Biomolecular Feedback Systems

Domitilla Del Vecchio
MIT

Richard M. Murray
Caltech

Version 1.0b, September 14, 2014

This is the electronic edition of *Biomolecular Feedback Systems*, available from
http://www.cds.caltech.edu/~murray/BFSwiki.

Printed versions are available from Princeton University Press,
http://press.princeton.edu/titles/10285.html.

Chapter 5
Biological Circuit Components

In this chapter, we describe some simple circuit components that have been constructed in *E. coli* cells using the technology of synthetic biology and then consider a more complicated circuit that already appears in natural systems to implement adaptation. We will analyze the behavior of these circuits employing mainly the tools from Chapter 3 and some of the tools from Chapter 4. The basic knowledge of Chapter 2 will be assumed.

5.1 Introduction to biological circuit design

In Chapter 2 we introduced a number of core processes and models for those processes, including gene expression, transcriptional regulation, post-translational regulation such as covalent modification of proteins, allosteric regulation of enzymes, and activity regulation of transcription factors through inducers. These core processes provide a rich set of functional building blocks, which can be combined together to create circuits with prescribed functionalities.

For example, if we want to create an inverter, a device that returns high output when the input is low and vice versa, we can use a gene regulated by a transcriptional repressor. If we want to create a signal amplifier, we can employ a cascade of covalent modification cycles. Specifically, if we want the amplifier to be linear, we should tune the amounts of protein substrates to be smaller than the Michaelis-Menten constants. Alternatively, we could employ a phosphotransfer system, which provides a fairly linear input/output relationship for an extended range of the input stimulation. If instead we are looking for an almost digital response, we could employ a covalent modification cycle with high amounts of substrates compared to the Michaelis-Menten constants. Furthermore, if we are looking for a fast input/output response, phosphorylation cycles are better candidates than transcriptional systems.

In this chapter and the next we illustrate how one can build circuits with prescribed functionality using some of the building blocks of Chapter 2 and the design techniques illustrated in Chapter 3. We will focus on two types of circuits: gene circuits and signal transduction circuits. In some cases, we will illustrate designs that incorporate both.

A gene circuit is usually depicted by a set of nodes, each representing a gene, connected by unidirectional edges, representing a transcriptional activation or a re-

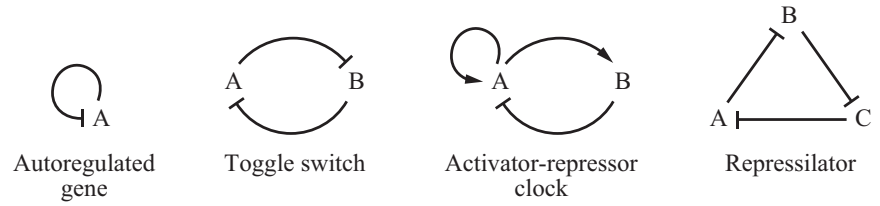


Figure 5.1: Early gene circuits that have been fabricated in bacteria *E. coli*: the negatively autoregulated gene [10], the toggle switch [30], the activator-repressor clock [6], and the repressilator [26].

pression. If gene z represses the expression of gene x, the interaction is represented by $Z \dashv X$. If instead gene z activates the expression of gene x, the interaction is represented by $Z \rightarrow X$. Inducers will often appear as additional nodes, which activate or inhibit a specific edge. Early examples of such circuits include an autoregulated circuit [10], a toggle switch obtained by connecting two inverters in a ring fashion [30], an activator-repressor system that can display toggle switch or clock behavior [6], and a loop oscillator called the repressilator obtained by connecting three inverters in a ring topology [26] (Figure 5.1).

Basic synthetic biology technology

Simple synthetic gene circuits can be constituted from a set of (connected) transcriptional components, which are made up by the DNA base pair sequences that compose the desired promoters, ribosome binding sites, gene coding region, and terminators. We can choose these components from a library of basic parts, which are classified based on biochemical properties such as affinity (of promoter, operator, or ribosome binding sites), strength (of a promoter), and efficiency (of a terminator).

The desired sequence of parts is usually assembled on plasmids, which are circular pieces of DNA, separate from the host cell chromosome, with their own origin of replication. These plasmids are then inserted, through a process called transformation in bacteria and transfection in yeast, in the host cell. Once in the host cell, they express the proteins they code for by using the transcription and translation machinery of the cell. There are three main types of plasmids: low copy number (5-10 copies), medium copy number (15-20 copies), and high copy number (up to hundreds). The copy number reflects the average number of copies of the plasmid inside the host cell. The higher the copy number, the more efficient the plasmid is at replicating itself. The exact number of plasmids in each cell fluctuates stochastically and cannot be exactly controlled.

In order to measure the amounts of proteins of interest, we make use of *reporter genes*. A reporter gene codes for a protein that fluoresces in a specific color (red, blue, green, or yellow, for example) when it is exposed to light of the appropriate

wavelength. For instance, green fluorescent protein (GFP) is a protein with the property that it fluoresces in green when exposed to UV light. It is produced by the jellyfish *Aequoria victoria* and its gene has been isolated so that it can be used as a reporter. Other fluorescent proteins, such as yellow fluorescent protein (YFP) and red fluorescent protein (RFP), are genetic variations of GFP.

A reporter gene is usually inserted downstream of the gene expressing the protein whose concentration we want to measure. In this case, both genes are under the control of the same promoter and are transcribed into a single mRNA molecule. The mRNA is then translated to protein and the two proteins will be fused together. This technique provides a direct way to measure the concentration of the protein of interest but can affect the functionality of this protein because some of its regulatory sites may be occluded by the fluorescent protein. Another viable technique is one in which the reporter gene is placed under the control of the same promoter that is also controlling the expression of the protein of interest. In this case, the production rates of the reporter and of the protein of interest are the same and, as a consequence, the respective concentrations should mirror each other. The reporter thus provides an indirect measurement of the concentration of the protein of interest.

Just as fluorescent proteins can be used as a readout of a circuit, inducers function as external inputs that can be used to probe the system. Two commonly used negative inducers are IPTG and aTc, as explained in Section 2.3, while two common positive inducers are arabinose and AHL. Arabinose activates the transcriptional activator AraC, which activates the pBAD promoter. Similarly, AHL is a signaling molecule that activates the LuxR transcription factor, which activates the pLux promoter.

Protein dynamics can usually be altered by the addition of a degradation tag at the end of the corresponding coding region. A degradation tag is a sequence of base pairs that adds an amino acid sequence to the functional protein that is recognized by proteases. Proteases then bind to the protein, degrading it into a non-functional molecule. As a consequence, the half-life of the protein decreases, resulting in an increased decay rate. Degradation tags are often employed to obtain a faster response of the protein concentration to input stimulation and to prevent protein accumulation.

5.2 Negative autoregulation

In this section, we analyze the negatively autoregulated gene of Figure 5.1 and focus on analyzing how the presence of the negative feedback affects the dynamics and the noise properties of the system. This system was introduced in Example 2.2.

Let A be a transcription factor repressing its own production. Assuming that the mRNA dynamics are at the quasi-steady state, the ODE model describing the

negatively autoregulated system is given by

$$\frac{dA}{dt} = \frac{\beta}{1+(A/K)^n} - \gamma A. \tag{5.1}$$

We seek to compare the behavior of this autoregulated system, which we also refer to as the closed loop system, to the behavior of the unregulated one:

$$\frac{dA}{dt} = \beta_0 - \gamma A,$$

in which β_0 is the unrepressed production rate. We refer to this second system as the open loop system.

Dynamic effects of negative autoregulation

As we showed via simulation in Example 2.2, negative autoregulation speeds up the response to perturbations. Hence, the time the system takes to reach its equilibrium decreases with negative feedback. In this section, we illustrate how this result can be analytically demonstrated by employing linearization. Specifically, we linearize the system about its equilibrium point and calculate the time response resulting from initializing the linearized system at an initial condition close to the equilibrium point.

Let $A_e = \beta_0/\gamma$ be the equilibrium of the unregulated system and let $z = A - A_e$ denote the perturbation with respect to such an equilibrium. The dynamics of z are given by

$$\frac{dz}{dt} = -\gamma z.$$

Given a small initial perturbation z_0, the response of z is given by the exponential

$$z(t) = z_0 e^{-\gamma t}.$$

The "half-life" of the signal $z(t)$ is the time $z(t)$ takes to reach half of z_0 and we denote it by t_{half}. This is a common measure for the speed of response of a system to an initial perturbation. Simple mathematical calculation shows that $t_{\text{half}} = \ln(2)/\gamma$. Note that the half-life does not depend on the production rate β_0 and only depends on the protein decay rate constant γ.

Now let A_e be the steady state of the negatively autoregulated system (5.1). Assuming that the perturbation z with respect to the equilibrium is small enough, we can employ linearization to describe the dynamics of z. These dynamics are given by

$$\frac{dz}{dt} = -\bar{\gamma} z,$$

where

$$\bar{\gamma} = \gamma + \beta \frac{n A_e^{n-1}/K^n}{(1+(A_e/K)^n)^2}.$$

In this case, we have that $t_{\text{half}} = \ln(2)/\bar{\gamma}$.

Since $\bar{\gamma} > \gamma$ (for any positive value of A_e), we have that the dynamic response to a perturbation is faster in the system with negative autoregulation. This confirms the simulation findings of Example 2.2.

Noise filtering

We next investigate the effect of the negative autoregulation on the noisiness of the system. In order to do this, we employ the Langevin modeling framework and determine the frequency response to the intrinsic noise on the various reactions. In particular, in the analysis that follows we treat Langevin equations as regular ordinary differential equations with inputs, allowing us to apply the tools described in Chapter 3.

We perform two different studies. In the first one, we assume that the decay rate of the protein is much slower than that of the mRNA. As a consequence, the mRNA concentration can be well approximated by its quasi-steady state value and we focus on the dynamics of the protein only. In the second study, we investigate the consequence of having the mRNA and protein decay rates in the same range so that the quasi-steady state assumption cannot be made. This can be the case, for example, when degradation tags are added to the protein to make its decay rate larger. In either case, we study both the open loop system and the closed loop system (the system with negative autoregulation) and compare the corresponding frequency responses to noise.

Assuming that mRNA is at its quasi-steady state

In this case, the reactions for the open loop system are given by

$$\texttt{R1}: \quad \mathrm{p} \xrightarrow{\beta_0} \mathrm{A} + \mathrm{p}, \qquad \texttt{R2}: \quad \mathrm{A} \xrightarrow{\gamma} \emptyset,$$

in which β_0 is the constitutive production rate, p is the DNA promoter, and γ is the decay rate of the protein. Since the concentration of DNA promoter p is not changed by these reactions, it is a constant, which we call p_{tot}.

