A noise-based method for finding bistabilities in biological reaction networks

Arvind Shankar Prabhakar
Department of Chemical Engineering, MIT

Abstract

Typically, in systems biology problems, a reasonable amount is known about network structure but very little is known about the parameters that govern the interactions in the network. One of the hard problems is to identify functional motifs in networks, such as bistabilities. This project describes a metropolis Monte Carlo method using to identify regions in parameter space for which a network is most likely to be bistable, using only the noise characteristics of the network.

1. INTRODUCTION

Biological reaction networks are typically understood in terms of functional modules called "motifs," each of which traditionally having a typical response to an input signal. Simple motifs are easy to identify just by looking at the network structure; however, for more complicated ones, it is not always obvious whether a given network structure can lead to a certain type of input-output relationship. Typical of this class are oscillators and bistable latches, which have many important cellular functions.

Bistabilities are arguably one of the most important features in cell signalling, and are typically thought to be used in situations where an "all-or-nothing" response is required to some input signal [1]. For inputs below a critical threshold, the signalling output remains near its basal state; for inputs above the threshold, the output increases to a high, active state. Bistable systems make use of hysteresis to remain in the active state, meaning that the input stimulus required to keep the system in the active state is lower than the input required for triggering the initial transition from basal to active state. Many cell signalling processes are thought to be bistable, including calcium spiking [2] and oocyte maturation [3]. In this work, we will look at two extremely common motifs which both lead to bistability: a positive feedback loop [4] and double phosphorylation [5].

Various methods have been proposed to determine whether a network with a given structure can exhibit bistability. It is possible, in certain situations, to tell from the network structure alone whether or not a network can exhibit multiple steady states [6], though with more complicated networks, this method may not yield any answer. If all the parameters are known, typically, one can try to determine all the fixed points and their stabilities; however, that is not always straightforward. It is also thought that noise, which is intrinsic to biological reaction networks, may change the stability properties of a system [7].

A method has been described in literature which describes a Metropolis Monte Carlo based parameter search to determine rate constants if some experimental data is known [8]. The method essentially involves randomly varying the parameters of the network, and computing the probability that the given data could have been drawn from this model as

$$P\left(Y\middle|\mu\right) \propto \exp\left[-\sum_i \frac{\left(Y_i-\mu_i\right)^2}{2\sigma_i^2}\right] \qquad (1)$$

In this project, instead of maximizing the probability that this model could yield some known data, we try to maximize the noise in the network, and hypothesize that this leads us to any bistability that may be present in the network.

2. THE METHOD

A. Basis for the method

Consider the biological reaction networks as a nonlinear dynamical system, say

$$\frac{dy_i}{dt} = f_i\left(\{y_j\};\{p_k\}\right) \qquad (2)$$

Here $\{y_j\}$ represents vector of all species in the system, $\{p_k\}$ is the set of all parameters of the system (which in this case would be a vector containing all the rate constants and initial concentrations of molecules present in this biological reaction network). We consider the case in which the network has reached a steady state, i.e.

$$f_i\left(\{y_j\};\{p_k\}\right) = 0 \qquad (3)$$

Assume that the system has a bistability, i.e. there exists a set of parameters $\{p_k\}$ for which this system has multiple steady states. A general method for finding bistabilities in a network would be to assume a particular set of parameters $\{p_k^o\}$ and solve for all the steady states of the network. However, if there is only one steady state, it is not clear how to choose another set of parameters to increase the chance of finding a bistability.

In reality, there are intrinsic stochastic fluctuations in the trajectories of molecule numbers with time. These fluctuations are due to random chance, and the trajectory that the system takes could be vastly different depending on which reactions stochastically occur first. If there is a bistability, we expect trajectories to reach one of two possible steady states, so if we define the noise as the variance in the trajectories, noise is maximum at a bistability. The question arises, does the noise increase as one approaches a bistability?

