Patrik Fortes

Analysing a bistable switch in systems biology with stochastic simulations

Turned in:
25-01-2012
Seminar given:

Supervisors:
Henrik Jönsson
Pawel Krupinski

Contents

Analysing a bistable switch in systems biology with stochastic simulations

Abstract

We analyse biological switches with the aid of computers and the Gillespie method, a stochastic solver. Stochastic calculations can lead to spontaneous switches which are, in part, determined by a subset of the parameters. In a deterministic description this is hidden. Initial analysis using a deterministic description of the system will reveal that by applying certain constraints on the system we can greatly simplify the calculations. This gives us a quick and simple way to calculate not only how many fixed points a system has for any specified set of parameters, but also the equilibrium concentrations. This is in great contrast to when the calculations are done stochastically, where you unlock certain behaviours and loose others, such as being able to accurately calculate equilibrium concentrations (number of molecules). We will also see that, as a rough estimate, the highest equilibrium concentration is dependent on a specific parameter, namely the quotient between the production and degradation rate. Finally we also look at a system in which an external signal is added, driving the system to switch.

Introduction

Biochemical systems can be described by either continuous (concentration) or discrete (number of molecules) variables. One common way of describing the dynamics of the systems is by treating the system as deterministic. However, on the cellular level where the concentration of chemicals can be as low as single molecules per unit volume, the randomness of chemical reactions becomes increasingly important. In such a case stochastic mathematics might be a better way of describing chemical reactions, especially if the system shows multi-stability or switch-behaviour.

Daniel T. Gillespie wrote a paper[1] back in '77, describing a new method of simulating chemical systems by using stochastic mathematics and computers. The method, called the Gillespie method, has become widely used when you wish to look at systems with coupled chemical reactions and is easy to implement, requiring at most four steps in the main body.

We will use the stochastic method to analyse a simple network presented in Gardner et al[2], dealing with what they call “Construction of a genetic toggle-switch in Escherichia-coli”. We use this simple synthetic switch system which has the potential to spontaneously switch from one state to another.

It has been discovered that nature sometimes takes advantage of gene constructions which act as switches, since they have been found to exhibit multi-stability and switching. Some examples of these are the bacteriophage λ switch[4] and the Cyano bacteria circadian oscillator[5]. These types of objects can be said to be addressable cellular memory units and so they can be used in bio-tech applications such as biocomputing or in the field of gene therapy[2]. Various types exist, both synthetic (engineered) and natural.

In the paper Gardner and his team construct a bistable switch and they analyse it deterministically. We will compare the deterministic analyses with the stochastic using the aforementioned method. The switch itself contains only two inducers, two promoters and two repressors which act on each others promoters (Fig. 0A). Two versions will be studied. In the first one the system will not have any switch inducers, whereas the second will include one inducer. This will then be analysed with the mass-action formalism using a system of differential equations as well as with stochastic methods.

Methods

The Gillespie method

In 1977 Daniel Gillespie wrote a paper describing a new way of simulating stochastic systems. The general problem discussed is a study of the time evolution of chemical reactions which are coupled together. The main approach is that time is continuous and at any point in time only one reaction can occur. With this idea it is now possible to construct a method for simulating systems. Time step is represented with a random variable according to a distribution, describing the time to the next reaction event to happen (see below). After the time step has been chosen it is added to the recent time point.

To this you generate a second random number which is going to represent the chosen chemical reaction at the above chosen time and the system is updated. The crucial steps of the algorithm can be summarised as follows

1: Generating a random time step.
2: Generating a random reaction.
3: Update system.
4: Repeat until final time is reached.

As a precaution one can add a step in between the first and second steps to check whether one has overstepped one's time limit or not.
As this was just a rough description of the method used, I will now describe in detail how it was implemented. First of all we know that every reaction known has its own reaction rate, propensity if you will, and so this has to be used in order to accurately describe the development of the system. This is done in the following manner. We know that a reaction with a larger rate will evolve at a higher speed. And so, the probability of choosing the reaction with the higher rate has to be higher than the probability for the reaction with a lower rate. This can be represented by introducing a factor in the probability,

$$\frac{1}{\sum_{\mu=1}^{n} H_\mu C_\mu}$$

The C_μ's are the stochastic reaction constant, i.e. the average probability that a specific reaction will happen in the next time interval dt. H_μ is the number of distinct molecular combinations in the state $(X_1,,,X_n)$, where X_i is the number of molecules of type i. When we work with the chosen time step we also have to use a probability distribution to accurately describe the flow of time. The Gillespie algorithm uses the logarithm distribution with the inverse of the random number, r, which is drawn from a uniform distribution in the interval (0,1)

$$\log(\frac{1}{r}) \quad .$$

The reason for this comes from statistical mechanics and is given as follows. To quote Gillespie directly[1], given that R_μ is the μ:th reaction in a set of reactions {R}, with state $(X_1,,,X_n)$ and defining a_μ as H_μ times C_μ:
“$P_0(\tau)$, the probability that, given the state $(X_1,,,X_n)$ at time t, no reaction will occur in the

time interval (t,t+τ); times a_μ dτ, the subsequent probability that an R_μ reaction will occur in the time interval (t+τ, t+τ+dτ): $P(\tau,\mu)d\tau=a_\mu P_0(\tau)d\tau$. To find an expression for $P_0(\tau)$, we first note that

$$1-\sum_\nu a_\nu d\tau'$$

is the probability that no reaction will occur in time dτ' from the state (X_1,,,X_n). Therefore

$$P_0(\tau'+d\tau')=P_0(\tau')(1-\sum_{\nu=1}^{M} a_\nu d\tau')$$

from which it is readily deduced that

$$P_0(\tau)=e^{-\sum_{\nu=1}^{M} a_\nu \tau} \quad ,,$$

This shows why we have to have the log distribution. With this in mind it now follows that when we take a random time step we should choose it in the following way

$$\tau=\frac{1}{\sum_{\nu=1}^{M} a_\nu}\log(\frac{1}{r})$$

When it comes to choosing the right reaction we only have to view the situation when we look at a bag filled with balls with the reaction numbers printed on them. Then we randomly pick one ball after another. To correctly give the probabilities we have to add weights to the balls so some reactions will be picked more often than others. This can be represented by simply adding a_μ:s and multiplying with a random number between 0 and 1 from a uniform distribution. After which we compare the product with the cumulated sum of the factors between the stochastic reaction constant and the number of distinct R_μ molecular combinations in the state (X_1,,,X_n) and choose the reaction which corresponds to the nearest lowest cumulated sum.

The Taylor approximation

According to Taylor, any function which is two times differentiable can be expanded around a point x_0 with Δx as a polynomial as

$$Y_{(x_0+\Delta x)}=Y_{(x_0)}+\Delta x\frac{dY_{(x_0)}}{dx}+\frac{\Delta x^2}{2}\frac{d^2Y_{(x_0)}}{dx^2} \quad . \qquad [i]$$

The extremum finding method

Because the simulations are stochastic, the statistics will not be smooth and so one has to allow for roughness which was accommodated by looking at larger areas, introducing averaging tools. For this we used a converted version of the simple standard numerical derivative given as.

$$\frac{\Delta Y}{\Delta X}=\frac{Y_{(i+1)}-Y_{(i-1)}}{2\Delta X}$$

With Y as a function of x, in our case meaning the number of molecules in the distribution of outputs, but instead of using the exact expression we used the expression.

$$Y_{(i+10)}-Y_{(i-10)} \qquad \text{[ii]}$$

In this version we average over twenty-one steps instead of three, then we checked the product between one of these differences and the one starting where the first one ends. If this was less than zero then the value of the distribution at that point and that point's position in *u* and *v* was logged, but not if the new position were smaller than the previous one.

Models

Mass-action formalism

In the mass-action formalism we look at chemical reactions and the formalism gives us a way of analysing the reactions deterministically. The main assumption in this formalism is that the reaction rate is proportional to the product between the concentrations of the reacting species, and the rate constant is the constant of proportionality. To find the time derivative of the concentration we have to account for the number of molecules of a specific type involved in the reaction, and determining the sign depending on if they are reactants(-) or products(+).

Chemical system

$$A+A\xrightarrow{k_1}B$$

Time derivative

$$\frac{dA}{dt}=-2\mathrm{k}_1A^2;\frac{dB}{dt}=k_1A^2$$

It's derived using the assumptions that the medium is well stirred and and that the concentrations of the partaking species are low. The assumption of mass-action formalism has the consequence that the probability for a reaction to occur is not dependent on the environment and in no part of the container is one reaction favoured over the others.

