwards moves on it towards the EP.

The slow manifold is attractive enough so that realizations of the stochastic system (25) will stay close to it. Due to this, it is sufficient to just take the oscillations on the slow manifold into account and therewith simplify the problem to two dimensions. We substitute the formula of the slow manifold (31) in the differential equations for z_2 and z_3 (29) and get a two dimensional reduced description of the system

$$\dot{z}_2 = \tilde{v}_{z_2} = \tilde{f}_2(z_2, z_3) + \tilde{\Sigma}_2(z_2, z_3)\Gamma \tag{32a}$$

$$\dot{z}_3 = \tilde{v}_{z_3} = \tilde{f}_3(z_2, z_3) + \tilde{\Sigma}_3(z_2, z_3)\Gamma, \tag{32b}$$

to which Theorem 1 can be applied.

2.5.4 Calculation of the Density Distribution

As a first approach to obtain the stationary density distribution $P(x)$, we make use of a linear approximation of the reduced system (32) around the EP. Taking this approach allows us to evaluate how well our method works with such an approximation, where the stationary density distribution can be obtained with minimal computational effort. For other systems, or if a high

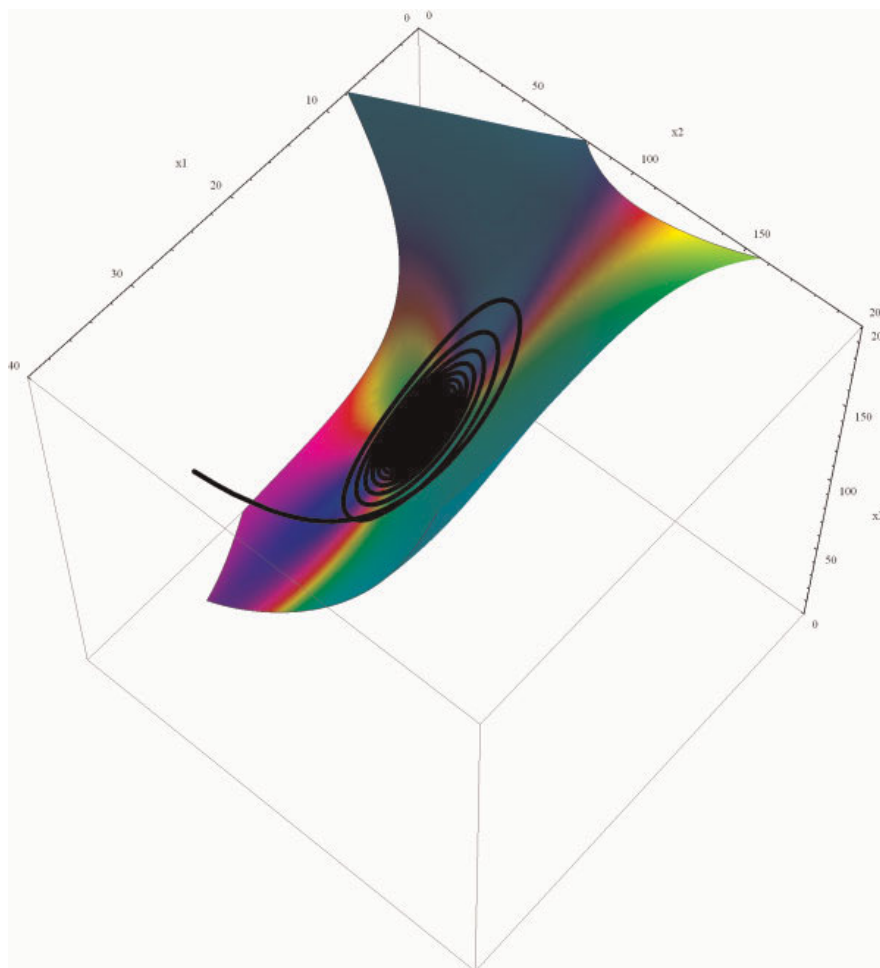


Figure 7
Slow Manifold of the MAP kinase signaling cascade. Slow manifold of (20) calculated up to the order of four. The black curve is a representative trajectory of the deterministic system.

precision of the result is required, it could however also be necessary to solve the nonlinear Fokker-Planck equations numerically.

The linear approximation of the reduced system (32) is given by

$$\dot{\eta} = A\eta + \Sigma_{\eta}\Gamma, \tag{33}$$

with the system matrix

$$A = \begin{bmatrix} -0.00636602 & -0.0901813 \\ 0.0117962 & 0.0054166 \end{bmatrix} \frac{1}{s} \tag{34}$$

and Γ the vector of the disturbances $\Gamma_i, i \in \{1, 2, 3\}$, of the original system (25).

