Single-variable reaction systems: Deterministic and stochastic models

M.N. Steijaert a, A.M.L. Liekens b, D. Bošnački c, P.A.J. Hilbers a, H.M.M. ten Eikelder a,*

a Department of Biomedical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
b VIB Department of Molecular Genetics, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium

corresponding author. Tel.: +31 40 247 5153.
E-mail address: h.m.m.t.eikelder@tue.nl (H.M.M. ten Eikelder).

Abstract
Biochemical reaction networks are often described by deterministic models based on macroscopic rate equations. However, for small numbers of molecules, intrinsic noise can play a significant role and stochastic methods may thus be required. In this work, we analyze the differences and similarities between a class of macroscopic deterministic models and corresponding mesoscopic stochastic models. We derive expressions that provide a clear and intuitive view upon the behavior of the stochastic model. In particular, these expressions show the dependence of both the dynamics and the stationary distribution of the stochastic model on the number of molecules in the system. As expected, most properties of the stochastic model correspond well with those in the deterministic model if the number of molecules is large enough. However, for some properties, both models are inconsistent, even if the number of molecules in the stochastic model tends to infinity. Throughout this paper, we use a bistable autophosphorylation cycle as a running example. For such a bistable system, we give an explicit proof that the rate of convergence to the stationary distribution (or the second eigenvalue of the transition matrix) depends exponentially on the number of molecules.

1. Introduction
In the past decades, new biochemical techniques have led to a large increase in knowledge about biology at the molecular level. Simultaneously, a growing awareness has emerged that reductionism alone is insufficient for unraveling the complex interactions within biochemical networks. This awareness has led to the rise of the holistic field of systems biology, in which wet-lab experiments are combined with ( multiscale ) computational modeling. Traditionally, most of the systems investigated in this field are described by deterministic models based on ordinary differential equations (ODEs). In addition, there is a growing interest in stochastic modeling techniques [1,2]. The advantage of those stochastic techniques is that they explicitly take into account the intrinsic noise that is present in real-life biochemical networks.

In this paper, we focus on the relations between stochastic and deterministic models for a certain class of biochemical systems. More specifically, we consider systems consisting of a number of similar molecules that can each be in two configurations. Suppose that the reversible interconversion between those configurations is defined by a macroscopic deterministic model based on kinetic rate laws. In principle, such a deterministic model describes the behavior of a reaction system for very large numbers of molecules. For instance, the behavior of a reaction system for very large numbers of molecules can be described by deterministic models based on macroscopic rate equations. However, for small numbers of molecules, intrinsic noise can play a significant role and stochastic methods may thus be required. In this work, we analyze the differences and similarities between a class of macroscopic deterministic models and corresponding mesoscopic stochastic models. We derive expressions that provide a clear and intuitive view upon the behavior of the stochastic model. In particular, these expressions show the dependence of both the dynamics and the stationary distribution of the stochastic model on the number of molecules in the system. As expected, most properties of the stochastic model correspond well with those in the deterministic model if the number of molecules is large enough. However, for some properties, both models are inconsistent, even if the number of molecules in the stochastic model tends to infinity. Throughout this paper, we use a bistable autophosphorylation cycle as a running example. For such a bistable system, we give an explicit proof that the rate of convergence to the stationary distribution (or the second eigenvalue of the transition matrix) depends exponentially on the number of molecules.

For some systems it is not trivial to explain how the expected behavior of the stochastic model relates to that of the deterministic model. This paper studies the relation between the stochastic and deterministic models for a specific class of reaction systems. As a running example, we study a bistable reaction system. More precisely, we use both a deterministic and a stochastic model to describe a highly idealized reaction module consisting of a phosphorylation reaction, a de-phosphorylation reaction and a trans-autophosphorylation reaction. This running example, further referred to as `autophosphorylation cycle', illustrates the paradoxical combination of bistability and stochasticity. After all, the long-term behavior of a stochastic model of a chemical reaction system

© 2010 Elsevier Inc. All rights reserved.

doi:10.1016/j.mbs.2010.06.006

0025-5564/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
is in general independent of its initial state, which intuitively conflicts with the required bistability.

In Section 3, we derive a potential function that provides a framework for the analysis of the stochastic model for both small and larger numbers of molecules. This function is used to describe the stationary distribution of the stochastic model and the expected transition times (see Section 4). The expected times of transitions between individual states are also related to the rate at which the probability distribution converges to the stationary distribution. This convergence clearly depends on the eigenvalues of the transition matrix of the stochastic model. We prove in Section 5 that under certain conditions, the transition matrix has an eigenvalue which converges to zero exponentially fast with an increasing number of molecules. In each of Sections 2–5, we first introduce the generic theory and subsequently exemplify this using the autophosphorylation cycle introduced in Section 2. A discussion of both the generic theory and the autophosphorylation cycle follows in Section 6. In this section, we also discuss the resemblances and differences between our work and some classic and more recent research papers in the field of statistical physics.

2. Two models

2.1. Model definitions

In this paper, we focus on systems in which each molecule can be in two 'configurations'. This type of system occurs in many different biological processes. For instance, many signaling proteins can be modified by reversible post-translational modifications such as phosphorylation or methylation [6]. Other examples include molecules that can switch between different conformations or localizations within a cell.

The generic system consists of a total of \( N \) molecules, which can be in either of the configurations \( X_0 \) and \( X_1 \), and the overall reactions

\[
X_0 \leftrightarrow X_1
\]

To allow a comparison between systems with various values of \( N \), we define the proportion \( x \) of the molecules that is in the \( X_1 \) configuration. This proportion can be derived from the number of molecules (denoted with \#) or from the concentrations (in square brackets):

\[
x = \frac{\#X_1}{\#X_0 + \#X_1} = \frac{\#X_1}{N} = \frac{[X_1]}{[X_0] + [X_1]}.\]

As mentioned before, we will compare two representations of this generic system.

The first representation of this system is the 'deterministic model'. In this representation, we use a system of ODEs to describe the time evolution of \( x \). The rate at which \( x \) changes due to the reaction \( X_1 \rightarrow X_0 \) is given by a real, smooth, non-negative function \( a(x) \); the rate of the opposite reaction is defined by a real, smooth, non-negative function \( b(x) \). If we assume \( a(x) > 0 \) for all \( 0 < x < 1 \), \( a(1) = 0 \), \( b(x) > 0 \) for all \( 0 < x < 1 \) and \( b(0) = 0 \). The dynamics of the deterministic model are given by

\[
\frac{dx}{dt} = a(x) - b(x). \tag{1}
\]

The second representation of the system described above is the 'stochastic model'. In this representation, we consider all possible 'microstates' (further 'states'), each of which corresponds with an integer number of molecules in configuration \( X_1 \). Hence, a model with \( N \) molecules has \( N+1 \) possible states, which are numbered according to the number of \( X_1 \) molecules (see Fig. 1). We allow only one reaction to occur at a time; hence, from each state only direct neighbor states can be reached in one reaction step. Due to this 'one-step-at-a-time property' and the Markovian properties of the model, the model is in fact a birth-and-death process [7] in which the forward and backward propensities \( x_0 \) and \( x_2 \) have the role of birth and death rate, respectively.