The Bistable switch

The chemical reactions describing the system of proteins with bound and unbound promoters are given as follows

$$\begin{aligned} u &\xrightarrow{k_1} \varnothing,\ \gamma u+p_v \underset{k_3}{\overset{k_2}{\Leftrightarrow}} p_{bv},\ p_u \xrightarrow{k_4} p_u+u \\ v &\xrightarrow{k_5} \varnothing,\ \beta v+p_u \underset{k_7}{\overset{k_6}{\Leftrightarrow}} p_{bu},\ p_v \xrightarrow{k_8} p_v+v \end{aligned} \qquad \text{[iii]}$$

Where *u* and *v* are proteins whereas the p_x is the promoter of the protein x and p_{bx} is the bound complex which is built out of p_x and y, the “other” protein. k_n are the rate constants in

the reaction. β, γ are seen in the reactions above as how many molecules of *u* and *v* are partaking in the two-way reaction. β and γ are important for the behaviour of the system regarding the number and type of fix-points. In this paper we will choose to set the β and γ to 2, meaning we will construct a complex which takes one promoter and a dimer or simply 2 molecules of either *u* or *v* depending on which promoter we use. This choice will allow for bi-stability in the system.
The first and the fifth reaction are two degradation reactions. In the fourth and the eighth reactions, the partaking promoters are generating new *u*:s and *v*:s. The second, third, sixth and the seventh reactions are two-way reactions, which tend to arrive at equilibrium concentrations in the partaking molecules' concentrations. When we derive the deterministic model, we will assume that the two-way reactions are fast and always at equilibrium, the time derivative of the concentrations of the promoters and the bound state, p_x and p_{bx}, are set to zero for all times. The properties of the system are such that whenever the system is in the bound state there will be no promoter to sustain the creation of the molecule which the specific promoter promotes and so other reactions will be conducted until the system becomes unbound again. Hence *u* and *v* acts as repressors of each others' production. A reaction is said to be fast if its speed of evolution is such that, relative to a system's complete set of reactions, it is greater than a majority of the system's reactions. The consequence of being a fast reaction is that the reaction will arrive at the equilibrium in such a way that the reaction can become a bottleneck in the system or fast to respond to any perturbation from the equilibrium.

Derivations

We will now derive the deterministic model of the system defined in the previous section and analyse it.
With the Michaelis-Menten or Hill description, the total derivatives in p_u, p_{bu}, p_v and p_{bv} are set to zero. This comes from the assumption that the binding-unbinding reaction is fast and always at equilibrium. It means that whenever we see a change in the amount of molecules of one species it will not be long before the system has undergone a change to accommodate the loss of concentration in one of the molecules, id est, the equilibrium will be restored in an instant. We also assume we have one promoter per gene

$$\begin{aligned} p_u + p_{bu} &= 1 \\ p_v + p_{bv} &= 1 \end{aligned} .$$

With these assumptions, *u*'s and *v*'s derivatives can be written as

$$\left\{ \begin{aligned} \frac{du}{dt} &= \frac{ac}{c+v^{\beta}} - eu \\ \frac{dv}{dt} &= \frac{bc}{c+u^{\gamma}} - ev \end{aligned} \right. . \qquad \text{[iv]}$$

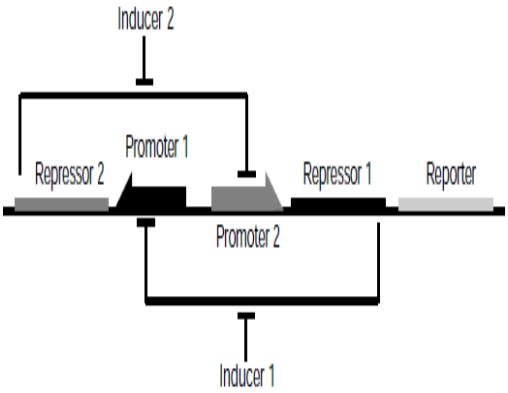
The parameters a, b, c, e are determined by the reaction constants in [v].

$$c = \frac{k_3}{k_2} = \frac{k_7}{k_6}, a = k_4, b = k_8, e = k_1 = k_5 \qquad \text{[v]}$$

Having completely defined the deterministic model allows us to analyse the dynamics of the system. In order to find the fix-points, the derivatives are set to zero and the system of equations can be simplified into

$$\left\{\begin{array}{l} 0 = e\left(\frac{ac}{e}\right)^{\gamma} v + ce(c+v^{\beta})^{\gamma} v - bc(c+v^{\beta})^{\gamma} \\ 0 = e\left(\frac{bc}{e}\right)^{\beta} u + ce(c+u^{\gamma})^{\beta} u - ac(c+u^{\gamma})^{\beta} \end{array}\right\} \qquad \text{[vi]}$$

where we will choose to use β=γ=2. We use these two equations ([iv]) to numerically search for parameter sets which give three positive values in *u* and *v* which result in derivatives with the value 0. Depending on the sign of the expressions of the derivatives, two of them will be stable with the other one being unstable or vice versa. Stability is defined as the behaviour a system has when it is perturbed around the fix-point. The fix-point is stable if, when the system is perturbed around the fix-point, the system will always be drawn to it. Unstable means that if, when the system is perturbed around the fix-point, the system will move a way from it.
For a given system, we can always start to calculate the so called null-clines. These are lines in which a single variable has zero derivative (Fig. 0B).



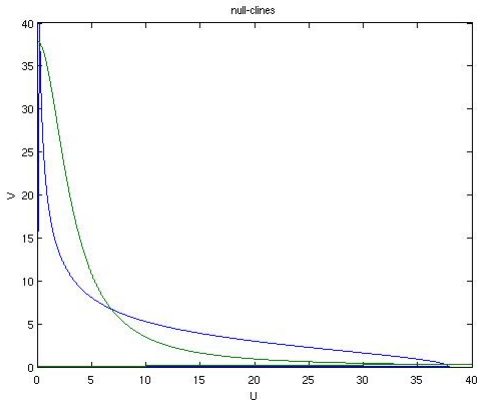


Fig 0A: Model of the switch. Promoter x is p_x in our system, Repressor y is *u* or v in our system. This picture was obtained with permission from Macmillan Magazines Ltd. Nature. Fig. 0B: Nullclines for the system where a=b=379.269, c=10, e,degradation, is equal to 10, The intersections are the points which results in zero derivatives in both *u* and v at the same time. For this specific symmetric, a=b, choice of parameters fix-points for *u* and v will be 0.2655 ,6.7785 and 37.6615.

Simulation results

We now display the results we obtained during our simulations. When the values are changed on the parameters the resulting change in number of fix-points can be quite visible as is seen in Fig. 1.

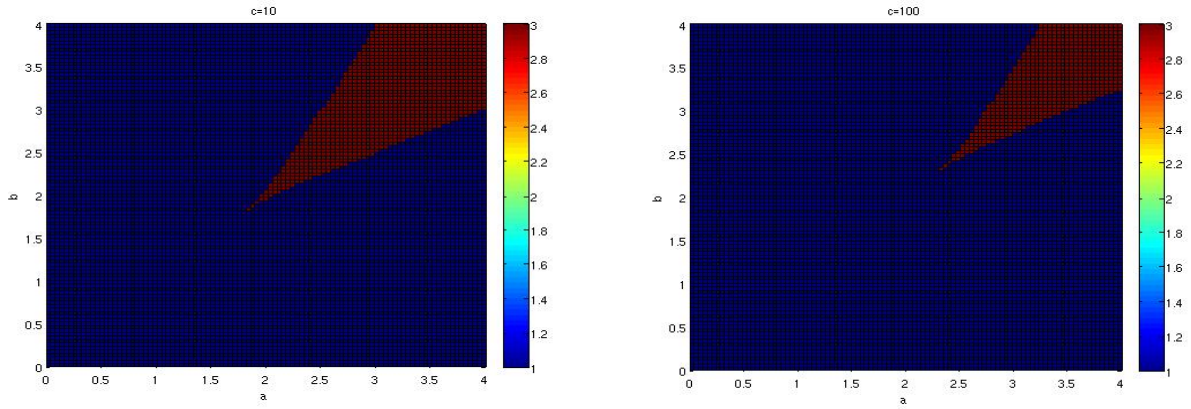


Fig. 1a Left: Parameter search when c has been set to 10, blue represents 1 equilibrium concentration of stability whereas red means 3, where one is unstable. Fig. 1b Right: Parameter search when c has been set to 100, Blue represents 1 equilibrium concentration of stability whereas red means 3 where one is in-stable. In both pictures you have a and b as the x and y axis. e, the degradation, was kept constant at 10. Please note logarithmic scales.

The red areas in Fig. 1 can be tabulated to give the different parameter sets which results in one or three equilibrium concentrations, [Tab. A1]. This allows us to do initial calculations in the deterministic environment, solving the system of differential equations [iv]. The method used for this is the second order Taylor approximation, [i]. The resulting calculations are shown in picture Fig. 2, in which we have chosen to display two evolutions of the same system but with different initial states. Meaning that *U* and *W* correspond to the same variable (*u*) and so do *V* and *Y* (*v*).

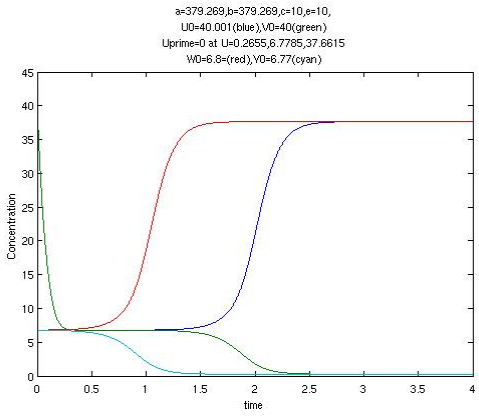


Fig. 2 Above: The initial state in the *U/V* system is U_0=40.001 and V_0=40.000 whereas the initial state in the *W/Y* system is W_0=6.8 whereas Y_0=6.77. a and b is 379.269, c=10, e,degradation, is equal to 10.