Employing the Langevin equation (4.11) of Section 4.1 and letting n_{A} denote the real-valued number of molecules of A and n_{p} the real-valued number of molecules of p, we obtain

$$\frac{dn_{\mathrm{A}}}{dt} = \beta_0 n_{\mathrm{p}} - \gamma n_{\mathrm{A}} + \sqrt{\beta_0 n_{\mathrm{p}}}\,\Gamma_1 - \sqrt{\gamma n_{\mathrm{A}}}\,\Gamma_2,$$

in which Γ_1 and Γ_2 depend on the noise on the production reaction and on the decay reaction, respectively. By letting $A = n_{\mathrm{A}}/\Omega$ denote the concentration of A and $p = n_{\mathrm{p}}/\Omega = p_{\text{tot}}$ denote the concentration of p, we have that

$$\frac{dA}{dt} = \beta_0 p_{\text{tot}} - \gamma A + \frac{1}{\sqrt{\Omega}}\left(\sqrt{\beta_0 p_{\text{tot}}}\,\Gamma_1 - \sqrt{\gamma A}\,\Gamma_2\right). \tag{5.2}$$

This is a linear system and therefore we can calculate the frequency response to any of the two inputs Γ_1 and Γ_2. In particular, the frequency response to input Γ_1 has magnitude given by

$$M^o(\omega) = \frac{\sqrt{\beta_0 p_{\text{tot}}/\Omega}}{\sqrt{\omega^2+\gamma^2}}. \tag{5.3}$$

We now consider the autoregulated system. The reactions are given by

$$\begin{aligned} &\texttt{R1}: \quad \text{p} \xrightarrow{\beta} \text{A} + \text{p}, && \texttt{R2}: \quad \text{A} \xrightarrow{\gamma} \emptyset, \\ &\texttt{R3}: \quad \text{A} + \text{p} \xrightarrow{a} \text{C}, && \texttt{R4}: \quad \text{C} \xrightarrow{d} \text{A} + \text{p}. \end{aligned}$$

Defining $p_{\text{tot}} = p + C$ and employing the Langevin equation (4.11) of Section 4.1, we obtain

$$\begin{aligned} \frac{dp}{dt} &= -aAp + d(p_{\text{tot}} - p) + \frac{1}{\sqrt{\Omega}}(-\sqrt{aAp}\,\Gamma_3 + \sqrt{d(p_{\text{tot}}-p)}\,\Gamma_4), \\ \frac{dA}{dt} &= \beta p - \gamma A - aAp + d(p_{\text{tot}} - p) + \frac{1}{\sqrt{\Omega}}(\sqrt{\beta p}\,\Gamma_1 - \sqrt{\gamma A}\,\Gamma_2 - \sqrt{aAp}\,\Gamma_3 \\ &\qquad + \sqrt{d(p_{\text{tot}}-p)}\,\Gamma_4), \end{aligned}$$

in which Γ_3 and Γ_4 are the noises corresponding to the association and dissociation reactions, respectively. Letting $K_{\text{d}} = d/a$,

$$N_1 = \frac{1}{\sqrt{\Omega}}(-\sqrt{Ap/K_{\text{d}}}\,\Gamma_3 + \sqrt{(p_{\text{tot}}-p)}\,\Gamma_4), \qquad N_2 = \frac{1}{\sqrt{\Omega}}(\sqrt{\beta p}\,\Gamma_1 - \sqrt{\gamma A}\,\Gamma_2),$$

we can rewrite the above system in the following form:

$$\begin{aligned} \frac{dp}{dt} &= -aAp + d(p_{\text{tot}} - p) + \sqrt{d}N_1(t), \\ \frac{dA}{dt} &= \beta p - \gamma A - aAp + d(p_{\text{tot}} - p) + N_2(t) + \sqrt{d}N_1(t). \end{aligned}$$

Since $d \gg \gamma,\beta$, this system displays two time scales. Letting $\epsilon := \gamma/d$ and defining $y := A - p$, the system can be rewritten in standard singular perturbation form (3.24):

$$\begin{aligned} \epsilon\frac{dp}{dt} &= -\gamma Ap/K_{\text{d}} + \gamma(p_{\text{tot}} - p) + \sqrt{\epsilon}\,\sqrt{\gamma}N_1(t), \\ \frac{dy}{dt} &= \beta p - \gamma(y + p) + N_2(t). \end{aligned}$$

By setting $\epsilon = 0$, we obtain the quasi-steady state value $p = p_{\text{tot}}/(1 + A/K_{\text{d}})$. Writing $\dot{A} = \dot{y} + \dot{p}$, using the chain rule for $\dot{p}$, and assuming that $p_{\text{tot}}/K_{\text{d}}$ is sufficiently small, we obtain the reduced system describing the dynamics of A as

$$\frac{dA}{dt} = \beta\frac{p_{\text{tot}}}{1 + A/K_{\text{d}}} - \gamma A + \frac{1}{\sqrt{\Omega}}\left(\sqrt{\beta\frac{p_{\text{tot}}}{1 + A/K_{\text{d}}}}\,\Gamma_1 - \sqrt{\gamma A}\,\Gamma_2\right) =: f(A,\Gamma_1,\Gamma_2). \tag{5.4}$$

The equilibrium point for this system corresponding to the mean values $\Gamma_1 = 0$ and $\Gamma_2 = 0$ of the inputs is given by

$$A_e = \frac{1}{2}\left(\sqrt{K_{\rm d}^2 + 4\beta p_{\rm tot} K_{\rm d}/\gamma} - K_{\rm d}\right).$$

The linearization of the system about this equilibrium point is given by

$$\left.\frac{\partial f}{\partial A}\right|_{A_e,\Gamma_1=0,\Gamma_2=0} = -\beta \frac{p_{\rm tot}/K_{\rm d}}{(1+A_e/K_{\rm d})^2} - \gamma =: -\bar{\gamma},$$

$$b_1 = \left.\frac{\partial f}{\partial \Gamma_1}\right|_{A_e,\Gamma_1=0,\Gamma_2=0} = \frac{1}{\sqrt{\Omega}}\sqrt{\frac{\beta p_{\rm tot}}{1+A_e/K_{\rm d}}}, \quad b_2 = \left.\frac{\partial f}{\partial \Gamma_2}\right|_{A_e,\Gamma_1=0,\Gamma_2=0} = -\frac{1}{\sqrt{\Omega}}\sqrt{\gamma A_e}.$$

Hence, the frequency response to Γ_1 has magnitude given by

$$M^c(\omega) = \frac{b_1}{\sqrt{\omega^2 + \bar{\gamma}^2}}. \tag{5.5}$$

In order to make a fair comparison between this response and that of the open loop system, we need the equilibrium points of both systems to be the same. In order to guarantee this, we set β such that

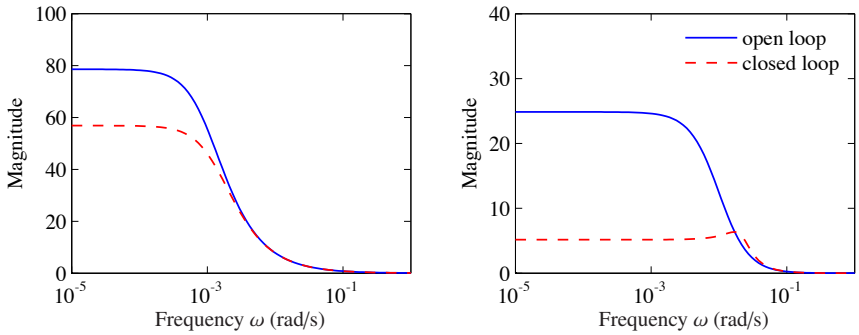
$$\frac{\beta}{1+A_e/K_{\rm d}} = \beta_0.$$

This can be attained, for example, by properly adjusting the strength of the promoter and of the ribosome binding site. As a consequence, we have that $b_1 = \sqrt{\beta_0 p_{\rm tot}/\Omega}$. Since we also have that $\bar{\gamma} > \gamma$, comparing expressions (5.3) and (5.5) it follows that $M^c(\omega) < M^o(\omega)$ for all ω. That is, the magnitude of the frequency response of the closed loop system is smaller than that of the open loop system at all frequencies. Hence, negative autoregulation attenuates noise at all frequencies. The two frequency responses are plotted in Figure 5.2a. A similar result could be obtained for the frequency response with respect to the input Γ_2 (see Exercise 5.1).

mRNA decay close to protein decay

In the case in which mRNA and protein decay rates are comparable, we need to model both the processes of transcription and translation. Letting $\mathrm{m_A}$ denote the mRNA of A, the reactions describing the open loop system modify to

$$\begin{aligned} &\texttt{R1}: \quad \mathrm{m_A} \xrightarrow{\kappa} \mathrm{m_A} + \mathrm{A}, && \texttt{R2}: \quad \mathrm{A} \xrightarrow{\gamma} \emptyset, \\ &\texttt{R5}: \quad \mathrm{p} \xrightarrow{\alpha_0} \mathrm{m_A} + \mathrm{p}, && \texttt{R6}: \quad \mathrm{m_A} \xrightarrow{\delta} \emptyset, \end{aligned}$$



(a) Slow protein decay (b) Enhanced protein decay

Figure 5.2: Effect of negative autoregulation on noise propagation. (a) Magnitude of the frequency response to noise $\Gamma_1(t)$ for both open loop and closed loop systems for the model in which mRNA is assumed at its quasi-steady state. The parameters are $p_{\text{tot}} = 10$ nM, $K_{\text{d}} = 10$ nM, $\beta = 0.001$ min^{-1}, $\gamma = 0.001$ min^{-1}, and $\beta_0 = 0.00092$ min^{-1}, and we assume unit volume Ω. (b) Frequency response to noise $\Gamma_6(t)$ for both open loop and closed loop for the model in which mRNA decay is close to protein decay. The parameters are $p_{\text{tot}} = 10$ nM, $K_{\text{d}} = 10$ nM, $\alpha = 0.001$ min^{-1}, $\beta = 0.01$ min^{-1}, $\delta = 0.01$ min^{-1}, $\gamma = 0.01$ min^{-1}, and $\alpha_0 = 0.0618$ min^{-1}.

while those describing the closed loop system become

$$\begin{aligned} &\texttt{R1}: \quad \mathrm{m_A} \xrightarrow{\kappa} \mathrm{m_A} + \mathrm{A}, && \texttt{R2}: \quad \mathrm{A} \xrightarrow{\gamma} \emptyset, \\ &\texttt{R3}: \quad \mathrm{A} + \mathrm{p} \xrightarrow{a} \mathrm{C}, && \texttt{R4}: \quad \mathrm{C} \xrightarrow{d} \mathrm{A} + \mathrm{p}, \\ &\texttt{R5}: \quad \mathrm{p} \xrightarrow{\alpha} \mathrm{m_A} + \mathrm{p}, && \texttt{R6}: \quad \mathrm{m_A} \xrightarrow{\delta} \emptyset. \end{aligned}$$

