

A computational method for the investigation of multi-stable systems and its application to genetic switches.

Miriam Leon, Mae Woods, Alex J. H. Fedorec and Chris P. Barnes

SUPPLEMENTARY INFORMATION

1 Biochemical kinetic models

In this section we provide the priors and equations used for the models given in the main text.

1.1 Gardner switch model priors

The priors used for the analysis of the Gardner switches using StabilityFinder are given below.

Table S1: Gardner switch priors in the deterministic and stochastic cases

Parameters				Species	
a_1	β	a_2	γ	s_1	s_2
0-60	0-5	0-60	0-5	0-100	0-100

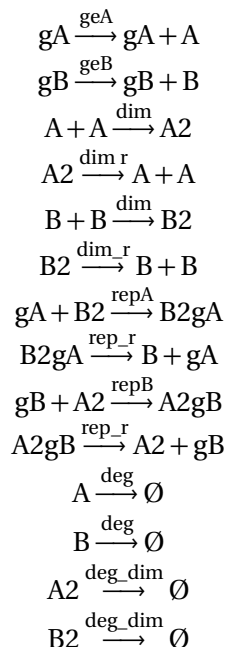
1.2 Lu switch models priors

Table S2: Priors of the classical(LU-CS), single positive (LU-SP) and double positive (LU-DP) Lu models.

Parameter	Symbol	CS	SP	DP
Production rate	gx	30-50	1-2	1-100
	gy	30-50	20-25	1-100
Degradation rate	kx	0-0.5	50-55	0-1
	ky	0-0.5	48-52	0-1
Hill coefficient	nxy	1-5	30-35	0-10
	nyx	1-5	0.1-0.2	0-10
Hill thresholds concentration	xyx	100-300	2-3	100-1000
	xyx	100-300	0.4-0.6	100-1000
Transcription rate fold change	lxy	0-0.5	0.02-0.04	0-1
	lyx	0-0.5	0.02-0.04	0-0.2
Hill coefficient	nXX	-	25-30	0-10
	nYY	-	0.01-0.02	0-10
Hill thresholds concentration	xXX	-	0.4-0.5	50-500
	xYY	-	1-3	50-500
Transcription rate fold change	lXX	-	65-72	1-20
	lYY	-	0.02-0.04	1-20

1.3 CS-MA

The classical toggle switch (CS-MA), under the assumption of mass action kinetics, is given by the following reactions:



The CS-MA switch was simulated using stochastic dynamics. The stoichiometry matrix and hazards defining the model are shown below:

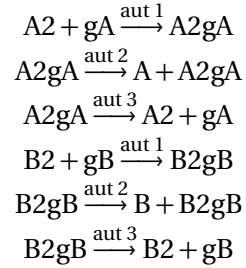
Table S3: CS-MA stoichiometry matrix

1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
-2.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
0.0	0.0	-2.0	0.0	0.0	1.0	0.0	0.0
0.0	0.0	2.0	0.0	0.0	-1.0	0.0	0.0
2.0	0.0	0.0	0.0	-1.0	0.0	0.0	0.0
0.0	-1.0	0.0	0.0	0.0	-1.0	0.0	1.0
0.0	1.0	0.0	0.0	0.0	1.0	0.0	-1.0
0.0	0.0	0.0	-1.0	-1.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	1.0	0.0	-1.0	0.0
-1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	-1.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	-1.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	-1.0	0.0	0.0

$$\begin{aligned}
h[1] &= geA \times gA \\
h[2] &= geB \times gB \\
h[3] &= dim \times A^2 \\
h[4] &= dim \times B^2 \\
h[5] &= dim_r \times B2 \\
h[6] &= dim_r \times A2 \\
h[7] &= geA \times gA \times B2 \\
h[8] &= rep_r \times B2gA \\
h[9] &= repB \times gB \times A2 \\
h[10] &= rep_r \times A2gB \\
h[11] &= deg \times A \\
h[12] &= deg \times B \\
h[13] &= deg_dim \times A2 \\
h[14] &= deg_dim \times B2
\end{aligned}$$

1.4 DP-MA

The DP-MA switch model has the following additional reactions:



The stoichiometry matrix and hazards defining the DP-MA model are given below.