The *U/V* system first makes an initial dip down towards the unstable equilibrium concentration. Then the curves increase their separation. The variable with higher concentration shoots up towards the high equilibrium concentration and the lower one moves down to the low equilibrium concentration. *W*'s initial state in the W/Y system is 0.022 above the unstable equilibrium concentration where as *Y*'s initial state is 0.008 below the equilibrium concentration. The curves began immediately to diverge from each other and they moved towards their respective stable equilibrium concentration.

The deterministic system gives us nothing more than how an initial state will evolve in time. But when we do these in a stochastic environment, then we open up new behaviours which were not allowed in the deterministic environment. The major difference between the two environments, deterministic and stochastic, is being able to “switch”, jump from one region to another. Initially it was noticed that the frequency of switching might depend on

one parameter only or several, this might not be visible or have a corresponding counterpart in the deterministic environment (Fig. 3).

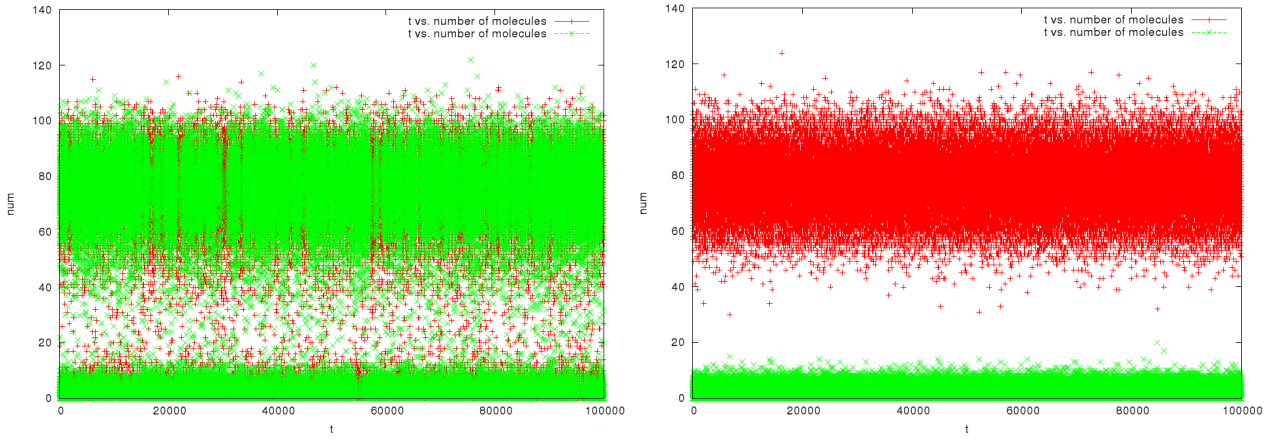


Fig. 3 Left and Right: has the reaction constants k_1, k_4, k_5 and k_8 in common, k_1 and k_5 as 12.74 and k_4 and k_8 1000. Fig. 3a Left: has c=100 with k_2 and k_6 as 1 and k_3 and k_7 as 100. Fig. 3b Right: has c=100 with k_2 and k_6 as 100 and k_3 and k_7 as 10000. e was kept constant and the same in both simulations.

We investigate when the stochastic model leads to spontaneous switching. In the first part, the model was analysed by varying the values of k_4 and k_8, corresponding to varying a and b in the Hill formalism, see Fig. 1.
After generating 100 models with varying values on k_4 and k_8, we implemented the Gillespie algorithm to conduct simulations with a project called organism[3]. To simplify the system and to study the effect of varying only k_4 and k_8, the rest of the reaction constants were kept fixed, k_1 and k_5 at 10 whereas k_2 and k_6 were kept at 10 and k_3 and k_7 at 100 to keep c fixed at 10. This was done in one of two simulations and in the other c was increased to 100 meaning that k_2 and k_6 were shifted to 1. From these 2 sets of simulations the number of switches, between the high state and the low, were in each simulation collected. As well as the concentration of molecules in the high state, which is situated where the largest cluster, at the highest number of molecules, of data points are sitting on the distribution. This corresponds to looking for the position of highest equilibrium concentration in the deterministic environment.

A stochastic model of the bistable switch leads to spontaneous switching

We now investigated the bistable switch with the Gillespie algorithm.
The parameters k_4 and k_8 were varied between 158 and 10000 logarithmically in 10 steps. The reason for the choice 158 is that when we work with the logarithm of the parameters a and b, k_4 and k_8, then it becomes easy to work with them if the lengths of the steps are 0.2 or 0.4. We mapped out the region of bi-stability for comparison with the stochastic simulations (cf. Fig. 4 with Fig. 1). The simulations in the deterministic system showed the same as the fix-point analysis.

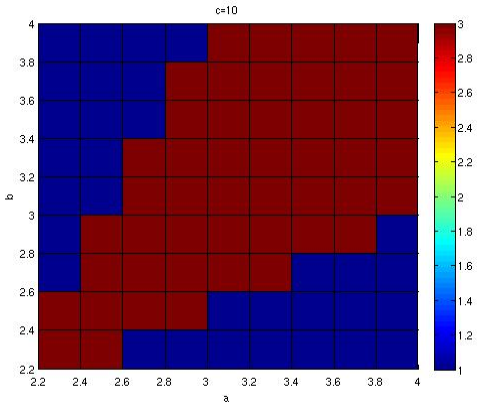


Fig. 4 Above: Number of equilibrium concentrations calculated by searching deterministically the domain of (u')'s function when c is 10. X-axis is a and Y-axis is b. The scales are logarithmic.

Next we investigated the number of spontaneous switches between the two states, high *u* and low *v* and low *u* and high *v*. As we move upwards in k_4 and k_8 the number of switches declined quite rapidly (Fig. 5a and Fig. 5b).

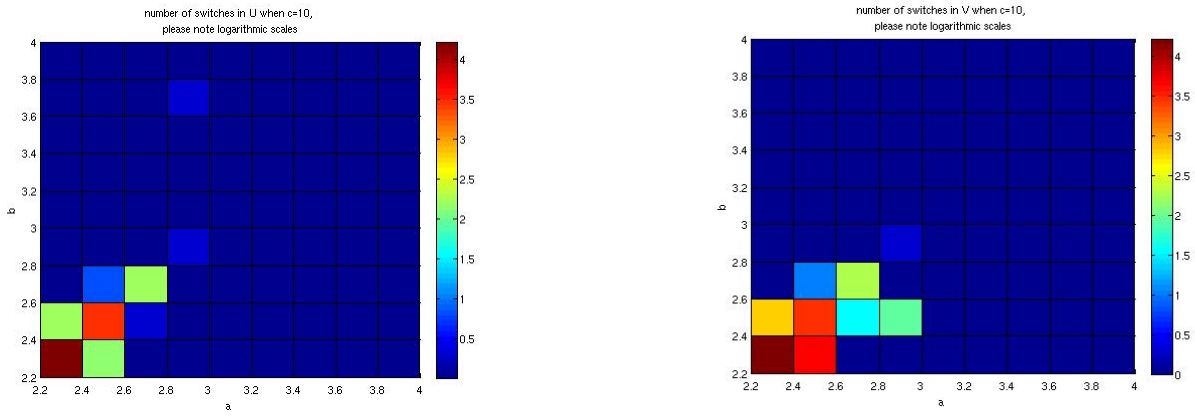


Fig. 5a Right: Number of switches in v. Fig. 5b Left: Number of switches in u. c=10, a and b is in [10^2.2 10^4] and e=10. Starting state is u_0=30 and v_0=0. The scales are logarithmic.

We also investigated where the system has its high stable state by analysing the distribution of concentration from the stochastic simulations (Fig. 6).

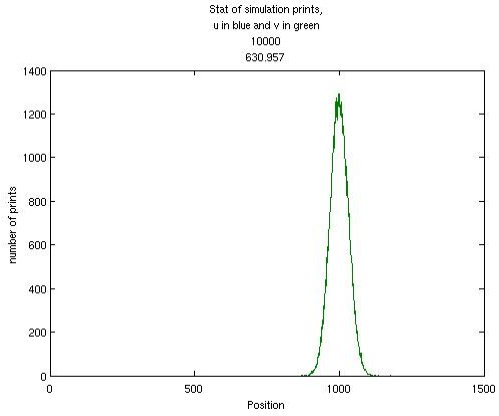


Fig. 6 Above: Example distribution of outputs in one simulation.

From the distribution, we then extracted the position of the highest equilibrium concentration by searching for the maximum at the highest concentration by using the

method [ii]. These equilibrium concentrations were then mapped out (Fig 7).

