Stochastic yield analysis of self-assembling, single-enzyme reaction networks

Yuliy Baryshnikov[1] Ed Coffman[2] Teddy Yimwadsana[2]

[1] Bell Labs, Lucent Technologies, Murray Hill, NJ 07974,
[2] Department of Electrical Engineering, Columbia University, NY 10027,

1st March 2005

1 Introduction

The goal of this note is to investigate yield patterns in a class of Reaction Networks (RN's) appearing naturally in many biochemical systems, and playing a crucial role in various designs of controllable self-assembling structures, nano-machines, and bio-computing.

In most of these schemes, a chain of reactions is triggered by a single activation agent, which may be interpreted in various settings as an enzyme, computational device, or a "motor" fueled by certain DNA-based complexes. The chemical kinetics of such RN's is extremely important and has been the subject of active research over the past few years in the context of DNA-based machines or computers, see e.g., [2, 10, 9, 7, 5, 6].

From the viewpoint of chemical kinetics, most of the research to date has centered on questions of sensitivity, response rates, and robustness of RN's, as they play the most notable role in applications of self-assembling structures in an ambient active system. By contrast, in this note, we concentrate on the "dual" question, namely that of the total yield of a reaction having certain well-defined products, or outputs, that result from the processing by an input agent. These yields have a central role in self assembly applications, where the final product is of utmost importance. As an example of such applications, we mention the bio-computing approaches to solving difficult (in particular, NP hard) computational problems using an exhaustive search, as epitomized by the seminal scheme proposed by Adleman [1]. For applications in this sector of nanotechnology the actual dynamics of the production has less significance than the amounts of the final product obtained.

The question of total yield in RN's has only scant coverage in the existing literature, which concentrates mostly on reversible networks with equilibria (or more general attractors) situated in the open positive orthant ([4] can be cited in this connection). Here, we offer an approach to computing total yield in single-enzyme networks using Markov-chain theory (see, e.g. [3]).

2 Single-enzyme reaction networks

As an archetypal example of such a reaction chain, we consider the catalytic DNA-based reaction loop described in [8]. Schematically, the RN is described by the equations

$$M + Q\overline{L} + L \underset{\mu}{\overset{\lambda}{\rightleftharpoons}} Q\overline{L}LM \tag{1}$$

$$Q\overline{L}LM \overset{\lambda^+}{\to} M + L\overline{L} + Q \tag{2}$$

$$Q\overline{L}LM \overset{\lambda^-}{\to} W \tag{3}$$

Here, M plays the role of an enzyme, or of a "motor" for which $Q\overline{L}$ and L serve as the fuel components, and $L\overline{L}$ and Q as the products. We include a denaturation reaction, which transforms the complex $Q\overline{L}LM$ into a "waste" product W.

The dynamics of this system is already nonlinear and does not yield to easy analysis. Numeric simulations, of course, do lead to an understanding of the basic metrics of the chain, but such an approach would require excessive computation for a multidimensional grid of parameters, where the main value of interest is the *limiting yield* of the reaction.

The system above is a special case of a much more general chain given by

$$X + Y \underset{\mu_1}{\overset{\lambda_o}{\rightleftharpoons}} C_1 \underset{\mu_2}{\overset{\lambda_1}{\rightleftharpoons}} C_2 \rightleftharpoons \ldots \rightleftharpoons C_{N-1} \underset{\mu_N}{\overset{\lambda_{N-1}}{\rightleftharpoons}} C_N \overset{\lambda_N}{\to} X + D \tag{4}$$

$$C_k \overset{\nu_k}{\to} W, \quad k = 1, \ldots, N \tag{5}$$

Here the catalyst is represented by X and the fuel by Y. The final product is D and we allow multiple intermediate stages $C_1, \ldots, C_N$, each of which has some rate of denaturation ν_k.

3 Exact results

In solving the general system of (4) and (5), we find the following formula for the limiting yield:

$$Z_\infty(x_0) = p \cdot \min\left(\frac{x_0}{1-p}, y_0\right) \tag{6}$$

where x_0 and y_0 are the respective initial values of X and Y, and p is the conditional probability that the continuous Markov chain shown in Figure 1 is absorbed in state $+$, given that it starts in state 0.