Table S4: DP-MA stoichiometry matrix

1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
-2.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	-2.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
0.0	0.0	2.0	0.0	0.0	-1.0	0.0	0.0	0.0	0.0
2.0	0.0	0.0	0.0	-1.0	0.0	0.0	0.0	0.0	0.0
0.0	-1.0	0.0	0.0	0.0	-1.0	1.0	0.0	0.0	0.0
0.0	1.0	0.0	0.0	0.0	1.0	-1.0	0.0	0.0	0.0
0.0	0.0	0.0	-1.0	-1.0	0.0	0.0	1.0	0.0	0.0
0.0	0.0	0.0	1.0	1.0	0.0	0.0	-1.0	0.0	0.0
-1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	-1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	-1.0	0.0	-1.0	0.0	0.0	1.0	0.0
0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	-1.0	0.0
0.0	-1.0	0.0	0.0	-1.0	0.0	0.0	0.0	0.0	1.0
1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	1.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	-1.0
0.0	0.0	0.0	0.0	-1.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	-1.0	0.0	0.0	0.0	0.0

$$h[1] = repA \times gA$$

$$h[2] = repB \times gB$$

$$h[3] = dim \times A^2$$

$$h[4] = dim_r \times B^2$$

$$h[5] = dim_r \times B2$$

$$h[6] = deg \times A2$$

$$h[7] = rep_r \times gA \times B2$$

$$h[8] = rep_r \times B2gA$$

$$h[9] = repB \times gB \times A2$$

$$h[10] = dim \times A2gB$$

$$h[11] = deg \times A$$

$$h[12] = deg \times B$$

$$h[13] = aut_1 \times B2 \times gB$$

$$h[14] = aut_2 \times B2gB$$

$$h[15] = aut_3 \times B2gB$$

$$h[16] = aut_1 \times A2 \times gA$$

$$h[17] = aut_2 \times A2gA$$

$$h[18] = aut_3 \times A2gA$$

$$h[19] = deg_dim \times A2$$

$$h[20] = deg^5_dim \times B2$$

The priors used for the mass action switches are given below.

Table S5: The priors used in the simple (CS-MA) and double positive (DP-MA) mass action models

Parameter	Symbol	Models			
		CS-MA		DP-MA	
		Bistable	Tristable	Bistable	Tristable
Gene expression	geA, B	0 - 10	0 - 10	0 - 10	0 - 10
Dimerization	dim	7 - 15	0 - 3	7 - 15	0 - 3
Monomerization	dim_r	0 - 10	0 - 10	0 - 10	0 - 10
Repression	repA, B	0 - 10	0 - 10	0 - 10	0 - 10
Dissociation of repression	rep_r	0 - 10	0 - 10	0 - 10	0 - 10
Monomer degradation	deg	0 - 10	0 - 10	0 - 10	0 - 10
Dimer degradation	deg_dim	0 - 1	0 - 1	0 - 1	0 - 1
Dimer promoter self-association	aut_1	-	-	0 - 10	0 - 10
Dimer promoter self-activation	aut_2	-	-	0 - 10	0 - 10
Dimer promoter self-dissociation	aut_3	-	-	0 - 10	0 - 10

1.5 Three node switch priors

Table S6: Priors used in the three-node switch

Parameter	Symbol	Range
Production rate	gx	3-5
Degradation rate	kx	0-0.2
Hill coefficient	nxy	0-2
Hill thresholds concentration	xxy	140-160
Transcription rate fold change	lxy	0-0.2
Hill coefficient	nxx	2-4
Hill thresholds concentration	xxx	90-110
Transcription rate fold change	lxx	8-12