One of the obvious differences between both models is that \( N \) is a parameter of the stochastic model but not an explicit part of the deterministic model. Consequently, the expected behavior of the stochastic model changes with growing \( N \), while the deterministic model is independent of \( N \). To allow a useful comparison, the stochastic models with different values of \( N \) must all have the same total concentration as is used in the deterministic model. This means that the volumes of the stochastic models scale linearly with \( N \).

The propensities of state change \( x_0 \) and \( x_2 \) in the stochastic model can be related to the deterministic production rates \( a(x) \) and \( b(x) \) as follows:

\[
x_0 = Na \left( \frac{k}{N} \right) \tag{2}
\]

\[
x_2 = Nb \left( \frac{k}{N} \right). \tag{3}
\]

Using those expressions, we ensure that both models act on the same time scale. The relations in (2) and (3) can be used for all unimolecular and pseudo-unimolecular reactions (i.e., reactions in which the total number of molecules does not change). In this way, assumptions about the deterministic model are transferred to the stochastic model. As a result, elementary steps that are hidden in the deterministic model remain hidden in the stochastic model. For the Michaelis–Menten reaction, this is discussed in [8].

Let \( z(t) \) be the probability distribution vector, which contains the probabilities \( z_k(t) \) that the system is in a state \( k \in \{0, \ldots, N\} \) at a time \( t \). The dynamics of the state probabilities in the stochastic model are described by the Chemical Master Equation [9]:

\[
\frac{dz}{dt} = M \cdot z, \tag{4}
\]

where the tridiagonal matrix \( M = (m_{ij}) \) (with \( i,j = 0, \ldots, N \)), is given by

\[
m_{ij} = \begin{cases} 
z_j & \text{if } j = i - 1, \\
\beta_j & \text{if } j = i + 1, \\
-\alpha_i - \beta_i & \text{if } j = i, \\
0 & \text{otherwise.}
\end{cases}
\]

The stationary distribution for such a birth-and-death process follows directly from its (stochastic) detailed balance property [10] and is given by:

\[
Z_k = \frac{\prod_{i=0}^{k-1} z_i}{\sum_{j=0}^{N} \prod_{i=0}^{j-1} z_i}. \tag{5}
\]

2.2. Trans-autophosphorylation

Throughout this paper we exemplify our methods with the bistable reaction system given by the following reactions:

\[
X + S \xrightleftharpoons{k_{1}} X_0 + S, \tag{6}
\]

\[
X + X_0 \xrightleftharpoons{k_2} 2X_0, \tag{7}
\]

\[
X_0 + U \xrightleftharpoons{k_3} X + U. \tag{8}
\]

Fig. 1. The \( N + 1 \) states of the stochastic model.
This system consists of proteins X that can be phosphorylated into X by a kinase S. In addition, protein molecules X are also subject to an inter-molecular (or trans-) autophosphorylation reaction in which phosphorylated kinase molecules X phosphorylate other molecules X. Finally, the phosphatase U de-phosphorylates X. We further refer to this system as the 'autophosphorylation cycle'.

For this specific case study, \( x = \frac{a_{xy}}{a_{xy} + a_{yx}} \) represents the proportion phosphorylated protein. We assume Michaelis–Menten kinetics with identical normalized Michaelis constants \( k \) for all three reactions, catalytic rate constants \( k_1 \) and \( k_2 \) for the phosphorylation reactions (6) and (7), respectively, and the maximum velocity \( k_3 \) for the de-phosphorylation reaction (8) with a constant phosphatase concentration [U]. Hence, the reaction rates for the deterministic model are given by

\[
a(x) = \frac{k_1(1-x)k_3}{k_1 + k_2} + \frac{k_2x(1-x)}{k_1 + k_2}
\]

\[
b(x) = \frac{k_3x}{k_1 + k_2}.
\]

For reasons of simplicity we consider dimensionless parameters. We further use \( k_1 = k_2 = k_3 = 1 \) and \( k = \frac{1}{x} \), although similar results can be obtained for a broader range of parameters. Note that the only remaining parameter is the 'input signal' \( s \), which represents the total amount of kinase S.

For both models of the autophosphorylation cycle, we can easily infer the long-term behavior. The deterministic model has one or more steady states that can either be stable or unstable. As shown in Fig. 2(a), the number of steady states depends on \( s \). For input signals \( s < s_1 = 0.651 \) the system has a stable steady state \( x = \gamma_1(s) \) with a low proportion of phosphorylated molecules, while for \( s \geq s_1 = 0.373 \) there is a stable state \( x = \gamma_2(s) \) with a high proportion of phosphorylated molecules. The most interesting parameter region is \( s_1 < s < s_2 \), where both stable steady states \( \gamma_1(s) \) and \( \gamma_2(s) \) exist, separated by an unstable steady state \( x = \gamma_3(s) \). If \( s \) is varied within this 'bistable range', the model displays hysteresis. In the remainder of this paper, we omit the argument \( s \) and write \( \gamma_1 \) instead of \( \gamma_1(s) \), etc.

For the stochastic model, a first indication of the long-term behavior is given by the stationary distribution \( \mathbf{z} \), which depends on both the input parameter \( s \) and the number of molecules \( N \). Typically, this distribution \( \mathbf{z} \) is bimodal for \( s \) within the bistable range (see for instance Fig. 2(b)) and unimodal for all other values of \( s \). In analogy with the \( \gamma_1, \gamma_2 \) and \( \gamma_3 \) for the deterministic model, we refer to the left local maximum, the right local maximum and the intermediate minimum of the bimodal distribution as \( G_1 \), \( G_2 \) and \( G_3 \), respectively. The maxima of the unimodal distributions are referred to as \( G_l \) for \( s < s_1 \) and \( G_3 \) for \( s > s_2 \). We are primarily interested in the behavior of this reaction system in the bistable range. In Section 3.2, we discuss how the corresponding bimodal stationary distributions depend on both \( s \) and \( N \).

3. Asymptotic form of the stationary distribution

3.1. The kinetic potential \( F \)

In this section, we rewrite the stationary distribution in such a way that its behavior for \( N \to \infty \) can easily be inferred. It is easily seen from (5) that the ratio of the stationary probabilities of states \( k \) and \( \ell \) (with \( k < \ell < N \)) equals

\[
Z_k \cdot Z_{\ell} = \prod_{i=k+1}^{\ell} \frac{a(i)}{a(i-1)}, \quad a(i) = a_i, \quad a(i) = \frac{a(i)}{a(i-1)} \prod_{i=k+1}^{\ell} b(i).
\]

We first rewrite this expression such that its behavior for large \( N \) becomes more clear. Define \( F \) and \( Q_\ell \) by

\[
F(x) = -\int_0^x \ln \left( \frac{a(y)}{b(y)} \right) dy,
\]

\[
Q_\ell = \prod_{i=k+1}^{\ell} \frac{a(i)}{b(i)}.
\]

Note that \( F \) does not depend on the number of molecules \( N \). In fact \( F \) has the role of a 'potential function'. As it is purely derived from the kinetic description of the reaction system, this function is further called the 'kinetic potential'. A further discussion about the use of this and related potential functions in the literature is given in Section 6.3.