From the coordinate transformation T applied to the original system, Σ_{η} is determined by

$$\Sigma_\eta \Gamma = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Gamma^{-1} \text{diag}(\sigma_i) \Gamma, \tag{35}$$

where $\text{diag}(\sigma_i)$ is a diagonal matrix with the diagonal elements σ_i being the standard deviations of the stochastic terms in the system (25). For the system (33) with $V = 1$ pl, Σ_η is obtained as

$$\Sigma_\eta = \begin{bmatrix} -0.0405545 & 0.063266 & 0.0482555 \\ -0.0499328 & -0.0234468 & 0.011101 \end{bmatrix} \frac{\text{nM}}{\text{s}}. \tag{36}$$

Following the same approach as in the previous example, the stationary density distribution is obtained as

$$P(\eta) \propto \exp\left(-\frac{1}{2} \eta^T \Xi^{-1} \eta\right) \tag{37}$$

with

$$\Xi = \begin{bmatrix} 18.0335 & -1.22879 \\ -1.22879 & 2.38375 \end{bmatrix} \frac{\text{nM}^2}{\text{s}} \tag{38}$$

for a cell with a volume of $V = 1$ pl (see Figure 8).

2.5.5 Determination of the Amplitude Distribution

With the preliminary work of the preceding sections we are now able to calculate the amplitude distribution according to Theorem 1. We consider the states $\mathbf{x}_\chi = (\chi, \alpha, \beta)^T$ with $\alpha, \beta \in \mathbb{R}$ not yet specified. These states are transformed into the z -coordinates by

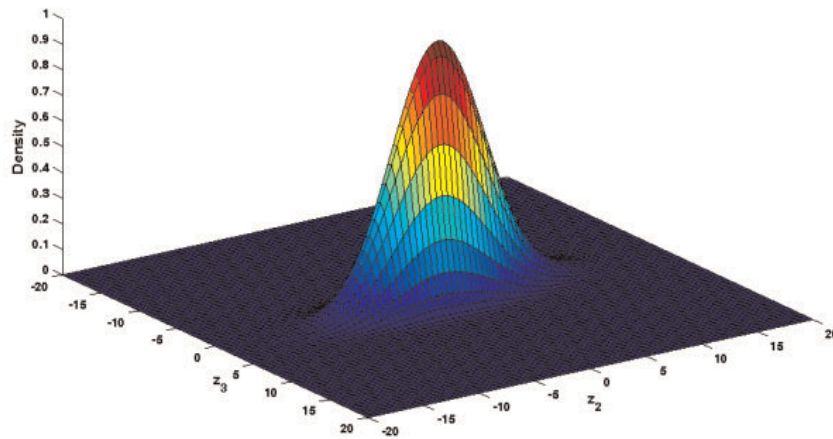


Figure 8
Linear distribution of the state of the MAP kinase signaling cascade. Linear density distribution of the MAP kinase signaling cascade on the slow manifold as calculated in (37). Because we linearized the system, the distribution is Gaussian.

$$\mathbf{z}_\chi = \mathbf{T}^{-1}(\mathbf{x}_\chi - \mathbf{x}_0). \tag{39}$$

The probability of the amplitude χ can now be calculated with the formula

$$P_A(\chi) = \max_{\alpha, \beta} \|\mathbf{v}_\eta\| \cdot P(z_{\chi,2}, z_{\chi,3}) \tag{40}$$

under the constraint that \mathbf{z}_χ has to lie on the slow manifold. This corresponds to that \mathbf{z}_χ has to satisfy (31), which gives us a dependency of β on α . This dependency can be obtained numerically by solving a convex optimization problem.

The resulting amplitude distributions in each of the three original coordinate directions x_i , $i \in \{1, 2, 3\}$ are shown in Figure 9. For comparison, the figure also contains measured amplitude distributions, which are the results of a long term simulation via the Direct Gillespie approach with Dizzy with a simulation over 900000 steps.

As can be seen in Figure 9, the simulation results fit the calculated predictions very well. However, it seems that we tend to underestimate the amplitude of the oscillations by a small amount. This can be explained as a result of Lemma 1, which states that the amplitude of a realization going through the state $\hat{\mathbf{x}}$ lies with a high probability in $[\hat{x}_1, \hat{x}_1 + \delta x_1]$, whereas we estimate its value by the lower end \hat{x}_1 of this interval (see appendix). The tendency to underestimate the amplitude seen in Figure 9 seems to be a direct result of neglecting the value of δx_1 . Further discussion of this point can be found in the conclusions.