1.6 LU-DP posterior distribution

1.7 Deterministic Gardner model basins of attraction

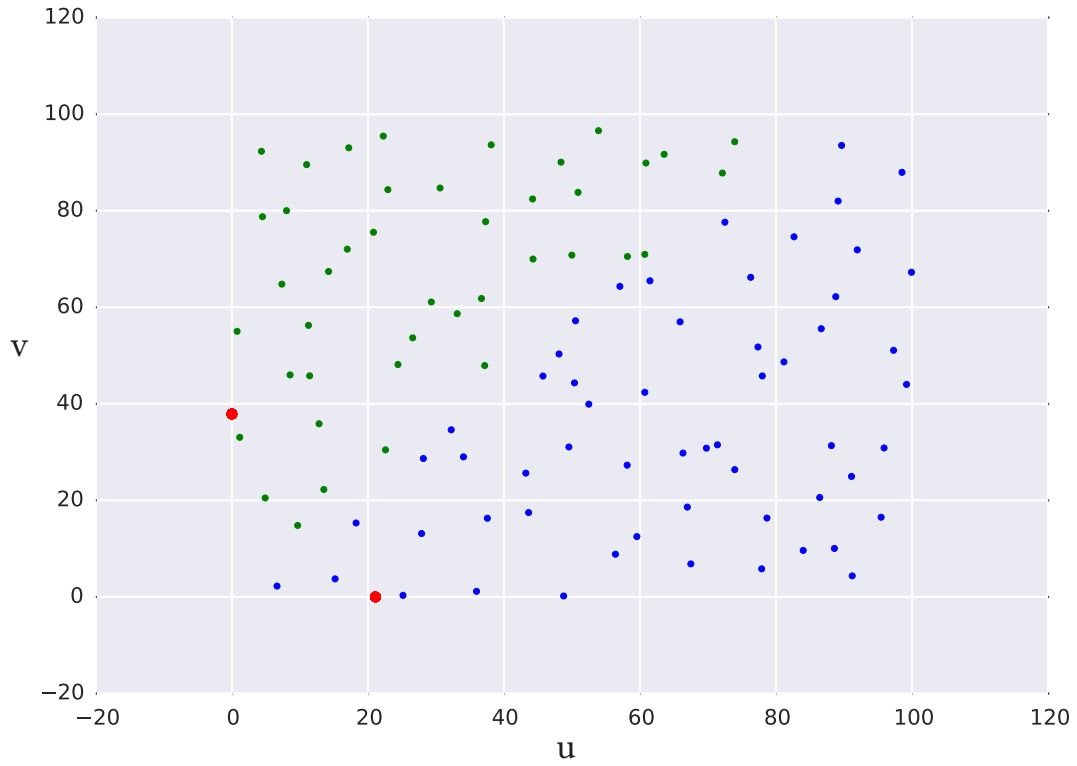


Figure S2: Basins of attraction of the deterministic Gardner model. The red points represent the steady state values. The green and blue points represent the initial conditions resulting to each steady state.

2 Clustering methods

2.1 Deterministic models

Algorithm 1 Clustering the steady state deterministic simulation results

```
1: for each data point do
2:   if first point then
3:     Make first cluster
4:     cluster counter = 1
5:   else
6:     for each cluster do
7:       if cluster within cluster means  $\pm$  delta then
8:         Add to existing cluster
9:         Update means of clusters
10:      end if
11:      if reached_end and not assigned to cluster then
12:        cluster counter += 1
13:        Add new cluster
14:      end if
15:    end for
16:  end if
17: end for
```

