
Waiting time distributions for clusters of complex molecules

R. THUL¹ and M. FALCKE²

¹ *School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK*

² *Hahn-Meitner-Institute, Dept. Theory, Glienicker Str. 100, D-14109 Berlin, Germany*

PACS 87.16.Ac – Theory and modelling
PACS 02.50.Ga – Markov processes
PACS 87.16.-b – Subcellular structure and processes

Abstract. - Waiting time distributions are in the core of theories for a large variety of subjects ranging from the analysis of patch clamp records to stochastic excitable systems. Here, we present a novel exact method for the calculation of waiting time distributions for state transitions of complex molecules with independent subunit dynamics. The absorbing state is a specific set of subunit states, i.e. is defined on the molecule level. Consequently, we formulate the problem as a random walk in the molecule state space. The subunits can possess an arbitrary number of states and any topology of transitions between them. The method circumvents problems arising from combinatorial explosion due to subunit coupling and requires solutions of the subunit master equation only.

The defining property of self amplifying reactions and reaction cascades is that a single initial event causes many subsequent reactions and product molecules. When the number of molecules causing the first event is small, initiation occurs stochastically. This is the case in the tiny volume of living cells, where such randomness may even be observable on the cell level. [1]. Binding of a ligand to a single G-protein coupled receptor leading to thousands of product molecules downstream in the signalling pathway is such an event [2]. Similarly, the crossing of the activation threshold of an excitable system causes strong amplification of the signal. That crossing may be preceded by a complex process consisting of many random elemental steps. Well known examples are puffs and sparks in intracellular Ca^{2+} release and randomly spiking neurons [1,3,4]. These random processes can be characterised by waiting time distributions, which are inter-puff and inter-spike interval distributions in biological terms.

The above examples have in common that only a few specific events out of the complex dynamics on one structural level drive the dynamics on the next higher level. For instance, among the many state transitions an ion channel may have, only the opening and closing events are relevant for membrane voltage dynamics. This suggests reducing the description of the micro- or mesoscopic level to the states relevant for the dynamics on the macroscopic level and the corresponding waiting time distributions. The complex processes on the microscopic level are subsumed

into the distributions. An excitable system defined that way is presented in [5,6]. Thus, waiting time distributions can serve as constituting functions of a theory of hierarchical stochastic systems.

Waiting time distributions have also been applied to model the behaviour of linear molecular motors such as myosin V and kinesin, or of rotary motors like the F1-ATPase [7–10]. A single motor step is perceived as a sequence of state transitions of the motor molecule. Since these motors always return to the same initial configuration after a stroke, the corresponding waiting time distribution defines a renewal process.

A revived interest in the description of oscillatory systems as a set of discrete states with cyclic, unidirectional state transitions arose recently spurred by the search for universality in the non-equilibrium phase transitions to synchronous oscillations [11] and the quest for the origin of oscillatory dynamics from random elements [5,11]. Waiting time distributions are one of the defining elements of these systems.

In summary, there is a growing interest for characterising systems that arise from direct applications and fundamental questions of physics, in terms of discrete states and waiting time distributions. In this letter, we will calculate waiting time distributions for state transitions of complex molecules which consist of several identical and independent subunits. The activated molecule state is a specific configuration of subunit states. This molecular state is the

absorbing state of our waiting time calculations. Let us assume for the moment that the activated molecule state is the one where all subunits are in the same state a . We consider the event when for the first time all but one of the subunits are already in the state a and the remaining subunit makes a transition into the state a . All individual subunits may have visited the state a many times before *all of them meet* there. Consequently and although all subunits are independent, the waiting time distribution we would like to compute is not a power of the single subunit waiting time distribution for a transition to a . The absorbing molecule state couples the trajectories of the individual subunits.

Methods based on Fokker-Planck-equations cannot be applied to these problems for two reasons. Firstly, there is no large parameter as e.g. the volume that allows for a proper derivation from microscopic dynamics, and secondly, the large number of dimensions of the phase space would render a Fokker-Planck equation intractable. We therefore need to solve the master equation for the complete molecule. However, this is often difficult to even write down due to combinatorial explosion of the number of states. Our approach provides a way out of that dilemma since it necessitates to solve the subunit master equations only. We will illustrate our ideas for clusters of molecules, but we start with a single molecule.

