2005-01-01

Optimal path to epigenetic switching

This work was made openly accessible by BU Faculty. Please share how this access benefits you. Your story matters.

| Version               | Accepted manuscript                                                                 |
|-----------------------|-------------------------------------------------------------------------------------|
| Citation (published version): | D.M. Roma, R.A. O'Flanagan, A.E. Ruckenstein, A.M. Sengupta, R. Mukhopadhyay. 2005. "Optimal path to epigenetic switching." PHYSICAL REVIEW E, Volume 71, Issue 1, pp. ? - ? (5). https://doi.org/10.1103/PhysRevE.71.011902 |

https://hdl.handle.net/2144/37971

Boston University
Optimal Path to Epigenetic Switching

David Marin Roma, Ruadhan A. O’Flanagan, Andrei E. Ruckenstein, and Anirvan M. Sengupta

Dept. of Physics and Astronomy and BioMaPS Institute, Rutgers University, Piscataway, NJ 08854

Ranjan Mukhopadhyay
Dept. of Physics, Clark University, Worcester, MA 01610

We use large deviation methods to calculate rates of noise-induced transitions between states in
multi-stable genetic networks. We analyze a synthetic biochemical circuit, the toggle switch, and
compare the results to those obtained from a numerical solution of the master equation.

PACS numbers: 02.50.Ey 05.70.Ln 82.39.-k

Fluctuations in bio-molecular networks have been the
subject of much research activity recently [1]. Studies
on noise in gene expression [2, 3, 4, 5, 6, 7], in signal
transduction [8] and in biochemical oscillators [9, 10, 11]
demonstrated that having a small number of molecules
affects, sometimes critically, the behavior of cellular
circuits. Stochastic aspects of the choice between lytic
and lysogenic developmental strategies of bacteriophage
lambda virus infection in E. coli were studied in an in-
fluential paper by Arkin, Ross and McAdams [12].

One of the interesting aspects developmental processes
is that one could get multiple heritable cell fates without
irreversible changes to the genetic information. Differ-
et cells with the same DNA sequence, showing different
phenotypes that are stably maintained through cell di-
visions, namely epigenetic phenomena, have been rep-
resented as multiple stable attractors in deterministic
descriptions of the biochemical dynamics. In this pa-
per, we are concerned with the robustness of such at-
tractors against spontaneous fluctuations which might
induce transitions from one stable state to another. Pre-
vious work in this area has modeled the effects of fluc-
tuations by adding Gaussian-distributed Langevin forces
in this approximation. We wish to test our methods on a sim-
pler system. We will consider the artificially constructed
toggle switch [18]. In this example, we find that the con-
tributions to the transition rate coming from corrections
to the Gaussian approximation can change the overall
rate by several orders of magnitude and, therefore, are
important for comparison with experimental results.

The theory of transition rates is a well developed sub-
ject (see [19] as well as references therein). For a bistable
system like the genetic switch we are considering, the
transition probability from one stable point to the other
is estimated by computing the probability of reaching
the saddle point between stable states, and, from there,
to follow the deterministic trajectory to the other stable
state, rather than to fall back to the initial state. The
transition rate is given by an expression of the form:

\[
\text{rate} = \frac{\lambda_+}{2\pi} \left[ \frac{\text{det} A_{fp}}{\text{det} A_{sp}} \right]^{1/2} \ast P(x_f, x_o),
\]

where \(\lambda_+\) is the positive eigenvalue of the matrix describ-
ing the linearized equations of motion around the saddle
point, \(A_{fp}\) and \(A_{sp}\) are the inverses of covariance matrices
appearing in the quasi-stationary Gaussian approxima-
tion of the probability distribution in the starting stable
point and in the saddle point respectively and \(P(x_f, x_o)\)
is the probability of finding the system at the saddle point
state in a quasi-stationary distribution centered around
the stable fixed point \(x_o\). Note that \(A_{sp}\) has one nega-
tive eigenvalue and the Gaussian distribution around the
saddle point is only a formal solution. A more precise
Much of the rest of the paper is devoted to the computation of $P(x_f, x_o)$ by large deviation methods. There are two related ways. In one approach, one keeps track of the trajectories in the space of numbers of different molecules, distributed according to a state dependent Poisson process, and computes time dependent transition probabilities as sum the probabilities of all paths connecting the initial and final points, which leads naturally to a path integral formulation of the stochastic process. In this way, the transition probability is evaluated as the exponential of the “action”. This action can be computed in a perturbation expansion (using the volume of the system as a parameter), in which the leading order correction is the line integral along the path that minimizes the action (optimal path) of a Lagrangian function. This calculation naturally gives rise to a Hamiltonian that corresponds to the evolution operator in the master equation [21], written in terms of numbers and raising and lowering operators expressed as exponentials of the phase variables conjugate to the numbers.

