Contents

1 Effect of feedback loops on switch robustness

1.1 Introduction

In this chapter I examine whether adding feedback loops to the genetic toggle switch increases its robustness to parameter fluctuations. To do this, I use ABC SMC to estimate the parameter values that allow the toggle switch model to behave like a switch. I then study the effect that adding feedback loops has to the toggle switch bistability, and finally I use model selection to select the most robust switch model out of the ones considered.

Structurally this chapter is organised as follows: In the first section I examine the genetic toggle switch with no added feedback loops. I use a parameter scan to find the parameter values that make it bistable and then use ABC SMC for parameter inference for a switch-like behaviour. In the following section I examine the effect that the addition of feedback loops to the genetic toggle switch has on its stability, and select the switch architectures that are capable of bistable behaviour. Finally, I use ABC-SysBio model selection to select the most robust model out of the bistable switches.

1.2 The bistable genetic toggle switch

Gardner, Cantor, & Collins (2000) constructed the first synthetic genetic toggle switch. Their model consisted of two mutually repressing transcription factors, and is defined by the following ODEs:

$$\frac{du}{dt} = \frac{a_1}{1 + v^\beta} - u \tag{1.1}$$

$$\frac{dv}{dt} = \frac{a_2}{1 + u^\gamma} - v,$$ (1.2)

where u is the concentration of repressor 1, v the concentration of repressor 2, a_1 and a_2 denote the effective rates of synthesis of repressors 1 and 2 respectively, β is the cooperativity of repression of promoter 1 and γ of repressor 2. This model is capable of bistable behaviour when a_1 and a_2 are balanced and when β, γ are > 1 (Gardner, Cantor, & Collins 2000).

1.2.1   Bifurcation analysis

First, I study the model given in Equations 1.1- 1.2 by conducting a bifurcation analysis in order to confirm that it is capable of bistable behaviour. A bifurcation analysis is used to determine the properties of a system in parameter space (Alon 2007). Here I used the PyDSTool (Clewley 2012), a python package used for the analysis of dynamical systems.

The parameters chosen here for the bifurcation analysis are within the range suggested by Gardner, Cantor, & Collins (2000). a_1 and a_2 are set to 10, and β, γ set to 2. A vector plot shows that the system has two steady states as shown in Figure 1.1B. Both states were found to be stable by examining the eigenvalues of the system at each steady state using Mathematica (Mathematica 2016).

I further study the system by conducting a bifurcation analysis, where all parameters remain constant to the values shown above, and only one parameter (a_1) is varied. The bifurcation analysis shows that by varying the parameter for the effective rate of synthesis of repressor 1 while all other parameters remain constant, the system is bistable when $5 \geq a_1 \leq 31$.
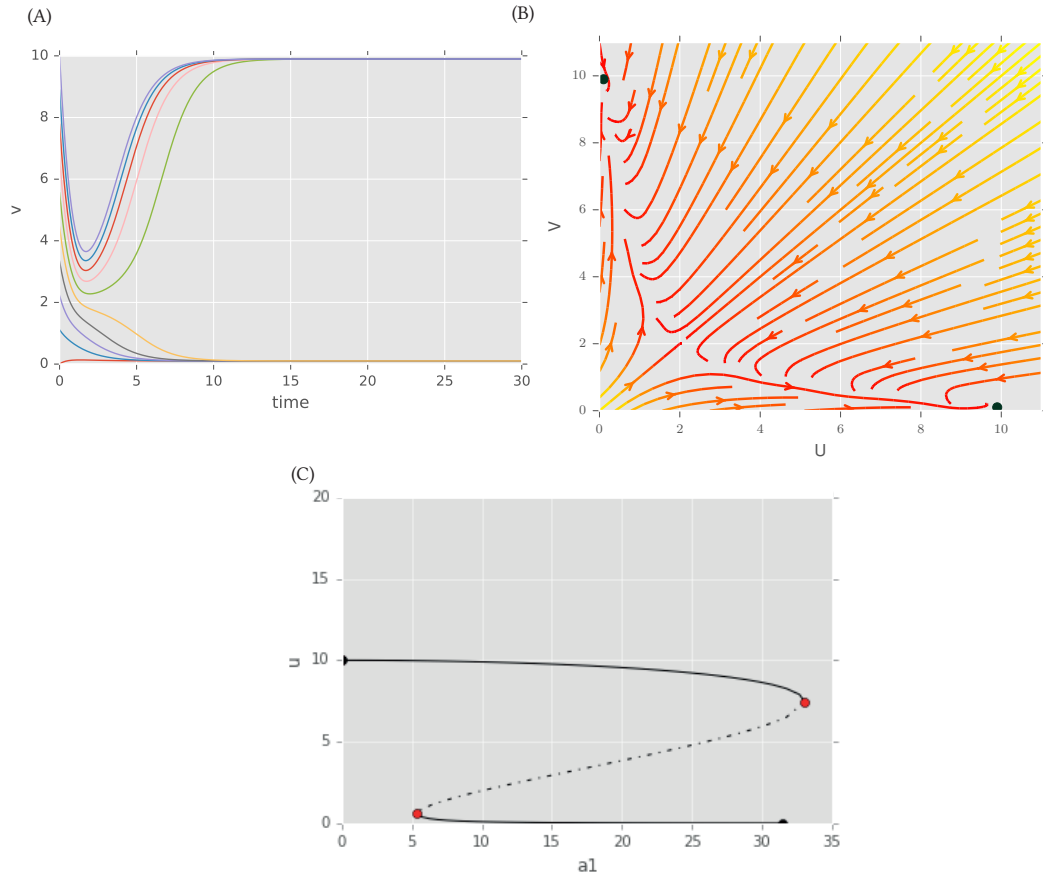
Figure 1.1   : The Gardner, Cantor, & Collins (2000) toggle switch is capable of bistable behaviour given a_1 and $a_2 = 10$, and β, $\gamma = 2$. (A) The timecourse of the simulated model using multiple initial conditions. (B) The vector plot of the Gardner switch shows there are two stable steady states and one unstable steady state. (C) A bifurcation diagram shows that the system is bistable when $5 \geq a_1 \leq 31$.
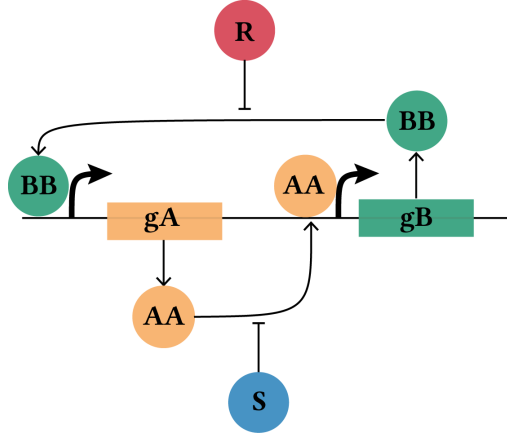
Figure 1.2   : An illustration of the model used in the parameter scan.

