Homework 3 (chapter 7 of Alon *"An Introduction to Systems Biology"* – and additional material covered in the lectures and available on the course web page)

Date delivered: May 9, 2014
Date due: May 19, 2014, 24.00
Mode of delivery: By e-mail, in PDF format.
Whereto: Mr Nicolas Innocenti <njain@kth.se>

This homework tests your ability to understand and evaluate systems biology models in the literature. In professional life you may have to identify and describe biological problems in situations where mathematical modeling and simulation are of added value.

The topic is the most recently published model of bacterial chemotaxis [1], and the goal is to check if this model has the property that the integrated response kernel is zero, and that it displays robust adaptation. A link to [1] is available on the course web page.

1 Background

The homework does not require taking into account the following survey of the pertinent literature, which is nevertheless given in this section, but for purpose of completeness and background only. The property that integrated response kernel in bacterial chemotaxis is (approximately) zero was descibed in Berg and co-workers in 1986 [2], that this is a theoretical problem was pointed out by P.G. de Gennes in 2004 [3]. An explanation of this property in terms of maximin strategies in bacteria playing "games against Nature" was given in [4], also partly repeated in Sect. 5 and Appendix of [1]. The bacterial response kernel to thermal flactuations (which does not have zero integral) was described in [5]. A mechanism for adaptation at the output of the chemotaxis pathway (the flagellar motor adapting to the level of CheY) was demonstrated by Berg and co-workers in [6]. This last paper shows that adaptation is possible without the methylation-demethylation pathway, albeit on fairly long time scales, and hence highlights the problem pointed out by de Gennes.

Part of the modelling approach of Celani, Shimizu and Vergassola is taken from an earlier study by Wingreen and collaborators [11] where there are more references to the literature. The earlier models on which the discussion in Alon Chapter 7 is based, and which do not consider the problem of zero-integral response, are Barkai & Leibler, 1997 [7], Alon, Surette, Barkai & Leibler, 1999 [8] and Yi, Huang, Simon & Doyle, 2000 [9].

The model

The model of [1] is given in eq. 11, where the various functions are listed in eq. 12. Most of the suggested parameter values are given in the paragraph below eq. 12.

The variable a is here activity of a receptor-CheA complex, which is assumed a function of ligand concentration L and methylation level m, $a = G(m, L)$. That this relationship is taken as a *function* means that it is assumed that the activity level of CheA quickly equilibrates to the given levels of L and m – which can both change, but more slowly. A comparison with eq. 11, which describes the result of the activity of CheA, and eq. 10 (third line), which describes how CheA is changed by the enzymatic activity of CheR and CheB, shows that the variable a is *the fraction of receptor-CheA complexes where CheA is phosphorylated*.

The function $f(m, L)$ in eq 12 is the sum of $f_m(m)$ defined in eq. 8 and $f_L(L)$ defined before eq. 4. The parameter n_α is the number of methyl-accepting receptors in a cluster that tend to be methylated and demethylated together; this value is supposed to be between 5 to 7 (the number of dimers in a cluster since receptors function as dimers).

The activity of CheY is described by an equation similar to the discussion in Alon

$$\frac{dy}{dt} = k_a a(1-y) - k_Z y \tag{1}$$

The variable y is *the fraction of CheY which is phosphorylated*. The parameter k_a is here the speed of phosphorylation of CheY by CheA following first-order kinetics, if all CheA would be phosphorylated ($a = 1$). The parameter k_z is analogously the dephosphorylation rate of CheY if all CheY would be be phosphorylated ($y = 1$). This rate is the sum of rate catalyzed by the phosphatase CheZ and spontaneous dephosphorylation of CheY. A way to write the above equation (1) for CheY more in line with the presentation in Alon would be ($y = [CheY - P]/[CheY^{TOT}]$ and $a = [CheA - P]/[CheA^{TOT}]$):

$$\begin{aligned} \frac{d[CheY-P]}{dt} &= k_+ a(1-y)[CheY^{TOT}][CheA^{TOT}] \\ &\quad -(k_-^{cat}[CheZ] + k_-^{auto}))[CheY^{TOT}]y \end{aligned} \tag{2}$$

Here $[CheY - P]$ is the concentration of phosphorylated CheY, $[CheY^{TOT}]$ is the total concentration of phosphorylated and unphosphorylated CheY, and $[CheA^{TOT}]$ is the total concentration of receptor-CheA complexes. All these concentrations would naturally be measured in M (mol/l), or in numbers/cell (1/cell being 10^{-9} M). The rate coefficients k_+ and k_-^{cat} would have dimension 1/time$*$concentration while k_-^{auto} would have dimension 1/time. The relationship between the combinations of rate coefficients in (2) and the notation used in Celani-Vergassola, and its suggested numerical values (after eq. 12), are

$$k_+[CheA^{TOT}] = k_a \approx 3s^{-1} \tag{3}$$

$$k_-^{cat}[CheZ] + k_-^{auto} = k_z \approx 2s^{-1} \tag{4}$$

The upshot of this discussion is that parameters k_a and k_z are not absolute constants. They will vary depending on the state of the cell, particularly the total concentration receptor-CheA complexes and the concentration of CheZ.