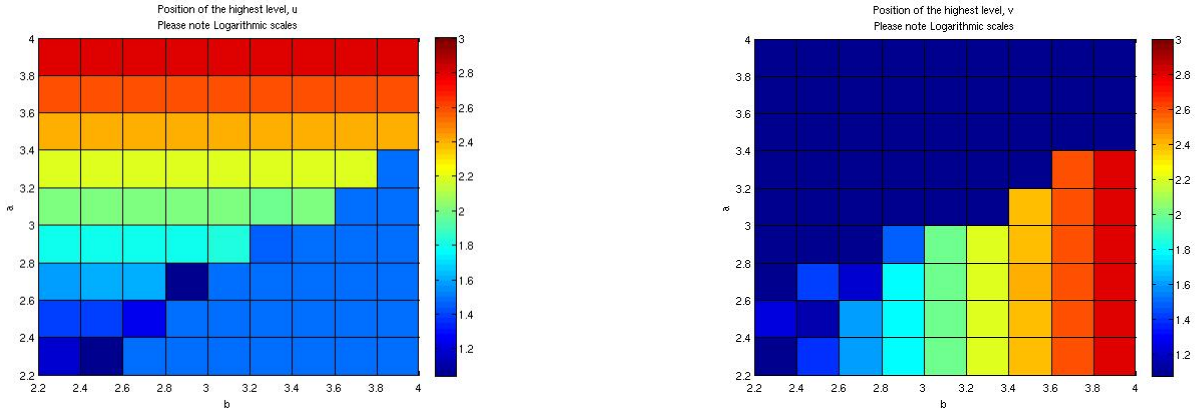


Fig. 7a Left: Extracted highest state in *u* from statistics, the scales are logarithmic. Fig. 7b Right: Extracted highest state in *v* from statistics, the scales are logarithmic. Fixed c at 10 in both cases and fixed degradation, e, at 10.

Fig. 7a and Fig. 7b show asymmetry which was investigated. The reason for the asymmetry was the initial states. The initial states were interchanged and we got the reverse effect (Fig. 8). Meaning, there exist regions in the a/b-graph where it's increasingly difficult for the species which is sitting in its low state to manage to jump upwards to its high region. This is indicated by the large triangular areas of blue colour.

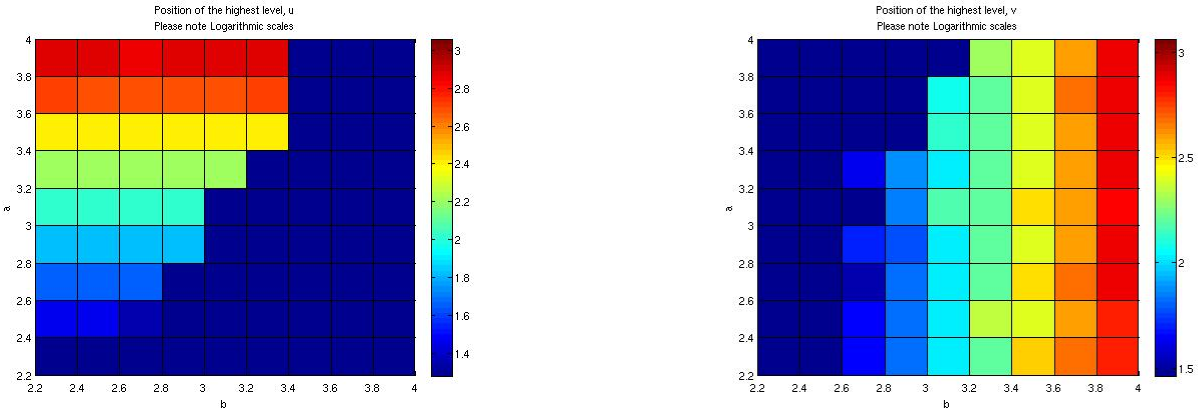


Fig 8a left and 8b right: Showing the reverse asymmetry in Fig. 7a and Fig. 7b. The scales are logarithmic.

It will later be shown that the system with c as 100 also exhibits the same asymmetry and there also the reason is the initial states.
To further analyse the switching behaviour, we investigated the time spent in each state by calculating the distributions of concentrations through a simulation.
We then analysed statistics for varying sets of parameters. Fig. 9 illustrates some features of switching behaviour, but the difference between the two needs a bit of explaining. Due to the fact that the highest equilibrium concentration in Fig. 9a is situated very low, it nearly becomes ambiguous when and where the system has switched. It can be shown with simulations that sometimes the statistics becomes so cluttered that the high state is indistinguishable from the low state. We can clearly see in Fig. 9b a difference in the amount of visits to the high state for *u* and *v*. Additional simulations at the stated values on a and b showed that these phenomena are statistical fluctuations (Fig. 10). Fig. 9c and Fig. 9d show two normal situations, 9c displaying a peak at 1000 and 9d showing a peak at 251 in rough accordance with one tenth of one of the parameters listed in the diagram. The reason for the tenth comes from the fact that the degradation is exactly 10.

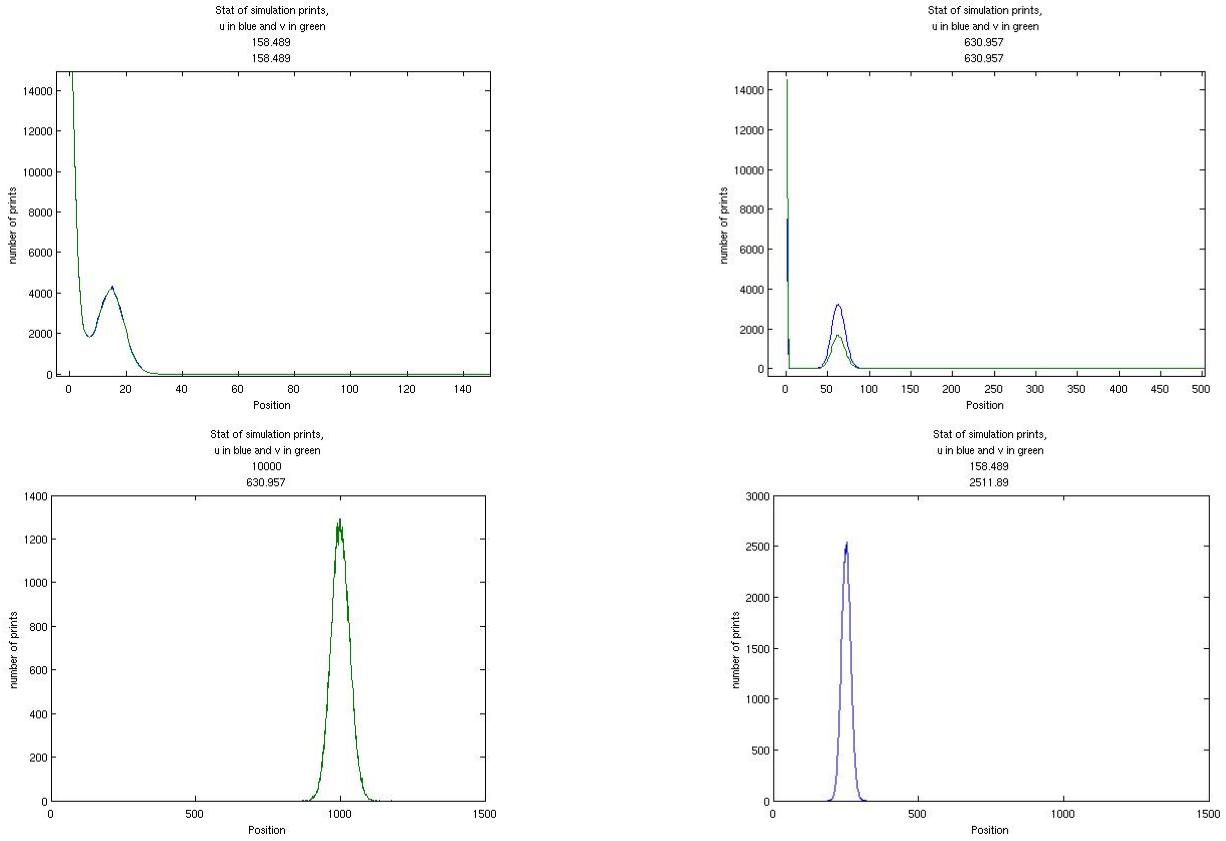


Fig. 9 These figures represent different scenarios with different sets of parameters. The upper number in the title is b's value and the lower is a's. Fig. 9a Top Left: Switching system with a low maximum equilibrium concentration. Fig. 9b Top Right: Switching system with a medium high maximum equilibrium concentration. Fig. 9c Bottom Left: Peak in *v* situated at 1000. Fig. 9d Bottom Right: Peak in *u* situated at 251.Please note that the overall maximum has been cut off in order to show the behaviour more clearly.

The difference in amplitude in the distributions in Fig. 9b was investigated with 100 simulations with the same parameter set (Fig. 10) and the means were calculated. The mean in *u* is at 2467 and 2454 in v. The standard deviation is at 67 and 68 in *u* and *v* respectively. The difference between the means were roughly 1%.

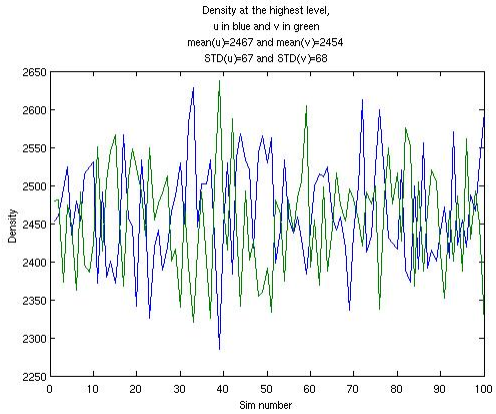


Fig. 10 Above: 100 simulations at a=b=630.957 with c=10, k_2 and k_6 at 0.1 and k_3 and k_7 at 1, and e at 10. The mean value in *u* is 2467 and *v* at 2454, the standard deviations are at 67 for *u* and 68 for v.