1.3   Designing a simple synthetic switch

The model used in Section 1.2.1 could be solved analytically and I demonstrated that it is bistable for the parameters given. Nevertheless, for a switch to be useful in synthetic biological applications, it must be capable of behaving like a switch over a range of parameter values, as these fluctuate within the cellular environment. Therefore, in this section I study the parameter ranges that can give rise to a bistable switch. This indicates whether the bistability of the switch models is robust to small parameter fluctuations.

In order to study the switch system in a more realistic way, I developed an extension to the Gardner, Cantor, & Collins (2000) switch. This new set of switches does not use the quasi-steady state approximation (QSSA) that is often used in modelling the toggle switch. Using mass action, this changes the two-equation system used in Equations 1.1 into a system of 18 equations. The equations describing the system are shown below and illustrated in Figure 1.2. The ODEs are given in Apeendix(XXX). The system consists of two genes, gA and gB. The products of the genes homodimerise and mutually repress each other. A symmetric model, where the parameters for equivalent reactions are set to be the same, was used for simplicity.

1.3.1   Parameter scan for model stability

In order to determine the range of parameter values for which the above model is bistable, I developed a parameter scanning algorithm. The algorithm is outlined in Algorithm 1 below. This method involves the scan of parameter values as well as

Table 1.1   Simple mass action switch

Equation	Description
$gA \xrightarrow{ge} gA + A$ $gB \xrightarrow{ge} gB + B$	gene expression
$A + A \xrightarrow{dim} A2$ $B + B \xrightarrow{dim} B2$	dimerization
$A2 \xrightarrow{dim\_r} A + A$ $B2 \xrightarrow{dim\_r} B + B$	monomerization
$gA + B2 \xrightarrow{rep} B2gA$ $gB + A2 \xrightarrow{rep} A2gB$	repression
$B2gA \xrightarrow{rep\_r} B + gA$ $A2gB \xrightarrow{rep\_r} A2 + gB$	dissociation
$A \xrightarrow{deg} \emptyset$ $B \xrightarrow{deg} \emptyset$	degradation

initial conditions for AA and BB. Parameter values are sampled randomly from a uniform distribution. Here, each set of samples will be referred to as a particle. For each particle, latin hypercube sampling is used to sample initial conditions (McKay, Beckman, & Conover 2000). The uniform priors of the two species in consideration represent a rectangle space, which is subdivided into equal parts. Then a random sample is drawn from each sub-part, as illustrated in Figure 1.3. This is used to ensure that the whole space is sampled uniformly. Latin hypercube sampling is done in two dimensions, in order to sample initial conditions for the two dimers, AA and BB.

Algorithm 1 Parameter scan algorithm

1: Select multiple sets of parameter values from a random uniform distribution between 0 and 10.
2: **for** each set of parameter values **do**:
3:     Select multiple sets of initial conditions for the two dimers via latin hyper-cube sampling
4:     **for** each set of initial conditions **do**
5:         Integrate ODEs of the model
6:         Solve system to steady state
7:     **end for**
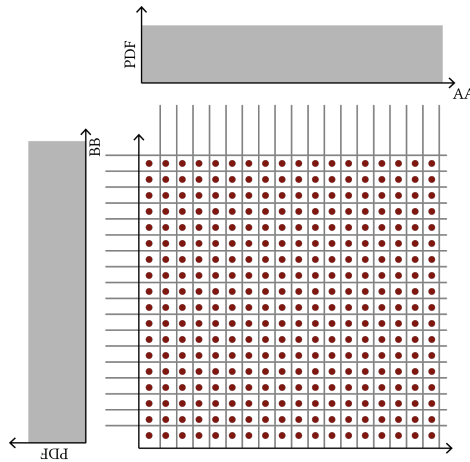8: **end for**

Figure 1.3 Latin hypercube sampling ensures that the whole space is sampled evenly. For the two species concerned, A and B, we assume uniform distributions, shown in grey. The joint space of the two distributions is divided into smaller equal parts and a random sample is drawn from within each subspace.

After each particle was solved to steady state, the value of each dimer at the last timepoint was taken, for all initial conditions sampled. These were used to make a phase diagram for each particle, which consists of the steady state value of one dimer plotted against the other. The parameter scan uncovered the presence of tristable, bistable and monostable systems given a different set of parameter values. An example of the phase plots for each of the three types of stabilities found during the scan are shown in Figure 1.4.
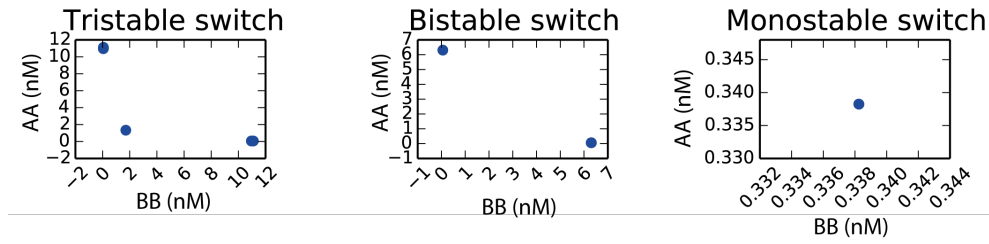

Figure 1.4   : An example for each type of switch found during the parameter scan. Each graph represents the steady state values of one dimer plotted against the other, from one parameter set and 100 initial conditions.

There were a total of 217 monostable, 16 bistable and 6 tristable switches out of 400 sampled parameter sets. The remaining samples did not reach steady state. The parameter values that produced each stability are shown in Figure 1.5.