For example, the limiting yield of the reaction network (1)-(3), that is, the limiting concentration of $L\overline{L}$, can be given explicitly as

$$Z_\infty(x_0) = \left(\frac{\lambda^+}{\lambda^+ + \lambda^-}\right) \min\left(\left(\frac{\lambda^+}{\lambda^-}\right) x_0, y_0\right) \tag{7}$$

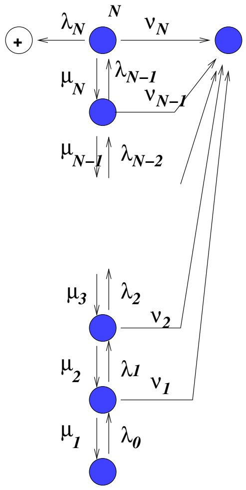

Figure 1: The generalized chain

In general, the probability, $p = p_0$, being sought is the first entry of the column-vector $P = (p_i),\ i = 0, \ldots, N$, of conditional probabilities of exiting at the $+$-state given the initial state. An elementary computation implies

$$P = A^{-1}B \tag{8}$$

where B is the column matrix all of whose elements are zero except the element in the last row, which is $-\lambda_N$, and where A is the generator given below:

$$A = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & . & . & . & 0 \\ \mu_1 & -\delta_1 & \lambda_1 & 0 & . & . & 0 \\ 0 & \mu_2 & -\delta_2 & \lambda_2 & 0 & . & 0 \\ . & . & . & . & . & . & . \\ 0 & . & 0 & 0 & \mu_{N-1} & -\delta_{N-1} & \lambda_{N-1} \\ 0 & . & . & 0 & 0 & \mu_N & -\delta_N \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ . \\ 0 \\ -\lambda_N \end{bmatrix}$$

where

$$\delta_k = \lambda_k + \mu_k + \nu_k.$$

System (8) yields

$$p_i = \frac{\alpha_i(N)}{\beta(N)} \qquad \text{for} \quad 0 \leq i \leq N \tag{9}$$

where α_i, β are the corresponding cofactors and determinant of A, respectively.

For example, for $N = 3$ the result is

$$\begin{aligned} \alpha_0(3) &= \lambda_0\lambda_1\lambda_2\lambda_3 \\ \alpha_1(3) &= \lambda_0\lambda_1\lambda_2\lambda_3 \\ \alpha_2(3) &= \lambda_0\lambda_1\lambda_2\lambda_3 + \lambda_0\nu_1\lambda_2\lambda_3 \\ \alpha_3(3) &= \lambda_0\lambda_1\lambda_2\lambda_3 + \lambda_0\nu_1\lambda_2\lambda_3 + \lambda_0\lambda_1\nu_2\lambda_3 + \lambda_0\nu_1\mu_2\lambda_3 \\ \beta(3) &= \lambda_0\lambda_1\lambda_2\lambda_3 + \lambda_0\nu_1\lambda_2\lambda_3 + \lambda_0\lambda_1\nu_2\lambda_3 + \lambda_0\nu_1\nu_2\lambda_3 + \lambda_0\lambda_1\lambda_2\nu_3 + \lambda_0\nu_1\lambda_2\nu_3 \\ &\quad + \lambda_0\lambda_1\nu_2\nu_3 + \lambda_0\nu_1\nu_2\nu_3 + \lambda_0\nu_1\mu_2\lambda_3\lambda_0\nu_1\mu_2\nu_3 + \lambda_0\lambda_1\nu_2\mu_3 + \lambda_0\nu_1\nu_2\mu_3 + \lambda_0\nu_1\mu_2\mu_3 \end{aligned} \tag{10}$$

4 Numerics

To verify our analysis, we performed numerical simulations of the catalysis reaction system, which can be rewritten as follows:

$$\begin{aligned} \dot{x} &= -\lambda_o xy + \mu_1 c_1 + \lambda_N c_N \\ \dot{y} &= -\lambda xy + \mu_1 c_1 \\ \dot{c_1} &= -(\lambda_1 + \mu_1 + \nu_1)c_1 + \lambda_o xy + \mu_2 c_2 \\ \dot{c_k} &= -(\lambda_k + \mu_k + \nu_k)c_k + \lambda_{k-1}c_{k-1} + \mu_{k+1}c_{k+1}, k = 2, \ldots, N-1 \\ \dot{c}_N &= -(\lambda_k + \mu_k + \nu_k)c_N + \lambda_{N-1}c_{N-1} \\ \dot{p} &= \lambda^N c_N \\ \dot{w} &= \sum_k \lambda_k c_k \end{aligned} \tag{11}$$