The methylation level m is supposed to change in time according to $\frac{dm}{dt} = F(m, a)$. The discussion in Section 2.4 (around eq. 10) shows that m is *the average number of methylated sites on a receptor-CheA complex*. It ranges from 0 to a maximum level M. For the *E coli* aspartate receptor – one of the most studied receptors and which mediates movement towards the attractants aspartate and maltose [10] – the number M is equal to 4. The function F, given in eq. 10 and eq. 12, hence describes the net effect of the methylation-demethylation pathway. It can generally be expected to be different for different receptors. We note that the two rate constants k_r and k_b have dimension 1/time $*$ concentration, but that the values given below eq. 12 are for the two products $k_r[CheR] = 0.1s^{-1}$ and $k_b[CheB] = 0.2s^{-1}$. As before, these two combinations are not absolute constants but will vary depending on the state of the cell, here on the concetrations of CheR and CheB. The particular form on how $F(a, m)$ depends on m is not discussed in detail in [1]. We note that if the two constants K_R and K_B in $F(m, a)$ are zero, then F does not depend on m. This is the essence of the Barkai-Leibler model, that the net methylation rate of receptor-CheA complexes only depends on the fraction of activated receptor-CheA complexes. The idea that there must be some dependence on m is taken from the earlier modelling study [11] (cited in [1]). Namely, if the average methylation level is m then there must be some chance (naturally depending on m) for a receptor to be already maximally methylated, and for such receptors the phosphorylation rate actually must be zero. Similarly, there must also be some chance (depending on m) for a receptor to not be methylated at all, and then the demethylation rate for such receptors must be zero. An model of the Barkai-Leibler type, embodied by a first-order kinetics equation

$$\frac{dm}{dt} = k_r[CheR](1 - a) - k_b[CheB]a \qquad \text{(similar to Alon eq P7.11)} \qquad (5)$$

therefore strictly cannot be true for m very close to 0 or very close to M. We note that Alon eq. 7.4.1 and eq. 7.4.14 assume that CheR works at saturation (zero-order kinetics) while CheB acts according to Michaelis-Menten kinetics. This does not change the argument that such an equation cannot strictly hold for m close to 0 or M. It should however be noted that Alon argues that the assumption that CheR works at saturation is essential while this assumption is replaced in [1] by the particular form of eq. 10.

The output variable p_r is the instantaneous probability that the bacterium is in the "running state" and is defined in eq. 14. In a homogeneous environment p_r equals $tau_r/(\tau_r + \tau_t)$ where τ_r is the average run duration and τ_t the average tumble duration. By eq. 13 (second line), eq 3 and eq 11 (last line) this equilibrium value is $p_r(y) = h(y) = (1 + (\frac{y}{y_0})^H)^{-1}$. The parameter y_0 is defined right after eq. 3 as CheY-P concentration at which the equilbrium probability that the motor is in the running state (the bias) is equal to $\frac{1}{2}$. Its numerical value $y_0 = 0.4$ is given after eq. 12. The parameter H is also defined right after eq. 3 as the steepness of the response (Hill coefficient) and its value is given after eq. 12 as 10. In an inhomogeneous environment the p_r will change in response to the CheY activity and tau_r and τ_t) will be averages over situations with different values of p_r. The time constant τ_t by which p_r changes is unfortunately not given in the paper, but one can guess that is should stand for the time of a tumble, which is $0.1s$.