Defining $p_{\text{tot}} = p + C$, employing the Langevin equation, and applying singular perturbation as performed before, we obtain the dynamics of the system as

$$\begin{aligned} \frac{dm_{\mathrm{A}}}{dt} &= F(A) - \delta m_{\mathrm{A}} + \frac{1}{\sqrt{\Omega}}\left(\sqrt{F(A)}\,\Gamma_5 - \sqrt{\delta m_{\mathrm{A}}}\,\Gamma_6\right), \\ \frac{dA}{dt} &= \kappa m_{\mathrm{A}} - \gamma A + \frac{1}{\sqrt{\Omega}}\left(\sqrt{\kappa m_A}\,\Gamma_1 - \sqrt{\gamma A}\,\Gamma_2\right), \end{aligned}$$

in which Γ_5 and Γ_6 model the noise on the production reaction and decay reaction of mRNA, respectively. For the open loop system we have $F(A) = \alpha_0 p_{\text{tot}}$, while for the closed loop system we have the Hill function

$$F(A) = \frac{\alpha p_{\text{tot}}}{1 + A/K_{\text{d}}}.$$

The equilibrium point for the open loop system is given by

$$m_e^o = \frac{\alpha_0 p_{\text{tot}}}{\delta}, \qquad A_e^o = \frac{\kappa \alpha_0 p_{\text{tot}}}{\delta \gamma}.$$

Considering Γ_6 as the input of interest, the linearization of the system at this equilibrium is given by

$$A^o = \begin{pmatrix} -\delta & 0 \\ \kappa & -\gamma \end{pmatrix}, \qquad B^o = \begin{pmatrix} \sqrt{\delta m_e^o/\Omega} \\ 0 \end{pmatrix}.$$

Letting $K = \kappa/(\gamma K_\mathrm{d})$, the equilibrium for the closed loop system is given by

$$A_e^c = \frac{\kappa m_e^c}{\gamma}, \qquad m_e^c = \frac{1}{2}\left(-1/K + \sqrt{(1/K)^2 + 4\alpha p_\mathrm{tot}/(K\delta)}\right).$$

The linearization of the closed loop system at this equilibrium point is given by

$$A^c = \begin{pmatrix} -\delta & -g \\ \kappa & -\gamma \end{pmatrix}, \qquad B^c = \begin{pmatrix} \sqrt{\delta m_e^c/\Omega} \\ 0 \end{pmatrix}, \tag{5.6}$$

in which $g = (\alpha p_\mathrm{tot}/K_\mathrm{d})/(1 + A_e^c/K_\mathrm{d})^2$ represents the contribution of the negative autoregulation. The larger the value of g—obtained, for example, by making K_d smaller (see Exercise 5.2)—the stronger the negative autoregulation.

In order to make a fair comparison between the open loop and closed loop system, we again set the equilibrium points to be the same. To do this, we choose α such that $\alpha/(1 + A_e^c/K_\mathrm{d}) = \alpha_0$, which can be done by suitably changing the strengths of the promoter and ribosome binding site.

The open loop and closed loop transfer functions are thus given by

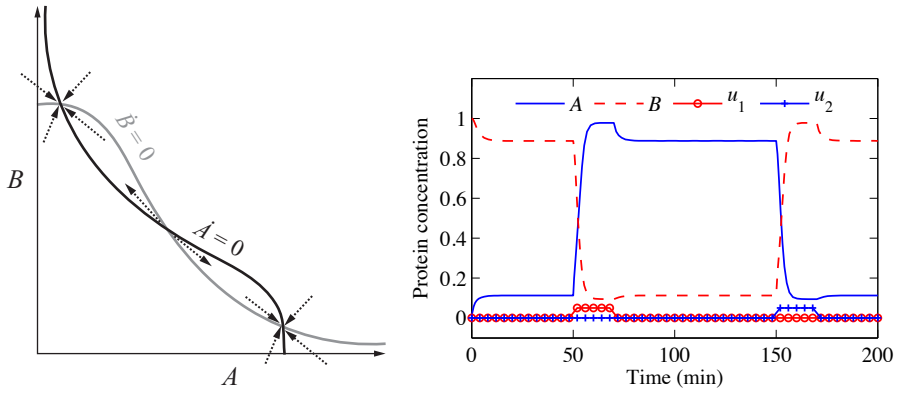
$$G^o_{A\Gamma_6}(s) = \frac{\kappa\sqrt{\delta m_e/\Omega}}{(s+\delta)(s+\gamma)}, \qquad G^c_{A\Gamma_6}(s) = \frac{\kappa\sqrt{\delta m_e/\Omega}}{s^2 + s(\delta+\gamma) + \delta\gamma + \kappa g}.$$

From these expressions, it follows that the open loop transfer function has two real poles, while the closed loop transfer function can have complex conjugate poles when g is sufficiently large. As a consequence, noise Γ_6 can be amplified at sufficiently high frequencies. Figure 5.2b shows the magnitude $M(\omega)$ of the corresponding frequency responses for both the open loop and the closed loop systems.

It follows that the presence of negative autoregulation attenuates noise with respect to the open loop system at low frequency, but it can amplify noise at higher frequency. This is a very well-studied phenomenon known as the "waterbed effect," according to which negative feedback decreases the effect of disturbances at low frequency, but it can amplify it at higher frequency. This effect is not found in first-order models, as demonstrated by the derivations performed when mRNA is at the quasi-steady state. This illustrates the spectral shift of the frequency response to intrinsic noise towards the high frequency, as also experimentally reported [7].

5.3 The toggle switch

The toggle switch is composed of two genes that mutually repress each other, as shown in the diagram of Figure 5.1. We start by describing a simple model with no



(a) Nullclines (b) Time traces

Figure 5.3: Genetic toggle switch. (a) Nullclines for the toggle switch. By analyzing the direction of the vector field in the proximity of the equilibria, one can deduce their stability as described in Section 3.1. (b) Time traces for $A(t)$ and $B(t)$ when inducer concentrations $u_1(t)$ and $u_2(t)$ are changed. The plots show a scaled version of these signals, whose absolute values are $u_1 = u_2 = 1$ in the indicated intervals of time. In the simulation, we have $n = 2$, $K_{d,1} = K_{d,2} = 1$ nM, $K = \sqrt{0.1}$ nM, $\beta = 1$ hrs^{-1}, and $\gamma = 1$ hrs^{-1}.

inducers. By assuming that the mRNA dynamics are at the quasi-steady state, we obtain a two-dimensional differential equation model given by

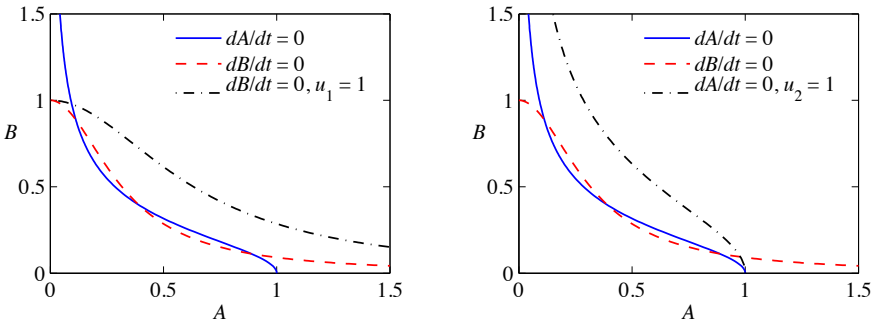
$$\frac{dA}{dt} = \frac{\beta}{1+(B/K)^n} - \gamma A, \qquad \frac{dB}{dt} = \frac{\beta}{1+(A/K)^n} - \gamma B,$$

in which we have assumed for simplicity that the parameters of the repression functions are the same for A and B.

Since the system is two-dimensional, both the number and stability of equilibria can be analyzed by performing nullcline analysis (see Section 3.1). Specifically, by setting $dA/dt = 0$ and $dB/dt = 0$ and letting $n \geq 2$, we obtain the nullclines shown in Figure 5.3a. The nullclines intersect at three points, which determine the equilibrium points of this system. The stability of these equilibria can be determined by the following graphical reasoning.

The nullclines partition the plane into six regions. By determining the sign of dA/dt and dB/dt in each of these six regions, we can determine the direction in which the vector field is pointing in each of these regions. From these directions, we can deduce that the equilibrium occurring for intermediate values of A and B (at which $A = B$) is unstable while the other two are stable (see the arrows in Figure 5.3a). Hence, the toggle switch is a bistable system.

The system trajectories converge to one equilibrium or the other depending on whether the initial condition is in the region of attraction of the first or the second equilibrium. The 45-degree line divides the plane into the two regions of attraction of the stable equilibrium points. Once the system's trajectory has converged to



(a) Adding inducer u_1 (b) Adding inducer u_2

Figure 5.4: Genetic toggle switch with inducers. (a) Nullclines for the toggle switch (solid and dashed lines) and how the nullcline $\dot{B} = 0$ changes when inducer u_1 is added (dash-dotted line). (b) Nullclines for the toggle switch (solid line) and how the nullcline $\dot{A} = 0$ changes when inducer u_2 is added (dotted line). Parameter values are as in Figure 5.3.

one of the two equilibrium points, it cannot switch to the other unless an external (transient) stimulation is applied.

In the genetic toggle switch developed by Gardner et al. [30], external stimulations were added in the form of negative inducers for A and B. Specifically, let u_1 be the negative inducer for A and u_2 be the negative inducer for B. Then, as we have seen in Section 2.3, the expressions of the Hill functions need to be modified to replace A by $A(1/(1+u_1/K_{d,1}))$ and B by $B(1/(1+u_2/K_{d,2}))$, in which $K_{d,1}$ and $K_{d,2}$ are the dissociation constants of u_1 with A and of u_2 with B, respectively. Hence, the system dynamics become

$$\frac{dA}{dt} = \frac{\beta}{1+(B/K_B(u_2))^n} - \gamma A, \qquad \frac{dB}{dt} = \frac{\beta}{1+(A/K_A(u_1))^n} - \gamma B,$$

in which we have let $K_A(u_1) = K(1+u_1/K_{d,1})$ and $K_B(u_2) = K(1+u_2/K_{d,2})$ denote the effective K values of the Hill functions. We show in Figure 5.3b time traces for $A(t)$ and $B(t)$ when the inducer concentrations are changed. The system starts from initial conditions in which B is high and A is low without inducers. Then, at time 50 the system is presented with a short pulse in u_2, which causes A to rise since it prevents B to repress A. As A rises, B is repressed and hence B decreases to the low value and the system state switches to the other stable steady state. The system remains in this steady state also after the inducer u_2 is removed until another transient stimulus is presented at time 150. At this time, there is a pulse in u_1, which inhibits the ability of A to repress B and, as a consequence, B rises, thus repressing A, and the system returns to its original steady state.