2.2 Stochastic models

2.2.1 Gap statistic

Algorithm 2 Choosing the optimal number of clusters

```
1: function WK(clusters, cluster_centres)
2:   for each cluster do
3:     for each point in cluster do
4:       a = matrix norm (cluster_centre – point)
5:     end for
6:      $dk = \sum ((a)^2) \times (2 \times \text{number of points in cluster})$ 
7:   end for
8:    $wk = \frac{\sum(dk)}{2 \times (\text{number of points in cluster})}$ 
9:   return wk
10: end function

11: function GAP_STATISTIC(data, cutoff)
12:   ks = [1,2,3,4]
13:   for k in ks do
14:     cluster_centres, clusters = KMEANS(data, k, cutoff)
15:      $Wk = \log(WK(\text{clusters}, \text{cluster\_centres}))$ 
16:     Create references datasets
17:     for each references dataset do
18:       cluster_centres, clusters = KMEANS(data, k, cutoff)
19:        $BWk = \log(WK(\text{clusters}, \text{cluster\_centres}))$ 
20:     end for
21:      $Wkb = \frac{\sum(BWk)}{10}$ 
22:      $sk = \sqrt{\sum(\frac{(BWk - Wkb)^2}{10})}$ 
23:   end for
24:    $sk = sk \times \sqrt{1 + \frac{1}{B}}$ 
25:   return ks, Wk, Wkb, sk, data_centres, clusters
26: end function

27: function DISTANCE(data, cutoff)
28:   ks, logWks, logWkbs, sk, clusters_means, clusts = GAP_STATISTIC(data, cutoff)
29:   gaps = logWks – logWkbs
30:   optimum number of clusters =  $gaps[i] \geq (gaps[i + 1] - sk[i + 1])$ 
31:   return cluster_counter, clusters_means
32: end function
```

2.2.2 K-means clustering

Algorithm 3 Clustering stochastic case

```
1: function KMEANS_CLUSTERING(data, k, cutoff)

2:   function UPDATE_CENTRES(old_centres, values)
3:     centre_coords = mean for each dimension
4:     shift = GETDISTANCE(centre_coords, old_centres)
5:     return shift, centre_coords
6:   end function

7:   function GETDISTANCE( $a, b$ )
8:      $dist = \sqrt{(a[x] - b[x])^2 + (a[y] - b[y])^2}$ 
9:     return  $dist$ 
10:  end function

11:  while True do
12:    for each point in data do
13:      for each cluster do
14:         $dist = \text{GETDISTANCE}(\text{point}, \text{cluster centre})$ 
15:      end for
16:      Find cluster with minimum distance
17:      Repopulate clusters
18:    end for
19:    biggest_shift  $\leftarrow$  0
20:    for as many times as there are clusters do
21:      shift, cluster_centres = UPDATE_CENTRES(old_centres, clusters)
22:      biggest_shift = max between shift, biggest_shift
23:    end for
24:    if biggest_shift  $\leq$  cutoff then
25:      break
26:    end if
27:  end while
28:  return cluster_centres, clusters
29: end function
```

3 Effect of variance on Gardner model posterior

The posterior of the deterministic model shown in Figure 2C in the main text, parameters a_1 and a_2 have an upper limit of values they can take and still create a bistable switch. In order to test this result, we find the roots of the system for large values of a_1 and a_2 in order to see if three roots are still found, two stable and one unstable. The results as shown in Figure S3 indicate that the system is still bistable for increasing values of a_1 and a_2 . This suggests that the upper limit found with StabilityFinder is an artefact of the variance limit imposed on the system. In order to find the steady states we impose an accepted distance from a given total variance for each model. When the two clusters of steady state values are too far apart, this increases total variance of the system and would consequently be rejected in StabilityFinder.

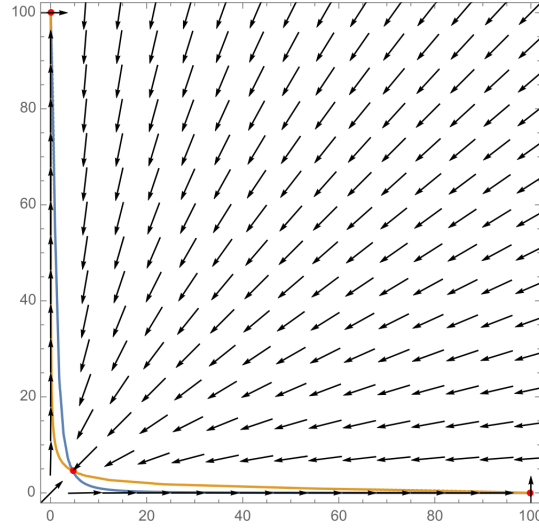


Figure S3: Solving the Gardner toggle switch. The parameters values used are: $a_1, a_2 = 100$ and $\beta, \gamma = 2$. The system has three roots, of which one was found to be unstable and the other two stable. This result disagrees with that found in StabilityFinder, that a_1 and a_2 have an upper limit of 30.