The full model considered by Celani, Shimizu and Vergassola is then

$$
\begin{aligned}
a &= G(m, L) && (6)\\
\frac{dm}{dt} &= F(a, m) && (7)\\
\frac{dy}{dt} &= k_a a(1-y) - k_z y && (8)\\
\frac{dp_r}{dt} &= \frac{1}{\tau_t}(1 - \frac{p_r}{h(y)}) && (9)
\end{aligned}
$$

where the functions are

$$
\begin{aligned}
G(m, L) &= (1 + e^{f(m,L)})^{-1} && (10)\\
f(m, L) &= n_\alpha \alpha_m (m_0 - m) + n_\alpha \log \frac{1 + L/K_{off}}{1 + L/K_{on}} && (11)\\
F(a, m) &= k_r[CheR](1-a)\frac{M-m}{M-m-K_R} - k_b[CheB]a\frac{m}{m+K_B} && (12)
\end{aligned}
$$

$$
h(y) = \left(1 + (\frac{y}{y_0})^H\right) \qquad (13)
$$

and the various parameters are as discussed in [1] and above.

The assignment

Consider the above model and answer the following questions:

1. Write down the full equations using the parameter values given in [1] and/or above. If you find any parameter missing, comment, and introduce a value for that parameter you find reasonable.

2. Implement the equations in a suitable computer program. Does the long-term limit of the solutions to these equations tend is a steady state, *i.e.* that all the concentrations tend to a limit for suitable long times? Test this with varying levels of the ligand concentration.

3. The same as the preceeding using two different values of the parameter k_B and two different values of the parameter k_R (in total four different parameter values). How do these fixed point compare to the analytical expressions given in [1] eq 13? If you find any differences, comment.

4. Verify that if ligand concentration changes quickly from one value l_1 to another value l_2, then the level of activation p_r changes. Describe how. Does the model have the property of *adaptation*, that the level level of activation in steady state (for long times) is independent of the ligand concentration?

5. Verify that the model displays *robust adaptation*. That is, show that the model still shows adaptation if some of the rate coefficients are changed. Do the tests with the same four different values of the parameters k_B and k_R as above. Show that *adaption time* is not robust in this model.

6. The paper of [1] focuses on linear response *i.e.* the effect of (relatively) small changes in ligand concentration. Describe what you think would be the effects of large changes in ligand concentration. Can you see these effects in a simulation? Comment.

7. According to the discussion above the rate of change dm/dt of the average methylation should depend on m when m is very close to zero or very close to its maximum value. Describe a modification of the standard model in which this could happen. Simulate the model with such parameter values. Does the model display adaptation at these parameter values?

   Hint: A simple way to get a low average methylation is to make the methylation rate sufficiently small, as would be the case if the concentration of $[CheR]$ in above is sufficiently low.

8. Translate the previous point into a prediction of what would happen if one would make a suitable mutatation of the E coli chemotaxis pathway. Do a literature search and see if you can find this mutation described. If so, comment.

   Hint: One approach would be keyword search on Pubmed, another would be to identify the protein of interest and check from a suitable data base. Other approaches are also possible. You are only expected to make a reasonable attempt at this question. An exhaustive search is not required.

References

[1] A. Celani, T. Shimizu, M. Vergassola, "Molecular and Functional Aspects of Bacterial Chemotaxis", *J. Statistical Physics* (2011) **144**:219-240.

[2] J. Segall, S. Block, H. Berg, "Temporal comparisons in bacterial chemotaxis" *PNAS* (1986) **83**: 8987-8991.

[3] P.-G. de Gennes, "Chemotaxis: the role of internal delays", *Eur Biophys J* (2004) **33**: 691-693.

[4] A. Celani, M. Vergassola, "Bacterial strategies for chemotaxis response" *PNAS* (2010) **107**:1391-96.

[5] E. Paster, W. Ryu, "The thermal impulse response of *Escherichia coli*", *PNAS* (2008) **105**:5373-5377.

[6] J. Yuan, R. Branch, B. Hosu, H. Berg, "Adaptation at the output of the chemotaxis signalling pathway", *Nature* (2012) **484**: 233.

[7] N. Barkai, S. Leibler, "Robustness in simple biochemical networks." *Nature* (1997) **387**:913-917.

[8] U. Alon, M. Surette, N. Barkai, S. Leibler, "Robustness in bacterial chemotaxis" *Nature* (1999) **397**:168-171.

[9] T.-M. Yi, Y. Huang, M.I. Simon, J. Doyle J. "Robust perfect adaptation in bacterial chemotaxis through integral feedback control." *PNAS* (2000) **97**:4649-4653.

[10] "Aspartate receptor", wikipedia, *http* : *//en.wikipedia.org/wiki/Aspartate_receptor*.

[11] Clinton H. Hansen, Robert G. Endres, Ned S. Wingreen "Chemotaxis in *Escherichia coli*: A molecular motor for robust precise adaptation", *PLoS Computational Biology* **4** e1 (2008).