Note that the effect of the inducers in this model is that of temporarily changing the shape of the nullclines by increasing the values of K_A and K_B. Specifically, high values of u_1 with $u_2 = 0$ will lead to increased values of K_A, which will shift the

point of half-maximal value of the Hill function $\beta/(1+(A/K_A)^n)$ to the right. As a consequence, the nullclines will intersect at one point only, in which the value of B is high and the value of A is low (Figure 5.4a). The opposite will occur when u_2 is high and $u_1 = 0$, leading to only one intersection point in which B is low and A is high (Figure 5.4b).

5.4 The repressilator

Elowitz and Leibler constructed an oscillatory genetic circuit consisting of three repressors arranged in a ring fashion and called it the "repressilator" [26] (Figure 5.1). The repressilator exhibits sinusoidal, limit cycle oscillations in periods of hours, slower than the cell-division time. Therefore, the state of the oscillator is transmitted between generations from mother to daughter cells.

A dynamical model of the repressilator can be obtained by composing three transcriptional modules in a loop fashion. The dynamics can be written as

$$\begin{aligned} \frac{dm_{\mathrm{A}}}{dt} &= F_1(C) - \delta m_{\mathrm{A}}, & \frac{dm_{\mathrm{B}}}{dt} &= F_2(A) - \delta m_{\mathrm{B}}, & \frac{dm_{\mathrm{C}}}{dt} &= F_3(B) - \delta m_{\mathrm{C}}, \\ \frac{dA}{dt} &= \kappa m_{\mathrm{A}} - \gamma A, & \frac{dB}{dt} &= \kappa m_{\mathrm{B}} - \gamma B, & \frac{dC}{dt} &= \kappa m_{\mathrm{C}} - \gamma C, \end{aligned} \tag{5.7}$$

where we take

$$F_1(P) = F_2(P) = F_3(P) = F(P) = \frac{\alpha}{1+(P/K)^n},$$

and assume initially that the parameters are the same for all the three repressor modules. The structure of system (5.7) belongs to the class of cyclic feedback systems that we have studied in Section 3.3. In particular, the Mallet-Paret and Smith Theorem 3.5 and Hastings et al. Theorem 3.4 can be applied to infer that if the system has a unique equilibrium point and this equilibrium is unstable, then the system admits a periodic solution. Therefore, to apply these results, we determine the number of equilibria and their stability.

The equilibria of the system can be found by setting the time derivatives to zero. Letting $\beta = (\kappa/\delta)$, we obtain

$$A_{eq} = \frac{\beta F_1(C_{eq})}{\gamma}, \qquad B_{eq} = \frac{\beta F_2(A_{eq})}{\gamma}, \qquad C_{eq} = \frac{\beta F_3(B_{eq})}{\gamma},$$

which combined together yield

$$A_{eq} = \frac{\beta}{\gamma} F_1\left(\frac{\beta}{\gamma} F_3\left(\frac{\beta}{\gamma} F_2(A_{eq})\right)\right) =: g(A_{eq}).$$

The solution to this equation determines the set of equilibria of the system. The number of equilibria is given by the number of crossings of the two functions

$h_1(A) = g(A)$ and $h_2(A) = A$. Since h_2 is strictly monotonically increasing, we obtain a unique equilibrium if h_1 is monotonically decreasing. This is the case when $g'(A) = dg(A)/dA < 0$, otherwise there could be multiple equilibrium points. Since we have that

$$\text{sign}(g'(A)) = \prod_{i=1}^{3} \text{sign}(F_i'(A)),$$

it follows that if $\Pi_{i=1}^{3}\text{sign}(F_i'(A)) < 0$ the system has a unique equilibrium. We call the product $\Pi_{i=1}^{3}\text{sign}(F_i'(A))$ the *loop sign*.

It follows that any cyclic feedback system with negative loop sign will have a unique equilibrium. In the present case, system (5.7) is such that $F_i' < 0$, so that the loop sign is negative and there is a unique equilibrium. We next study the stability of this equilibrium by studying the linearization of the system.

Letting P denote the equilibrium value of the protein concentrations for A, B, and C, the Jacobian matrix of the system is given by

$$J = \begin{pmatrix} -\delta & 0 & 0 & 0 & 0 & F_1'(P) \\ \kappa & -\gamma & 0 & 0 & 0 & 0 \\ 0 & F_2'(P) & -\delta & 0 & 0 & 0 \\ 0 & 0 & \kappa & -\gamma & 0 & 0 \\ 0 & 0 & 0 & F_3'(P) & -\delta & 0 \\ 0 & 0 & 0 & 0 & \kappa & -\gamma \end{pmatrix},$$

whose characteristic polynomial is given by

$$\det(sI - J) = (s+\gamma)^3(s+\delta)^3 - \kappa^3 \prod_{i=1}^{3} F_i'(P). \tag{5.8}$$

The roots of this characteristic polynomial are given by

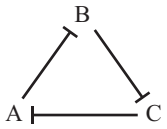
$$(s+\gamma)(s+\delta) = r,$$

in which $r \in \{\kappa F'(P), -(\kappa F'(P)/2)(1 - i\sqrt{3}), -(\kappa F'(P)/2)(1 + i\sqrt{3})\}$ and $i = \sqrt{-1}$ represents the imaginary unit. In order to invoke Hastings et al. Theorem 3.4 to infer the existence of a periodic orbit, it is sufficient that one of the roots of the characteristic polynomial has positive real part. This is the case if

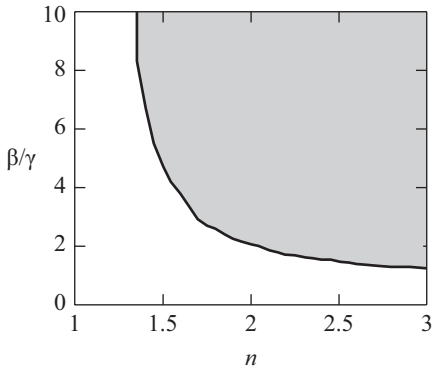
$$\kappa|F'(P)| > 2\gamma\delta, \qquad |F'(P)| = \alpha \frac{n(P^{n-1}/K^n)}{(1 + (P/K)^n)^2},$$

in which P is the equilibrium value satisfying the equilibrium condition

$$P = \frac{\beta}{\gamma} \frac{\alpha}{1 + (P/K)^n}.$$



(a) Repressilator



(b) Parameter space

Figure 5.5: Parameter space for the repressilator. (a) Repressilator diagram. (b) Space of parameters that give rise to oscillations. Here, we have set $K = 1$ for simplicity.

One can plot the pair of values $(n,\beta/\gamma)$ for which the above two conditions are satisfied. This leads to the plot of Figure 5.5b. When n increases, the existence of an unstable equilibrium point is guaranteed for larger ranges of β/γ. Of course, this "behavioral" robustness does not guarantee that other important features of the oscillator, such as the period, are not changed when parameters vary.

A similar result for the existence of a periodic solution can be obtained when two of the Hill functions are monotonically increasing and only one is monotonically decreasing:

$$F_1(P) = \frac{\alpha}{1+(P/K)^n}, \qquad F_2(P) \quad = \frac{\alpha(P/K)^n}{1+(P/K)^n}, \qquad F_3(P) = \frac{\alpha(P/K)^n}{1+(P/K)^n}.$$

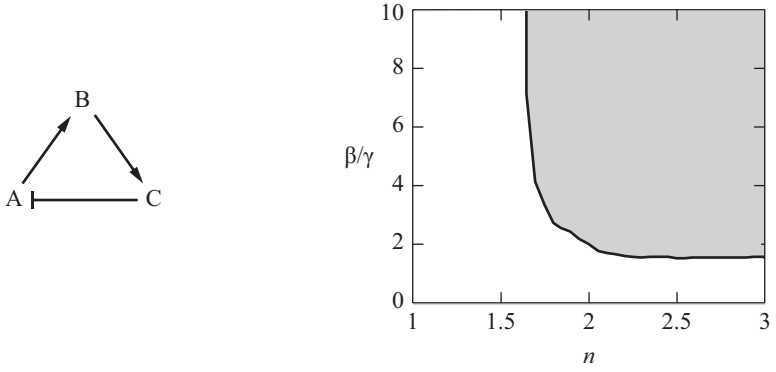
That is, two interactions are activations and one is a repression. We refer to this as the "non-symmetric" design. Since the loop sign is still negative, there is only one equilibrium point. We can thus obtain the condition for oscillations again by establishing conditions on the parameters that guarantee that at least one root of the characteristic polynomial (5.8) has positive real part, that is,

$$\kappa(|F_1'(P_3)F_2'(P_1)F_3'(P_2)|)^{(1/3)} > 2\gamma\delta, \tag{5.9}$$

in which P_1, P_2, P_3 are the equilibrium values of A, B, and C, respectively. These equilibrium values satisfy:

$$P_2 = \frac{\beta}{\gamma}\frac{(P_1/K)^n}{1+(P_1/K)^n}, \qquad P_3 = \frac{\beta}{\gamma}\frac{(P_2/K)^n}{1+(P_2/K)^n}, \qquad P_1(1+(P_3/K)^n) = \frac{\beta}{\gamma}.$$

Using these expressions numerically and checking for each combination of the parameters $(n,\beta/\gamma)$ whether (5.9) is satisfied, we can plot the combinations of n and β/γ values that lead to an unstable equilibrium. This is shown in Figure 5.6b.



(a) Loop oscillator (b) Parameter space

Figure 5.6: Parameter space for a loop oscillator. (a) Oscillator diagram. (b) Space of parameters that give rise to oscillations. As the value of n is increased, the range of the other parameter for which a periodic cycle exists becomes larger. Here, we have set $K = 1$.

From this figure, we can deduce that the qualitative shape of the parameter space that leads to a limit cycle is the same in the repressilator and in the non-symmetric design. One can conclude that it is then possible to design the circuit such that the parameters land in the filled region of the plots.