We also want to mention that we compare the results obtained from our method to stochastic Gillespie simulations, and not to realizations of the Langevin equation (24). The

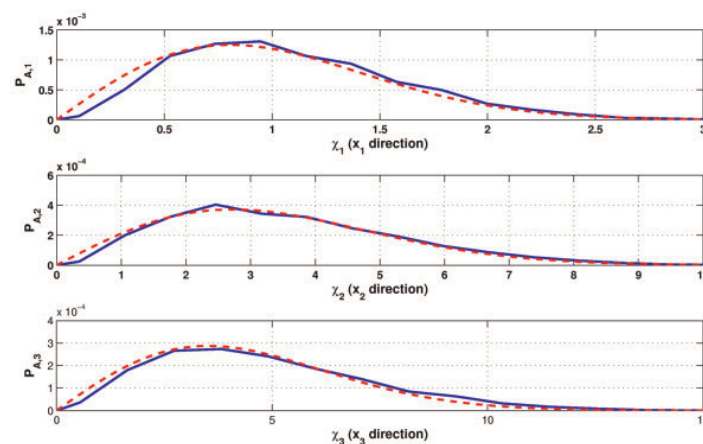


Figure 9
Linear amplitude distribution of the oscillations of the MAP kinase signaling cascade. Linear amplitude distribution of the oscillations of the MAP kinase signaling cascade. The solid curve corresponds to data determined by stochastic simulation of the nonlinear system (24), the dashed curve is the prediction according to the calculations in this section.

reason herefor is that Gillespie simulations better predict the behavior of biochemical networks and are thus the method of choice. However in the derivation of the amplitude distribution we approximated the stochasticity of the system using white noise terms. The results of both methods to describe the intrinsic noise of the system may under certain circumstances lead to different results, which may be a further explanation of the small bias between the measured and the theoretical predicted amplitude distribution in Figure 9. Furthermore some bias may be explained due to the linearization of the system around its steady state.

2.5.6 Numerical Approach

Sometimes it is not adequate anymore to analyze the linearized system and one has to analyze the original nonlinear one. Although there might be several special cases where the probability distribution P of a system of nonlinear stochastic differential equations is analytically computable, this is not possible in the general case. As a consequence there is only the possibility to obtain the solutions numerically. To give an example how to solve a problem with the numerical approach, we decided to analyze the same system as in the preceding section, as defined in (24), except of changing the cell volume to $V = 0.017$ pl, which corresponds to $\frac{1}{60}$ of the original cell volume. Because the variance of the disturbances is highly dependent on the amount of entities of the different protein species, the shape of the solution changes dramatically in a way that we cannot get good results anymore by analyzing the linearized system.

To calculate the density distribution of the nonlinear system we have to solve the nonlinear Fokker-Planck equation (11). We therefore applied the algorithm developed in [20]. The shape of the density distribution, which is non-Gaussian, is shown in Figure 10.

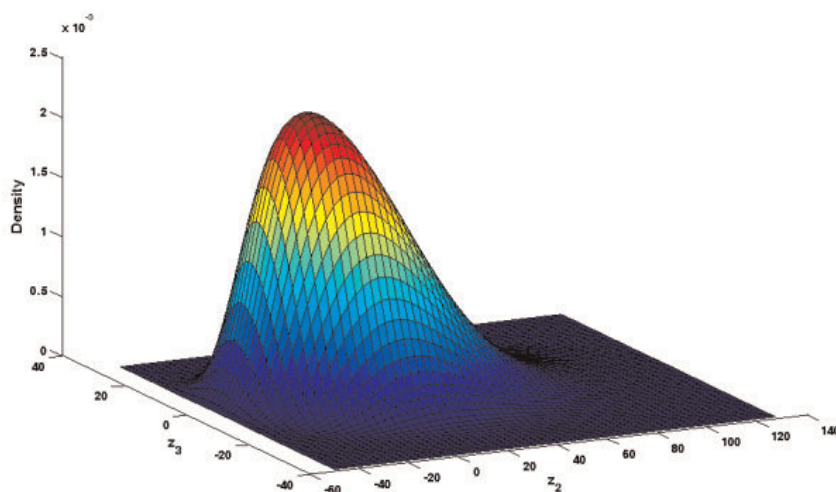


Figure 10
Nonlinear Distribution of the state of the MAP kinase signaling cascade. Density distribution of the MAP kinase cascade on the slow manifold (31).