Trivially

\[
\frac{1}{N} \ln Q_\ell = -\frac{1}{N} \sum_{i=k+1}^{\ell} \ln \left( \frac{a(i)}{b(i)} \right) = -\frac{1}{N} \sum_{i=k}^{\ell} c \left( \frac{1}{N} \right)
\]

with the function \( c \) defined by

\[
c(x) = \ln \left( \frac{a(x)}{b(x)} \right).
\]

The right-hand side of (12) can be seen as a Riemann sum, i.e., as an approximation of an integral. The rectangle rule states that for sufficiently smooth functions \( c \)

\[
\int_s^\ell c(x) \, dx = h \sum_{i=1}^m c(x_i + ih) + \Delta(x, \beta, h),
\]

in which \( h = \frac{\ell - s}{m} \) is the step size and the error \( \Delta(x, \beta, h) \) has the asymptotic behavior

\[
\Delta(x, \beta, h) = \frac{h}{2} \left( c(x) - c(\beta) \right) + O(h^2).
\]
see for instance [11]. Using this relation we can rewrite (12) as
\[
\frac{1}{N} \ln Q_{\ell k} = \int \frac{c(x)dx}{N} \left( c(k) - c(\ell) \right) + O(N^{-2})
\]
\[
= -F(\ell) + F(k) - \frac{1}{2N} \left( c(k) - c(\ell) \right) + O(N^{-2}).
\]
(14)

where \( \ell \equiv \ell/N \) and \( k \equiv k/N \). Eliminating \( c \) with (13) finally results in
\[
Q_{\ell k} = e^{\Delta F(k)-\Delta F(\ell)} \frac{a(\ell)b(k)}{a(k)b(\ell)} (1 + O(N^{-1})).
\]
(15)

The relations (14) and (15) only hold for \( 0 < k, \ell < N \). In other cases, (14) does not hold since either \( c(k) \) or \( c(\ell) \) is undefined. Substitution of (15) in (9) gives
\[
\frac{k}{z} \rightarrow k \quad \text{and} \quad \frac{\ell}{z} \rightarrow \ell \quad \text{if} \quad N \rightarrow \infty.
\]

Using (16) we see that the behavior of \( Z/nz \) is determined by the sign of the difference \( F(k) - F(\ell) \). If this difference is negative, the quotient \( Z/nz \) will increase exponentially with \( N \); if this difference is positive, the quotient will decrease exponentially with increasing \( N \). Only if \( F(\ell) = F(k) \), then \( Z/nz \) will tend to a finite constant value if \( N \) increases. This implies that, for sufficiently large \( N \), the maxima of the stationary distribution correspond with the minima of \( F \) and the minima of the stationary distribution correspond with maxiima of \( F \). Clearly, the kinetic potential \( F \) determines the behavior of the system for large values of \( N \).

For the deterministic system, it is easily verified that \( \frac{dF(x)}{dx} < 0 \) for solutions \( x(t) \) of (1) and that stable/unstable steady states of (1) correspond with internal minima/maxima of \( F \). This means that \( F \) can be used to construct a Lyapunov function for (1).

Above we have shown that (for large enough \( N \)) the locations of the minima/maxima of the stationary distribution correspond with the locations of the maxima/minima of \( F \). As discussed above, the unstable/stable steady states of the deterministic model also correspond with the maxima/minima of \( F \). Consequently, if \( N \) is large enough, the locations of the maxima/minima of the stationary distribution of the stochastic model correspond with the stable/ unstable states of the deterministic model.

3.2. Asymptotic behavior of the autophosphorylation cycle

Now we apply the methods from Section 3.1 on the autophosphorylation cycle. Fig. 3 shows the kinetic potential \( F \) for input \( s = 0.52 \). Typically, within the bistable range, this function has two local minima. For all \( s \) in the bistable range, the function \( F \) has two local minima in \( \gamma_1 \) and \( \gamma_2 \), separated by a local maximum in \( \gamma_2 \). This also means that for large enough \( N \) the locations of the extrema \( G_1, G_2 \) and \( G_3 \) of the stationary distribution correspond with \( \gamma_1, \gamma_2 \) and \( \gamma_3 \), respectively.

The relative order of the two maxima of the stationary distribution in \( G_1 \) and \( G_3 \) can also easily be found. For \( s \leq \sigma_0 \approx 0.5297 \) the global minimum of \( F \) is in \( \gamma_1 \), while for \( s > \sigma_0 \) the global minimum of \( F \) is in \( \gamma_2 \). Now consider the case \( s < \sigma_0 \) (as shown in Fig. 3). In that case, \( F(\gamma_1) < F(\gamma_2) \) and thus (16) implies that the global maximum of the stationary distribution is in \( G_1 \), if \( N \) is sufficiently large. Moreover, the quotient of the probabilities in \( G_1 \) and \( G_1 \) tends to zero exponentially fast in \( N \). Hence, although the local maximum of the stationary distribution in \( G_1 \) exists for all \( N \), this maximum becomes exponentially small compared to the dominant maximum in \( G_1 \). In fact, for large \( N \) and \( s < \sigma_0 \) most probability in the stationary distribution is concentrated around \( G_1 \). This means that, for \( t \rightarrow \infty \), the system will almost always be near \( \gamma_1 \). Similarly, if the input signal \( s < \sigma_0 \), the dominant maximum of the stationary distribution will be in \( G_2 \). In this case, the system will almost always be near \( \gamma_2 \) for large \( t \). Only for \( s = \sigma_0 \) both minima of \( F \) have the same value and both local maxima of the stationary distribution have non-vanishing (but not necessarily equal) probabilities. The predicted dependence of the stationary distribution on \( s \) and \( N \) can indeed be observed in Fig. 4. To allow a better comparison of various values of \( N \), this figure shows \( NZ_0F(k) \) as a function of the proportion phosphorylated proteins \( k/N \).

For \( s = \sigma_0 \) the previous considerations imply that the stationary distribution is ‘hardly bimodal’ if \( N \) is large. This contrasts with the deterministic model, which has two stable steady states for all \( s \) in the interval \( (\sigma_1, \sigma_2) \). However, note that we are considering the stationary distribution, i.e., the distribution for \( t \rightarrow \infty \). The system may still exhibit bistable behavior for large (finite) times. In the next section, we show that the time needed before the system loses its bistability increases exponentially with \( N \).

4. Expected transition times


It is often interesting to see how fast one state in the stochastic model is reached from another state. Consider for example the aforementioned autophosphorylation cycle with \( s \) is in the bistable range. The stochastic model predicts that, in the long run, the system can most likely be found in one of the maxima of the stationary distribution. However, the stationary distribution provides no information about the time required to reach one of those maxima from a given initial state. In this section, we show how the kinetic potential \( F \) can be used to calculate the expected time \( T_{k \rightarrow \ell} \) for a transition from state \( k \) to state \( \ell \).