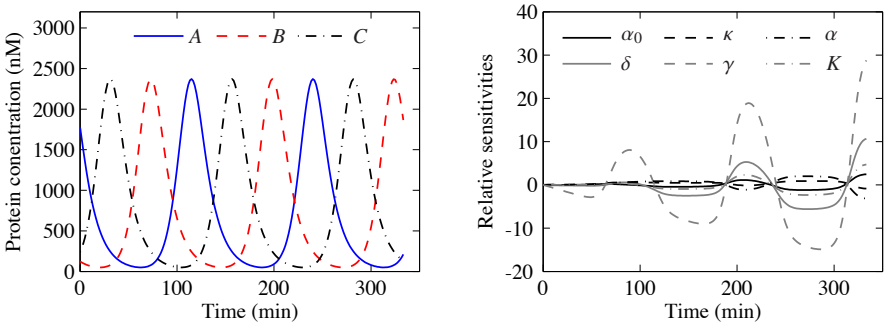
In practice, values of the Hill coefficient n between one and two can be obtained by employing repressors that have cooperativity higher than or equal to two. There are plenty of such repressors, including those originally used in the repressilator design [26]. However, values of n greater than two may be hard to reach in practice. To overcome this problem, one can include more elements in the loop. In fact, it is possible to show that the value of n sufficient for obtaining an unstable equilibrium decreases when the number of elements in the loop is increased (see Exercise 5.6). Figure 5.7a shows a simulation of the repressilator.

In addition to determining the space of parameters that lead to periodic trajectories, it is also relevant to determine the parameters to which the system behavior is the most sensitive. To address this question, we can use the parameter sensitivity analysis tools of Section 3.2. In this case, we model the repressilator Hill functions adding the basal expression rate as it was originally done in [26]:

$$F_1(P) = F_2(P) = F_3(P) = \frac{\alpha}{1 + (P/K)^n} + \alpha_0.$$

Letting $x = (m_A, A, m_B, B, m_C, C)$ and $\theta = (\alpha_0, \delta, \kappa, \gamma, \alpha, K)$, we can compute the sensitivity $S_{x,\theta}$ along the limit cycle corresponding to nominal parameter vector θ_0 as illustrated in Section 3.2:

$$\frac{dS_{x,\theta}}{dt} = M(t, \theta_0) S_{x,\theta} + N(t, \theta_0),$$



(a) Protein concentration (b) Sensitivity

Figure 5.7: Repressilator parameter sensitivity analysis. (a) Protein concentrations as functions of time. (b) Sensitivity plots. The most important parameters are the protein and mRNA decay rates γ and δ. Parameter values used in the simulations are $\alpha = 800$ nM/s, $\alpha_0 = 5 \times 10^{-4}$ nM/s, $\delta = 5.78 \times 10^{-3}$ s^{-1}, $\gamma = 1.16 \times 10^{-3}$ s^{-1}, $\kappa = 0.116$ s^{-1}, $n = 2$, and $K = 1600$ nM.

where $M(t, \theta_0)$ and $N(t, \theta_0)$ are both periodic in time. If the dynamics of $S_{x,\theta}$ are stable then the resulting solutions will be periodic, showing how the dynamics around the limit cycle depend on the parameter values. The results are shown in Figure 5.7, where we plot the steady state sensitivity of A as a function of time. We see, for example, that the limit cycle depends strongly on the protein degradation and dilution rate δ, indicating that changes in this value can lead to (relatively) large variations in the magnitude of the limit cycle.

5.5 Activator-repressor clock

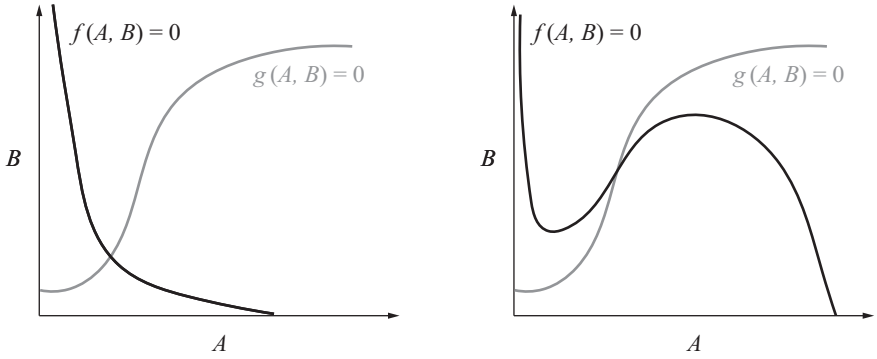
Consider the activator-repressor clock diagram shown in Figure 5.1. The activator A takes two inputs: the activator A itself and the repressor B. The repressor B has the activator A as the only input. Let $\mathrm{m_A}$ and $\mathrm{m_B}$ represent the mRNA of the activator and of the repressor, respectively. Then, we consider the following four-dimensional model describing the rate of change of the species concentrations:

$$\begin{aligned} \frac{dm_\mathrm{A}}{dt} &= F_1(A,B) - \delta_A m_\mathrm{A}, & \frac{dm_\mathrm{B}}{dt} &= F_2(A) - \delta_B m_\mathrm{B}, \\ \frac{dA}{dt} &= \kappa_A m_\mathrm{A} - \gamma_A A, & \frac{dB}{dt} &= \kappa_B m_\mathrm{B} - \gamma_B B, \end{aligned} \tag{5.10}$$

in which the functions F_1 and F_2 are Hill functions and given by

$$F_1(A,B) = \frac{\alpha_A (A/K_A)^n + \alpha_{A0}}{1 + (A/K_A)^n + (B/K_B)^m}, \qquad F_2(A) = \frac{\alpha_B (A/K_A)^n + \alpha_{B0}}{1 + (A/K_A)^n}.$$

The Hill function F_1 can be obtained through a combinatorial promoter, where there are sites both for an activator and for a repressor. The Hill function F_2 has the



(a) $n = 1$ (b) $n = 2$

Figure 5.8: Nullclines for the two-dimensional system (5.11). Graph (a) shows the only possible configuration of the nullclines when $n = 1$. Graph (b) shows a possible configuration of the nullclines when $n = 2$. In this configuration, there is a unique equilibrium, which can be unstable.

form considered for an activator when transcription can still occur at a basal level even in the absence of an activator (see Section 2.3).

We first assume the mRNA dynamics to be at the quasi-steady state so that we can perform two-dimensional analysis and invoke the Poincaré-Bendixson theorem (Section 3.3). Then, we analyze the four-dimensional system and perform a bifurcation study.

Two-dimensional analysis

We let $f_1(A,B) := (\kappa_A/\delta_A)F_1(A,B)$ and $f_2(A) := (\kappa_B/\delta_B)F_2(A)$. For simplicity, we also define $f(A,B) := -\gamma_A A + f_1(A,B)$ and $g(A,B) := -\gamma_B B + f_2(A)$ so that the two-dimensional system is given by

$$\frac{dA}{dt} = f(A,B), \qquad \frac{dB}{dt} = g(A,B). \tag{5.11}$$

To simplify notation, we set $K_A = K_B = 1$ and take $m = 1$, without loss of generality as similar results can be obtained when $m > 1$ (see Exercise 5.7).

We first study whether the system admits a periodic solution for $n = 1$. To do so, we analyze the nullclines to determine the number and location of steady states. Let $\bar{\alpha}_A = \alpha_A(\kappa_A/\delta_A)$, $\bar{\alpha}_B = \alpha_B(\kappa_B/\delta_B)$, $\bar{\alpha}_{A0} = \alpha_{A0}(\kappa_A/\delta_A)$, and $\bar{\alpha}_{B0} = \alpha_{B0}(\kappa_B/\delta_B)$. Then, $g(A,B) = 0$ leads to

$$B = \frac{\bar{\alpha}_B A + \bar{\alpha}_{B0}}{(1+A)\gamma_B},$$

which is an increasing function of A. Setting $f(A,B) = 0$, we obtain that

$$B = \frac{\bar{\alpha}_A A + \bar{\alpha}_{A0} - \gamma_A A(1+A)}{\gamma_A A},$$

which is a monotonically decreasing function of A. These nullclines are displayed in Figure 5.8a.

We see that we have only one equilibrium point. To determine the stability of the equilibrium, we calculate the linearization of the system at such an equilibrium. This is given by the Jacobian matrix

$$J = \begin{pmatrix} \dfrac{\partial f}{\partial A} & \dfrac{\partial f}{\partial B} \\ \dfrac{\partial g}{\partial A} & \dfrac{\partial g}{\partial B} \end{pmatrix}.$$

In order for the equilibrium to be unstable and not a saddle, it is necessary and sufficient that $\mathrm{tr}(J) > 0$ and $\det(J) > 0$. Graphical inspection of the nullclines at the equilibrium (see Figure 5.8a) shows that

$$\left.\frac{dB}{dA}\right|_{f(A,B)=0} < 0.$$

By the implicit function theorem (Section 3.5), we further have that

$$\left.\frac{dB}{dA}\right|_{f(A,B)=0} = -\frac{\partial f/\partial A}{\partial f/\partial B},$$

so that $\partial f/\partial A < 0$ because $\partial f/\partial B < 0$. As a consequence, we have that $\mathrm{tr}(J) < 0$ and hence the equilibrium point is either stable or a saddle.

To determine the sign of $\det(J)$, we further inspect the nullclines and find that

$$\left.\frac{dB}{dA}\right|_{g(A,B)=0} > \left.\frac{dB}{dA}\right|_{f(A,B)=0}.$$

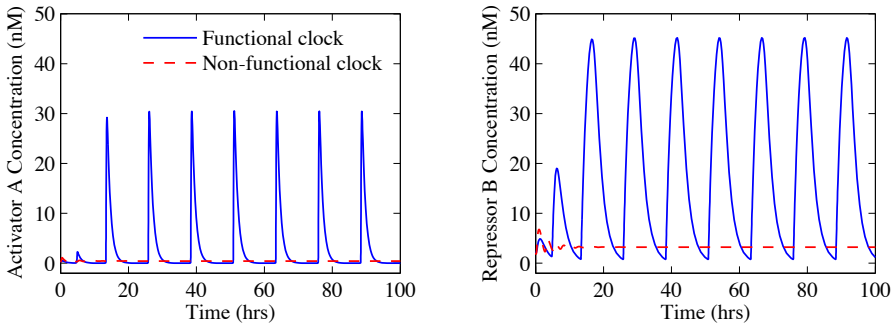
Again using the implicit function theorem we have that

$$\left.\frac{dB}{dA}\right|_{g(A,B)=0} = -\frac{\partial g/\partial A}{\partial g/\partial B},$$

so that $\det(J) > 0$. Hence, the ω-limit set (Section 3.3) of any point in the plane is necessarily not part of a periodic orbit. It follows that to guarantee that any initial condition converges to a periodic orbit, we need to require that $n > 1$.