This work was inspired by problems from intracellular Ca^{2+} dynamics. Calcium release from storage compartments is controlled by ion channels, which we will use to exemplify our concept. The open probability of the channel depends on the calcium concentration on the outside of the storage compartment (see [1] for a review). It increases with increasing Ca^{2+} concentration, i.e. Ca^{2+} release is self-amplifying. Very high Ca^{2+} concentrations inhibit the channel and terminate release. Activation and inhibition together create a bell-shaped dependence of the open probability on $[\text{Ca}^{2+}]$.

Release channels are grouped into clusters, each containing 1–40 channels [1]. The spontaneous Ca^{2+} release through channels of a single cluster, called a puff, is the elemental random event of intracellular Ca^{2+} dynamics. The dependence of the open probability on the Ca^{2+} concentration entails that channels in a cluster communicate via the local Ca^{2+} concentration. However, this feedback is eliminated when the Ca^{2+} concentration is constant throughout a cluster so that channels behave independently. As a good approximation, this is true when all channels are closed or at most one channel is open. The absorbing state can be defined in a way that the waiting time problem describes one of the transitions between these two situations [13].

We consider a channel molecule consisting of h identical subunits. They possess an arbitrary number of states and any topology of transitions between them. The states of the channel molecule can be described as points in a multidimensional state space. Let $q(X, t|Y, \tau)$ denote the probability of the channel to be in state X at time t when

it was in state Y at time τ , and $F(X, t|Y, 0)dt$ the probability of arriving in state X for the first time in the interval $[t, t+dt]$ when the initial state was Y . These two quantities satisfy the relation [12]

$$q(X, t|Y, 0) = \delta_{X,Y}\delta(t) + \int_0^t d\tau q(X, t|X, \tau)F(X, \tau|Y, 0). \quad (1)$$

The probability of being in X at time t equals the probability of arriving there for the first time in some interval $[\tau, \tau + d\tau]$, $\tau \leq t$, times the probability of being in X at time t given the initial condition (X, τ) . The integral includes all τ between 0 and t , and the delta functions take care of initial values. In the case of a time-homogeneous Markov process, $q(X, t|Y, \tau) = q(X, t - \tau|Y, 0)$, and the Laplace transform of eq. (1) for $X \neq Y$ reads as

$$\hat{q}(X, s|Y, 0) = \hat{F}(X, s|Y, 0)\hat{q}(X, s|X, 0). \quad (2)$$

Since $F(X, t|X, 0) = \delta(t)$, we obtain finally

$$\hat{F}(X, s|Y, 0) = \frac{\hat{q}(X, s|Y, 0)}{\hat{q}(X, s|X, 0)}(1 - \delta_{X,Y}) + \delta_{X,Y}. \quad (3)$$

The state of a subunit is defined by the binding of ligands to binding sites on the subunit. We model the stochastic association and dissociation of ligands by a master equation. Let n denote the number of states of a single subunit, and $p(i, t|j, 0)$ the probability of it being in state i at time t if the initial state was j . The time evolution of $P(t|j, 0) := (p(1, t|j, 0), \dots, p(n, t|j, 0))$ is governed by the master equation $d_t P(t|j, 0) = WP(t|j, 0)$. The entries w_{kl} , $k \neq l$ of the matrix $W \in \mathbb{R}^{n \times n}$ represent the transition probability densities from state l to state k . The diagonal elements w_{ii} are such that each column of W sums to 0.

We will use 3-state and 4-state subunit schemes in this report. Four states correspond to 2 binding sites on a subunit. Three states could result e.g. from lumping of the two inhibited states [13] leading to unique rest, activated and inhibited states of a subunit. The transition matrices are given by

$$W_4 = \begin{pmatrix} -15 & 2 & 0 & 3 \\ 5 & -4 & 3 & 0 \\ 0 & 2 & -9 & 3 \\ 10 & 0 & 6 & -6 \end{pmatrix}, \quad (4a)$$

for the 4-state model and by

$$W_3 = \begin{pmatrix} -6 & 10 & 12 \\ 2 & -15 & 3 \\ 4 & 5 & -15 \end{pmatrix}, \quad (4b)$$

for the 3-state model. Since our approach holds for any subunit dynamics with constant transition rates, the entries in eq. (4) are chosen arbitrarily. The only restriction is that they comply with detailed balance. Note that the

entries have units of inverse time. We denote the 4 states as 00, 10, 11 and 01, where each digit indicates the state of a binding site as occupied (1) or empty (0). The 3 states read as i , a and r . The state of a channel molecule will be described by the numbers of subunits in each of the subunit states, i.e. (n_i, n_a, n_r) for a channel consisting of 3-state subunits and $(n_{00}, n_{10}, n_{11}, n_{01})$ for 4-state subunits.