The important thing is to note that the fluctuations are dependant on the shape of the potential energy minimum that governs the current fixed point. The shallower the fixed point, the more the noise. "Shallowness" of fixed points is quantitatively described by the eigenvalue of the Jacobian of the system at its steady state, and, as we approach (in parameter space) a bistability from a point where is only one fixed point, we expect to go through a bifurcation. Since, at the bifurcation, one eigenvalue is zero, we expect the noise in the system to increase as we approach the bifurcation.

B. Algorithm

The algorithm is as follows:

1. Start at a random point in parameter space $\{p\}$.
2. Run a set of trajectories using the Gillespie method
3. Calculate the variance of these trajectories at some point in time.
4. Randomly change $\{p\}$ to $\{p'\}$ by drawing from a probability distribution around $\{p\}$.
5. Run a new set of trajectories using Gillespie and calculate the new variance.
6. If the new variance is larger than the previous, keep the new set of parameters. If not, throw a random number to determine whether to keep the new parameters. If the random number < $g\left(\{p'\}\right)/g\left(\{p\}\right)$, keep the new parameter set.

A few details are worth pointing out. Because of the nature of rate constants as having an exponential effect on the system, new rate constants have been chosen from a distribution which is Gaussian distributed over several orders of magnitude centred on the current value. Care also has to be taken to ensure that the rate constants do not go below zero. The simulations have been set up so that rate constants cannot go above or below fixed limits, which are 10^{-5} and 10^{4} in the units used. These are wide limits and most measured rate constants for protein-protein interactions in biology have been found to fall within these limits.

Ideally, we should select a function g to reflect the probability that the simulation results are drawn from a bimodal distribution, rather than a unimodal one. However, this has not been implemented yet; as of now, an artificial function using the exponential of the noise has been used, as heuristically it seemed to work best in the case of the Ras-Sos model.

For each point in parameter space, a hundred trajectories have been run to get statistics about the behaviour of the network. This number is arbitrary; more runs will yield better statistics but take a longer time. Also, as this is just an initial exploration of whether a method is feasible, each simulation is limited to 10,000 Monte Carlo steps.

Each trajectory is simulated using the Gillepsie algorithm [9], with the help of a tool written by Lis and Artomov [submitted].

3. RESULTS

We shall examine this method in the light of two biological reaction networks that possess a bistability: the classical Ras-Sos positive feedback loop and the ERK double phosphorylation bistability.

A. The Ras-Sos positive feedback loop

The Ras-Sos positive feedback loop is a key gateway in the processing of signals in many cell types, including T-cells, which are a part of the human immune system. Ras activation is important for the development of T and B lymphocytes and for their effector functions directed against invading pathogens. Ras activation (the conversion of GDP-bound Ras to a GTP-bound form) is modulated by Sos, which is brought to the membrane under the action of external stimulus. The Sos molecule has two pockets to which Ras can bind, an allosteric and a catalytic site. Upon binding of RasGDP to the allosteric site, Sos has a low level of catalytic activity for the conversion of RasGDP to RasGTP, but if RasGTP is bound to the catalytic site, the catalytic activity of Sos for the same reaction is much higher. This positive feedback loop leads to a bistability. The molecule RasGAP converts RasGTP to RasGDP.

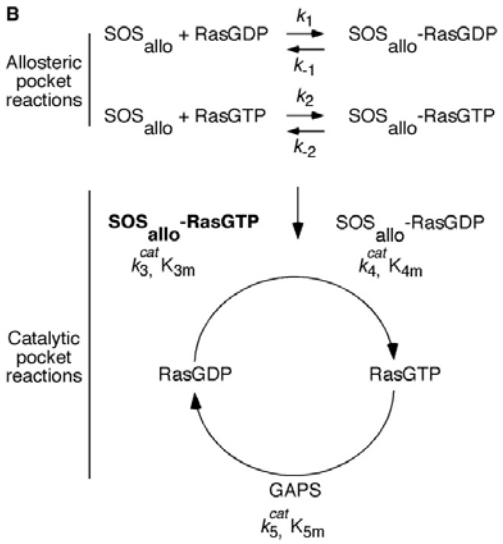

Figure 1: The Ras-Sos network [4].