We now study the case $n = 2$. In this case, the nullcline $f(A, B) = 0$ changes and can have the shape shown in Figure 5.8b. In the case in which, as in the figure, there is only one equilibrium point and the nullclines both have positive slope at the intersection (equivalent to $\det(J) > 0$), the equilibrium is unstable and not a saddle if $\mathrm{tr}(J) > 0$. This is the case when

$$\frac{\gamma_B}{\partial f_1/\partial A - \gamma_A} < 1.$$



(a) Activator time traces (b) Repressor time traces

Figure 5.9: Effect of the trace of the Jacobian on the stability of the equilibrium. The above plots illustrate the trajectories of system (5.11) for both a functional ($\mathrm{tr}(J) > 0$) and a non-functional ($\mathrm{tr}(J) < 0$) clock. The parameters in the simulation are $\delta_A = 1 = \delta_B = 1$ hrs^{-1}, $\alpha_A = 250$ nM/hrs, $\alpha_B = 30$ nM/hrs, $\alpha_{A0} = .04$ nM/hrs, $\alpha_{B0} = .004$ nM/hrs, $\gamma_A = 1$ hrs^{-1}, $\kappa_A = \kappa_B = 1$ hrs^{-1}, $K_A = K_B = 1$ nM, $n = 2$ and $m = 4$. In the functional clock, $\gamma_B = 0.5$ hrs^{-1}, whereas in the non-functional clock, $\gamma_B = 1.5$ hrs^{-1}.

This condition reveals the crucial design requirement for the functioning of the clock. Specifically, the repressor B time scale must be sufficiently slower than the activator A time scale. This point is illustrated in the simulations of Figure 5.9, in which we see that if γ_B is too large, the trace becomes negative and oscillations disappear.

Four-dimensional analysis

In order to deepen our understanding of the role of time scale separation between activator and repressor dynamics, we perform a time scale analysis employing the bifurcation tools described in Section 3.4. To this end, we consider the following four-dimensional model describing the rate of change of the species concentrations:

$$\begin{aligned} \frac{dm_{\mathrm{A}}}{dt} &= F_1(A,B) - (\delta_A/\epsilon)\, m_{\mathrm{A}}, & \frac{dm_{\mathrm{B}}}{dt} &= F_2(A) - (\delta_B/\epsilon)\, m_{\mathrm{B}}, \\ \frac{dA}{dt} &= \nu\big((\kappa_A/\epsilon)\, m_{\mathrm{A}} - \gamma_A A\big), & \frac{dB}{dt} &= (\kappa_B/\epsilon)\, m_{\mathrm{B}} - \gamma_B B. \end{aligned} \tag{5.12}$$

This system is the same as system (5.10), where we have explicitly introduced two parameters ν and ϵ that model time scale differences, as follows. The parameter ν determines the relative time scale between the activator and the repressor dynamics. As ν increases, the activator dynamics become faster compared to the repressor dynamics. The parameter ϵ determines the relative time scale between the protein and mRNA dynamics. As ϵ becomes smaller, the mRNA dynamics become faster compared to protein dynamics and model (5.12) becomes close to the two-dimensional model (5.11), in which the mRNA dynamics are considered at

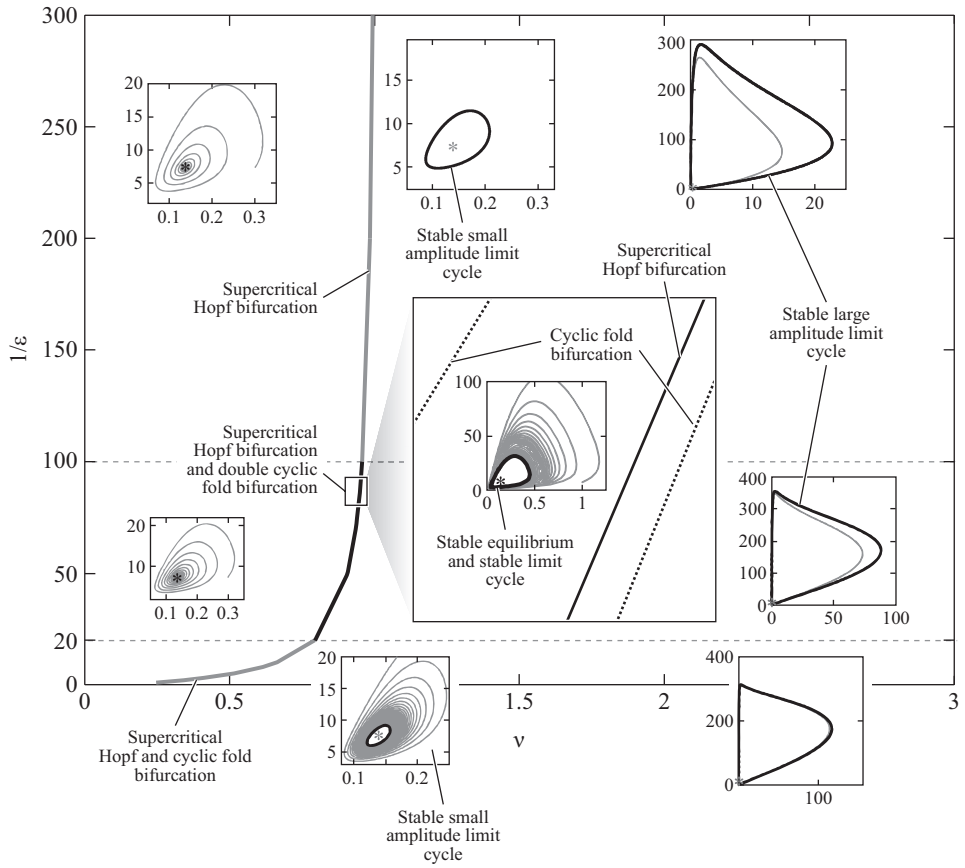
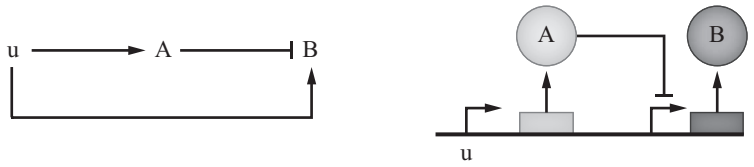


Figure 5.10: Design chart for the activator-repressor clock. We obtain sustained oscillations past the Hopf bifurcation point for values of ν sufficiently large independent of the difference of time scales between the protein and the mRNA dynamics. We also notice that there are values of ν for which a stable equilibrium point and a stable limit cycle coexist and values of ν for which two stable limit cycles coexist. The interval of ν values for which two stable limit cycles coexist is too small to be able to numerically set ν in such an interval. Thus, this interval is not practically relevant. The values of ν for which a stable equilibrium and a stable periodic orbit coexist are instead relevant. Figure adapted from [21].

the quasi-steady state. Thus, ϵ is a singular perturbation parameter. In particular, equations (5.12) can be taken to standard singular perturbation form by considering the change of variables $\bar{m}_\mathrm{A} = m_\mathrm{A}/\epsilon$ and $\bar{m}_\mathrm{B} = m_\mathrm{B}/\epsilon$. The details on singular perturbation can be found in Section 3.5.

The values of ϵ and of ν do not affect the number of equilibria of the system but they do determine the stability of the equilibrium points. We thus perform bifurcation analysis with ϵ and ν as the two bifurcation parameters. The bifurcation analysis results are summarized by Figure 5.10. In terms of the ϵ and ν parameters, it is thus possible to design the system as follows: if the activator dynamics are



(a) Incoherent feedforward loop (b) Genetic implementation

Figure 5.11: The incoherent feedforward loop (a) with a possible implementation (b). The circuit is integrated on a DNA plasmid denoted u. Protein A is under the control of a constitutive promoter in the DNA plasmid u, while B is repressed by A. Protein B, in turn, is also expressed by a gene in the plasmid u. Hence B is also "activated" by u.

sufficiently sped up with respect to the repressor dynamics, the system undergoes a Hopf bifurcation (Hopf bifurcation was introduced in Section 3.3) and a stable periodic orbit arises. The chart illustrates that for intermediate values of $1/\epsilon$, more complicated dynamic behaviors can arise in which a stable equilibrium coexists with a stable limit cycle. This situation corresponds to the *hard excitation* condition [60] and occurs for realistic values of the separation of time scales between protein and mRNA dynamics. Therefore, this simple oscillator motif described by a four-dimensional model can capture interesting dynamic behaviors, including features that lead to the long-term suppression of a rhythm by external inputs.

From a fabrication point of view, the activator dynamics can be sped up by adding suitable degradation tags to the activator protein. Similarly, the repressor dynamics can be slowed down by adding repressor DNA binding sites (see Chapter 6 and the effects of retroactivity on dynamic behavior).

5.6 An incoherent feedforward loop (IFFL)

In Section 3.2, we described various mechanisms to obtain robustness to external perturbations. In particular, one such mechanism is provided by incoherent feedforward loops. Here, we describe an implementation that was proposed for making the equilibrium values of protein expression robust to perturbations in DNA plasmid copy number [14]. In this implementation, the input u is the amount of DNA plasmid coding for both the intermediate regulator A and the output protein B. The intermediate regulator A represses the expression of the output protein B through transcriptional repression (Figure 5.11). The expectation is that the equilibrium value of B is independent of the concentration u of the plasmid. That is, the concentration of B should adapt to the copy number of its own plasmid.

In order to analyze whether the adaptation property holds, we write the differential equation model describing the system, assuming that the mRNA dynamics

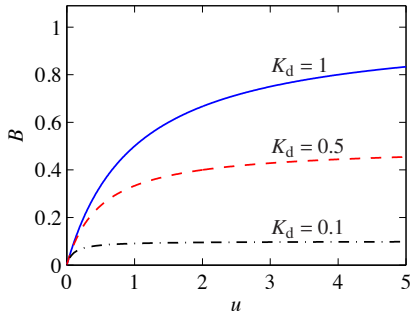


Figure 5.12: Behavior of the equilibrium value of B as a function of the input u. Concentration is in μM.

are at the quasi-steady state. This model is given by

$$\frac{dA}{dt} = k_0 u - \gamma A, \qquad \frac{dB}{dt} = \frac{k_1 u}{1 + (A/K_d)} - \gamma B, \tag{5.13}$$

in which k_0 is the constitutive rate at which A is expressed and K_d is the dissociation constant of the binding of A with the promoter. This implementation has been called the sniffer in Section 3.2. The equilibrium of the system is obtained by setting the time derivatives to zero and gives

$$A = \frac{k_0}{\gamma} u, \qquad B = \frac{k_1 u}{\gamma + k_0 u / K_d}.$$

From this expression, we can see that as K_d decreases, the denominator of the right-hand side expression tends to $k_0 u / K_d$ resulting in the equilibrium value $B = k_1 K_d / k_0$, which does not depend on the input u. Hence, in this case, adaptation would be reached. This is the case if the affinity of A to its operator sites is extremely high, resulting also in a strong repression and hence a lower value of B. In practice, however, the value of K_d is nonzero, hence the adaptation is not perfect. We show in Figure 5.12 the equilibrium value of B as a function of the input u for different values of K_d. As expected, lower values of K_d lead to weaker dependence of B on the u variable.