With a constant matrix W which can be diagonalised, a solution of the master equation is given by

$$P(t|j, 0) = \sum_{i=1}^n c_{ji} V_i \exp(\lambda_i t). \quad (5)$$

The coefficients $c_{ji} \in \mathbb{R}$ are determined by initial conditions. $\lambda_i \in \mathbb{R}$ and $V_i = (v_{1i}, \dots, v_{ni})$ denote an eigenvalue and the corresponding eigenvector of W , respectively.

In a first step, we focus on channels that are activated when all h subunits are in a unique activated state labelled by the index a . Hence, the activated channel state A is $n_a = h$, $n_i = 0$, $i \neq a$. We consider an initial condition I that has $m_i, i = 1, \dots, n$ subunits in state i . The probability $q := q(A, t|I, 0)$ of a channel being activated at time t given I at time 0 is

$$\begin{aligned} q(A, t|I, 0) &:= \prod_{i=1}^n p(a, t|i, 0)^{m_i} \\ &= \prod_{i=1}^n \left(\sum_{j=1}^n c_{ij} v_{aj} \exp(\lambda_j t) \right)^{m_i} \\ &= \prod_{i=1}^n \left(\sum_{\{k_i\}} \frac{m_i!}{k_{i1}! \dots k_{in}!} \prod_{j=1}^n (c_{ij} v_{aj} e^{\lambda_j t})^{k_{ij}} \right) \\ &= \sum_{\{k_1\} \dots \{k_n\}} \frac{m_1! \dots m_n!}{k_{11}! \dots k_{nn}!} \prod_{j=1}^n \prod_{i=1}^n (c_{ij} v_{aj})^{k_{ij}} e^{l_j \lambda_j t}. \end{aligned} \quad (6)$$

$\{k_i\}$ refers to all decompositions of m_i into n summands, i.e. all n -tuples of positive integers obeying $\sum_j k_{ij} = m_i$. We can regroup the last line of eq. (6) into

$$q = \sum_{\{l\}} M(\{l\}, m) \exp\{(l_1 \lambda_1 + \dots + l_n \lambda_n) t\}, \quad (7)$$

with $m := (m_1, \dots, m_n)$. $\{l\}$ denotes all decompositions of h into n summands ($\sum_{j=1}^n l_j = h$), and

$$M(\{l\}, m) := \sum'_{\{k_1\} \dots \{k_n\}} \frac{m_1! \dots m_n!}{k_{11}! \dots k_{nn}!} \prod_{i,j=1}^n [c_{ij} v_{aj}]^{k_{ij}}. \quad (8)$$

The prime indicates the restriction $\sum_j k_{ji} = l_i$. The ansatz for q in eq. (6) reflects the fact that we consider the subunits to be identical and to gate independently [14]. Since there are $r := \binom{h+n-1}{n-1}$ terms in eq. (7) corresponding to r ways to decompose h into n summands — note at

the same time that there are exactly r channel states —, we cast eq. (7) into the form

$$q = \sum_{j=1}^r M_j(m) \exp(\eta_j t). \quad (9)$$

$\eta_j = l_1 \lambda_1 + \dots + l_n \lambda_n$ is a linear combination of the eigenvalues of W , and $M_j(m)$ equals $M(\{l\}, m)$ evaluated at the j th decomposition of h . The Laplace transform of eq. (9) reads as

$$\hat{q}(A, s|I, 0) = \sum_{j=1}^r \frac{M_j(m)}{s - \eta_j}. \quad (10)$$

The common denominator of $\hat{q}(A, s|I, 0)$ is independent of the initial condition. Therefore, it cancels when we insert eq. (10) into eq. (3). Setting $\tilde{q}(A, s|I, 0) := \sum_j M_j(m) \prod_{k \neq j} (s - \eta_k)$ leads to