An alternative but exactly equivalent approach is to start directly from the master equation and solve it in the Eikonal approximation [22]. We will present our arguments in this article using this approach, which is easier to explain mathematically, and provides an easier way to compute the next order correction in the volume expansion, a term which has not been computed before in relation to these genetic switches. As we will see, in the case of the toggle switch, the next order term in $\ln P(x_f, x_o)$ makes an important contribution to the overall rate of transition.

The general ideas are developed in the context of the simple example of the toggle switch. This artificially realized switch consists of two genes that repress each other’s expression, placed in a high copy plasmid in E. coli. Once expressed, each protein can bind particular DNA sites upstream of the gene which codes for the other protein, thereby repressing its transcription. If we denote the $i$-th protein concentration by $x_i$, the deterministic system is described by the equations:

\begin{align*}
\dot{x}_1 &= \frac{a_1}{1 + (x_2/K_2)^m} x_1 - \frac{x_1}{\tau} \quad (2) \\
\dot{x}_2 &= \frac{a_2}{1 + (x_1/K_1)^m} x_2 - \frac{x_2}{\tau} \quad (3)
\end{align*}

the constants $a_1$ and $a_2$ incorporate all aspects of transcription and translation reactions. The Hill exponents, $m$ and $n$, represent the degree of cooperative binding of proteins to DNA, and $\tau^{-1}$ is the protein degradation/dilution rate (assumed equal for the two proteins). $K_i$ is the effective dissociation constant for binding of protein 1 in the promoter of gene 2. $K_2$ is the corresponding parameter for protein 2. For some regions of parameter space, the system has three stationary points: two stable ones and a saddle point [18].

For the purposes of this discussion, we model the stochastic evolution of the protein concentrations in the system by a birth-death process in which protein $i$ is made in short-lived bursts of size $b_i$ and proteins are diluted or degraded at a rate $\tau^{-1}$. A more detailed description involving proteins and RNA will be published elsewhere. It is worth noting that, while both the burst size $b_i$ and the RNA production rate show up as parameters in the stochastic modeling, only their product, $a_i$, shows up in the effective deterministic equations [2] for the protein levels.

To compute the rate of transition from one fixed point to the other, we must solve the master equation [21], which describes the time evolution of the probability distribution of protein concentrations. The qualitative behavior of the stationary solution for the bistable system can be described in simple intuitive terms: the solution displays two peaks centered around the stable points. If we start with probability one around one of the stable points, rare transitions lead to a long tail which leaks into the domain of attraction of the other stable point, in very much the same way in which the probability amplitude extends beyond the classically allowed region in quantum mechanical tunneling through a barrier. This analogy motivates the Eikonal approximation to the solution of the master equation [22]. The master equation is given by,

$$\frac{\partial P}{\partial t} = \Omega \sum_\epsilon [W_\epsilon(x - \hat{\epsilon}/\Omega)P(x - \hat{\epsilon}/\Omega, t) - W_\epsilon(x)P(x, t)]$$

(4)

where $\Omega$ is the volume of the system, $\hat{\epsilon}/\Omega = \Delta \vec{x}$ is the concentration change associated with individual reaction events, the rate of which is given by $\Omega W_\epsilon(x)$. Assuming that the distribution is quasi-stationary in the region of interest, we consider solutions of the WKB form:

$$P(x, t) = C \exp[-\Omega S(x)], \quad S(x_o) = 0. \quad \text{(5)}$$

$x_o$ being the initial stable point. In the same way the wave function in quantum mechanics is computed using an expansion in powers of $\hbar$, it customary to find the
probability $P(x, t)$ by expanding $S(x)$ in powers of inverse volume, which plays the same role as $\hbar$ in quantum mechanics, since the bigger the volume, the less likely are fluctuations to happen. Then, to first order in $\Omega^{-1}$, we write:

$$S(x) = S_0(x) + \Omega^{-1} S_1(x) + O(\Omega^{-2}).$$

Assuming that the scaled transition rates $W_\varepsilon(x)$ are smooth functions of $x$, and expanding $S$ to first order, $S(x - \hat{e}/\Omega) = S(x) - \hat{e} \sum_i \frac{\partial S}{\partial x_i} S(x)$, collecting the terms which do not contain powers of $\Omega$ we have:

$$\frac{\partial P(x, t)}{\partial t} = HP(x, t) \tag{6}$$

$$H(x, p) = \sum_\varepsilon [W_\varepsilon(x)(e^{\hat{e}p} - 1)] \tag{7}$$

where $H$ is the Hamiltonian describing the time evolution of the probability distribution, and we define the momentum $p_i$ as:

$$p_i = \frac{\partial}{\partial x_i} S_0(x) \tag{8}$$

If we expand the Hamiltonian $H$ in $p$ and keep terms up to second order in $p$ we recover the Gaussian approach used in [13, 14]. Since we are considering a situation where the transitions are so rare that the probability does not change much in time, the Hamiltonian will be very small.

The main contribution to the transition probability is obtained by evaluating $P$ along a particular trajectory $W_\varepsilon(x)$. This trajectory, called the optimal path, is the solution to Hamilton’s equations derived from Eq.7:

$$\dot{x}_i = \frac{\partial H(x, p)}{\partial p_i} = \sum_\varepsilon \hat{e}_i W_\varepsilon(x) e^{(\hat{e}_i p_a)} \tag{9}$$

$$\dot{p}_i = -\frac{\partial H(x, p)}{\partial x_i} = -\sum_\varepsilon \hat{e}_i \frac{\partial W_\varepsilon(x)}{\partial x_i} (e^{(\hat{e}_i p_a)} - 1) \tag{10}$$

For the toggle switch example we have four $\hat{e}_i$-s describing jumps to the right, left, up or down, given by $b_1 x_1, -\hat{x}_1, b_2 \hat{x}_2,$ and $-\hat{x}_2$, respectively. The relevant Hamiltonian defined on times long compared to the inverse binding/unbinding rates of proteins at the two promoters is given by:

$$H = \frac{a_1/b_1}{(1 + (x_2/K_1)^n)} (e^{b_1 p_1} - 1) + \frac{a_2/b_2}{(1 + (x_1/K_2)^m)} (e^{b_2 p_2} - 1) + \frac{x_1}{\tau} (e^{p_1} - 1) - \frac{x_2}{\tau} (e^{-p_2} - 1). \tag{11}$$

As already mentioned above, $K_{1,2}$ are the effective dissociation constants for binding of proteins 1, 2 at the promoter of gene 2,1, respectively, $b_i$ is the burst size of protein $i$ and the ratio $a_i/b_i$ is a measure of the RNA production rate associated with the transcription of the gene $i$.

To extract the values of the burst size parameters, the spontaneous transition rate has to measured experimentally for more than one conditions. Since this has not yet been done, we will compare the results of the Eikonal approximation to the solution obtained by direct diagonalization of the Hamiltonian [11]. For simplicity we will set the parameters $K_i = 1, b_i = 1$.

The optimal path for the transition from one stable point to the other starts near one stable point, proceeds to the saddle point and from there it follows the deterministic trajectory to the other stable point. Thus we must first find solutions of Eqs. 9 and 10 which start at (near) the initial stable point and end at the saddle point. At the end points we have $p_1 = p_2 = 0$, and $H = 0$. This also implies that if the system is at the stable point it will remain there. So, the optimal path must instead start at a point very close to but not exactly at the fixed point. In this case, the Hamiltonian will be a very small number (and constant). In what follows, we will make the approximation $H = 0$. The initial conditions for the momentum equations can be obtained by approximating the probability around the stable point by a Gaussian distribution $P = e^{-\Omega S_g}$ with $S_g = \frac{1}{2} A_{ij} \delta x_i \delta x_j$ (note that we use summation convention, i.e., repeated indices are summed over). Then $p_i = \frac{\partial S_g}{\partial x_i} = A_{ij} \delta x_j$, and we expand the equation $H = 0$ around the stable point to find $A_{ij}$. Then we have a two point boundary value problem which can be solved by various methods [23]. The solution of
the equations of motion \[9\] and \[10\] for a set of parameters, projected to concentration space, is shown in Figure \[1\]. We integrate equations \[9\] along the optimal path \(C\) to obtain \(S_0 = \int_C p(t) dx_1\).