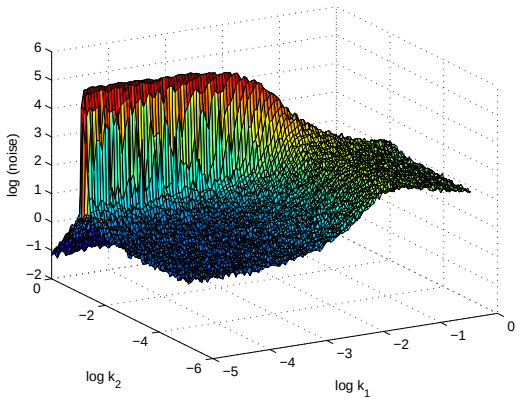

Figure 2 : A “noise map” of the parameter space

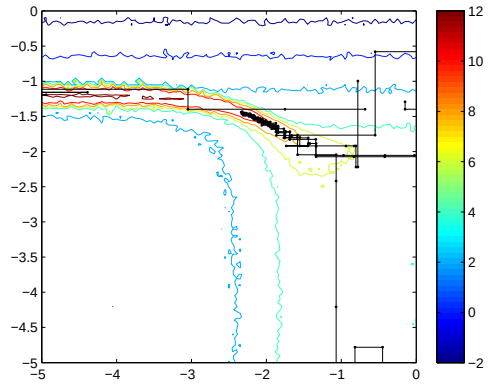

Figure 3: A contour plot of the noise map of the parameter space with trajectories superimposed.

The Ras-Sos bistability was explored by assuming two of the parameters (the two catalytic rates of Sos with RasGDP and RasGTP bound to the allosteric pocket) were unknown; the other parameters were taken from [4]. To measure the performance of the method, a “noise map” was made of the system by running trajectories and calculating their variance at each point in parameter space. This “noise map” corresponds to a free energy landscape and the peak in the noise map corresponds to the bistable region.

A set of ten trajectories were then run using the algorithm that has been described. All the trajectories went to bistable the region in parameter space (Figure 2).

A similar set of ten simulations was run by allowing four parameters instead of two to vary. Out of these ten, in two cases it reached a bistability; in three others, it went to a region in parameter space where activation of the network takes a time which is comparable to the time for which the network is simulated, so that some of the trajectories end up

fully activated and others end up not yet activated. This, obviously, would change if the system were run for a longer simulation time. In the rest of the simulations, it did not reach a high-noise state.

B. Double phosphorylation ERK bistability

The mitogen-activated protein kinase (MAPK) cascade is a key component of intracellular signalling with many functions. The typical motif in this cascade is the double phosphorylation of a MAPK by its activator, a MAP kinase kinase (MAPKK). This motif can lead to a bistability, as shown by Kholodenko [5] and others. We shall examine this bistability in the specific case of MEK as the MAPKK and ERK as the MAPK. This network involves phosphorylation of the MAPK by the MAPKK on two sites, and dephosphorylation of both sites by a phosphatase. All reactions follow enzyme kinetics. A schematic of the network is shown in Figure 4, where M stands for ERK and MKP3 is the phosphatase of ERK..

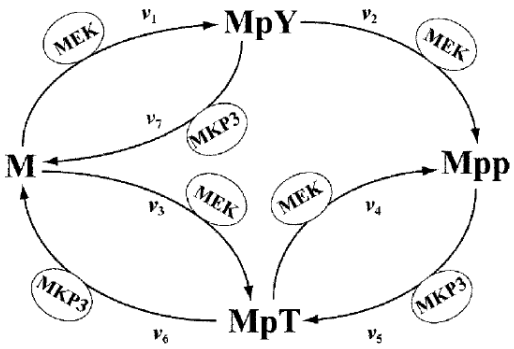

Figure 4: The MEK-ERK network.