$$\hat{F}(A, s|I, 0) = \frac{\hat{q}(A, s|I, 0)}{\hat{q}(A, s|A, 0)} = \frac{\tilde{q}(A, s|I, 0)}{\tilde{q}(A, s|A, 0)}, \quad (11)$$

for $I \neq A$. The first passage time distribution in the time domain follows readily as

$$F(A, t|I, 0) = \sum_{j=1}^{r-1} \frac{\tilde{q}(A, s_j|I, 0)}{\tilde{q}'(A, s_j|A, 0)} \exp(s_j t), \quad I \neq A, \quad (12)$$

using contour integration. The prime indicates the derivative with respect to s , and $\{s_j\}$ is the set of all zeros of $\tilde{q}(A, s|A, 0)$. Note that the time scales of the first passage time distribution are solely governed by the properties of the return probability of the activated state, and the determining polynomial is of order $(r-1)$. Although $F(A, t|I, 0)$ is a sum of decaying exponentials, it can exhibit a maximum, because the coefficients $\tilde{q}(A, s_j|I, 0)/\tilde{q}'(A, s_j|A, 0)$ can have different signs for different values of s_j .

So far, we used an activated molecule state A with all subunits in the same subunit state a . Hence, there is only one realisation of A in terms of subunit states. However, experiments have demonstrated that channels are conducting in more than one configuration. That is true for the IP₃R channel in particular, which is deemed open when 3 out of 4 subunits are in the active state [1]. Therefore, we consider the general case that the molecule is activated whenever it is in one of the states A_i , $i = 1, \dots, r_a < r$. Let $F(A_i A, t|I, 0) dt$ denote the joint probability of being in the state A_i and in the activated state $A := \sum_i A_i$ for the first time in the interval $[t, t+dt]$ given an initial state I at time 0. The probability $F(A_i A, t|I, 0) dt$ subsumes all trajectories in phase space that arrive at A_i for the first time without having visited any of the other states A_j , $j \neq i$ before. Consequently, the probability density $F(A, t|I, 0)$ for the molecule of arriving in A for the first time is $F(A, t|I, 0) = \sum_i F(A_i A, t|I, 0)$,

since the A_i are mutually exclusive [15]. This leads directly to a generalisation of equation (1) as

$$q(A_i, t|I, 0) = \delta_{A_i, I} \delta(t) + \sum_{j=1}^{r_a} \int_0^t d\tau F(A_j A, \tau|I, 0) q(A_i, t|A_j, \tau), \quad (13)$$

$i = 1, \dots, r_a$, which is equivalent to

$$\hat{q}(A_i, s|I, 0) = \sum_{j=1}^{r_a} \hat{F}(A_j A, s|I, 0) \hat{q}(A_i, s|A_j, 0), \quad (14)$$

for $I \neq A_i$. Equation (14) is a system of r_a equations for the r_a unknowns $\hat{F}(A_j A, s|I, 0)$. Let $B \in \mathbb{C}^{r_a \times r_a}$ be a matrix with the entries $b_{ij} := \hat{q}(A_i, s|A_j, 0)$, then $\hat{F}(A_j A, s|I, 0) = \sum_i B_{ji}^{-1} \hat{q}(A_i, s|I, 0)$.

To illustrate the preceding scheme, we analyse a molecule that consists of three subunits with three states each. We choose the initial state with all subunits in the third state, $I = (0, 0, 3)$, and the activated state as $A = A_1 \cup A_2 = (1, 2, 0) \cup (0, 2, 1)$. This leads to

$$B = \begin{pmatrix} \hat{q}(A_1, s|A_1, 0) & \hat{q}(A_1, s|A_2, 0) \\ \hat{q}(A_2, s|A_1, 0) & \hat{q}(A_2, s|A_2, 0) \end{pmatrix}, \quad (15)$$

so that e.g. $\hat{F}(A_1 A, s|I, 0) = N_1 / \det B$ with

$$N_1 := \hat{q}(A_2, s|A_2, 0) \hat{q}(A_1, s|I, 0) - \hat{q}(A_1, s|A_2, 0) \hat{q}(A_2, s|I, 0) \quad (16)$$