From the histograms in Figure 1.5 it can be seen that the parameter for gene expression (ge) tends to be relatively high when bistability arises whereas the parameter for the reverse reaction of repression (rep_r) tends to be low. Degradation (deg) also tends to be low. From this analysis I showed that the toggle switch is capable of bistable behaviour for a range of parameter values.
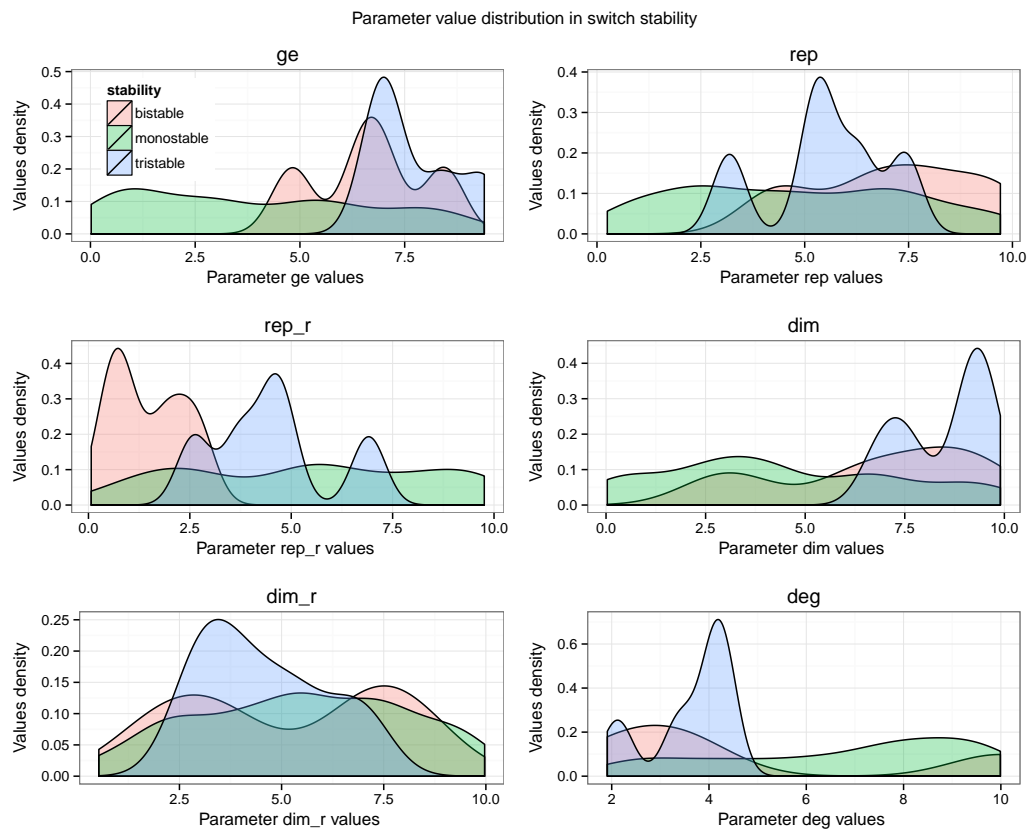
Figure 1.5 The distribution of parameter values that resulted in monostable, bistable and tristable switches in the parameter scan. Each graph represents the distribution of the values of one parameter.

1.3.2 Toggle switch parameter inference

In this section, I extend the analysis carried out in Section 1.3.1 to study the problem of how to rationally design a synthetic biological system to perform a behaviour of choice. In order to address this question I use a Bayesian approach, known as Approximate Bayesian Computation and described in Section(XXX), implemented in a software package, ABC-SysBio (Liepe et al. 2010).

This approach is capable of approximating the posterior distribution that gives rise to the behaviour of choice (Toni et al. 2009). By simulating the model in question, this approach can identify an approximate posterior distribution via a series of intermediate distributions. This method can be used for the rational design of synthetic biological systems by defining some design objectives to which the model is fitted to (Barnes et al. 2011). By specifying the inputs to the system and the outputs required, the posterior of the model that can produce this behaviour can be identified.

Here I use ABC-SysBio to fit a model to the design objectives of a switch-like behaviour. I extend the model used in Section 1.3.1 by adding two inducers to the system, S and R. S removes the A homodimer (AA) from the system by binding to it thus removing the repression on gene B. R removes BB from the system with the same mechanism. These inducers represent the stimuli that will turn the switch ON and OFF in a biological setting. The following equations are added to the existing set of equations for the mass action switch:

The range of parameter values shown to produce a bistable switch in Section 1.3.1 were used as priors and are shown in Table 1.3.

Table 1.2   Toggle switch inducer equations

Equation	Description
$S + A2 \overset{\text{rep\_dim}}{\rightarrow} SA2$ $R + B2 \overset{\text{rep\_dim}}{\rightarrow} RB2$	Inducer repression
$SA2 \overset{\text{rep\_dim\_r}}{\rightarrow} S + A2$ $RB2 \overset{\text{rep\_dim\_r}}{\rightarrow} R + B2$	Inducer dissociation
$R \overset{\text{deg}}{\rightarrow} \varnothing$ $S \overset{\text{deg}}{\rightarrow} \varnothing$	Inducer degradation

Table 1.3   The prior distributions used for the standard toggle switch parameter inference. The values indicate the lower and upper limits of a uniform distribution.

ge	rep	rep_r	dim	dim_r	deg	rep dim	rep_dim_r	deg_sr
6-9	4-10	1-4	4-10	2-7	2-5	0.05-0.1	0.01-0.05	0.01-0.05

1.3.3   Design specifications

The following were defined as the design specifications for a bistable genetic toggle switch. The two inducers, S and R, are used as inputs to flip the switch ON and OFF respectively. The required output is the switch flipping from the OFF state to the ON state and then to the OFF state again (Figure 1.6).
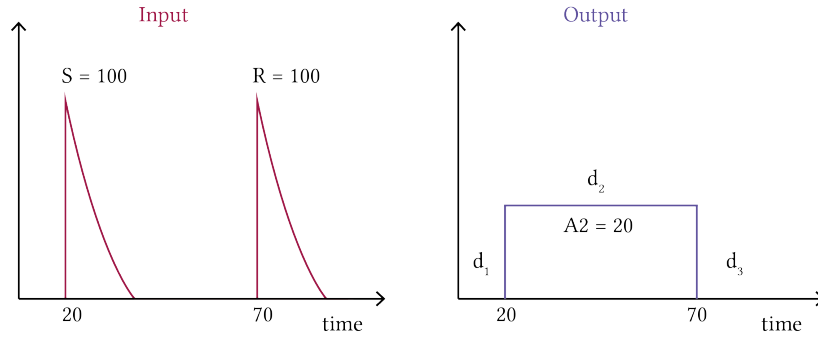

Figure 1.6   : Design specification for ABC parameter inference. The input to the system consists of an event turning on the stimulus (S) at t=20 and another turning on the repressor (R) at t=70. The output specification is the response to the switch to these stimuli.