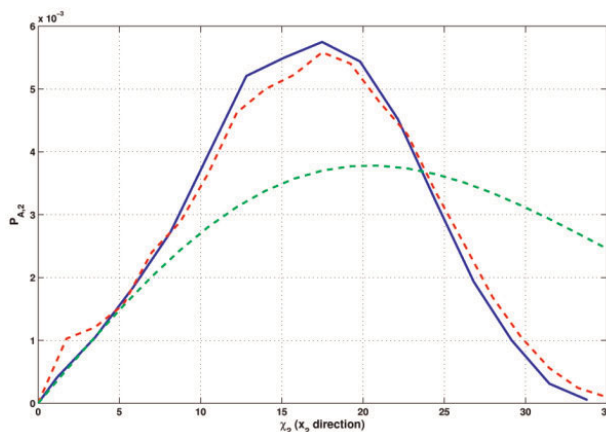


Figure 11
Nonlinear amplitude distribution of the oscillations of the MAP kinase signaling cascade. Nonlinear amplitude distribution of the oscillations of the MAP kinase signaling cascade in the negative x_2 -direction. The solid curve corresponds to data experimentally measured, the dashed red curve is the prediction according to the calculations in this section and the dash-dotted green curve is the outcome we get by calculating the amplitude distribution of the linearized system.

From the density distribution $P(x)$, we can compute the amplitude distribution $P_A(\chi)$ with the method proposed in this paper. In this example, we analyze the amplitude distribution $P_A(\chi)$ of x_2 for peak amplitudes below the equilibrium point values, i.e. $\vec{e} = (0, -1, 0)^T$, because then the effect of the nonlinearities is even higher than in the other direction. The reason for this is quite simple: The amount of entities of each species is not allowed to become negative. As can be seen in Figure 11 the amplitude distribution, as obtained by applying Theorem 1, quite well fits the predicted results gained from a long time simulation while the predicted values of the linearized system are far away from these values.

3. Conclusion

We introduce a method to determine the amplitude distribution of a wide class of linear and nonlinear stochastic systems given by (1), which display sustained stochastic oscillations. The method is applicable to systems where a stationary density distributions exists and can be computed either analytically or numerically.

The method is based on computing the flux density of realizations for the states where the tangent on the level curve of the density distribution is normal to the direction in which the oscillations are measured. We showed that under certain conditions this flux density is directly proportional to the probability of an oscillation with an adequate amplitude to occur. Our results can be used in the analysis of systems being influenced by strong internal or external noise, as we often find them in biophysical problems.

As already discussed at the end of the second example, for certain systems the calculated amplitude distribution may contain a small bias depending on the exact structure of the system

and on how well the the assumption necessary for the application of our method are satisfied. For the wide class of nonlinear systems analyzed we could obtain as a result that realizations being at state \hat{x} will most likely have an amplitude lying in $[\hat{x}_1, \hat{x}_1 + \delta x_1]$ as stated in Lemma 1. By requiring δx_1 to be small we justified to approximate the amplitude by \hat{x}_1 . However, if for specific systems the distribution in $[\hat{x}_1, \hat{x}_1 + \delta x_1]$ can be calculated or estimated, it would be possible to reduce the bias for systems where δx_1 is small but not negligible. Such an extension would lead to a more refined approximation of the amplitude distribution.

Stochastic oscillations may occur in biological systems not only as a disturbing side effect, but also in a constructive manner [21,22]. This is not because stochastic oscillations have great benefits over more "traditional" types of oscillations like deterministic limit cycles, but due to the fact that there seems to be no reason for a preference of deterministic limit cycles. In biophysical systems, the type of oscillations we study in this paper often occurs at parameter values in the vicinity of a Hopf bifurcation in the deterministic part of the model [23]. This is important in the field of robustness analysis of biological networks [24], because stochastic oscillations can possibly improve the robustness of oscillations in a network. In this respect, our method to compute the oscillation amplitude may be helpful in order to characterize a parameter region in which either deterministic or stochastic oscillations of a comparable amplitude occur robustly.

Appendix

Proof for Lemma 1

In this section we give a short proof for Lemma 1. We therefore consider the level curve defined by $\hat{P}(\mathbf{x}) = \hat{P}_0$ of the density distribution going through the state $\hat{x} \in \hat{X}$.