Let us first focus on the expected transition times between neighboring states. The expected time for a transition from state \( k \) to state \( k+1 \) is given by
\[
T_{k \rightarrow k+1} = \sum_{\ell \neq k} \frac{1}{\beta_\ell} \prod_{j=1}^{k-1} \frac{\beta_j}{\gamma_j},
\]
(17)

see for instance [7,10]. This ‘single-step time’ \( T_{k \rightarrow k+1} \) takes into account all possible paths from \( k \) to \( k+1 \), including paths that visit \( \gamma_1 \).
states $\ell < k$. We first analyze how $T_{k-k+1}$ changes if $N$ tends to infinity. Rewriting (17) in terms of the functions $a(x)$ and $b(x)$ yields

$$
T_{k-k+1} = \sum_{i=0}^{k} \frac{1}{Na(i)} \prod_{j=1}^{k-i} a(\hat{k}) = \sum_{i=0}^{k} \frac{1}{Na(i)} Q_{t,k},
$$

in which $Q_{t,k}$ is defined by (11). We would like to use the expression for $Q_{t,k}$ given in (15). However, that expression does not hold for $i=0$. Therefore, for $0 < k < N$, we separate the summation as follows:

$$
T_{k-k+1} = \sum_{i=0}^{k} \frac{1}{Na(i)} b(\hat{k}) Q_{t,k} + \sum_{i=0}^{k} \frac{1}{Na(i)} a(\hat{k}) Q_{t,k}.
$$

Since $b(\hat{k}) = O(N^{-1})$, the first term is $O(N^{-1})$ times the second term. Hence we can write

$$
T_{k-k+1} = \left( \sum_{i=0}^{k} \frac{1}{Na(i)} Q_{t,k} \right) 1 + O(N^{-1}).
$$

Using (15), we can rewrite (18) as

$$
T_{k-k+1} = \sum_{i=0}^{k} \frac{1}{Na(i)} e^{N(i(i)-i(\hat{k}))} \left( a(\hat{k}) b(\hat{k}) \right) 1 + O(N^{-1}).
$$

We now study the behavior of $T_{k-k+1}$ in the limit $N \to \infty$, $k \to k$ for various points $k$. From (19) follows that

$$
T_{k-k+1} = O(N^{-1})
$$

if $F(x) > F(\hat{k})$ for all $x < \hat{k}$,

$$
T_{k-k+1} = Ce^{N}$

for positive constants $C$ and $\gamma$

if $F(x) < F(\hat{k})$ for any $x < \hat{k}$.

The first case will be further referred to as ‘non-exponential-single-step times’, while the second will be referred to as ‘exponential single-step times’. A similar dependence on $F$ can be found for single-step times $T_{k-k-1}$ in the opposite direction.

The reason for the occurrence of exponential single-step times is that a path from state $k$ to $k+1$ may include ‘detours’ via states $m$ with $m < k$. Both exponential and non-exponential single-step times include this kind of detours, for which the probability decreases with $N$ while the expected time increases with $N$. The difference between the two types of single-step times is that for the non-exponential type the decreasing probability of detours cancels out their growing expected times, while for the exponential type the expected times of some rare paths increase faster in $N$ than their probabilities decrease.

For the non-exponential single-step times, a more precise analysis yields that

$$
T_{k-k+1} = \frac{1}{Na(k) - b(k)} + O(N^{-2}) \text{ for } N \to \infty.
$$

Note that this behavior corresponds with that of the deterministic model, which moves from any $y = k$ with $a(y) > b(y)$ to $y + \Delta y = k + \frac{1}{2}$ in approximately a time $\frac{\pi}{\sqrt{2y+\Delta y}}$.

Next, we consider the expected time for a transition from state $k$ to a state $\ell$ with $\ell > k$, which is given by

$$
T_{k-\ell} = \sum_{i=0}^{\ell} T_{i-i+1}.
$$

The combination of this expression with (19) yields an expression of the expected transition time (‘multi-step time’) between states $k$ and $\ell$ (with $\ell > k$) in terms of the kinetic potential $F$. The multi-step time $T_{k-\ell}$ for a transition with approximately constant $k$ and $\ell$, can be found by summing $O(N)$ single-step times. If all those single-step times are of the non-exponential type, we find $T_{k-\ell} = O(1)$. However, if one or more of those single-step times are of the exponential type, we find $T_{k-\ell} > Ce^{\gamma}$ (with positive constants $C$ and $\gamma$). Hence, the way multi-step times depend on $N$ is determined by the shape of $F$ on the interval $[0, \hat{k}]$.

4.2. Transition times in the autophosphorylation cycle

As discussed above, the shape of the kinetic potential $F$ determines how single-step and multi-step times depend on $N$. We now apply this knowledge to the autophosphorylation cycle. Fig. 5 shows $F$ for the autophosphorylation cycle with $s = 0.55$. In this figure, five regions (I–V) can be defined, which (as discussed below) determine the behavior of both the stochastic and deterministic model. Since $F$ has a similar shape for all $s$ in $(\sigma_0, \sigma_2)$, the regions I–V can be defined for each $s$ in this interval. This is shown in Fig. 6. For all $s$ in $(\sigma_0, \sigma_2)$, both models show the same qualitative behavior in the same regions. Similar, but mirrored, results can be obtained for all $s$ in $(\sigma_1, \sigma_0)$.

The region to which $\hat{k}$ belongs determines the asymptotic behavior of $T_{k-k+1}$. For $k$ in regions I and IV, there are no $x < \hat{k}$, for which $F(x) < F(\hat{k})$. Therefore, all $T_{k-k+1}$ are of the non-exponential type. However, for each $k$ in regions II, III and V, there

![Fig. 4](image-url) The scaled stationary distributions $N_{a,k}$ for $s = 0.52$ and 0.55, and various $N$ as a function of the proportion phosphorylated proteins $k/N$.

![Fig. 5](image-url) Function $F$ for $s = 0.55$ with regions I–V. The symbol $^{-\uparrow}$ denotes the value of $x$ (with $\gamma_2 < x < \gamma_1$) for which $F(x) = F(\gamma_1)$.
always is an \( x < k \), for which \( F(x) < F(k) \). Hence, in those regions, all single-step times \( T_{k \to k+1} \) are of the exponential type. Remarkably, region III exemplifies that even single-step times in the direction of increasing stationary probability can grow exponentially with \( N \).

Obviously, the shape of \( F \) also determines how single-step times in the opposite direction depend on \( N \). For \( k \) in region \( V \), there is no \( x > k \), for which \( F(x) < F(k) \) and \( T_{k \to k-1} \) is of the non-exponential type. For \( k \) in any of the other regions, there is always a \( x > k \) with \( F(x) < F(k) \). Consequently, all corresponding single-step times \( T_{k \to k-1} \) are of the exponential type.

We now compare the single-step times of the stochastic model with the behavior of the deterministic model. Table 1 shows for each region the types of single-step times in both directions and the sign of the derivative of the deterministic model. Recall that the deterministic model moves from any \( y = k \) with \( a(y) > b(y) \) to \( y + \Delta y = k + \frac{1}{c} \) in a time \( O(N^{-1}) \); analogously, it moves from any \( y = k \) with \( b(y) > a(y) \) to \( y - \Delta y = k - \frac{1}{c} \) in a time \( O(N^{-1}) \).

In regions I, IV and \( V \) the deterministic model moves in the direction for which the single-step time in the stochastic model is non-exponential. On average, the stochastic model makes a step of size \( 1/N \) in this direction in a time \( O(N^{-1}) \). This is in good agreement with the deterministic model, which moves over the same distance in a time \( O(N^{-1}) \). Contrastingly, for regions II and III, single-step times in both directions are of the exponential type. For both of these regions the deterministic model moves in the direction of the local minimum of \( F \); again, the deterministic model moves over an interval of size \( 1/N \) in that direction in a time \( O(N^{-1}) \). This marks an inconsistency between the stochastic and the deterministic model.