Again, all but two rate constants in the MEK-ERK network are assumed to be known (and equal to the values reported in [5]); the other two are allowed to vary and regions of maximum noise are found according to the algorithm described. Again, we make a "noise map" of the system (Figure 5) and use that to look at trajectories obtained from the Monte Carlo algorithm (Figure 6).

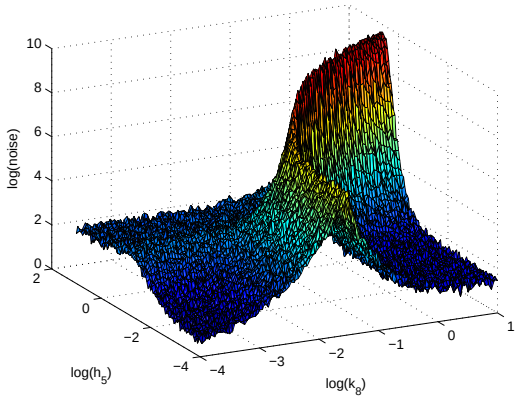

Figure 5: A noise map of the MEK-ERK system

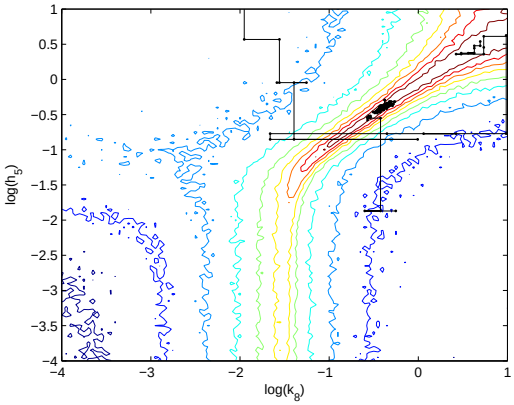

Figure 6: Monte Carlo trajectories for the MEK-ERK system

Again, for the MEK-ERK system, Monte Carlo trajectories go to the high noise region; however, just like in the multi-parameter Ras-Sos case, it is not entirely clear that the high-noise region represents the bistability (but the bistability is part of the nigh-noise region, again).

4. DISCUSSION

The method seems to successfully find high-noise regions in parameter space, especially in cases where only a few parameters are involved. For higher numbers of parameters, it does not seem to work so quickly. A big question that is not fully satisfactorily answered is, does the maximum noise actually correspond to a bistability? In both systems tested, the parameters given in literature (for which the networks are bistable) correspond to the high-noise regions as shown in the noise maps; however, it is not clear whether all high-noise regions are actually

bistable. On simulating some of these points, it is seen that the time taken for trajectories to activate is similar to the total time of simulation, which means that only some trajectories do “fire” during the simulation, leading to a high “noise” which is false.

Many of the details remain to be worked out, but the method shows promise in determining regions of bistability in complex biological reaction networks. The more general question of how much we can say about a network from its noise properties remains largely unanswered.

5. REFERENCES

[1]. Brandman, O., Meyer, T., *Science*, 322, 17 October 2008.

[2]. T. Meyer, L. Stryer, *PNAS*, 85, 5051 (1988).

[3]. W. Xiong, J. E. Ferrell Jr., *Nature* 426, 460 (2003).

[4]. Das, J., Ho, M., Zikherman, J., Govern, C., Yang, M., Weiss, A., Chakraborty, A.K., Roose, J.P., *Cell*, 136, 337–351, (2009)

[5]. Markevich, N.I., Hoek, J.B., Kholodenko, B.N., *J. Cell Bio*, Volume 164, Number 3, 353–359, (2004)

[6]. Craciun, G., Tang, Y., Feinberg, M., *PNAS*, 103, 23, 8697–8702 (2006).

[7]. Scott, M., Hwa, T., Ingalls, B., *PNAS*, 104, 18, 7402–7407 (2007)

[8]. Battogtokh, D., Asch, D.K., Case, M.E., Arnold, J., Schüttler, H.B., *PNAS*, 99:16904-16909 (2002)

[9]. Gillespie, D.T. *J. Phys. Chem,* Vol. 8 1, No. 25, 1977