In the following we argue that there exists a δx_1 such that for a high probability an arbitrarily chosen realization of (1) being at state \hat{x} at time t won't reach a state \mathbf{x} with $x_1 > \hat{x}_1 + \delta x_1$ until the next oscillation and therefore the measured amplitude will lie in $[\hat{x}_1, \hat{x}_1 + \delta x_1]$ with δx_1 positive and small but yet not further determined (see Figure 2). We first calculate the Taylor series expansion for P up to the order of two:

$$\begin{aligned}
 P(x_1 + \delta x_1, x_2 + \delta x_2) = & P(x_1, x_2) + \left. \frac{\partial P}{\partial x_1} \right|_{\mathbf{x}=\hat{\mathbf{x}}} \delta x_1 \\
 & + \left. \frac{\partial P}{\partial x_2} \right|_{\mathbf{x}=\hat{\mathbf{x}}} \delta x_2 + \left. \frac{\partial^2 P}{2\partial x_1^2} \right|_{\mathbf{x}=\hat{\mathbf{x}}} \delta x_1^2 + \left. \frac{\partial^2 P}{\partial x_1 \partial x_2} \right|_{\mathbf{x}=\hat{\mathbf{x}}} \delta x_1 \delta x_2 \\
 & + \left. \frac{\partial^2 P}{2\partial x_2^2} \right|_{\mathbf{x}=\hat{\mathbf{x}}} \delta x_2^2 + O(\|\delta \mathbf{x}\|^3).
 \end{aligned}
 \tag{41}$$

After substituting the definitions made in the equations (8) and utilizing the knowledge that $\left. \frac{\partial P}{\partial x_2} \right|_{\mathbf{x}=\hat{\mathbf{x}}}$ vanishes because of the definition of \hat{x} , setting

$$\delta P = P(x_1 + \delta x_1, x_2 + \delta x_2) - P(x_1, x_2), \text{ we get}$$

$$\delta P = a\delta x_1 + \frac{\partial^2 P}{2\partial x_1^2} \Big|_{\mathbf{x}=\hat{\mathbf{x}}} \delta x_1^2 + \frac{\partial^2 P}{\partial x_1 \partial x_2} \Big|_{\mathbf{x}=\hat{\mathbf{x}}} \delta x_1 \delta x_2 + \frac{\kappa}{2} \delta x_2^2 + O(\|\delta \mathbf{x}\|^3). \tag{42}$$

If we stay on a level curve, it must hold that $\delta P = 0$ and δx_1 has to be a function of δx_2 locally around $\hat{\mathbf{x}}$. This function is approximated with another Taylor series expansion up to order two as

$$\delta x_1 = k_1 \delta x_2 + k_2 \delta x_2^2 + O(\|\delta x_2\|^3), \tag{43}$$

with k_1, k_2 yet unknown constants. By substituting (43) in (42) we get

$$0 = k_1 a \delta x_2 + \left(k_2 a + k_1^2 \frac{\partial^2 P}{2\partial x_1^2} \Big|_{\mathbf{x}=\hat{\mathbf{x}}} + k_1 \frac{\partial^2 P}{\partial x_1 \partial x_2} \Big|_{\mathbf{x}=\hat{\mathbf{x}}} + \frac{\kappa}{2} \right) \delta x_2^2 + O(\|\delta \mathbf{x}\|^3). \tag{44}$$

Because the coefficients of δx_2 and δx_2^2 must vanish independently for (44) to hold, it follows that $k_1 = 0$ and $k_2 = -\frac{\kappa}{2a}$, giving

$$\delta x_1 = -\frac{\kappa}{2a} \delta x_2^2 + O(\|\delta x_2\|^3) \tag{45}$$

This means that we may approximate a level curve of the density distribution locally by a parabola. For small deviations and small times t it is possible to make some approximations for (1). First we may approximate $\Sigma(\mathbf{x})$ in an area around $\hat{\mathbf{x}}$ by the constant

$$\hat{\Sigma} := \Sigma(\mathbf{x}) \Big|_{\mathbf{x}=\hat{\mathbf{x}}}. \tag{46}$$