**FIG. 1:** Optimal path for the parameters, \(a_1 = 156, a_2 = 30, n = 3, m = 1, K_1 = K_2 = 1, b_1 = b_2 = 1\) and \(\tau = 1\). \(x_i\) are dimensionless. The ellipsoid indicates the orientation of the Gaussian spread around the stable point. The size of the spread scales like \(\Omega^{-1}\).

The \(S_1\) factor can be viewed as a correction due to fluctuations around the optimal path and could be calculated following references \[24\] and \[25\]. Collecting coefficients of powers of \(\Omega\) in the \(\Omega^{-1}\) expansion we derive an equation for \(S_1\):

\[
\sum_i W_e \hat{e}_i \frac{\partial S_1}{\partial x_i} = \frac{1}{2} W_e \hat{e}_j \frac{\partial p_j}{\partial x_i} \hat{e}_i + \sum_i \hat{e}_i \frac{\partial W_e (x)}{\partial x_i} \hat{e}_i \hat{e}_i = 0 \quad (12)
\]

In turn, after using the equations of motion to rewrite the first term as derivative along the optimal path \(x_{opt}(t')\), Eq. \[12\] can be transformed into:

\[
\frac{d}{dt} S_1 = \sum_i \frac{1}{2} W_e (x) \hat{e}_i \frac{\partial p_j}{\partial x_i} \hat{e}_j \hat{e}_i \hat{e}_i + \sum_i \hat{e}_i \frac{\partial W_e (x)}{\partial x_i} \hat{e}_i \hat{e}_i \quad (13)
\]

To proceed we need \(\frac{\partial p_j}{\partial x_i}\) along the path. From Hamilton’s equations \[16\], it follows that \(\delta p(t)_a = M(t)_{ab} \delta x(t)_b\), and thus we can use the components of the matrix \(M\) in place of the derivative \(\frac{\partial p_j}{\partial x_i}\) in \[12\]. Moreover, \[15\] also implies that:

\[
\begin{align*}
\delta \dot{x}^a &= \frac{\partial^2 H}{\partial p_a \partial x^i} \delta x^i + \frac{\partial^2 H}{\partial \dot{p}_a \partial p_i} \delta p_i \quad (14) \\
\delta \dot{p}_a &= -\frac{\partial^2 H}{\partial x^a \partial x^i} \delta x^i - \frac{\partial^2 H}{\partial x^a \partial \dot{p}_i} \delta \dot{p}_i \quad (15)
\end{align*}
\]

Combining this together with the time derivative of \(\delta p(t)\),

\[
\delta \dot{p} = M \delta x + M \delta \dot{x} \quad (16)
\]

leads to the following set of coupled differential equations for \(M\):

\[
\begin{align*}
M_{ab} + M_{ac} \frac{\partial^2 H}{\partial x^b \partial p_c} + M_{ac} \frac{\partial^2 H}{\partial p_c \partial p_d} M_{db} \\
+ \frac{\partial^2 H}{\partial x^a \partial p_c} M_{cb} + \frac{\partial^2 H}{\partial x^a \partial x^b} &= 0 \quad (17)
\end{align*}
\]

with initial conditions: \(M_{ij}(t = 0) = A_{ij}\) (defined below equation \[11\]). Finally, solving these equations together with equations \[9\] and \[10\] we integrate equation \[13\] to obtain \(S_1\). Given the above values of \(S_0\) and \(S_1\) we compute the transition probability, \(P(x_f, x_o)\), from the starting stable point, \(x_o\), to the saddle point, \(x_f\). Using equation \[9\] we can, therefore, find the transition rate for any large value of \(\Omega\). We now compare this calculation to the direct estimation of transition rates as described below.