1.3.3.1   Distance function

In order to fit the switch model to the behaviour specified above, a distance function must be defined. The distance function defines the quantity that is minimized at each successive iteration of ABC SMC. Three distances are measured, one for each state of the switch, OFF-ON-OFF. Each distance is the sum of the distances of the simulated protein levels to the desired protein levels at each timepoint. All three distances must reach the minimum threshold for the process to be complete.

$$d_1 = \sum_{i=0}^{20} (s_i - t_1)^2 \tag{1.3}$$

$$d_2 = \sum_{i=21}^{70} (s_i - t_2)^2 \tag{1.4}$$

$$d_3 = \sum_{i=71}^{100} (s_i - t_3)^2, \tag{1.5}$$

where i represents the timepoints, s_i the simulation result at each timepoint and $t_1 = 0, t_2 = 20, t_3 = 0$ represent each target behaviour. t_1, t_2, t_3 were set to 0, 20, 0 respectively.

1.3.4 Results

The results of the parameter inference of the toggle switch are shown in Figure 1.7. The model was shown to successfully behave like a switch within the parameter range used here. The resulting timecourse of the last population matches the design specifications. It can be seen from the posterior distribution that gene expression (ge) must be high relative to the prior. Repression (rep) and degradation must both be low and the rate of dimerisation (dim) must be high relative to the prior. The posterior constitutes the specifications that can be used when building a synthetic switch in the lab, as the appropriate components can be tweaked for a successful circuit. More generally, the posterior distribution shows that the parameter space that can give rise to this behaviour is limited. A very small portion of the prior is capable of producing the desired design specifications. In the next section I will examine wether the addition of feedback loops can increase the volume of the posterior that produces the required behaviour.
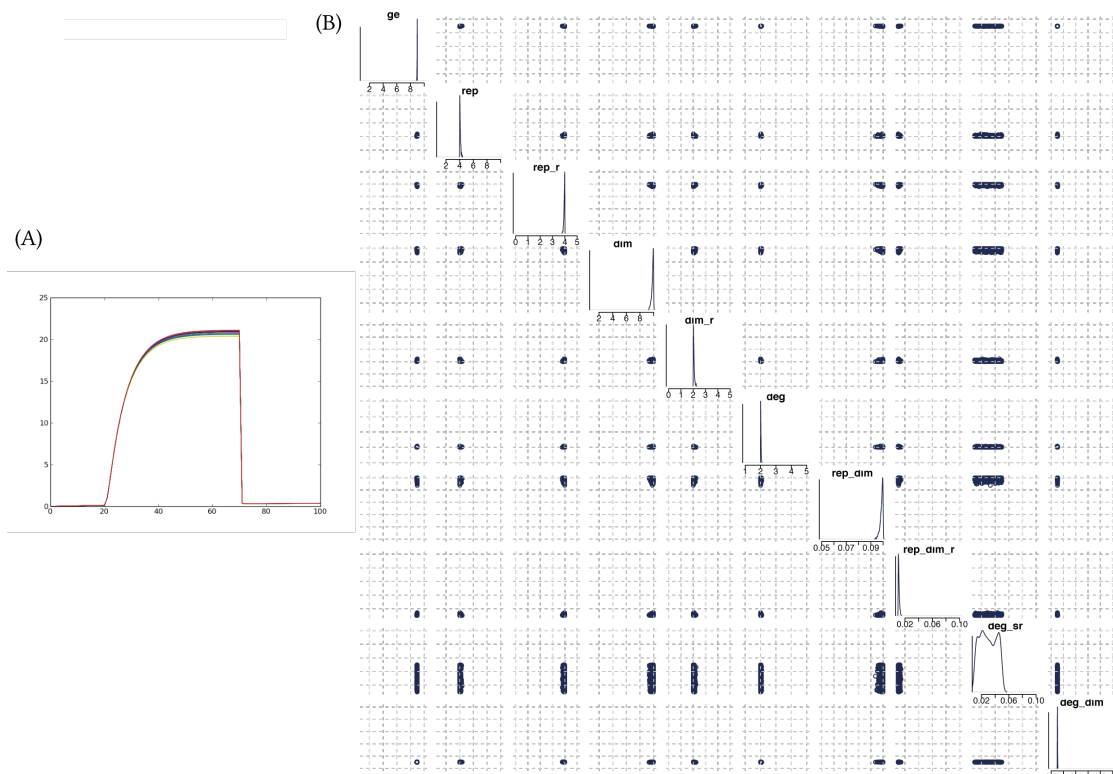
Figure 1.7 (A) The time series of the final population (for final $\epsilon = 2$) of the standard toggle switch ABC-SMC parameter inference. The stimulus, that represses A2, is added at t=20 and the repressor, that represses B2 is added at t=70. (B) The posterior distribution of the toggle switch.

1.4   Designing a more robust genetic toggle switch

In this section I examine whether the addition of feedback loops to the toggle switch can increase its robustness to parameter fluctuations. Here I define a robust system as a device that can withstand fluctuations in parameter values and still produce the desired behaviour (parametric robustness). Feedback loops are well known key regulatory motifs (Brandman et al. 2005).

As was shown in Section 1.3, the posterior distribution of the simple toggle switch was narrow, and thus the behaviour of choice will not be robust. Here I examine whether adding feedback loops to the genetic toggle switch can increase parametric robustness for the desired design specifications.