We can rewrite the norm of the first row elements of $\hat{\Sigma}$ by

$$\sqrt{\hat{\Sigma}_{11}^2 + \hat{\Sigma}_{12}^2} \approx \frac{L}{a}, \tag{47}$$

which can be easily validated by the definitions of L and a given in the equations (8). We furthermore approximate the second element of $\mathbf{f}(\mathbf{x})$ in an area around $\hat{\mathbf{x}}$ by

$$v := f_2(\mathbf{x}) \Big|_{\mathbf{x}=\hat{\mathbf{x}}}. \tag{48}$$

Putting everything together we get a one dimensional Wiener process for the movement of (1) in the direction of x_2 for small times Δt :

$$\Delta \hat{x}_2 \approx v + \sqrt{\hat{\Sigma}_{21}^2 + \hat{\Sigma}_{22}^2} \Gamma, \tag{49}$$

with $\Delta \mathbf{x}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}$. We approximate the average displacement in the direction of x_1 by combination of the average displacement of this Wiener process ($\langle x_2(t) \rangle = vt$) with equation (45):

$$\langle \Delta x_1(t) \rangle \approx -\frac{v^2 \kappa}{2a} t^2. \tag{50}$$

Because the stochastic part of (49) may be neglected (see Assumption 4), we can get a linear one dimensional stochastic equation for the movement of (1) in the direction of x_1 for small times t :

$$\Delta \dot{x}_1 \approx -\frac{v^2 \kappa}{a} t + \frac{L}{a} \Gamma. \tag{51}$$

It can be shown [10, p. 129 ff.] that the variance of (51) evolves as

$$\text{var}(\Delta x_1(t)) \approx \frac{L^2}{a^2} t. \tag{52}$$

We may think of the solution of (51) as a growing Gaussian distribution moving along the trajectory of the deterministic system (see Figure 2). We now determine the maximal value of $\Delta x_1(t)$ that a realization of (1) starting at \hat{x} may have if it evolves in an area around the average displacement determined by the standard deviation $\sqrt{\text{var}(\Delta x_1(t))}$. The maximal value of $\Delta x_1(t)$ is time dependent and grows in the beginning due to the stochastic part of (51), but afterwards shrinks due to the deterministic part. It reaches its maximal value when the time derivatives of the mean value and of the standard deviation have the same absolute value:

$$-\frac{\partial}{\partial t} \langle \Delta x_1(t_{max}) \rangle = \frac{\partial}{\partial t} \sqrt{\text{var}(\Delta x_1(t_{max}))} \tag{53}$$

$$\Rightarrow t_{max} = \left(\frac{3 \sqrt{\frac{L}{2v^2 \kappa}}}{3} \right)^2. \tag{54}$$

We can now calculate δx_1 to

$$\delta x_1 = -\frac{v^2 \kappa}{2a} t_{max}^2 + \frac{L}{a} \sqrt{t_{max}} \tag{55}$$

$$= \frac{3 L}{4 a} \sqrt[3]{\frac{L}{2v^2 \kappa}}. \tag{56}$$

For a Gaussian distribution it is true that 68.3% of all realizations stay in an area around the average displacement determined by the standard deviation. This means that only $\frac{100\% - 68.3\%}{2} = 15.9\%$ of the realizations reach an amplitude greater than $\hat{x}_1 + \delta x_1$ after passing \hat{x} . Because it is also possible for a realization to have its maximal value in the direction of x_1 before passing \hat{x} , the overall probability for a realization to have an amplitude lower than $\hat{x}_1 + \delta x_1$ is

$$\Pi(\text{amplitude} \leq x_1 + \delta x_1) = (100\% - 15.9\%)^2 = 70.8\%. \tag{57}$$

In the same way it can be shown that 95.5% of all realizations have an amplitude lower than $\hat{x}_1 + 2\sqrt[3]{2}\delta x_1$.

Proof for Lemma 2

Assume a small unit cell with the edge lengths dr and $d\phi$ with its center being \mathbf{x} and the level curve of P through \mathbf{x} given and going through the center of the edges with the lengths dr (see Figure 3). If dr and $d\phi$ are small, we may assume the density distribution P and the values $\mathbf{f}(\mathbf{x})$ and $\Sigma(\mathbf{x})$ of (1) as constant inside the unit cell. Due to the small size of the Lie-derivative of P along f , the movement normal to the level curve is dominated by the stochastic part of (1). The net fluxes φ_1 and φ_3 trough the edges 1 and 3 are vanishing and therefore the overall flux is tangential to the level curve. The net flux φ_2 (φ_4), trough the edges 2 (the edge 4) can be determined by integration (see, for example, [[25], p. 622]):

$$d\phi_4 = \left\langle \int_{\text{edge } 4} P(\mathbf{x})(\mathbf{f}(\mathbf{x}) + \Sigma(\mathbf{x})\Gamma) \cdot d\vec{n}_x \right\rangle \tag{58}$$

$$\approx Pvd r, \tag{59}$$

$$d\phi_2 = -d\phi_4 \tag{60}$$

By letting dr and $d\phi$ go to zero the absolute value of the net flux density $\|\vec{\varphi}(\mathbf{x})\|$ follows:

$$\|\vec{\varphi}(\mathbf{x})\| = \lim_{dr \rightarrow 0} \left\| \frac{d\phi_{2,4}(\mathbf{x})}{dr} \right\| = P(\mathbf{x})v(\mathbf{x}). \tag{61}$$