The computation of $\hat{q}(A_i, s|I, 0)$ and $\hat{q}(A_i, s|A_j, 0)$ requires an extension of our notion of eq. (6). Now, the transition from A_j to A_i comprises all trajectories *from any* subunit configuration in A_j to *any* configuration in A_i . Take e.g. the transition from A_1 to A_2 . Since the subunit that is initially in the inhibited (first) state can either move to the active (second) or to the resting (third) state, we need to consider both possible transitions with appropriate combinatorial factors:

$$q(A_2, t|A_1, 0) = p(2, t|2, 0)^2 p(3, t|1, 0) + 2p(3, t|2, 0)p(2, t|2, 0)p(2, t|1, 0). \quad (17)$$

Analogously, $q(A_i, t|I, 0)$ must take into account all combinatorial possibilities to get from I to A_i , so that e.g.:

$$q(A_1, t|I, 0) = 3p(2, t|3, 0)^2 p(1, t|3, 0). \quad (18)$$

These state probabilities are sums of exponentials, so that the Laplace transforms \hat{q} are of the form of eq. (10). It entails that $\hat{F}(A_i A, s|I, 0)$ are fractions of polynomials of order $2r - 2$ and are readily back transformed in analogy to eq. (12). The left panel in fig. 1 depicts the outcome of this concept. The analytical solution is almost indistinguishable from results of stochastic simulations.

We now move one structural level up to channel clusters, in order to describe events like puffs in intracellular

Ca^{2+} dynamics. A puff starts when the first channel in the cluster opens. The probability for the first channel activation depends on the initial states of all channels. Let $u := (u_1, \dots, u_r)$ with $\sum_{i=1}^r u_i = N$ denote the primary distribution of the N channels among the r channel states. Then, assuming that channels in a cluster behave independently, the probability distribution for a cluster to be activated for the first time at time t reads as

$$F_u(t) := \sum_{i=1}^r u_i F(A, t|I_i, 0) G_i^{u_i-1} \prod_{j \neq i} G_j^{u_j}, \quad (19)$$

$$G_i := G(A, t|I_i, 0) := 1 - \int_0^t F(A, \tau|I_i, 0) d\tau. \quad (20)$$

Here, G_i is the probability that a channel that was originally in the channel state I_i has not yet reached A at time t ; note that $G(A, t|A, 0) = \delta(t)$. Equation (19) states that a cluster is activated for the first time when one channel opens for the first time while all other channels are still closed.

Since subunits represent the defining dynamical elements, initial data are often given as the number of subunits in each of the n states. However, this does not determine the distribution of the Nh subunits among the N channels. Consequently, we need to average over all possible initial cluster states with $b_i, i = 1, \dots, n$ subunits in the i th state. This leads to the probability distribution $\bar{F}_b(t) = (1/\Gamma) \sum_u f_u F_u(t)$, where the sum runs over all r -tuples u obeying

$$\sum_{i=1}^r u_i = N, \quad \sum_{j=1}^r m_{ij} u_j = b_i, \quad i = 1, \dots, n, \quad (21)$$

and where

$$f_u := N! \prod_{j=1}^r \frac{1}{u_j!} \left[\frac{h!}{m_{1j}! \dots m_{nj}!} \right]^{u_j}. \quad (22)$$

The constants m_{ij} represent the number of subunits in the state i given the j th channel state ($\sum_i m_{ij} = h$). Hence, a single square bracket equals the number of possibilities to generate a distinct channel state, so that f_u corresponds to the number of ways to realise a cluster state where u_i channels are in the state i . The normalization, $\Gamma := \sum_u f_u$, can be shown to evaluate to $(Nh)! / (b_1! \dots b_n!)$.

The right panel in fig. 1 compares results for \bar{F}_b to stochastic simulations of an entire cluster. The agreement between analytic results and simulations is excellent. Note the resolution of the delta peak at $t = 0$. Although the graph illustrates only small fractions of the probability distributions, the concord persists throughout the tails.