For both regions II and III it can be shown that the occurrence of exponential single-step times is due to rare detours that pass the local maximum in \( F \). It is possible to rewrite (17) in such a way that single-step times can be expressed in terms of distances of possible paths from \( k \) to \( k + 1 \). In particular, we are interested in paths that visit some state \( \ell \) (with \( \ell < k \)), but not states \( \ell - 1 \). We refer to the probability for such a path as \( P_{\ell \to k+1}^{(\ell)} \). The expected time to go from state \( k \) to state \( \ell \) (without visiting state \( k + 1 \)) and then to state \( k + 1 \) (without visiting state \( \ell - 1 \)) is further referred to as \( T_{k \to k+1}^{(\ell)} \). The single-step time from \( k \) to \( k + 1 \) can now be expressed as follows:

\[
T_{k \to k+1} = \sum_{\ell=0}^{\infty} P_{k \to k+1}^{(\ell)} T_{k \to k+1}^{(\ell)}.
\]

As shown in Fig. 7 for the case \( k = N/2 \), this separation according to ‘path diameter’ clearly uncovers the occurrence of the aforementioned rare paths. With increasing \( N \), the probability of visiting states with \( \ell < \gamma_2 \) decreases (see Fig. 7(a)), while the corresponding expected time increases (see Fig. 7(b)). The net effect is an (approximately exponential) increase in the product of the probability and expected times for some \( \ell < \gamma_2 \) (see Fig. 7(c) and the detailed view in Fig. 7(d)). By omitting all paths with \( \ell < \gamma_2 \) from the calculation, a non-exponential single-step time is obtained which approximates the time of the deterministic model.

By applying the expression (20) for multi-step times, we find that \( T_{k \to k} = O(1) \), if both \( \ell \) and \( \ell \) are in the same region with non-exponential single-step times. All other multi-step times contain exponential single-step times and therefore grow exponentially fast with \( N \). Fig. 8(a) shows the expected times for transitions between the local maxima \( G_1 \) and \( G_2 \) of the stationary distribution for \( s = 0.55 \). A path from one local maximum to the other includes a large number of exponential single-step times in the direction of decreasing stationary probability. Logically, the expected time for such a path grows exponentially in \( N \). However, as discussed before, even expected times of single steps in the direction of increasing stationary probability may grow exponentially in \( N \). As a result, also multi-step times in such a direction may be grow exponentially with \( N \). In Fig. 8(b) such an exponential dependence is shown for transitions from \( (G_1 + G_2)/2 \) to \( G_1 \) and from \( (G_2 + G_3)/2 \) to state \( G_3 \).

5. Eigenvalue analysis

5.1. Eigenvalues of \( M \)

In the previous section, we have seen that, although every state can be reached from every other state, the expected time for such a transition can be extremely long and may even grow exponentially with \( N \). This means that the probability vector \( \mathbf{z}(t) \) may converge very slowly to the stationary distribution. In this section, we show how the eigenvalues of \( M \) provide more insight into this behavior and its dependence on \( N \).

The Perron–Frobenius theorem [12] (applied to the matrix \( I + \mathbf{cM} \) for a suitable \( \mathbf{c} \)) yields that there is one simple eigenvalue \( \chi_0 = 0 \) with a stochastic eigenvector \( \mathbf{v}_0 \). Furthermore, Gershgorin’s circle theorem [13,14] states that each of the eigenvalues of \( M \) is in at least one of the discs with center \( -x_i - \beta_i \) and radius \( x_i + \beta_i \) (with \( i = 0, 1, \ldots, N \)). Consequently, all eigenvalues have a non-positive real part.

It is easily verified that if \( \mathbf{M} \) is similar to the symmetric tridiagonal matrix \( \mathbf{V} = (v_{ij}) \) (with \( i,j = 0, 1, \ldots, N \)), with elements

\[
v_{ij} = \begin{cases} \sqrt{x_i + \beta_i}, & \text{if } j = i + 1, \\ \sqrt{x_i - \beta_i}, & \text{if } j = i - 1, \\ -(x_i + \beta_i), & \text{if } j = i, \\ 0, & \text{otherwise.} \end{cases}
\]

Table 1

<table>
<thead>
<tr>
<th>Region</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_{k \to k+1} )</td>
<td>non-exp</td>
<td>exp</td>
<td>exp</td>
<td>non-exp</td>
<td>exp</td>
</tr>
<tr>
<td>( T_{k \to k-1} )</td>
<td>exp</td>
<td>exp</td>
<td>exp</td>
<td>exp</td>
<td>non-exp</td>
</tr>
<tr>
<td>Deterministic model</td>
<td>( \frac{1}{c} \times 0 )</td>
<td>( \frac{1}{c} \times &lt; 0 )</td>
<td>( \frac{1}{c} \times &gt; 0 )</td>
<td>( \frac{1}{c} \times &gt; 0 )</td>
<td>( \frac{1}{c} \times &lt; 0 )</td>
</tr>
</tbody>
</table>

2 When the determinant of a tridiagonal matrix \( \mathbf{A} \) is expressed in terms of its elements \( A_{ij} \), the off-diagonal elements only appear in products of the form \( A_{ij}, A_{ji}, A_{ki}, A_{ik} \). For the tridiagonal matrices \( \mathbf{M} \) and \( \mathbf{V} \) these products are identical. \( \mathbf{M} \) and \( \mathbf{V} \) also have the same elements on their diagonals. As a result, the determinants det(\( \mathbf{M} - \rho \mathbf{I} \)) and det(\( \mathbf{V} - \rho \mathbf{I} \)) for any \( \rho \) are identical. This means that \( \mathbf{M} \) and \( \mathbf{V} \) have the same eigenvalues.

Fig. 6. Steady states of the deterministic model (see also Fig. 2) plotted together with the curves that separate the ‘exponential’ and ‘non-exponential’ regions in the stochastic model (dotted lines).
Hence the eigenvalues of $V$ and $M$ are real. In summary this means that there is one eigenvalue $\lambda_0 = 0$; all other eigenvalues $\lambda_1, \ldots, \lambda_N$ are negative and real. We will order the eigenvalues such that $0 < |\lambda_1| < |\lambda_2| < \cdots < |\lambda_N|$. Moreover, $V$ has an orthonormal set of $N + 1$ eigenvectors. The corresponding eigenvectors of $M$ will be written as $v_0, \ldots, v_N$. The (stochastic) eigenvector with eigenvalue $\lambda_0 = 0$ is the stationary distribution: $v_0 = Z$.