Proof for Theorem 1

Because the density distribution P is smooth and the curvature of its level curves does not change its sign, there exists exactly one state $\hat{\mathbf{x}} \in \hat{X}$ with $\hat{x}_1 = \chi$ for every $\chi \in (0, \infty)$. \hat{X} is a smooth simple curve in the state space which can be parameterized by

$$\hat{X} : (0, \infty) \ni \chi \rightarrow \hat{\mathbf{x}}(\chi) = \begin{bmatrix} \chi \\ \hat{x}_2(\chi) \end{bmatrix}, \tag{62}$$

with $\hat{x}_2(\chi)$ an unknown but smooth function in χ . We can get a term proportional to the average number of oscillations of all realizations per time unit $\frac{\Delta N}{\Delta t}$ if we determine the total net flux of realizations φ_t through this curve:

$$\frac{\Delta N}{\Delta t} \propto \phi_t = \int_{\mathbf{x} \in X} \langle \vec{\varphi}(\mathbf{x}) | \vec{n}_0(\mathbf{x}) \rangle, \tag{63}$$

with $\vec{n}_0(\mathbf{x})$ the normalized normal vector on \hat{X} in $\hat{\mathbf{x}}$. We may determine the probability $P_A(\chi)$ of the amplitude $\chi \in (0, \infty)$ by multiplying the flux density of realizations $\vec{\varphi}(\mathbf{x})$ with Ω

(\hat{x}, χ) , the probability that a oscillation of a realization contributing to the flux density at \hat{x} has the amplitude, and integrating over all $\hat{x} \in \hat{X}$:

$$P_A(\chi) \propto \int_{\mathbf{x} \in X} \langle \Omega(\mathbf{x}, \chi) \bar{\varphi}(\mathbf{x}) | \bar{n}_0(\mathbf{x}) \rangle. \tag{64}$$

From the condition $\langle \frac{\partial \hat{x}(\chi)}{\partial \chi} | \bar{n}_0 \rangle = 0$ we get

$$\bar{n}_0 = \frac{1}{\sqrt{1 + \left(\frac{\partial \hat{x}_2(\chi)}{\partial \chi}\right)^2}} \begin{bmatrix} -\frac{\partial \hat{x}_2(\chi)}{\partial \chi} \\ 1 \end{bmatrix}. \tag{65}$$

Because of Lemma 2 we know that the flux density $\bar{\varphi}(\mathbf{x})$ is directed along the x_2 -axis and has the absolute value $P(\hat{x}) \nu(\hat{x})$. Therefore we can calculate (64) to

$$P_A(\chi) \propto \int_0^\infty \Omega(\mathbf{x}(\tilde{\chi}), \chi) \begin{bmatrix} 0 \\ P(\mathbf{x}(\tilde{\chi})) \nu(\mathbf{x}(\tilde{\chi})) \end{bmatrix} \cdot \begin{bmatrix} -\frac{\partial x_2}{\partial \chi}(\tilde{\chi}) \\ 1 \end{bmatrix} d\tilde{\chi} \tag{66}$$

$$= \int_0^\infty \Omega(\mathbf{x}(\tilde{\chi}), \chi) P(\mathbf{x}(\tilde{\chi})) \nu(\mathbf{x}(\tilde{\chi})) d\tilde{\chi}. \tag{67}$$

Although we do not know the exact shape of $\Omega(\hat{x}, \chi)$, which may be complicated to calculate, we know a lower and an upper bound for Ω . The lower bound is given due to the fact that Ω must be zero for $\chi < \hat{x}_1$, because if (1) goes through the state \hat{x} , it has at least the amplitude of this state. An upper bound for almost all realizations is given by Lemma 1 with $\hat{x}_1 + 2\sqrt[3]{2}\delta x_1$. We may therefore conclude that Ω is approximately zero for χ being outside of the interval $[\hat{x}_1, \hat{x}_1 + 2\sqrt[3]{2}\delta x_1]$, which simplifies (67) to

$$P_A(\chi) \propto \int_{-2\sqrt[3]{2}\delta x_1}^0 \nu \left(\begin{bmatrix} \chi + \xi \\ x_2(\chi + \xi) \end{bmatrix} \right) P \left(\begin{bmatrix} \chi + \xi \\ x_2(\chi + \xi) \end{bmatrix} \right) \Omega \left(\begin{bmatrix} \chi + \xi \\ \hat{x}_2(\chi + \xi) \end{bmatrix}, \chi \right) d\xi. \tag{68}$$