From the master equation \[4\], it follows that the eigenvalues of \(H\) measure the decay rates of non-stationary states corresponding to eigenvectors of \(H\) with nonzero eigenvalues. The equilibrium state is represented by the “zero mode”, i.e., the eigenvector of \(H\) with zero eigenvalue, the existence of which is guaranteed by the transition matrix character of the Hamiltonian and conservation of probability. To compute the eigenvalues of the Hamiltonian, we write the master equation in discrete form, replacing the continuous concentration variables \((x_1, x_2)\) with a lattice with lattice parameter \(1/\Omega\). Although the system displays infinitely many states, typically, the gap between the real parts of the eigenvalues for first and second excited states is much larger than the absolute value of the real part of the first eigenvalue. This is because the gap between the first excited state and the second or the third excited states are governed by local relaxation rate around the two fixed points, but, the gap between the ground state and the first excited state is governed by the transition rate between the two stable fixed points. The local relaxation rates are order one in \(\Omega\), whereas, the transition rate is exponentially small for large \(\Omega\) (in practice, we find the ratios of the real parts to be about \(10^3\)). Thus an arbitrary probability distribution rapidly decays into a linear combination of the stationary state and the first excited state. Equivalently, the state could be described as a linear combination of two states, each representing a quasi-stationary distribution around a stable fixed point. From then on, we can
project the evolution to this two state system. If we start
with probability $p_o$ of being in the state $(1, 0)^T$, then the
Master equation gives:

$$\frac{d}{dt} \begin{pmatrix} p_o \\ p_f \end{pmatrix} = \begin{pmatrix} -r_{12} & r_{21} \\ r_{12} & -r_{21} \end{pmatrix} \begin{pmatrix} p_o \\ p_f \end{pmatrix}$$

The two-by-two effective transition matrix has columns
which sum to zero ensuring probability conservation.
Also, the trace $0 + \epsilon_1 = r_{12} + r_{21}$, where $\epsilon_1$ is the
eigenvalue of the first excited state. Therefore the first excited
eigenvalue will be the sum of the forward and backward
rates. In the case of the asymmetric systems, one rate
is usually far greater than the other. Consequently the
larger rate among $r_{12}$ and $r_{21}$ will be approximately given
by $\epsilon_1$, which we computed numerically using the Matlab
routine “eigs” for sparse matrices as well as by Lanczos
algorithm [26]. For a symmetric choice of parameters for
the two proteins, each rate is just $\epsilon_1/2$.

To explicitly extract the $S_0$ and $S_1$ contributions to the
rate from the Lanczos results, we re-scale the volume of
the system $\Omega \rightarrow \nu \Omega$ which, in turn, leads to a re-scaling
of rates of individual reaction events as $f(x) \rightarrow \nu f(x)$. As
a function of volume scale factor, $\nu$, the logarithm of
the rate has the form: $ln(rate) = S_0 \nu + b$, where $b$ includes
both $S_1$ and the logarithm of the pre-factor of $P(x_f, x_o)$
in Eq.4. The results and comparison with the Eikonal ap-
proximation are shown in Fig.2. The dotted line is a fit
to the data points obtained from calculation of the eigen-
values, and we see that the slope and intercept computed
from equations [15] are in good agreement with these
values. Note that, in this example, $S_1$ and the pre-factor
are significant contributions to the transition rate.

When we perform these calculations for the “standard”
model of the lambda switch [14, 27], we find a rate three
orders of magnitude higher than the observed rate of $10^{-7}$
per generation [28]. In retrospect, it is clear that accounting
for the stability of the lysogenic state requires a more
complex model which should include the effect of DNA
looping [17]. Whether the stability is due to suppression
of fluctuation or due to disappearance of the lytic “fixed”
point [20] remains an open question.

Optimal path methods are routinely used for studying
rare events related to failure of communication net-
works modeled as birth and death processes [31]. Such
large deviation methods are likely to be important in the
context of robustness and adaptability of biological net-
works. This paper illustrates the power of an approach to
fluctuations based on the Eikonal approximation to so-
lutions of the master equation. The scheme incorporates
large deviations in a natural way and provides a quanti-
tative method scalable to large networks. We also hope
that beyond being an efficient computational tool, this
method will provide further insight into to the stability
of epigenetic states of complex genetic networks.