We now study the behavior of the solution of the Master equation (4), with a stochastic vector $z(0)$ as initial value at $t = 0$. The initial vector can be written as

$$z(0) = \sum_{i=0}^{N} c_i v_i.$$ 

Since $z(0)$ and $v_0 = Z$ are stochastic vectors and all other eigenvectors of $M$ have a zero element sum, the coefficient $c_0 = 1$. This leads to the solution

$$z(t) = Z + \sum_{i=1}^{N} c_i v_i \exp(\lambda_i t). \quad (22)$$

Clearly, the eigenvalues $\lambda_1, \ldots, \lambda_N$ determine how fast $z(t)$ converges to $Z$. In particular $\lambda_1$ is important for the time required for this convergence and how this time depends on $N$. Since $\lambda_1(0)$ is the second largest eigenvalue of the symmetric matrix $V$, its value can be found from the Rayleigh quotient

$$\lambda_1 = \max_{y \cdot w} \frac{y^T \cdot V \cdot y}{y^T \cdot y}. \quad (23)$$

where $w$ is the eigenvector of $V$ with eigenvalue $\lambda_0 = 0$. It is easily verified that $w$ is related to the corresponding eigenvector $Z$ of $M$ by $w_i = \sqrt{Z_i}$. By using an appropriate vector $y$ perpendicular to $w$ (23) gives a lower bound for $\lambda_1$:

$$\frac{y^T \cdot V \cdot y}{y^T \cdot y} \leq \lambda_1 < 0. \quad (24)$$

5.2. Eigenvalue analysis of a bistable system

The construction of an appropriate vector $y$ to use in (24) depends on the specific system that is considered. We will consider the case of a bistable system, i.e., a system with a kinetic potential $F$ with two local minima. Hence for sufficiently large $N$, the stationary distribution $Z$ has two local maxima. This holds for the considered autophosphorylation cycle if the parameter $s$ satisfies $\alpha_1 < s < \alpha_2$. For this network the first two eigenvectors of $M$ are shown in Fig. 9. From that figure we conclude that the corresponding eigenvector $v_1$ for $M$ looks like an adapted form of $v_0 = Z$ in which the first half is multiplied by a positive factor and the second half is multiplied by a negative factor. In this way, an approximation for $v_1$ can be constructed. However, in view of the Rayleigh quotient (23) we are more interested in an approximation $y$ of the eigenvector of $V$ with eigenvalue $\lambda_1$. Therefore, we use a similar approach to construct an approximate eigenvector $y$. In other words, we use two parts of $w$ (the eigenvector of $V$ with eigenvalue $\lambda_0 = 0$), multiplied by suitable coefficients, to construct the approximate eigenvector $y$ (with eigenvalue $\lambda_1$).
Recall that V has an eigenvector w corresponding to eigenvalue \( \lambda_0 = 0 \). This vector w has elements \( w_i = \sqrt{Z_i} \). We will use this vector to construct parts of the vector y, which is used as an approximation for the eigenvector corresponding to \( \lambda_1 \).

Let \( r \) be an index with \( 0 \leq r < N \). Define

\[
y_i = (w_0, \ldots, w_r, 0, \ldots, 0)^T = \left( \sqrt{Z_0}, \ldots, \sqrt{Z_r}, 0, \ldots, 0 \right)^T.
\]

\[
y_r = (0, 0, \ldots, w_{r+1}, \ldots, w_N)^T = \left( 0, 0, \ldots, \sqrt{Z_{r+1}}, \ldots, \sqrt{Z_N} \right)^T.
\]

The numerator of the Rayleigh quotient in (24) is given by

\[
y^\top \cdot V \cdot y = \sum_{i=0}^{N} y_i^2 = (1 - \omega)^2 \sum_{i=0}^{r} Z_i + \omega^2 \sum_{i=r+1}^{N} Z_i.
\]

Using (26), this can be written as

\[
y^\top \cdot y = (1 - \omega)^2 \omega + \omega^2 (1 - \omega) = \omega (1 - \omega).
\]

Combining the numerator and denominator we obtain the following lower bound for \( \lambda_1 \):

\[
-\frac{\alpha Z_r}{\omega (1 - \omega)} \leq \lambda_1 < 0.
\]

Note that this bound holds for any value of \( r \). Moreover, besides \( Z_r \), also \( \omega \) depends on \( r \), see (26).

We now use the fact that we consider a system with a bimodal stationary distribution \( Z \), with local maxima in \( G_1 \) and \( G_2 \) and an intermediate local minimum in \( G_2 \). We select \( r = G_2 \) such that \( k \) is (approximately) the position of the intermediate maximum \( \gamma_2 \) of the function \( F \). For this choice we compute \( Z_r \) and \( \omega \). Now \( \omega \) is the total probability in the left part of the distribution, i.e., up to \( r \). That part is centered around \( G_1 \) and approximately Gaussian. In Appendix A is shown that the sum over all probabilities \( Z_i \) with \( i \leq G_2 \) can be approximated by:

\[
\omega = \sum_{i=0}^{G_2} Z_i = \sqrt{\frac{2\pi N}{F' (\gamma_2)}}.
\]

The right part of the distribution is centered around the point \( G_3 \). Using the same method as the one leading to (28), we obtain

\[
1 - \omega \approx G_3 = \sqrt{\frac{2\pi N}{F' (\gamma_3)}}.
\]

Substitution of (28) and (29) in (27) gives

\[
-\frac{\alpha G_2 G_3 \sqrt{F' (\gamma_1) F' (\gamma_3)}}{2 \pi N G_3 G_2} \leq \lambda_1 < 0.
\]

This bound for \( \lambda_1 \) still contains the stationary probabilities in the points \( G_2 \) and \( G_3 \). We assume for the moment that the stationary probabilities \( Z_i \) have their global maximum in \( k = G_2 \) (as it is the case for the autophosphorylation cycle with \( s \) in \( (c_0, \sigma_2) \)). We now rewrite the bound to find its asymptotic dependence on \( N \). Firstly, we know \( \alpha G_2 / N \approx \rho (\gamma_2) \). Secondly, since \( G_2 \) has its global maximum at \( G_3 \), we know from Appendix A that

\[
Z_{G_2} \approx \sqrt{\frac{F' (\gamma_2)}{2 \pi N}}.
\]

Thirdly, using (16) we find that the quotient of \( Z_{G_2} \) and \( Z_{G_3} \) satisfies:

\[
\frac{Z_{G_2}}{Z_{G_3}} = O \left( e^{\rho (\gamma_1) - F (\gamma_3)} \right).
\]

Combining the above, we rewrite the lower bound in (30) as:

\[
-\frac{\alpha G_2 Z_{G_2} \sqrt{F' (\gamma_1) F' (\gamma_3)}}{2 \pi N Z_{G_3} G_2} = -O \left( a_n \sqrt{\frac{Z_{G_2}}{Z_{G_3}}} \right) = -O \left( a_n \sqrt{\frac{1}{2 \pi}} \right).
\]

Hence, we obtain a lower bound for \( \lambda_1 \) with asymptotic behavior for \( N \rightarrow \infty \) given by

\[
-\left( \frac{N^{1/2} N^{1/4} (F (\gamma_2))}{} \right) \leq \lambda_1 < 0.
\]

Of course a similar bound can be derived if not \( G_3 \) but \( G_1 \) is the global maximum of the stationary distribution. This results in (31) with \( \gamma_1 \) replaced by \( \gamma_2 \). In both cases, \( \lambda_1 \) tends to zero exponentially fast if \( N \rightarrow \infty \). As a result, the convergence of the solution of (22) to
the stable stationary \( Z \) (which depends on a time constant \( 1/\lambda_1 \)) can take extremely long for a large number of molecules \( N \).

Note that, up to \((27)\), the argument above holds for any function \( F \) and any point \( r \). However, the selection of \( r \) to correspond with the minimum \( \gamma_2 \) of \( F \), that lies between two maxima in \( \gamma_1 \) and \( \gamma_3 \), is of course only possible if \( F \) has this behavior.