If δx_1 is small, we can assume P and ν being approximately constant in the integration interval, which simplifies (68) to

$$P_A(\chi) \propto \int_{-2\sqrt[3]{2}\delta x_1}^0 \nu \left(\begin{bmatrix} \chi \\ x_2(\chi) \end{bmatrix} \right) P \left(\begin{bmatrix} \chi \\ x_2(\chi) \end{bmatrix} \right) \Omega \left(\begin{bmatrix} \chi + \xi \\ x_2(\chi + \xi) \end{bmatrix}, \chi \right) d\xi \tag{69}$$

$$= P(\mathbf{x}(\chi)) \nu(\mathbf{x}(\chi)) \int_{-2\sqrt[3]{2}\delta x_1}^0 \Omega \left(\begin{bmatrix} \chi + \xi \\ x_2(\chi + \xi) \end{bmatrix}, \chi \right) d\xi$$

For small δx_1 the integral of Ω over $(-2\sqrt[3]{2}\delta x_1, 0)$ is approximately 1. The result of this last approximation is given by Theorem 1.

References

1. Leloup JC and Goldbeter A: *Proc Natl Acad Sci USA* 2003, **100(12):7051–7056**.

2. Tiedemann HB, Schneltzer E, Zeiser S, Rubio-Aliaga I, Wurst W, Beckers J, Przemeczek GKH and de Angelis MH: *J Theor Biol* 2007, **248**:120–129.
3. van Kampen NG: *Stochastic processes in Physics and Chemistry* Amsterdam, The Netherlands: Elsevier; 32007.
4. Muratov CB and Vanden-Eijnden E: *Chaos* 2008, **18**:0151111-1–0151111-11.
5. El-Samad H and Khammash M: *Proceedings of the 45th IEEE Conference on Decision and Control* 2006, 2382–2387.
6. Beuter A, Glass L, Mackey MC and Titcombe MS: *Nonlinear Dynamics in Physiology and Medicine, of Interdisciplinary Applied Mathematics* New York: Springer; 2003, **25**.
7. Kuske R, Gordillo LF and Greenwood P: *Journal of Theoretical Biology* 2007, **245**:459–469.
8. Yu N, Kuske R and Li Y: *Physical Review E* 2006, **73**:056205-1–056205-8.
9. Aparicio JP and Solari HG: *Mathematical Biosciences* 2001, **169**:15–25.
10. Gillespie DT: *Markov Processes* San Diego, CA: Academic Press, INC; 1992.
11. Gillespie DT: *Journal of Chemical Physics* 2000, **113**:297–306.
12. Bagheri N, Stelling J and Doyle FJ: *Bioinformatics* 2007, **23(3)**:358–364.
13. Gardiner CW: *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* Springer; 32004.
14. Pearson G, Robinson F, Gibson TB, Xu BE, Karandikar M, Berman K and Cobb MH: *Endocrine Reviews* 2001, **22(2)**:153–183.
15. Kholodenko B: *European Journal of Biochemistry* 2000, **67(6)**:1583–1588.
16. Hilioti Z, Sabbagh W, Paliwal S, Bergmann A, Goncalves MD, Bardwell L and Levchenko A: *Curr Biol* 2008, **18(21)**:1700–06.
17. Ramsey S, Orrell D and Bolouri H: *Journal of Bioinformatics and Computational Biology* 2005, **3**:415–436.
18. Khalil HK: *Nonlinear Systems* Upper Saddle River, NJ, USA: Prentice-Hall; 32001.
19. Jacobs SJ: *Journal of Atmospheric Sciences* 1991, **48(7)**:893–901.
20. Sjöberg P: **Numerical Solution of the Fokker-Planck Approximation of the Chemical Master Equation.** PhD thesis Uppsala Universitet; 2005.
21. Rao CV, Wolf DM and Arkin AP: *Nature* 2002, **420(6912)**:231–237.
22. Steuer R, Zhou C and Kurths J: *Biosystems* 2003, **72(3)**:241–251.
23. Waldherr S and Allgöwer F: *International Journal of Systems Science* 2009, **40**:769–782.
24. Kitano H: *Nat Rev Genet* 2004, **5(11)**:826–837.
25. Bronstein IN and Semendjajew KA: *Taschenbuch der Mathematik* Frankfurt/Main, Germany: Verlag Harri Deutsch; 201983.