Sometimes when we studied the statistics in one of the scenarios we could see one stable state for *u* and a different one for *v* (Fig. 11), this is obvious since for every set of unequal a and b it is probable that if one were to switch the a and b we would find a new set of parameters which also has three equilibrium concentrations. The logic behind it is that if one were to switch a and b then this is the same as switching *u* and v, meaning *u* becomes

v and vice versa.

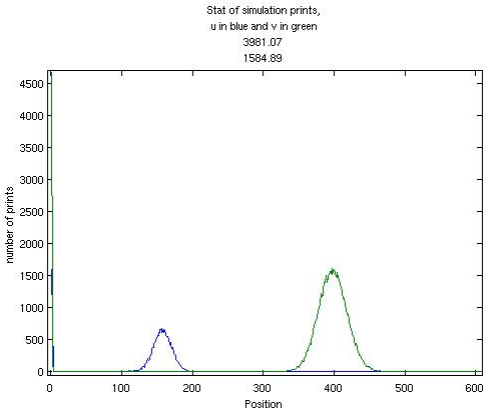


Fig. 11 Left: Statistics for one set of parameters, please note the two peaks. Please note also that because the majority of cases has the bulk of the statistics lying at the first point id est zero it drives scales on the axis upwards causing the rest of picture to drown and so the first point has been excluded to show these two peaks. The upper number in the title is b's value and the lower is a's.

Slowing down the binding process

The procedure was now repeated but with c shifted by a factor 10. We start with the deterministic chart telling when we have more than one equilibrium concentration (Fig. 12).

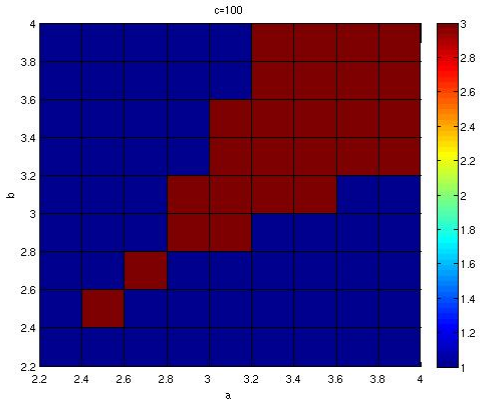


Fig. 12 Left: Number of equilibrium concentrations calculated by searching the domain of u's function when c is 100 X-axis is a and Y-axis is b. The scales are logarithmic.

Due to the fact that c has been shifted by a factor of 10, it becomes increasingly difficult to find parameter sets which show three equilibrium concentrations (Fig. 12). But as it was seen when c was 10, only a few sets show the behaviour of being prone to spontaneously switch from the low state to high and vice versa, making it even more difficult to find candidate sets which result in switches. So we expect higher equilibrium concentrations in the parameter set of a and b which correspond to the same parameter set when c is 10. With this we expect to find more molecules and less relative noise.

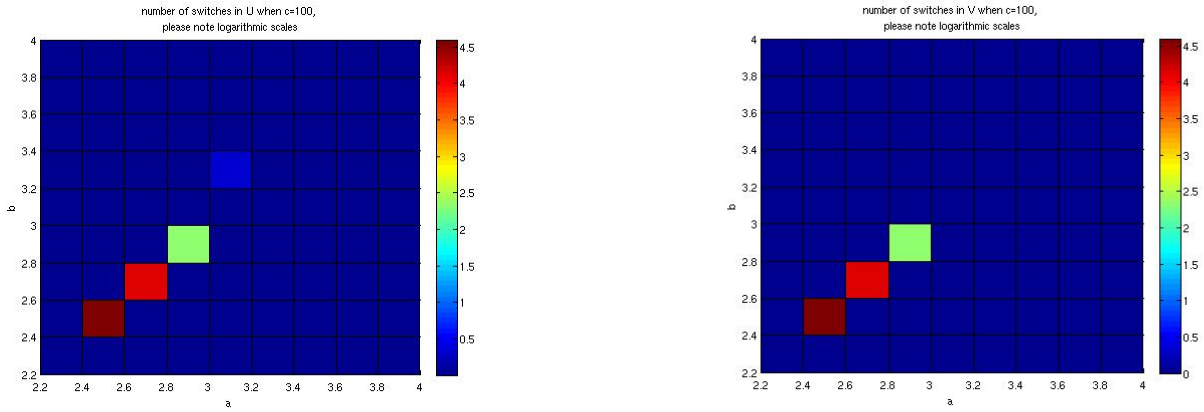


Fig. 13a Right: Number of switches in v. Fig. 13b Left: Number of switches in u. c=100, a and b is in [10^2.2 10^4] and e=10. Starting state is u0=30 and v0=0.The scales are logarithmic.

We analysed the statistics with equation [ii], (cf. Fig. 7).

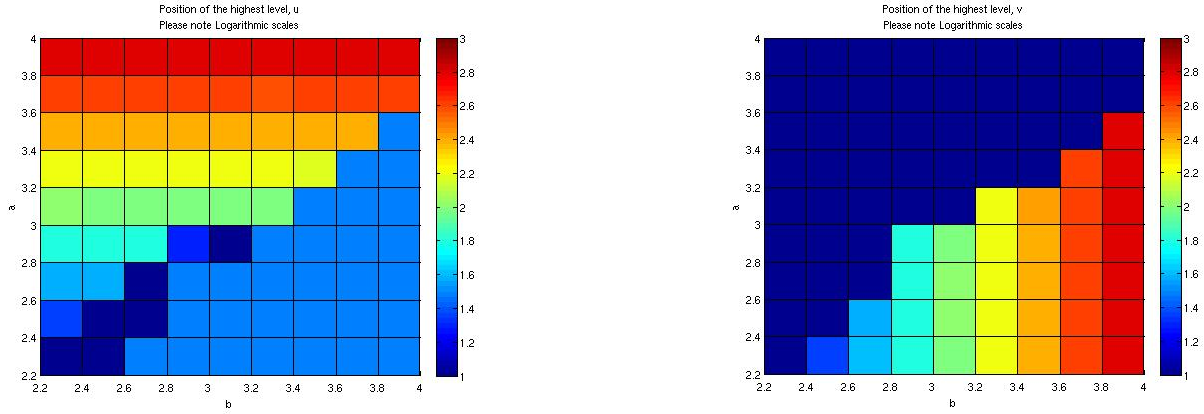


Fig. 14a Left: Extracted highest state in *u* from statistics, the scales are logarithmic. Fig. 14b Right: Extracted highest state in *v* from statistics, the scales are logarithmic. Fixed c at 100 in both cases and fixed degradation at 10.

As it was predicted when c equals 10, the same asymmetry was seen with the same reason, the initial states(Fig. 15).

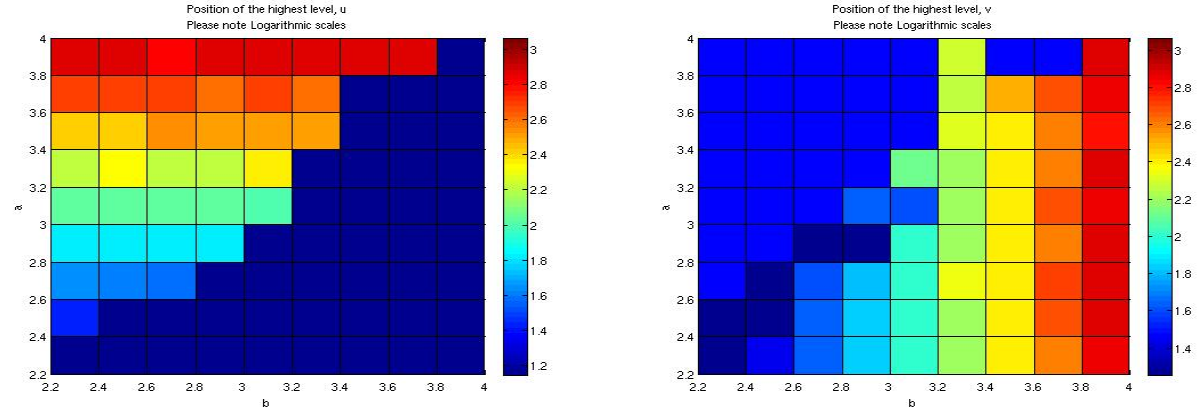


Fig. 15a left and 15b right: Showing the reverse asymmetry in the system represented by Fig. 14a and Fig 14b. The scales are logarithmic.

As with c equals 10, In Fig. 16 we see different outcomes of 4 scenarios when the statistics had been sorted. There is no major difference between this case and the case when c equals 10.
However it wasn't possible to find a similar case when the maximum equilibrium concentrations differed between *u* and *v* in this set of parameters. In Fig. 16a we see a

system which does not show any particular switching behaviour. This is also confirmed by looking at the system's parameters in Fig. 16a and comparing them to the deterministic map (Fig. 12) which shows the amount of equilibrium concentrations for each set of a and b.