Acknowledgment: AMS thanks John Little and John
Reinitz for useful discussions. AER and AMS acknowl-
edge the hospitality of the Kavli Institute of Theoretical
Physics and of Aspen Center for Physics while this work
was being done. This work was supported in part by an
NIH P20 Grant GM64375.

[1] C.V. Rao, D.M. Wolf, and A.P. Arkin Nature 420, 237
(2002). Erratum in: Nature 421, 190 (2003).
[2] T.B. Kepler and T.C. Elston, Biophys. J. 81, 3116
(2001).
[3] M. Thattai M and A. van Oudenaarden, Proc. Natl.
Acad. Sci. U.S.A. 98, 8614 (2001).
[4] E.M. Ozbudak, M. Thattai, I. Kurtser, A.D. Grossman,
and A. van Oudenaarden, Nat. Genet. 31, 69 (2002).
[5] M.B. Elowitz, A.J. Levine, E.D. Siggia, and P.S. Swain, Science 297, 1183 (2002).
[6] P.S. Swain, M.B. Elowitz, and E.D. Siggia, Proc. Natl. Acad. Sci. U.S.A. 99, 12795 (2002).
[7] W.J. Blake, M. Kærn, C.R. Cantor, and J.J. Collins, Nature 422, 633 (2003).
[8] P.B. Detwiler, S. Ramanathan, A. Sengupta, and B.I. Shraiman, Biophys. J. 79, 2801 (2000).
[9] N. Barkai and S. Leibler, Nature 403, 267 (2000).
[10] J.M. Vilar, H.Y. Kueh, N. Barkai and S. Leibler, Proc. Natl. Acad. Sci. U.S.A. 99, 5988 (2002).
[11] D. Gonze, J. Halloy, and P. Gaspard, J. Chem. Phys. 116, 10997 (2002).
[12] A.P. Arkin, J. Ross, and H.H. McAdams, Genetics 149, 1633 (1998).
[13] W. Bialek. in Advances in Neural Information Processing 13, edited by T.K. Leen, T.G. Dietterich, and V. Tresp, (MIT Press, Cambridge, 2001), p. 103.
[14] E. Aurell and K. Sneppen, Phys. Rev. Lett. 88, 048101 (2002).
[15] F.J. Isaacs, J. Hasty, C.R. Cantor, J.J. Collins, Proc. Natl. Acad. Sci. U. S. A. 100, 7714 (2003).
[16] X.M. Zhu, L. Yin, L. Hood, and P. Ao, Funct. Integr. Genomics, e-publication Feb 5 (2004).
[17] I.B. Dodd, K.E. Shearwin, A.J. Perkins, T. Burr, A. Hochschild, and J.B. Egan, Genes Dev. 18, 344 (2004).
[18] T.S. Gardner, C.R. Cantor, and J.J. Collins, Nature 403, 339 (2000).
[19] Hänngi, Talkner, and Borokovec. Rev. Mod. Phys. 62, 250 (1990).
[20] J.S. Langer, Phys. Rev. Lett. 21, 973 (1968).
[21] N.G. van Kampen, Stochastic Processes in Physics and Chemistry (North Holland, Amsterdam, 1992) 2nd ed.
[22] M.I. Dykman, E. Mori, J. Ross, and P. M. Hunt, J. Chem. Phys. 100, 5735 (1994).
[23] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical Recipes in C: The Art of Scientific Computing, (Cambridge University Press, New York, 1992) 2nd ed.
[24] B.J. Matkowsky, Z. Schuss, C. Knessl, C. Tier, and M. Mangel, Phys. Rev. A 29, 3359 (1984).
[25] R.S. Maier and D. L. Stein, Phys. Rev. E 48, 931 (1993).
[26] J.K. Cullum, R.A. Willoughby, Lanczos algorithms for large symmetric eigenvalue computations, (Birkhäuser, Boston, 1985).
[27] M.A. Shea and G.K. Ackers, J. Mol. Biol. 18, 211 (1985).
[28] J. W. Little, D.P. Shepley, and D.W Wert, EMBO J. 18, 4299-4307 (1999).
[29] M. Santillan and M.C. Mackey, Biophys. J. 86, 75 (2004).
[30] A. Shwartz and A. Weiss, Large Deviations for Performance Analysis: QUEUES, Communication, and Computing, Chapman and Hall, London, 1995.