In summary, we have presented an exact method to calculate waiting time distributions for clusters of stochastic molecules from master equations. Molecule subunits should be coupled by the absorbing state only but do not need to be identical (as we assumed for convenience here)

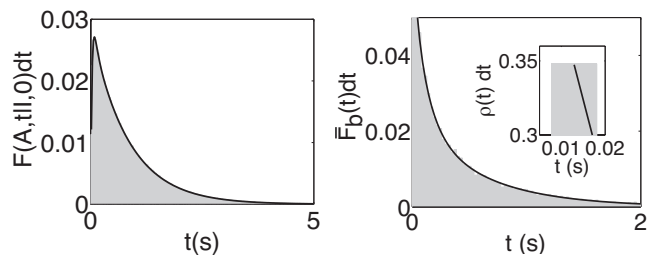


Fig. 1: Left: Escape probability F for a channel with 3 3-state subunits using eqs. (12) and (14). The initial state is $I = (0, 0, 3)$, and the absorbing state is $A = (0, 2, 1) \cup (1, 2, 0)$. Right: Escape probability \bar{F}_b for a cluster of 7 trimeric channels. Each subunit is modelled by 4 states. 10 is the activated state of the subunit. Initial conditions are $(n_{00}, n_{10}, n_{11}, n_{01}) = (2, 8, 4, 7)$, and the absorbing state is when a channel has 3 subunits in 10 for the first time. The inset shows a blow up of the delta-peak at $t = 0$. Lines - analytic results, shaded areas - histograms from stochastic simulations (500000 trials).

in order for the method to be applicable. Constructing the Laplace transform of the waiting time distributions from the Laplace transform of solutions of the master equation of single subunits circumvents combinatorial explosion arising from the coupling of the subunit master equations by the absorbing state. Note, that the method even calculates splitting probabilities for first passage to different absorbing states along the way.

New approaches for solving non-factorisable master equations have been suggested in recent years. They mainly focus on numerical schemes and approximations or involve complex computation for subunits with more than two states (see e.g. [16,17]). The advantage of our method lies in utilising explicitly the independence of the subunits despite the coupling by the absorbing state. Hence, we start directly from the master equation of a subunit, and not from the full molecule master equation. This drastically reduces the dimension of the phase space.

An effective method to calculate waiting time distributions, as we have demonstrated here, paves the way to a description of hierarchical stochastic systems. Such a theory is needed in particular for the modelling of processes in living cells, where we have small numbers of active elements that exhibit discrete states.

REFERENCES

- [1] FALCKE M., *Advances in Physics*, **53** (2004) 255.
- [2] ALBERTS B., BRAY D., LEWIS J., M. RAFF, K. ROBERTS and WATSON J., *Molecular Biology of the Cell* (Garland Publishing, Inc. New York & London) 1994.
- [3] GUTKIN B. S. and ERMENTROUT G. B., *Neural Comput.*, **10** (1998) 1047.
- [4] VERECHTCHAGUINA T., SCHIMANSKY-GEIER L. and SOKOLOV I. M., *Phys. Rev. E*, **70** (2004) 031916.

- [5] PRAGER T., NAUNDORF B. and SCHIMANSKY-GEIER L., *Physica A*, **325** (2003) 176.
- [6] PRAGER T., FALCKE M., SCHIMANSKY-GEIER L. and ZAKS M.A., *Phys. Rev. E*, (2007) in press.
- [7] SVOBODA K., MITRA P. and BLOCK S., *Proc. Natl. Acad. Sci. USA*, **91** (1994) 11782.
- [8] CAPPELLO G., BADOUAL M., OTT A., PROST J. and BUSONI L., *Phys. Rev. E*, **68** (2003) 021907.
- [9] SANTOS J., FRANOSCH T., PARMEGGIANI A. and FREY E., *Phys. Biol.*, **13** (2005) 207.
- [10] ZELDOVICH K., JOANNY J.-F. and PROST J., *Eur. Phys. J. E*, **17** (2005) 155.
- [11] WOOD K., DEN BROECK C. V., KAWAI R. and LINDENBERG K., *Phys. Rev. Lett.*, **96** (2006) 145701.
- [12] REDNER S., *A Guide to First-Passage Processes* (Cambridge University Press) 2001.
- [13] THUL R. and FALCKE M., *Phys. Rev. E*, **73** (2006) 061923.
- [14] HILLE B., *Ion Channels of Excitable Membranes* (Sinauer Associates Sunderland, MA USA) 2001.
- [15] FELLER W., *An Introduction to Probability Theory and Its Applications*, Vol. **1**. (Wiley New York) 1968.
- [16] CAO Y., GILLESPIE D. T. and PETZOLD L. R., *J. Chem. Phys.*, **122** (2005) 014116.
- [17] NGUYEN V., MATHIAS R. and SMITH G. D., *Bull. Math. Biol.*, **67** (2005) 392.