In this analysis, we have not modeled the cooperativity of the binding of protein A to the promoter. We leave as an exercise to show that the adaptation behavior persists in the case cooperativity is included (see Exercise 5.8).

For engineering a system with prescribed behavior, one has to be able to change the physical features so as to change the values of the parameters of the model. This is often possible. For example, the binding affinity ($1/K_d$ in the Hill function) of a transcription factor to its site on the promoter can be weakened by single or multiple base-pair substitutions. The protein decay rate can be increased by adding degradation tags at the end of the gene expressing protein B. Promoters that can

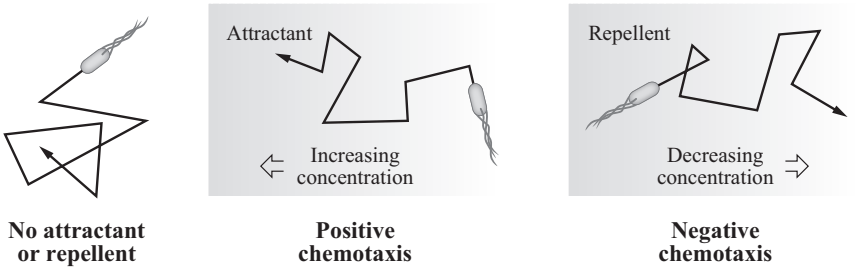


Figure 5.13: Examples of chemotaxis. In the absence of attractant or repellent, the bacterium follows a random walk. In the presence of an attractant, the random walk is biased in the direction in which the concentration of attractant increases (positive chemotaxis), while in the presence of a repellent the random walk is biased in the direction in which the concentration of the repellent decreases (negative chemotaxis). Figure adapted from Phillips, Kondev and Theriot [78].

accept multiple transcription factors (combinatorial promoters) can be realized by combining the operator sites of several simple promoters [46]. Finally, the overall protein production rate can be tuned by controlling a number of different system properties, including promoter's and the ribosome binding site's strength.

5.7 Bacterial chemotaxis

Chemotaxis refers to the process by which microorganisms move in response to chemical stimuli. Examples of chemotaxis include the ability of organisms to move in the direction of nutrients or move away from toxins in the environment. Chemotaxis is called *positive chemotaxis* if the motion is in the direction of the stimulus and *negative chemotaxis* if the motion is away from the stimulant, as shown in Figure 5.13. Many chemotaxis mechanisms are stochastic in nature, with biased random motions causing the average behavior to be either positive, negative or neutral (in the absence of stimuli).

In this section we look in some detail at bacterial chemotaxis, which *E. coli* use to move in the direction of increasing nutrients. The material in this section is based primarily on the work of Barkai and Leibler [9] and Rao, Kirby and Arkin [83].

Control system overview

The chemotaxis system in *E. coli* consists of a sensing system that detects the presence of nutrients, an actuation system that propels the organism in its environment, and control circuitry that determines how the cell should move in the presence of chemicals that stimulate the sensing system.

The actuation system in *E. coli* consists of a set of flagella that can be spun using a flagellar motor embedded in the outer membrane of the cell, as shown in

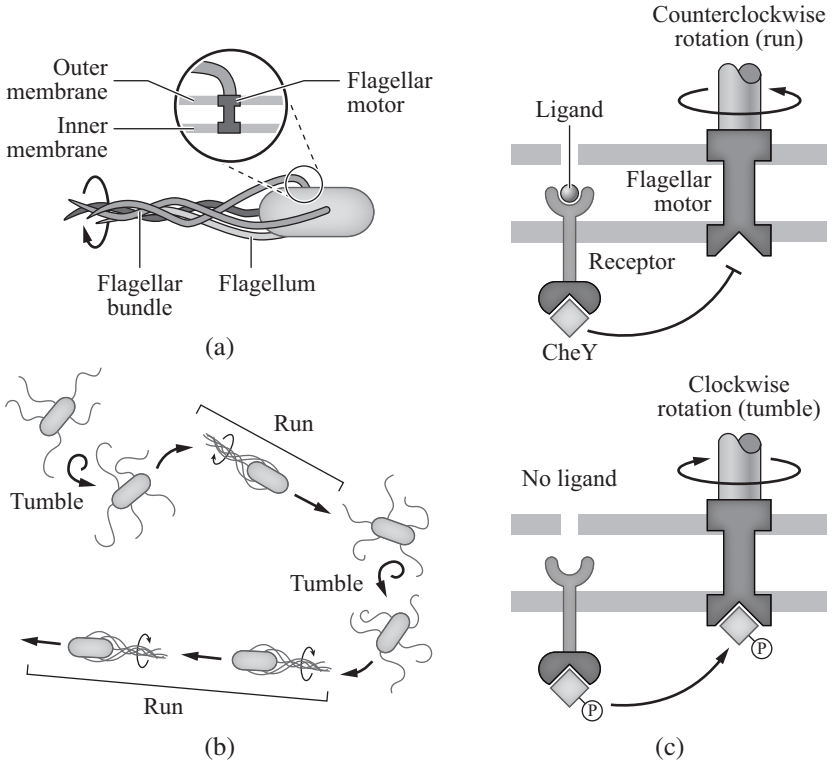
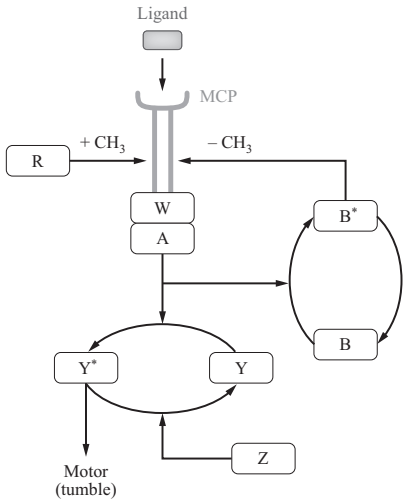


Figure 5.14: Bacterial chemotaxis. (a) Flagellar motors are responsible for spinning flagella. (b) When flagella spin in the clockwise direction, the organism tumbles, while when they spin in the counterclockwise direction, the organism runs. (c) The direction in which the flagella spin is determined by whether the CheY protein is phosphorylated. Figures adapted from Phillips, Kondev and Theriot [78].

Figure 5.14a. When the flagella all spin in the counterclockwise direction, the individual flagella form a bundle and cause the organism to move roughly in a straight line. This behavior is called a "run" motion. Alternatively, if the flagella spin in the clockwise direction, the individual flagella do not form a bundle and the organism "tumbles," causing it to rotate (Figure 5.14b). The selection of the motor direction is controlled by the protein CheY: if phosphorylated CheY binds to the motor complex, the motor spins clockwise (tumble), otherwise it spins counterclockwise (run) (Figure 5.14c).

A question that we address here is how the bacterium moves in the direction in which the attractant concentration increases. Because of the small size of the organism, it is not possible for a bacterium to sense gradients across its length. Hence, a more sophisticated strategy is used, in which the temporal gradient, as opposed to the spatial gradient, guides the organism motion through suitable combination of run and tumble motions. To sense temporal gradients, *E. coli* compares the cur-



(a) Sensing system

Non-ligand bound			Ligand bound		
Species	p	ΔE (kcal/mol)	Species	p	ΔE (kcal/mol)
	0.017	2.37		0.003	3.55
	0.125	1.18		0.017	2.37
	0.500	0.00		0.125	1.18
	0.874	–1.18		0.500	0.00
	0.997	–3.55		0.980	–2.37

(b) Receptor complex states

Figure 5.15: Control system for chemotaxis. (a) The sensing system is implemented by the receptor complex, which detects the presence of a ligand in the cell's outer environment. The computation part is implemented by a combined phosphorylation/methylation process, which realizes a form of integral feedback. The actuation component is realized by the CheY phosphorylated protein controlling directly the direction in which the motor spins. Figure from Rao et al. [83] (Figure 1A). (b) Receptor complex states. The probability of a given state being in an active configuration is given by p. ΔE represents the difference in energy levels from a reference state. Figure adapted from [72].

rent concentration of attractant to the past concentration of attractant and if the concentration increases, then the concentration of phosphorylated CheY protein is reduced. As a consequence, less phosphorylated CheY will bind to the motor complex and the tumbling frequency is reduced. The net result is a biased random walk in which the bacterium tends to move toward the direction in which the gradient of attractant concentration increases.

A simple model for the molecular control system that regulates chemotaxis is shown in Figure 5.15a. We start with the basic sensing and actuation mechanisms. A membrane-bound protein MCP (methyl-accepting chemotaxis protein) that is capable of binding to the external ligand serves as a signal transducing element from the cell exterior to the cytoplasm. Two other proteins, CheW and CheA, form a complex with MCP. This complex can either be in an active or inactive state. In the active state, CheA is autophosphorylated and serves as a phosphotransferase for two additional proteins, CheB and CheY. The phosphorylated form of CheY then binds to the motor complex, causing clockwise rotation of the motor (tumble).

The activity of the receptor complex is governed by two primary factors: the binding of a ligand molecule to the MCP protein and the presence or absence of

up to four methyl groups on the MCP protein. The specific dependence on each of these factors is somewhat complicated. Roughly speaking, when the ligand L is bound to the receptor then the complex is less likely to be active. Furthermore, as more methyl groups are present, the ligand binding probability increases, allowing the gain of the sensor to be adjusted through methylation. Finally, even in the absence of ligand the receptor complex can be active, with the probability of it being active increasing with increased methylation. Figure 5.15b summarizes the possible states, their free energies and the probability of activity.

Several other elements are contained in the chemotaxis control circuit. The most important of these are implemented by the proteins CheR and CheB, both of which affect the receptor complex. CheR, which is constitutively produced in the cell, methylates the receptor complex at one of the four different methylation sites. Conversely, the phosphorylated form of CheB demethylates the receptor complex. As described above, the methylation patterns of the receptor complex affect its activity, which affects the phosphorylation of CheA and, in turn, phosphorylation of CheY and CheB. The combination of CheA, CheB and the methylation of the receptor complex forms a negative feedback loop: if the receptor is active, then CheA phosphorylates CheB, which in turn demethylates the receptor complex, making it less active. As we shall see when we investigate the detailed dynamics below, this feedback loop corresponds to a type of integral feedback law. This integral action allows the cell to adjust to different levels of ligand concentration, so that the behavior of the system is invariant to the absolute ligand levels.