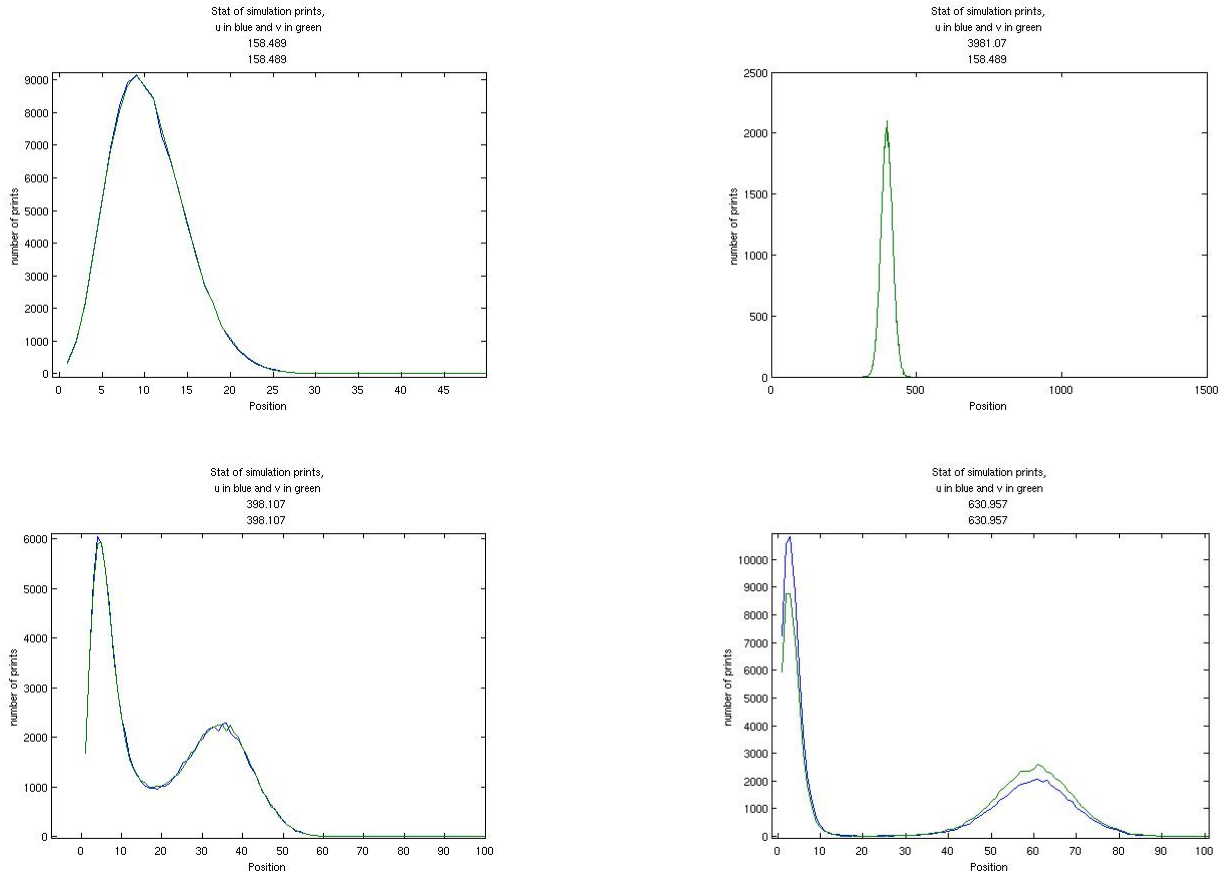


Fig. 16 These figures represent different scenarios with different sets of parameters. The upper number in the title is b's value and the lower is a's. Fig. 16a Top Left: Switching system with a low maximum equilibrium concentration. Fig. 16b Top Right: Non-switching system with a medium high maximum equilibrium concentration. Fig. 16c Bottom Left: Switching system with peak at 36. Fig. 16d Bottom Right: Switching system with peak at 63.

The system showed in Fig. 16b has the same properties as the one in Fig. 16a but has four to five times wider peak as the peak in Fig. 16a. In Fig. 16c and Fig. 16d we see two clear pictures of when we have entered a parameter region where the system will exhibit switching.

Analysing the behaviour by adjusting protein-DNA binding dynamics, altering spontaneous switching

We will show that statistics can be improved upon by changing parameters while retaining the equilibrium concentrations.

In order to improve the statistics (Fig. 5 and Fig. 13) we made further simulations. This time decreasing the binding/unbinding constants (k_2, k_3, k_6 ,k_7) by a factor of 100 while keeping c fixed at 10 and as it was anticipated it did have its effect (Fig. 17) as it was predicted in Fig. 3. Even though we did not do the same simulations for c at 100, we now predict that the same will happen if we do the same shift when c equals 100.

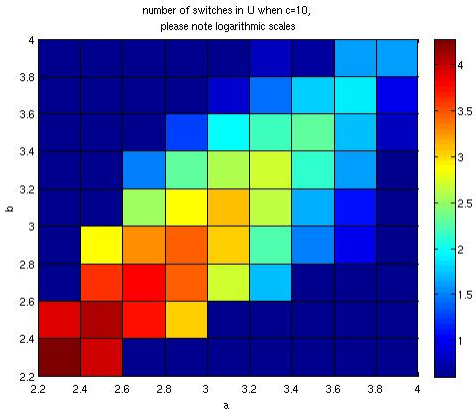


Fig. 17 Above: Number of switches in *u* when c is 10. The scales are logarithmic.

From these statistics, the maximum equilibrium concentrations were then mapped out in Fig 18.

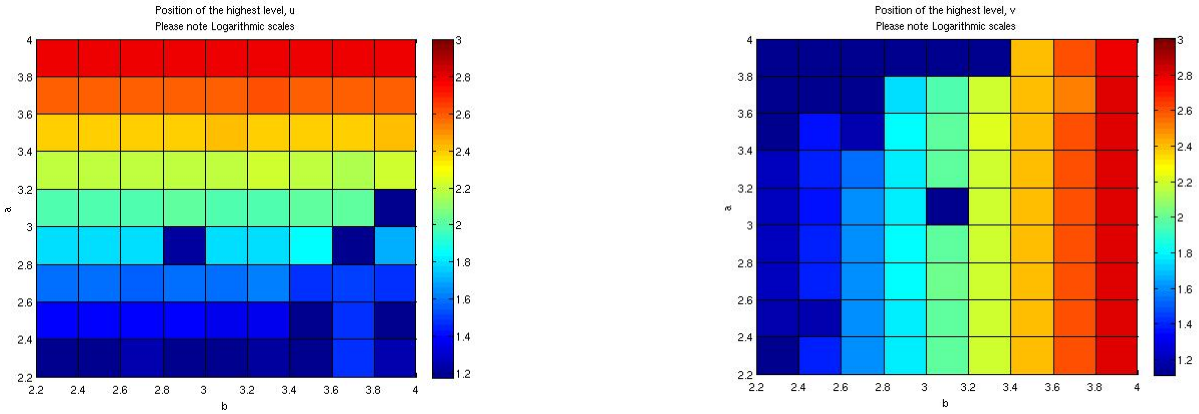


Fig. 18a Left: The position of u's high equilibrium concentration. Fig. 18b Right: The position of v's high equilibrium concentration. The scales are logarithmic.

From picture Fig. 17, it is clear that the largest amount of switches occurs when k_4 and k_8, id est a and b, have low values, but now the region with spontaneous switching has increased. Then the amount declines as we go upwards.

Adding an inducer to the system

Next we wanted to see if adding a switch inducer could drive the switching.
In Gardner et al.[2] it was shown that he could make a system switch by adding IPTG, an agent which acts as a switch inducer. We then explored the possibility of doing the same. Because it was not a system switching by random chance, there had to be modifications to the system. The system was altered by adding two species and two reactions.

$$u \xrightarrow{k_1} \emptyset \,, 2u + p_v \underset{k_3}{\overset{k_2}{\Leftrightarrow}} p_{bv} \,, p_u \xrightarrow{k_4} p_u + u$$

$$v \xrightarrow{k_5} \emptyset \,, 2v + p_u \underset{k_7}{\overset{k_6}{\Leftrightarrow}} p_{bu} \,, p_v \xrightarrow{k_8} p_v + v$$

$$\emptyset \xrightarrow{k_9} IPTG \,, u + IPTG \xrightarrow{k_{10}} uI_c$$

These two species were then given their own initial states. The first added reaction was a creation reaction which creates IPTG. The second added molecule was a complex

consisting of one IPTG and one u, represented by the last reaction. The idea was that the reaction will, with the aid of the inducer, start to bind *u* and the inducer into a complex. We used 1 parameter set for all 100 simulations (Fig. 19). Because of the chosen parameter set and the need to only run the simulations during a fraction of the normal total time, which before was 10^5, the simulations were fast. In this system, it was also vital to express a way to evaluate if and when the system had switched. Because the simulations were done with Gillespie, we chose specific time points. These were not the same for all simulations, meaning that we sometimes had to accept that the data point were only the nearest to the chosen time point. After that, the states were summed together to give a mean of the total amount of molecules.