1.4.1   Models of the genetic toggle switch

Both positive and negative feedback loops were considered. Therefore 7 models were examined for their capability to behave like a switch. The simple toggle switch, switches with positive autoregulation in either or both nodes and switches with negative autoregulation in either or both nodes. The models considered are illustrated in Figure 1.8.

In order to study each model analytically I built extensions to the (Gardner, Cantor, & Collins 2000) toggle switch in order to incorporate positive and negative feedback to the system. The models for the double autoregulation models are shown below. For the single autoregulation models, the unnecessary autoregulation term is set to 0.

Double positive autoregulation:

$$\frac{du}{dt} = \frac{a_1}{1 + v^\beta + k_u^{\delta_u}} - u \tag{1.6}$$

$$\frac{dv}{dt} = \frac{a_2}{1 + u^\gamma + k_v^{\delta_v}} - v, \tag{1.7}$$

Double negative autoregulation:

$$\frac{du}{dt} = \frac{a_1 + k_u^{\delta_u}}{1 + v^\beta} - u \tag{1.8}$$

$$\frac{dv}{dt} = \frac{a_2 + k_v^{\delta_v}}{1 + u^\gamma} - v, \tag{1.9}$$

where k represents the effective binding of the transcription factor to its own promoter and γ represents the polymerisation of the bound transcription factor.
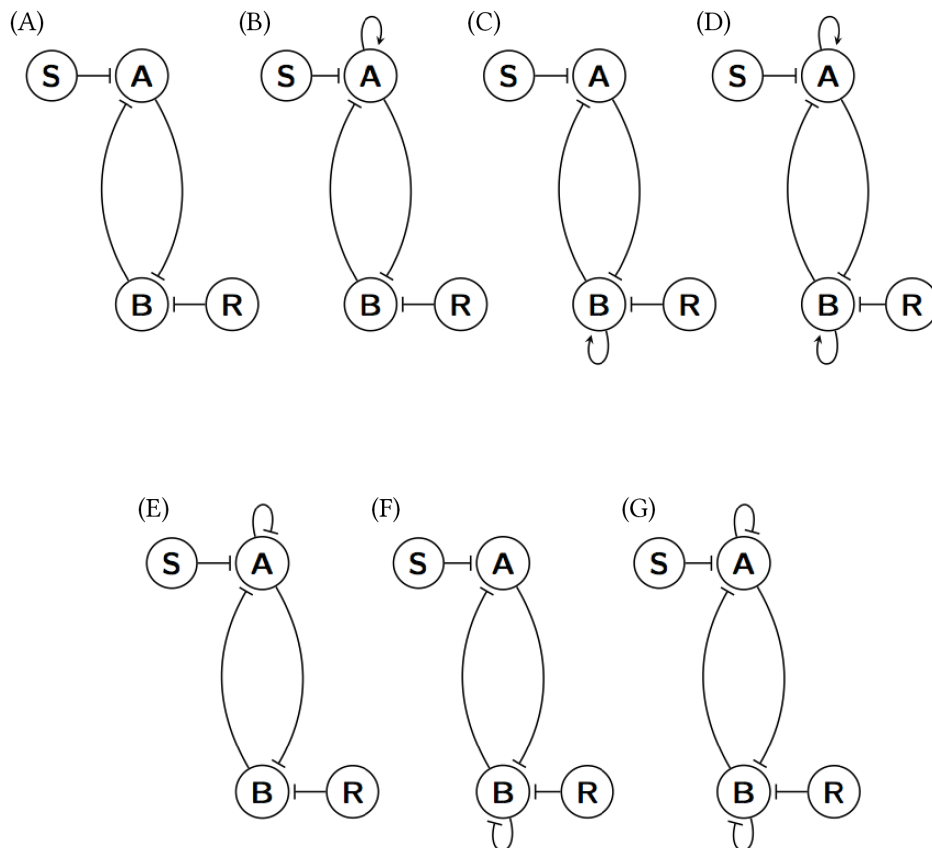
Figure 1.8   : Toggle switch designs considered for model selection. 7 models were used: The simple toggle switch (A), the switches with positive autoregulation on either or both nodes (B, C, D) and the switches with negative autoregulation on either or both nodes (E, F, G)

1.4.1.1 Bifurcation analyses

In order to use these models for model selection, it must first be determined whether they are capable of behaving like a switch. ABC SMC model selection is used to select models that can produce the same behaviour, over a greater range of parameter values. If a model is not capable of producing the desired behaviour for the prior range, then it will not be used for model selection.

I used the PyDSTool (Clewley 2012) in order to determine whether each of the 7 switches is capable of bistable behaviour. The same analysis as Section 1.2.1 was used to identify the steady states of the systems and their stabilities. As shown in Figure 1.9, both single and double positive autoregulation is consistent with bistable behaviour. Two stable states were found for both cases when $k = 2$ and $\delta = 1$.
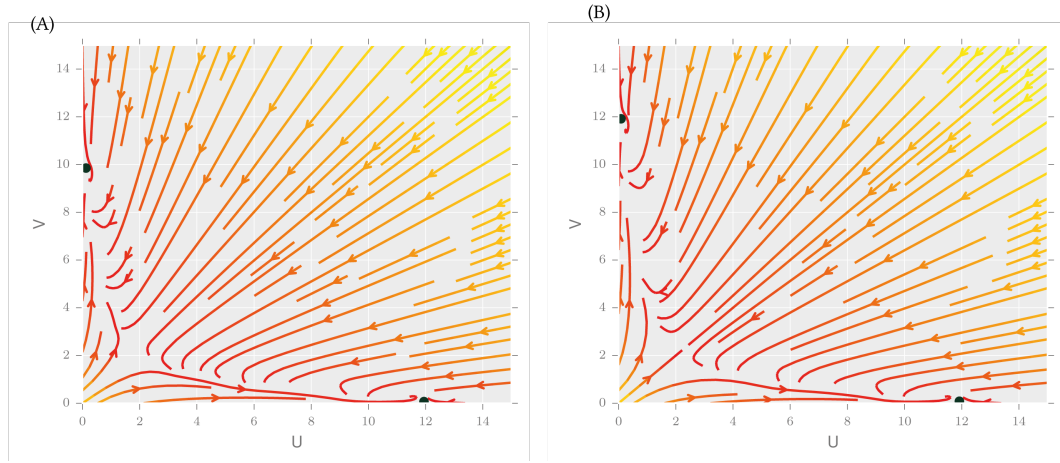

Figure 1.9 : Both single and positive autoregulation is consistent with bistable behaviour. (A) Single positive autoregulation. Two stable steady states are found in the single positive autoregilation model at (u, v) = (0.069, 11.94) and (u, v) = (9.85,0.122). (B) Double positive autoregulation. The model with double positive autoregulation has two stable steady states at (u, v)=(0.122, 9.85) and (u,v)=(11.94, 0.069).