### 5.3. Eigenvalue analysis for the autophosphorylation cycle

For the autophosphorylation cycle, the dependence of eigenvalues \( \lambda_1 \) and \( \lambda_2 \) on \( N \) is shown in Fig. 10. This figure shows that \( \lambda_1 \) indeed tends exponentially to zero if \( N \to \infty \), while \( \lambda_2 \) converges to a negative constant value. This has a number of implications for the dynamics of the system. As \( \lambda_2, \ldots, \lambda_N \) are negative and do not vanish with increasing \( N \), the corresponding terms \( c_2v_2\exp(\lambda_2t), \ldots, c_Nv_N\exp(\lambda_Nt) \) in \((22)\) decrease exponentially in \( t \), for each \( N \). However, \( \lambda_1 \) tends to zero exponentially fast if \( N \to \infty \). As a result, the corresponding term in \((22)\) decreases with a time constant \( 1/\lambda_1 \) that grows exponentially in \( N \).

Roughly, the dynamics for a constant input \( s \) can be regarded as three separate phases: (a) the initial phase, in which the terms corresponding to \( \lambda_2, \ldots, \lambda_N \) decrease exponentially over time; (b) the ‘quasi-stationary’ phase, in which the system can be approximated by the terms corresponding to \( \lambda_0 \) and \( \lambda_1 \); and (c) the stationary phase, which occurs if the term \( c_1v_1\exp(\lambda_1t) \) has also vanished. In contrast with phase (c), both phases (a) and (b) depend on the initial conditions. The length of phase (b) depends on both the initial distribution \( x(0) \) and eigenvalue \( \lambda_1 \) and grows exponentially with \( N \).

Fig. 9 shows the elements of eigenvectors \( v_0 \) and \( v_1 \) for the first two eigenvalues of \( M \). Recall that \( v_0 \) corresponds with the stationary distribution \( Z \). All quasi-stationary distributions depend on both \( v_0 \) and \( v_1 \) and are of the form \( v_0 + \varphi v_1 \). For a given initial condition \( x(0) \), the system will converge fast to \( v_0 + \varphi v_1 \), with \( \varphi \) depending on \( x(0) \). On a much longer timescale, \( \varphi \) will tend to zero, eventually resulting in the stationary solution.

### 6. Conclusions and discussion

This paper shows that a potential function can be a useful tool for the comparison of stochastic and deterministic models of chemical reaction systems. The kinetic potential used in this paper provides both intuitive and precise information about the way the behavior of the stochastic model depends on the number of molecules \( N \). This can be used to study the convergence (and in some respects the lack of convergence) of the behavior of this model to that of the corresponding deterministic model with increasing \( N \). As discussed in Section 6.3, potential functions have been used before to analyze the stationary distribution and expected times of stochastic models. Nevertheless, our work provides several important novelties.

![Fig. 10. Eigenvalues \( \lambda_1 \) (solid line) and \( \lambda_2 \) (dashed line) of \( M \) as a function of \( N \) for \( s = 0.55 \).](image-url)

Previously, potential functions have been used because they provide information about the stationary distribution and expected times. We show that our kinetic potential provides additional information in ways we have not encountered in the literature before. In bistable systems, the kinetic potential can be used to give an explicit proof that the second eigenvalue of the transition matrix \( M \) depends on \( N \) in an exponential fashion. The kinetic potential also allows to distinguish between regions with characteristic behavior. For the autophosphorylation cycle, five regions can be identified with characteristic behavior of both the stochastic and deterministic model (as summarized in Fig. 5 and Table 1).

Our approach provides a generic and self-contained framework for a biologically interesting class of systems. In this approach, we particularly focus on the dependence of the stochastic model on \( N \) and its relation with the deterministic model. The analysis is purely based on the deterministic rate equations, without further reference to thermodynamic notions. The approach discussed in this paper is only suitable for single-variable systems, but can be generalized to a larger class of systems.

### 6.1. Autophosphorylation cycle

Throughout this paper, we have exemplified our approach using an idealized autophosphorylation cycle. The relative simplicity of this autophosphorylation cycle and the chosen Michaelis–Menten kinetics allows us to analytically study both the deterministic and the stochastic dynamics. Analysis of this relatively simple system provides insight into its behavior to an extent that is generally impossible for more complex systems. We do not claim that Michaelis–Menten kinetics provide a realistic description for a real-life network with this topology. In fact, although a phosphorylation cycle without autophosphorylation can realistically be modeled with Michaelis–Menten kinetics [15], it is known that this generally is not the case for more complex networks derived from such a cycle [16]. Nevertheless, a detailed analysis of the validity of Michaelis–Menten kinetics in this particular case is outside the scope of this paper.

In some respects, the behavior of the deterministic model reflects that of the stochastic model for very large \( N \). For instance, the local maxima of the stationary distribution tend to the stable steady states of the deterministic model for \( N \to \infty \). On the other hand, the deterministic model only represents the expected behavior of the stochastic model during the initial and quasi-stationary phase. Indeed, the length of the quasi-stationary phase grows exponentially in \( N \). Hence, in the limit \( N \to \infty \), the subsequent stationary phase is never reached.

For transition times between different proportions of phosphorylated proteins the comparison is more complicated. If a transition is impossible in the deterministic model, this is reflected in the stochastic model by an expected time that grows exponentially in \( N \) for \( N \to \infty \). If a transition is possible in the deterministic model, the corresponding expected time in the stochastic model can either tend to the transition time in the deterministic model or—Remarkably—grow exponentially for \( N \to \infty \), depending on the shape of \( F \).

The exponential dependency of expected single-step times on \( N \) is explained by the separation of those times into probabilities and times by path diameter (see Fig. 7). In the deterministic model, the influence of extremely long paths is ignored, since these paths occur rarely for large \( N \). However, in this way the deterministic model also ignores the enormous influence of such paths in cases in which they do occur. In real-life biochemical reaction networks, \( N \) is often relatively small and these paths may still have a significant role. When modeling such systems with a deterministic model, one should be aware of the existence of those paths and the fact that they are neglected in the model.
6.2. Bistability in real-life networks

Bistability is a common feature of biochemical reaction networks. It can be found in the various types of networks that are present in a cell. For instance, autoregulatory protein–DNA feedback loops in gene regulatory networks cause bistability [17]. Also in signaling networks [18] and in networks that regulate cellular processes, bistable subnetworks (or ‘motifs’) can be identified. The autophosphorylation cycle that is used as a running example in this paper, is an idealized representation of one of the smallest phosphorylation modules that can theoretically yield bistability [19]. We observed that 143 out of 292 kinases in the Human Protein Reference Database [20] show autophosphorylation. 3

A well-know example of bistability in intracellular signaling and regulation is the protein Ca2+/calmodulin-dependent protein kinase II (CaMKII) [21,22], which is involved in synaptic plasticity (i.e., the cellular process that is required for the storage of memory in the brain). Although the exact mechanism is much more complicated, the CaMKII network shares some characteristics with our autophosphorylation cycle. CaMKII proteins form rings on which each protein can be phosphorylated and dephosphorylated. In addition, phosphorylated CaMKII proteins can autophosphorylate other CaMKII proteins. This results in two switch states: one with a large and one with a small proportion of phosphorylated proteins. The highly phosphorylated switch state can be stable for a long time and induce other processes which potentiate the synaptic connection between two neurons.

Another example is the eukaryotic cell cycle [23]. This cycle is divided into distinct phases that each contain processes involved in the duplication of cell contents and division of the cell. Transitions between those phases require robust and irreversible switching events, since unwanted phase transitions can be harmful. Also these switching events are due to bistable motifs in the reaction network (see for instance [24]).