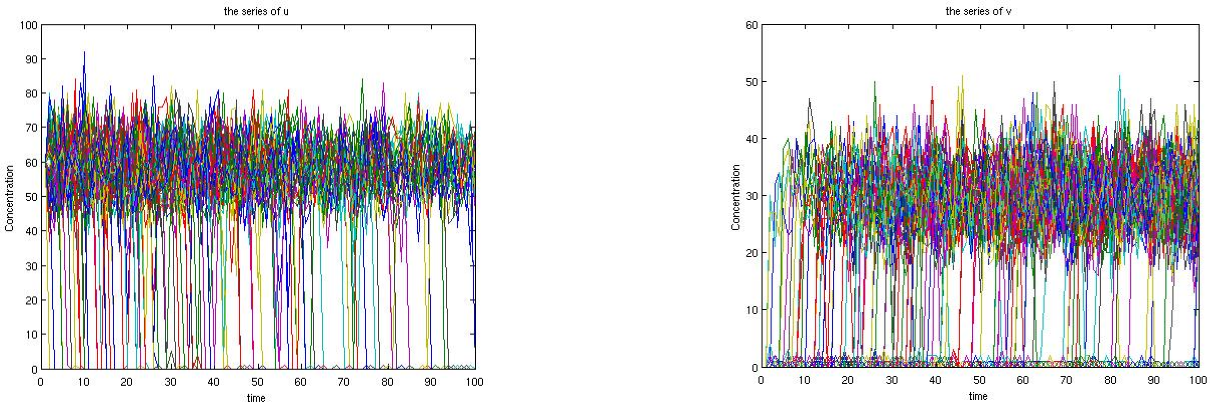


Fig. 19a Left: Time evolution of the total system in *u* with the parameters k_1=k_5=10, k_2=k_6=2, k_3=k_7=20, k_4=600, k_8=300, with k_9 as the creation of the IPTG with the value 13.5 and k_{10} as the new included mass action reaction with the value 10. Fig. 19b Right: The same but instead v.

We saw that, if *u* started in the high state, *u* began to decrease and that v, which started in the low state, grew in number of molecules. The number of molecules in IPTG grew proportional to time (Fig. 20).

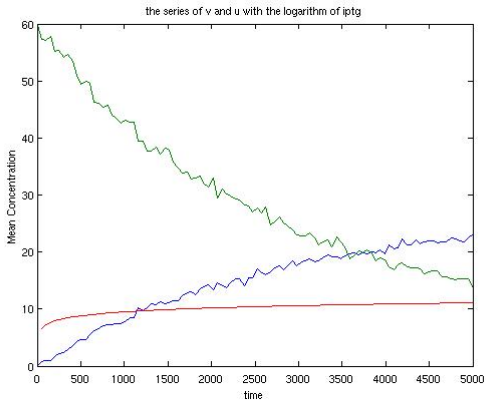


Fig. 20: Average time evolution from 100 simulations and with 100 steps between 0 and 5000 in every simulation. The parameters were the same as the above system. Green represents *u* and blue represents v. Red is the logarithm of the concentration of IPTG, the switch inducer.

Conclusions and discussions

We have implemented the Gillespie method and simulated a small stochastic system. When it comes to the performance of the Gillespie method, we found that for our system the Gillespie method is efficient and the implementation was straightforward. It should however be noted that while a majority of our simulations were fast (approx. 1min.), for some we saw an increase in simulation time more than an order of magnitude (20-25 min).

We could correlate this to large production rates which cause larger number of molecules and shorter time steps between each reaction. Interestingly we also saw a large decrease in simulation time when we lowered the binding and unbinding rates, although the number of molecules did not increase (cf. Figs. 7 and 18). This effect might be a limiting factor when one does simulations in larger systems. This can be circumvented by using methods which are derivatives of the Gillespie method, like the tau-leaping methods[6].

For our investigations we used a well known system of a bistable switch. Despite the simple construction of using only two repressors the system can have one or two stable states.

Initially we analysed the deterministic description of the system. This was used to map out parameter regions where the system shows a bistable behaviour, focusing on the change in production rates. We could see a trend of getting bi-stability with increased production rates (Fig. 1). This guided the choice of parameters in our further stochastic analysis.

We did stochastic simulations in parameter regions where the deterministic model exhibited bi-stability to investigate whether the stochasticity can lead to spontaneous switch behaviour. We could indeed see spontaneous switching, highlighting the differences between the stochastic and deterministic approach to simulations of systems involving bistable states. Still, when we analysed the frequency of switching we found large variations (Fig. 5). The region in which we saw a higher switch behaviour was limited to the region where the production rates are low but still high enough to get bi-stability. In biology spontaneous switching relates to the robustness of the decision making mechanism. A system that stays in one of the two states is considered to be robust. For example differentiation is often driven by a bistable switch and here it is important that a selected state is maintained[7], but for other systems spontaneous switches might be favourable.

To investigate the spontaneous switching further we studied the dependence of the number of spontaneous switches, on the parameter c which corresponds to the quotient between the unbinding and binding rates. Larger c means increased unbinding. We saw that increasing c lead to decreased number of spontaneous switches (Fig. 5 and Fig. 13). We relate this to increased number of molecules in the system with larger c which makes switching less likely.

Interestingly we found that changing the binding and unbinding rates while keeping the ratio (c) fixed also affects the switching behaviour of the system. Decreasing the binding and unbinding rates by factor 100 resulted in a larger region of parameter space in which switching was pronounced (Fig. 5 and Fig. 17). The consequence of a slow binding/unbinding rate is that a promoter gets locked in a repressing (or non-repressing) state for longer times, leading to larger deviations in protein production (and numbers of molecules). This locking behaviour hence promotes spontaneous switching. This hints that the internal dynamics of the system is important for its behaviour in the stochastic case. This change is not available in the deterministic Hill description and so the effect is not visible. This shows an increased possibility for analysing the bistable switch network in our stochastic simulations compared to if a Hill description is used.

In our simulations we saw that the stable equilibrium concentrations were dependent on a and b but not on c or k_2 and k_3 (cf. Fig. 7 and Fig. 14). Equation [iv] together with definitions

[v] show that we can eliminate the dependence of the system equations on the parameter c by rescaling the variables. The consequence of this is that the position of the stable fix-points depend solely on production and degradation rates but they do not depend on c or k_2, k_3.

Finally we investigated the behaviour of the system when switching was driven by an external signal. The signal inactivates one of the repressors. We found that it is indeed possible to use IPTG as a switch inducer. We found that if *u* had a high initial state the number of molecules decreased. We relate this to when *v* and the u-promoter (p_u) binds together to form the bound complex (p_{bu}). If this happened, the creation of *u* would be shut off and only degradation and the IPTG-complex reaction would affect the number of molecules in u. This allowed *v* to increase in number of molecules and switch. Our model had a slower switching response relative to the Gardner et al.[2] but we found a similar behaviour (Fig. 20).

Supplement and references

[1]: Daniel T. Gillespie, Exact Stochastic simulation of coupled chemical reactions, The Journal of Physical chemistry, 81, 2340-2361 (1977)

[2]: Timothy S. Gardner, Charles R. Cantor and James J. Collins, Construction of a genetic toggle switch in Escherchia Coli, Nature 403,339-342 (2000)

[3]:http://home.thep.lu.se/~henrik/software.html → http://dev.thep.lu.se/organism/
See Appendix part 2 for more details regarding contributors.

[4]: Ptashne, M. A. Genetic Switch: Phage λ and Higher Organisms,Cell Press, Cambridge, MA (1992)

[5]: Ishiura, M. et al. Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria. Science 281, 1519-1523 (1998)

[6]: David T. Gillespie, Stochastic simulation of Chemical Kinetics, Annu. Rev. Phys.Chem. 58:35-55 (2007)

[7]: Chickarmane V, Troein C, Nuber UA, Sauro HM, Peterson C.
PLoS Comput Biol. 2006 Sep 15;2(9):e123.

Appendix

1

[Tab. A1]:

A	B	C	D	max	Num	A	B	C	D	max	Num	A	B	C	D	max	Num	A	B	C	D	max	Num
885.87	885.87	10	10	7.54	3	6158.48	6158.48	12.92	16.68	28.13	3	2335.72	2335.72	10	27.83	6.96	3	6158.48	6158.48	16.68	21.54	16.11	3
1438.45	1438.45	10	10	13.66	3	6158.48	10000	12.92	16.68	27.26	3	3792.69	3792.69	10	27.83	12.86	3	10000	10000	16.68	21.54	27.22	3
2335.72	2335.72	10	10	22.93	3	10000	6158.48	12.92	16.68	46.32	3	6158.48	6158.48	10	27.83	21.68	3	3792.69	3792.69	16.68	27.83	4.19	3
2335.72	3792.69	10	10	22.08	3	10000	10000	12.92	16.68	46.14	3	6158.48	10000	10	27.83	20.76	3	6158.48	6158.48	16.68	27.83	11.87	3
3792.69	2335.72	10	10	37.83	3	2335.72	2335.72	12.92	21.54	6.37	3	10000	6158.48	10	27.83	35.84	3	10000	10000	16.68	27.83	20.75	3
3792.69	3792.69	10	10	37.67	3	3792.69	3792.69	12.92	21.54	12.61	3	10000	10000	10	27.83	35.66	3	6158.48	6158.48	16.68	35.94	8.26	3
3792.69	6158.48	10	10	37.21	3	6158.48	6158.48	12.92	21.54	21.54	3	2335.72	2335.72	10	35.94	4	3	10000	10000	16.68	35.94	15.62	3
3792.69	10000	10	10	35.78	3	6158.48	10000	12.92	21.54	20.25	3	3792.69	3792.69	10	35.94	9.51	3	10000	10000	16.68	46.42	11.46	3
6158.48	3792.69	10	10	61.53	3	10000	6158.48	12.92	21.54	35.81	3	6158.48	6158.48	10	35.94	16.54	3	10000	10000	16.68	59.95	7.89	3
6158.48	6158.48	10	10	61.43	3	10000	10000	12.92	21.54	35.58	3	6158.48	10000	10	35.94	15.06	3	2335.72	2335.72	21.54	10	8.23	3
6158.48	10000	10	10	61.15	3	3792.69	3792.69	12.92	27.83	9.15	3	10000	6158.48	10	35.94	27.7	3	3792.69	3792.69	21.54	10	16.29	3
10000	3792.69	10	10	99.99	3	6158.48	6158.48	12.92	27.83	16.35	3	10000	10000	10	35.94	27.47	3	6158.48	6158.48	21.54	10	27.82	3
10000	6158.48	10	10	99.97	3	6158.48	10000	12.92	27.83	13.8	3	3792.69	3792.69	10	46.42	6.68	3	6158.48	10000	21.54	10	26.15	3
10000	10000	10	10	99.9	3	10000	6158.48	12.92	27.83	27.66	3	6158.48	6158.48	10	46.42	12.47	3	10000	6158.48	21.54	10	46.25	3
885.87	885.87	10	12.92	4.76	3	10000	10000	12.92	27.83	27.36	3	10000	10000	10	46.42	21.07	3	10000	10000	21.54	10	45.95	3
1438.45	1438.45	10	12.92	10.16	3	3792.69	3792.69	12.92	35.94	6.03	3	3792.69	3792.69	10	59.95	3.25	3	3792.69	3792.69	21.54	12.92	11.81	3
2335.72	2335.72	10	12.92	17.52	3	6158.48	6158.48	12.92	35.94	12.22	3	6158.48	6158.48	10	59.95	9.19	3	6158.48	6158.48	21.54	12.92	21.12	3
2335.72	3792.69	10	12.92	16.21	3	10000	10000	12.92	35.94	20.93	3	10000	10000	10	59.95	16.06	3	6158.48	10000	21.54	12.92	17.83	3
3792.69	2335.72	10	12.92	29.24	3	6158.48	6158.48	12.92	46.42	8.81	3	6158.48	6158.48	10	77.43	6.39	3	10000	6158.48	21.54	12.92	35.72	3
3792.69	3792.69	10	12.92	29.03	3	10000	10000	12.92	46.42	15.87	3	10000	10000	10	77.43	12.09	3	10000	10000	21.54	12.92	35.33	3
3792.69	6158.48	10	12.92	28.4	3	6158.48	6158.48	12.92	59.95	5.69	3	10000	10000	10	100	8.88	3	3792.69	3792.69	21.54	16.68	7.79	3
3792.69	10000	10	12.92	26.09	3	10000	10000	12.92	59.95	11.83	3	1438.45	1438.45	12.92	10	9.83	3	6158.48	6158.48	21.54	16.68	15.78	3
6158.48	3792.69	10	12.92	47.61	3	10000	10000	12.92	77.43	8.48	3	2335.72	2335.72	12.92	10	17.34	3	10000	10000	21.54	16.68	27.03	3
6158.48	6158.48	10	12.92	47.48	3	10000	10000	12.92	100	5.32	3	2335.72	3792.69	12.92	10	15.29	3	6158.48	6158.48	21.54	21.54	11.38	3
6158.48	10000	10	12.92	47.12	3	1438.45	1438.45	16.68	10	5.7	3	3792.69	2335.72	12.92	10	29.21	3	10000	10000	21.54	21.54	20.5	3
10000	3792.69	10	12.92	77.41	3	2335.72	2335.72	16.68	10	12.69	3	3792.69	3792.69	12.92	10	28.92	3	6158.48	6158.48	21.54	27.83	7.34	3
10000	6158.48	10	12.92	77.38	3	3792.69	3792.69	16.68	10	21.98	3	3792.69	6158.48	12.92	10	28.09	3	10000	10000	21.54	27.83	15.28	3
10000	10000	10	12.92	77.3	3	3792.69	6158.48	16.68	10	20.19	3	3792.69	10000	12.92	10	24.03	3	10000	10000	21.54	35.94	10.95	3
1438.45	1438.45	10	16.68	7.25	3	6158.48	3792.69	16.68	10	36.75	3	6158.48	3792.69	12.92	10	47.59	3	10000	10000	21.54	46.42	6.86	3
2335.72	2335.72	10	16.68	13.25	3	6158.48	6158.48	16.68	10	36.47	3	6158.48	6158.48	12.92	10	47.42	3	3792.69	3792.69	27.83	10	11.14	3
3792.69	3792.69	10	16.68	22.29	3	6158.48	10000	16.68	10	35.63	3	6158.48	10000	12.92	10	46.95	3	6158.48	6158.48	27.83	10	20.8	3
3792.69	6158.48	10	16.68	21.41	3	10000	6158.48	16.68	10	59.85	3	10000	3792.69	12.92	10	77.41	3	10000	10000	27.83	10	35.15	3
6158.48	3792.69	10	16.68	36.82	3	10000	10000	16.68	10	59.67	3	10000	6158.48	12.92	10	77.37	3	3792.69	3792.69	27.83	12.92	5.42	3
6158.48	6158.48	10	16.68	36.65	3	2335.72	2335.72	16.68	12.92	8.99	3	10000	10000	12.92	10	77.26	3	6158.48	6158.48	27.83	12.92	15.33	3
6158.48	10000	10	16.68	36.18	3	3792.69	3792.69	16.68	12.92	16.6	3	1438.45	1438.45	12.92	12.92	6.7	3	10000	10000	27.83	12.92	26.79	3
10000	6158.48	10	16.68	59.89	3	6158.48	6158.48	16.68	12.92	27.99	3	2335.72	2335.72	12.92	12.92	13.01	3	6158.48	6158.48	27.83	16.68	10.66	3
10000	10000	10	16.68	59.79	3	6158.48	10000	16.68	12.92	26.81	3	3792.69	3792.69	12.92	12.92	22.16	3	10000	10000	27.83	16.68	20.17	3
1438.45	1438.45	10	21.54	4.41	3	10000	6158.48	16.68	12.92	46.29	3	3792.69	6158.48	12.92	12.92	20.93	3	10000	10000	27.83	21.54	14.81	3
2335.72	2335.72	10	21.54	9.83	3	10000	10000	16.68	12.92	46.06	3	6158.48	3792.69	12.92	12.92	36.79	3	10000	10000	27.83	27.83	10.19	3
3792.69	3792.69	10	21.54	17.02	3	2335.72	2335.72	16.68	16.68	5.17	3	6158.48	6158.48	12.92	12.92	36.57	3	6158.48	6158.48	35.94	10	14.69	3
3792.69	6158.48	10	21.54	15.63	3	3792.69	3792.69	16.68	16.68	12.28	3	6158.48	10000	12.92	12.92	35.94	3	10000	10000	35.94	10	26.47	3
6158.48	3792.69	10	21.54	28.46	3	6158.48	6158.48	16.68	16.68	21.36	3	10000	6158.48	12.92	12.92	59.87	3	6158.48	6158.48	35.94	12.92	9.48	3
6158.48	6158.48	10	21.54	28.24	3	6158.48	10000	16.68	16.68	19.45	3	10000	10000	12.92	12.92	59.74	3	10000	10000	35.94	12.92	19.73	3
6158.48	10000	10	21.54	27.59	3	10000	6158.48	16.68	16.68	35.77	3	2335.72	2335.72	12.92	16.68	9.48	3	10000	10000	35.94	16.68	14.14	3
10000	6158.48	10	21.54	46.34	3	10000	10000	16.68	16.68	35.47	3	3792.69	3792.69	12.92	16.68	16.84	3	10000	10000	35.94	21.54	8.86	3
10000	10000	10	21.54	46.2	3	3792.69	3792.69	16.68	21.54	8.62	3	3792.69	6158.48	12.92	16.68	14.58	3	10000	10000	46.42	10	19.12	3
												6158.48	3792.69	12.92	16.68	28.42	3	10000	10000	46.42	12.92	13.16	3
																		10000	10000	59.95	10	11.45	3

Fig. Tabulated sets of parameters C is quotient between k_2 and k_3 or k_6 and k, D is the degradation or k_1 and k_5 A and B are C times k_4 or k_8. Max is where the highest equilibrium concentration is sitting and Num is the amount of equilibrium concentrations for that set of parameters.

17/8-11

2

Project Organism is being run and maintained by Henrik Jönsson Ass. Prof., Jérémy Gruel (post doc), Pawel Krupinski (post doc), Behruz Bozorg (PhD student), Patrik Fortes (bachelor student) with Pontus Melke (PhD student) and Patrik Sahlin (PhD student) as former contributors. The project is written in C++ with library support from Boost. 17/8-11