On the other hand, negative autoregulation was not consistent with bistable behaviour for the parameter values used here. The vector plot of the switch with single negative autoregulation show that there is one stable steady state when the levels of the unregulated protein are high and the levels of the negatively regulated protein low. The vector plot for the switch with double negative autoregulation shows one stable steady state when the levels of both proteins are low.

Negative autoregulation occurs when the protein binds to its own promoter and

stops further production. This means that the protein levels rise to a lower level than what would be expected for an unregulated system (Alon 2007). This means that, in the case of the double negative autoregulation, the concentration levels of either protein are not sufficient to dominate over the other and create bistability. In the case of the single negative autoregulation, only the protein without autoregulation is able to reach sufficient levels to dominate the system, and thus the only steady state of the system is when the unregulated protein is high.
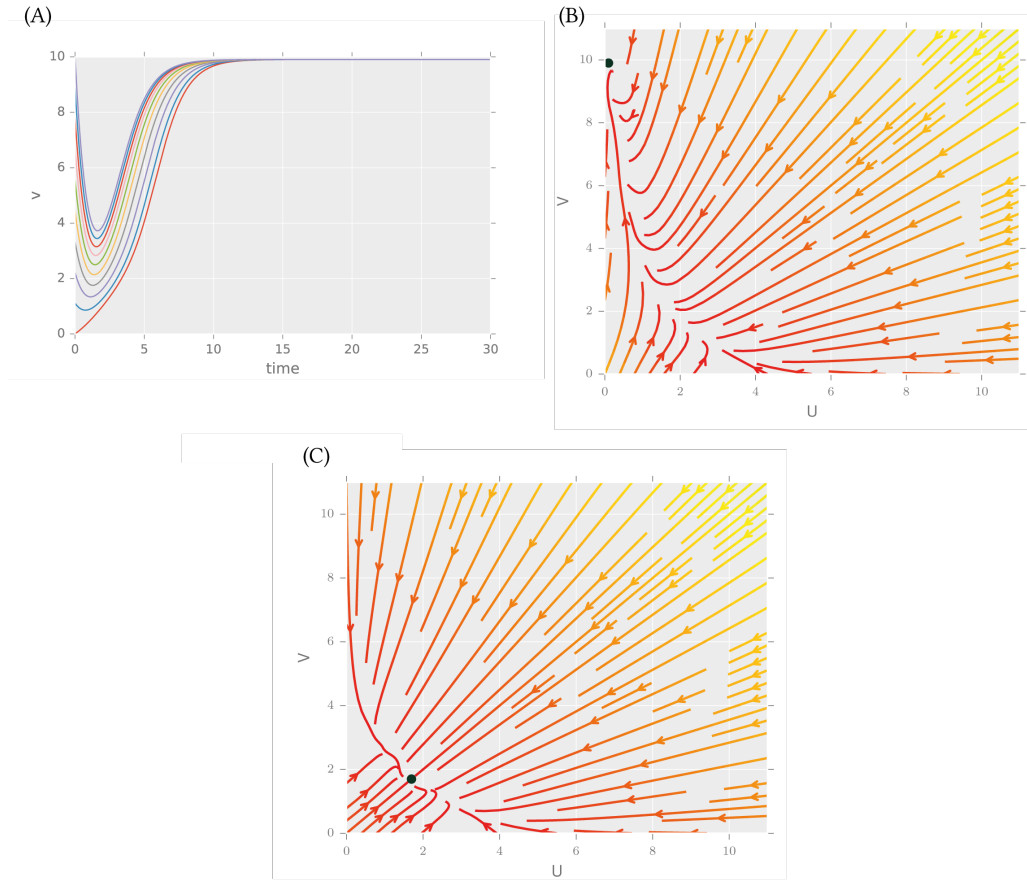

Figure 1.10   :(A) Timecourse of the single negative autoregulation switch, when k_u=2 with varying initial conditions. The system is monostable. (B) One stable steady state is found for the single negative autoregulation switch when k_u=2 at u=0.099 and v=9.903. (C) The vector plot of the switch with double negative autoregulation has one steady state at u=1.699 and v=1.699.

The models shown here to be capable of bistable behaviour will be used in the following section for model selection, to determine whether the addition of positive feedback loops increases the robustness of the system to parameter fluctuations.