The CaMKII system and the regulation of the cell cycle regulation have clearly different functions but similar requirements. The bistable modules in those networks need to switch fast between two states at the correct moment, while they should avoid switching at other moments. As those modules are relatively complex, they cannot straightforwardly be analyzed with our single-variable approach. We can, however, gain some qualitative understanding of such modules from our autophosphorylation cycle. Recall that the shape of the kinetic potential $F$ provides information about both the deterministic and the stochastic description of this network. Hence, manipulation of the shape of $F$ affects the properties of the system for both large and small numbers of molecules.

In the case of the autophosphorylation cycle, the shape of $F$ can be changed by the input kinase concentration $s$. For instance, a temporary increase in $s$ can change the kinetic potential from a bimodal to a unimodal function and by that increase the probabilities for states with large proportions of phosphorylated proteins. A more complex reaction system may involve more sophisticated mechanisms to reshape the kinetic potential and by that manipulate the properties of the switch. For instance, deeper minima in the kinetic potential may lead to a more stable switch for smaller numbers of molecules.

6.3. Related work

The kinetic potential $F$ has been used before to describe extinction times for birth-and-death processes [25–27]. However, it is different from the potential functions generally used to describe stochastic reaction systems [10,28]. The difference between the two types of potential functions is that the first type is defined on a macroscopic scale and the latter on a mesoscopic scale. The advantage of the first type is that it does not require reformulation when the size of the system (i.e., the number of molecules $N$) is changed. This makes it more suitable for the applications we are interested in.

For classic mass-action systems with known microscopic parameters the kinetic potential $F$ is linearly related to the Gibbs free energy. However, this is generally not the case if the mesoscopic model is derived from a macroscopic model. The advantage of using $F$ instead of the Gibbs free energy is that the former does not require information about thermodynamic properties such as temperature. In many deterministic models of biochemical reaction networks such properties are not explicitly taken into account. Moreover, $F$ can also straightforwardly be applied to non-elementary reaction kinetics.

The quasi-stationary behavior of the autophosphorylation cycle is reminiscent of Keizer’s paradox described by [29]. In that system, there is one reaction species of interest, which is involved in two reactions: an autocatalysis reaction and a first-order degradation reaction. The deterministic model has two stable steady states: total extinction (i.e., zero molecules) and a non-zero value for which degradation and autocatalysis are balanced. The stochastic model has one true steady state (total extinction) and a quasi-stationary distribution. In analogy with our autophosphophorylation cycle, the quasi-stationary distribution in the system of [29] vanishes after exponentially long times. Numerical results discussed in [29] indicate that the second eigenvalue of the transition matrix $M$ tends to zero in an exponential fashion. An important novelty in this paper is that we provide an actual proof that this behavior is indeed exponential.

In the field of statistical physics, other bistable reaction systems have been studied with interesting results that partially overlap with ours, e.g. [30,31]. There are, however, some clear differences. In statistical physics the dependence of the stationary distribution on the number of molecules $N$ is usually studied by describing the dynamics of the stochastic system with a Fokker–Planck equation [32] and then studying the stationary solutions of this equation. Note, however, that the Fokker–Planck equation is only an approximation that is obtained by omitting higher order terms. In our approach we start with the exact stationary distribution (5) and manipulate that to obtain the exact expression (9) for the quotient of two probabilities. This leads to the asymptotic behavior as given in (16).

A well-known theoretical reaction system that can show bistability is the so-called Schögl model, introduced by [30]. In a recent paper [33], the relations between the stochastic and deterministic model of this reaction system are discussed. This paper also describes the differences between mathematical and chemical detailed balance. Many results in this paper can be related to our results. The authors discuss these subjects from a thermodynamic point of view, while our focus is purely kinetic. Therefore, both papers provide complementary insights. Note that this paper also observes the exponential growth of the second eigenvalue of the transition matrix, which is confirmed by our explicit proof.

Another recent paper, [27], shows that also for the Schögl model, there are exponentially long transition times, which can easily be found from the shape of the potential function. This is in good agreement with our findings. In the discussion section of [27], the authors mention some open questions regarding the most probable and most dominant paths between two states. The separation of paths according to path diameter (discussed in Section 4.2) provides the information required to answer those questions.
Moreover, Fig. 7 provides a clear example for the fact that the most probable path is not always the most dominant path. For many chemical reaction systems, it is practically impossible to study the stochastic dynamics analytically. In most of those cases, the simulation of individual trajectories with the exact stochastic simulation algorithm introduced by Gillespie [34] provides a useful tool to obtain insight into the behavior of the system. However, one should keep in mind that a solid stochastic analysis in some cases may require an extremely large number of simulations. For example, in the autophosphorylation cycle the expected time for a transition from one given state to another may depend heavily on rare paths. A reasonable estimation of this expected time requires enough simulations to sufficiently sample all those rare paths. For increasingly large values of N, the probability of those paths (but not necessarily their influence) decreases exponentially, leading to an exponential increase in the number of required simulations.

Acknowledgments

This work is supported by the European Commission through the Evolving Cell Signaling Networks in Silico (ESIGNET) project of the Sixth Framework Programme.

Appendix A. Stationary distributions near local maxima

The relation (16) allows an easy computation of the form of the stationary distribution near a local maximum. Let k be a local maximum of the distribution Z and let 0 be a nearby state for which we want to express the probability. If k and N are large, the maximum of the distribution in k corresponds with a minimum of the function F in k = k = 1. For points  not too far away from k, we can approximate the difference F( k) − F( 0) by a second order Taylor series:

\[ F( k) − F( 0) ∼ \frac{F'( k)( k − 0)^2}{2} \frac{F''( k)( k − 0)^2}{2N^2}. \]

The first-order term in this Taylor series vanishes since F( k) = 0. Substitution in (16) gives

\[ Z_i ≈ Z_k \exp \left( \frac{−F'( k)( k − 0)^2}{2N} \right) \frac{a(k) b(k)}{a(0) b(0)}. \]

with  = 0/N. The exponential term gives rise to a normal distribution around k, with a variance \( \sigma^2 = N/F'( k) \). The square root term equals 1 for = k. Moreover, the influence of this term diminishes with increasing N, hence we can omit this term for large values of N. This leads to

\[ Z_i ≈ Z_k \exp \left( \frac{−F'( k)( k − 0)^2}{2N} \right). \] (32)

Now we assume the maximum at k is the only maximum in the stationary distribution. Hence, all probabilities Z_i will be very small, except the Z_k with  near k. In that situation, the sum of all probabilities Z_i can be computed by using the Gaussian approximation (32). Instead of summing over all i, it is sufficient to sum over all i with \([k − d, k + d]\), where d is a few times the standard deviation \( \sigma \). This yields

\[ \sum_{i=k-d}^{k+d} Z_i ≈ Z_k \int_{k-d}^{k+d} \exp \left( −\frac{y^2}{2N} F'( k) \right) dy ≈ Z_k \sqrt{\frac{2\pi N}{F'( k)}}. \]

Since the sum of all probabilities in the distribution must be one, we conclude that in this situation the maximum of the probability distribution is approximately given by

\[ Z_k ≈ \frac{F'( k)}{2\pi N}. \] (33)

If F has two or more local minima, then the probability distribution Z will also have two or more local maxima for large enough N. It is easily verified that, in that situation, (33) is still an upper bound for the local maximum of the probability distribution in k.

References