Solving the system of differential equations in (11) for $N = 1$ numerically, using the initial condition $X(0) = x_0$ and $Y(0) = y_0 = 1$ where x_0 is between 0 and 1.0, we obtain Figures 2 and 3 in which the parameter values are

$$\lambda = 0.9 \ \ \mu = 0.1 \ \ \lambda^+ = 0.4 \ \ \lambda^- = 0.1$$

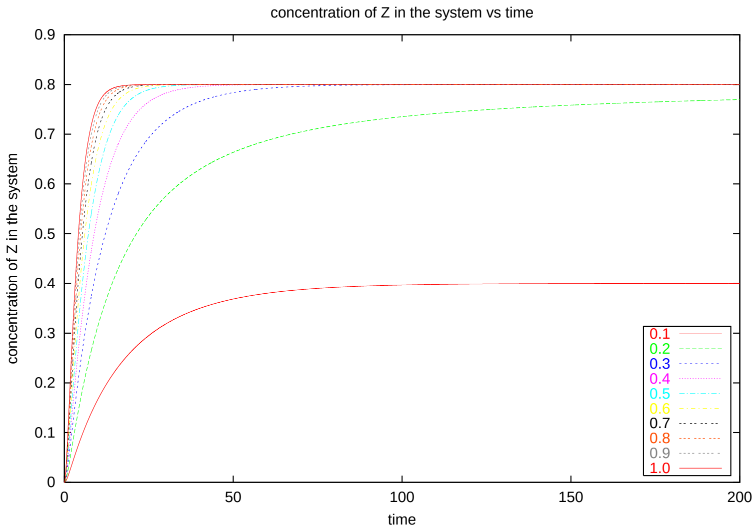

Figure 2: The relationship between the concentration of Z and x_0 for $0.1 \leq x_0 \leq 1$.

We can see a clear consistency of our exact results (7) with the numerics.

5 Final remark

We proposed a novel scheme based on the theory of Markov chains with absorbing states which allows us to predict the final yields for a class of reaction chain networks characterized by a single enzyme property. We conjecture that one can extend significantly the scope of this approach.

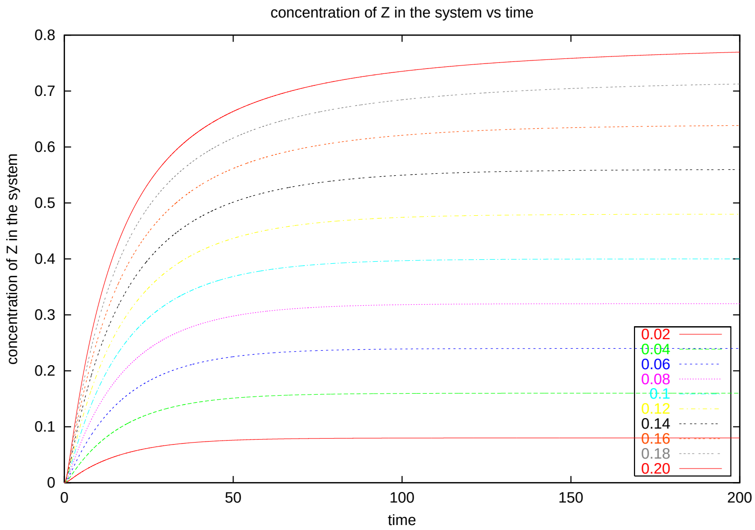

Figure 3: The relationship between the concentration of Z and x_0 for $0.02 \leq x_0 \leq 0.2$.

References

[1] L. Adleman. Molecular computation of solutions to a combinatorial problem. *Science*, 266:1021–1024, 1994.

[2] Bernard Yurke and Allen P. Mills, JR. Using DNA to Power Nanostructures. In *Genetic Programming and Evolvable Machines, 4*, June 2003.

[3] Erhan Cinlar. Introduction to stochastic processes, 1975.

[4] Martin Feinberg. Lectures on chemical reaction networks, Autumn 1979.

[5] D. Gracias, J. Tien, T. Breen, C. Hsu, and G. Whitesides. Forming electrical networks in three dimensions by self-assembly. *Science*, 289:1170–1173, 2000.

[6] C. Mao, T. LaBean, J.H. Reif, and N.C. Seeman. Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. 407(6803):493–496, 2000. also erratum 408(6813): 750, 2000.

[7] C. M. Niemeyer and M. Adler. Nanomechanical devices based on DNA. *Angewandte Chemie International Edition*, 41:3779–3783, 2002.

[8] A. J. Turberfield, J. C. Mitchell, B. Yurke, Jr. A. P. Mills, M. I. Blakey, and F. C. Simmel. DNA Fuel for Free-Running Nanomachines. In *Physical Review Letters 90(11)*, March 2003.

[9] E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman. Design and self assembly of two-dimensional DNA crystals. *Nature*, 394:539–544, 1998.

[10] Bernard Yurke, Andrew J. Turberfield, Allen P. Mills, Friedrich C. Simmel, and Jennifer L. Neumann. A DNA-fuelled molecular machine made of DNA. *Nature (Letters to Nature)*, 406:605–608, 2000.