1.4.2   ABC for model selection

Bayesian model selection is used to directly compare the competing designs using ABC-SysBio (Liepe et al. 2010). This method enables the simultaneous inference of the kinetic parameters and the structure of the models in order to select the model producing the same behaviour over a greater parameter range. ABC SMC can be used for model selection by adding the model as a parameter to the selection process. The algorithm samples models as well as parameters at each iteration. When the last ε is reached, the algorithm will have concluded to a posterior distribution of parameters for each model, a subset of the prior distribution that can give the best rise to the data. The model that performs over a greater posterior parameter range is selected as the most robust (Toni et al. 2009). The process is outlined in Figure 1.11.
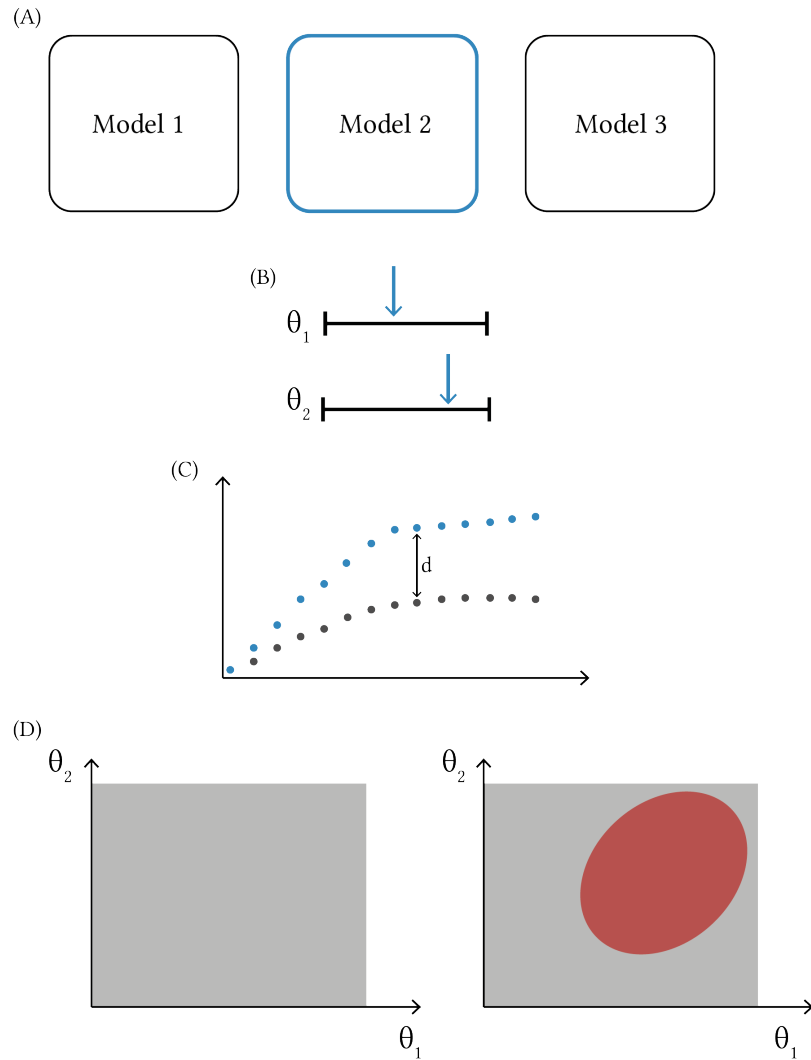
Figure 1.11 : ABC SMC model selection. (A) A model is sampled, and the (B) parameter values for that model are sampled. The system is simulated and the result compared to the desired behaviour. (C) If the distance between the two is smaller than the current accepted threshold, then the samples are retained. If not they are discarded. (D) This method can select the model that can produced the desired behaviour over a greater parameter range (red ellipse).

In order to select for the more robust toggle switch, the standard toggle switch was compared to switches with positive autoregulation in one or both nodes, which were shown to be capable of bistable behaviour in Section 1.4.1.1. The mass action models were used for model selection, in order to represent the system in a more realistic way. The equations shown in Table 1.4 are added to the equations of the simple toggle switch used in Section 1.3.2. Toggle switches with autoregulation on both nodes have both sets of equations added. The same design specifications as used in Section 1.3.2 were used.

Table 1.4   Autoregulated switches additional equations

Equations	Description
Positive autoregulation A	
$A2 + gA \xrightarrow{aut\_1} A2gA$	dimer self-association
$A2gA \xrightarrow{aut\_2} A + A2gA$	self-induced expression
$A2gA \xrightarrow{aut\_3} A2 + gA$	dimer self-dissociation
Positive autoregulation B	
$B2 + gB \xrightarrow{aut\_1} B2gB$	dimer self-association
$B2gB \xrightarrow{aut\_2} B + B2gB$	self-induced expression
$B2gB \xrightarrow{aut\_3} B2 + gB$	dimer self-dissociation

Given the models shown above and the parameter priors shown in Table 1.5, ABC SMC model selection was carried out.

Table 1.5   The prior distributions used for model selection. The values indicate the lower and upper limits of a uniform distribution.

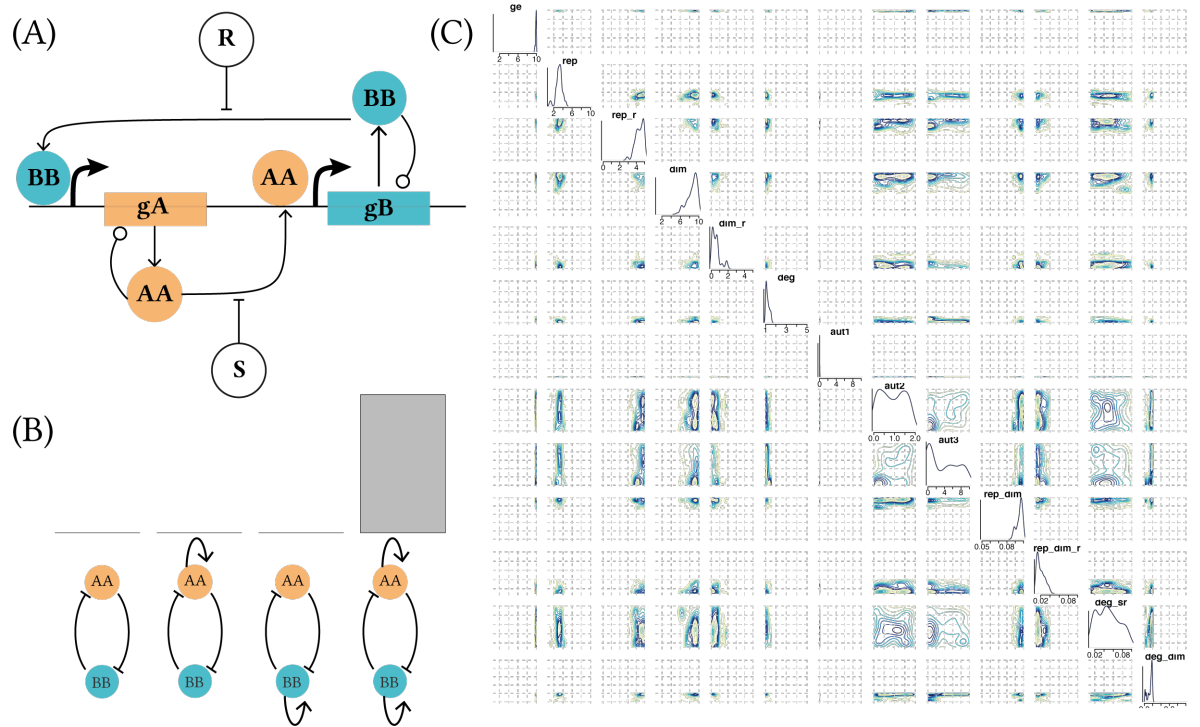| ge	rep	rep_r	dim	dim_r	deg	rep dim	rep_dim_r	deg_sr	aut_1	aut_2	aut_3
1-10	1-10	0-5	1-10	0-5	1-5	0.05-0.1	0.01-0.1	0-1	0-10	0-2	0-10

Figure 1.12 : (A) The toggle switch models considered. Interactions denoted by a circular arrow head can be positive of zero. (B) The toggle switch with positive autoregulation on A was found to be the most robust to parameter fluctuations. Three repeats of the model selection were carried out, and the median values are shown here. Upper and lower quartile error bars are included but are too small to be visible. (C) The posterior distribution of the toggle switch with positive autoregulation on A and B.

The results are shown in Figure 1.12. The toggle switch with positive autoregulation on both nodes was found to be the most robust model. This indicates that although all the models considered are capable of behaving like a switch, the model with double positive autoregulation could do that over a greater parameter range.

1.5   Discussion

Here I developed a more realistic model for the genetic toggle switch, using mass action. This model does not use the QSSA. In Chapter(XXX) I explore this model further and show that the QSSA is not necessary. I showed that this model is capable of bistable behaviour, within a given parameter range. I further studied this by using ABC SMC parameter inference to fit the toggle switch to a switching behaviour of choice. The parameter inference revealed the range of parameter values that can produce the behaviour of choice. These parameter values can be used to design a synthetic toggle switch that will behave in the specified manner.

Further, I showed that negative autoregulation is not consistent with bistable behaviour for the parameter values examined here. Adding small levels of single negative autoregulation to the system caused to revert to monostability. The only steady state was when the unregulated protein was high and the negatively regulated protein was low. This makes sense, as the negatively regulated protein cannot reach high enough levels to repress the other protein and dominate the system, whereas the unregulated protein is free to reach a higher steady state. In the case of double negative autoregulation, neither protein is free to reach sufficient levels to dominate the other, and the only steady state was when the levels of both proteins are low. This indicates that the system reaches a deadlock situation where both proteins are repressed and cannot reach a higher steady state. The models used here are deterministic and simplified down to two equations. Stochastic dynamics where noise is added to the system or a more complex model could be capable of overcoming this deadlock situation. More specifically, if transcriptional or translational bursting is included, the protein that receives the first boost can dominate the system on time and escape the deadlock situation (Strasser, Theis, & Marr 2012).

Finally, I showed that the addition of positive feedback loops make the genetic toggle switch more robust to parameter fluctuations. This means that the model was capable of producing the desired behaviour over a greater parameter range. This indicates that small fluctuations in parameters in the cellular environment will not affect the system's ability to be bistable, and thus makes it more suitable for use

in synthetic biological applications where a very constrained parameter set can be too restrictive. This makes it a better candidate for building new synthetic devices based on the toggle switch design.

The volume of the posterior distribution of even the most robust switch out of the ones examined here was still not large. This means that the behaviour of choice is still constrained, even after the addition of the positive feedback loops. A caveat of the analysis used here that has to be considered is that the parameter space is not searched for simply combinations that can produce a bistable switch. The behaviour that is required is very specific, and it is probable that the plethora of constraints put to the system result in the discard of parameter combinations that create a bistable switch. Firstly, a specific steady state level is required, for both the ON and OFF stages. When the switch is OFF, the protein levels must be as close to zero as possible, and when the switch is ON, the protein levels are required to approach 20nM. This requirement will discard any switches that have a higher or lower ON state, or a slightly higher OFF state. Additionally, the design specifications used here dictate that the time to reach steady state has to be quick. There is not much transition time allowed between the ON and OFF states, and the protein is required to reach steady state within a few timepoints. This also results in the exclusion of systems that reach steady state slower, but still act as a bistable switch. Therefore, the switches examined here, follow a very constrained behaviour. In the next Chapter I develop a method that is more flexible in the type of switches it can detect.

1.6   Summary

In this chapter I studied the Gardner, Cantor, & Collins (2000) toggle switch and showed that it is bistable. I further studied the genetic toggle switch model using mass action. I identified the parameter ranges that produce a bistable behaviour and could be used as prior distributions for parameter inference. Further, I studied the effect of adding feedback loops has on the robustness of the genetic toggle switch. I found that the switch with double positive autoregulation is the most robust to parameter fluctuations. In the next chapter I address some of the shortcomings of the method used here by developing a new algorithm, StabiliyFinder.

Bibliography

Alon, U. (2007). *An Introduction To The Systems Biology*. Chapman & Hall/CRC.

Barnes, C. P., Silk, D., Sheng, X., & Stumpf, M. P. H. (2011). 'Bayesian design of synthetic biological systems.' *Proceedings of the National Academy of Sciences of the United States of America* **108**(37), 15190–15195.

Brandman, O., Ferrell, J. E., Li, R., & Meyer, T. (2005). 'Interlinked fast and slow positive feedback loops drive reliable cell decisions.' *Science* **310**(5747), 496–498.

Clewley, R. (2012). 'Hybrid models and biological model reduction with PyDSTool.' *PLoS Computational Biology* **8**(8), e1002628–e1002628.

Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). 'Construction of a genetic toggle switch in Escherichia coli'. *Nature* **403**(6767), 339–342.

Liepe, J., Barnes, C., Cule, E., Erguler, K., Kirk, P., Toni, T., & Stumpf, M. P. H. (2010). 'ABC-SysBio–approximate Bayesian computation in Python with GPU support.' *Bioinformatics (Oxford, England)* **26**(14), 1797–1799.

Mathematica (2016). *version 10.3*. Wolfram Research, Inc.

McKay, M. D., Beckman, R. J., & Conover, W. J. (2000). 'A comparison of three methods for selecting values of input variables in the analysis of output from a computer code'. *Technometrics* **42**(1), 55–61.

Strasser, M., Theis, F. J., & Marr, C. (2012). 'Stability and multiattractor dynamics of a toggle switch based on a two-stage model of stochastic gene expression.' *Biophysical Journal* **102**(1), 19–29.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., & Stumpf, M. P. H. (2009). 'Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems.' *Journal of the Royal Society, Interface / the Royal Society* **6**(31), 187–202.