Modeling

From a high level, we can view the chemotaxis as a dynamical system that takes the ligand L concentration as an input and produces the phosphorylated CheY bound to the motor complex as an output, which determines the tumbling frequency. We let T represent the receptor complex and we write A, B, Y and Z for CheA, CheB, CheY and CheZ, respectively. As in previous chapters, for a protein X we let X^* represent its phosphorylated form.

Each receptor complex can have multiple methyl groups attached and the activity of the receptor complex depends on both the amount of methylation and whether a ligand is attached to the receptor site (Figure 5.15b). Furthermore, the binding probabilities for the receptor also depend on the methylation pattern. We let T_i^x represent a receptor that has i methylation sites filled and ligand state x (which can be either u if unoccupied or o if occupied). We let m represent the maximum number of methylation sites ($m = 4$ for *E. coli*). Using this notation, the transitions between the states correspond to the following reactions (also shown in Figure 5.16):

$$T_i^u + L \rightleftharpoons T_i^o,$$

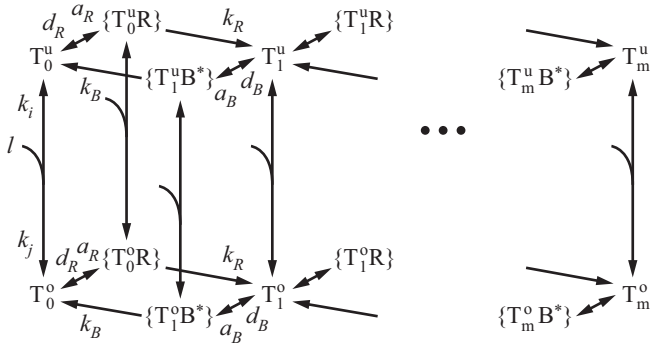


Figure 5.16: Methylation model for chemotaxis. Figure adapted from [9] (Box 1).

$$\begin{aligned} &\mathrm{T}_i^{\mathrm{x}} + \mathrm{B}^* \rightleftharpoons \mathrm{T}_i^{\mathrm{x}}{:}\mathrm{B}^* \longrightarrow \mathrm{T}_{i-1}^{\mathrm{x}} + \mathrm{B}^*, && i > 0,\ \mathrm{x} \in \{\mathrm{u},\mathrm{o}\}, \\ &\mathrm{T}_i^{\mathrm{x}} + \mathrm{R} \rightleftharpoons \mathrm{T}_i^{\mathrm{x}}{:}\mathrm{R} \longrightarrow \mathrm{T}_{i+1}^{\mathrm{x}} + \mathrm{R}, && i < m,\ \mathrm{x} \in \{\mathrm{u},\mathrm{o}\}, \end{aligned}$$

in which the first reaction models the binding of the ligand L to the receptor complex, the second reaction models the demethylation of methylated receptor complex by B^*, and the third reaction models the methylation of the receptor complex by R.

We now must write reactions for each of the receptor complexes with CheA. Each form of the receptor complex has a different activity level that determines the extent to which CheA can be phosphorylated. Therefore, we write a separate reaction for each form, which for simplicity we assume to be a one-step process for each $\mathrm{T}_i^{\mathrm{o}}$ and $\mathrm{T}_i^{\mathrm{u}}$ species:

$$\mathrm{T}_i^{\mathrm{x}} + \mathrm{A} \xrightarrow{k_i^x} \mathrm{T}_i^{\mathrm{x}} + \mathrm{A}^*,$$

where $\mathrm{x} \in \{\mathrm{o},\mathrm{u}\}$ and $i = 0,\ldots,m$. As a consequence, the production rate of A^* by all the above reactions is given by

$$A \sum_{i=1}^{4} \left(k_i^o T_i^o + k_i^u T_i^u \right).$$

Considering that the ligand-receptor binding reaction is at its quasi-steady state because it is very fast, and letting $T_i = T_i^u + T_i^o$ represent the total amount of receptors with i sites methylated, we further have that

$$T_i^u = \frac{1}{1 + L/K_L} T_i, \qquad T_i^o = \frac{L/K_L}{1 + L/K_L} T_i,$$

in which K_L is the dissociation constant of the receptor complex-ligand binding. It follows that if we let T_i^A denote the "effective" concentration of T_i that phosphory-

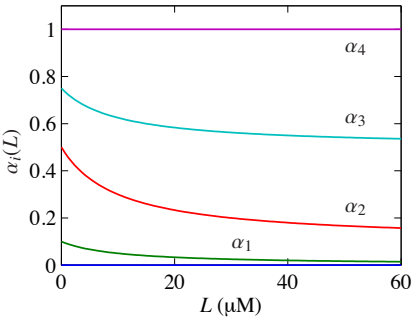


Figure 5.17: Probability of activity $\alpha_i(L)$ as a function of the ligand concentration L.

lates A, we have

$$k_i^a T_i^A = k_i^o T_i^o + k_i^u T_i^u = T_i \left(\frac{k_i^u + k_i^o (L/K_L)}{1 + L/K_L} \right),$$

for a suitably defined constant k_i^a. Hence, we can write

$$T_i^A = \alpha_i(L) T_i, \quad \text{with} \quad \alpha_i(L) = \frac{\alpha_i^o (L/K_L)}{1 + L/K_L} + \frac{\alpha_i^1}{1 + L/K_L},$$

and $T_i^I = T_i - T_i^A$, for suitable constants α_i^1 and α_i^o. The coefficients α_i^o and α_i^1 capture the effect of presence or absence of the ligand on the activity level of the complex and $\alpha_i(L)$ can be interpreted as the probability of activity. Following [83], we take the coefficients to be

$$\begin{aligned} &\alpha_0^1 = 0, \quad \alpha_1^1 = 0.1, \quad \alpha_2^1 = 0.5, \quad \alpha_3^1 = 0.75, \quad \alpha_4^1 = 1, \\ &\alpha_0^o = 0, \quad \alpha_1^o = 0, \quad \alpha_2^o = 0.1, \quad \alpha_3^o = 0.5, \quad \alpha_4^o = 1, \end{aligned}$$

and choose $K_L = 10$ μM. Figure 5.17 shows how each $\alpha_i(L)$ varies with L.

The total concentration of active receptors can now be written in terms of the receptor complex concentrations T_i and the activity probabilities $\alpha_i(L)$. We write the concentration of activated complex $\mathrm{T^A}$ and inactivated complex $\mathrm{T^I}$ as

$$T^A = \sum_{i=0}^{4} \alpha_i(L) T_i, \qquad T^I = \sum_{i=0}^{4} (1 - \alpha_i(L)) T_i.$$

These formulas can now be used in our dynamics as an effective concentration of active or inactive receptors. In particular, letting k_0 represent a lumped reaction rate constant for the phosphorylation of A by $\mathrm{T^A}$, for which we employ a one-step reaction model, we have

$$\mathrm{T^A + A} \xrightarrow{k_0} \mathrm{T^A + A^*}. \tag{5.14}$$

We next model the transition between the methylation patterns on the receptor. For this, we assume that CheR binds only inactive receptors and phosphorylated CheB binds only to active receptors [83, 72]. This leads to the following reactions:

$$\begin{aligned}
&\mathrm{T}_i^{\mathrm{A}} + \mathrm{B}^* \underset{d_B}{\overset{a_B}{\rightleftharpoons}} \mathrm{T}_i^{\mathrm{A}}{:}\mathrm{B}^* \xrightarrow{k_B} \mathrm{T}_{i-1}^{\mathrm{A}} + \mathrm{B}^*, && i \in \{2,3,4\},\\
&\mathrm{T}_1^{\mathrm{A}} + \mathrm{B}^* \underset{d_B}{\overset{a_B}{\rightleftharpoons}} \mathrm{T}_1^{\mathrm{A}}{:}\mathrm{B}^*,\\
&\mathrm{T}_i^{\mathrm{I}} + \mathrm{R} \underset{d_R}{\overset{a_R}{\rightleftharpoons}} \mathrm{T}_i^{\mathrm{I}}{:}\mathrm{R} \xrightarrow{k_R} \mathrm{T}_{i+1}^{\mathrm{I}} + \mathrm{R}, && i \in \{1,2,3\},\\
&\mathrm{T}_4^{\mathrm{I}} + \mathrm{R} \underset{d_R}{\overset{a_R}{\rightleftharpoons}} \mathrm{T}_4^{\mathrm{I}}{:}\mathrm{R},
\end{aligned}$$

in which we accounted for the fact that R can still bind to $\mathrm{T}_4^{\mathrm{I}}$ even without methylating it and B^* can still bind $\mathrm{T}_1^{\mathrm{A}}$ even without demethylating it. Assuming the complexes' concentrations are at their quasi-steady state values, and letting $K_R = (d_R + k_R)/a_R$ and $K_B = (d_B + k_B)/a_B$ denote the Michaelis-Menten constant for the above enzymatic reactions, we have the following relations:

$$[\mathrm{T}_i^{\mathrm{R}}{:}\mathrm{R}] = \frac{T_i^I R}{K_R}, \qquad [\mathrm{T}_i^{\mathrm{A}}{:}\mathrm{B}^*] = \frac{T_i^A B^*}{K_B},$$

in which we have approximated d_R/a_R (d_B/a_B) by K_R (K_B) accounting for the fact that $k_R \ll d_R$ ($k_B \ll d_B$). We can now write the differential equation for T_i considering the fact that

$$\frac{dT_i}{dt} = \frac{dT_i^I}{dt} + \frac{dT_i^A}{dt}.$$

Specifically, we have that

$$\begin{aligned}
\frac{dT_i^I}{dt} &= -k_R \frac{T_i^I R}{K_R} + k_R \frac{T_{i-1}^I R}{K_R},\\
\frac{dT_i^A}{dt} &= -k_B \frac{T_i^A B^*}{K_B} + k_B \frac{T_{i+1}^A B^*}{K_B},
\end{aligned}$$

with the conservation laws $R_{\mathrm{tot}} = R + \sum_{i=1}^4 [\mathrm{T}_{\mathrm{i}}^{\mathrm{I}}{:}\mathrm{R}]$ and $B^*_{\mathrm{tot}} = B^* + \sum_{i=1}^4 [\mathrm{T}_{\mathrm{i}}{:}\mathrm{B}^*]$, in which B^*_{tot} represents the total amount of phosphorylated CheB protein. Considering the quasi-steady state expressions for the complexes and the fact that $\sum_{i=1}^4 T_i^A = T^A$ and $\sum_{i=1}^4 T_i^I = T^I$, these conservation laws lead to

$$R = \frac{R_{\mathrm{tot}}}{1 + T^I/K_R}, \qquad B^* = \frac{B^*_{\mathrm{tot}}}{1 + T^A/K_B}.$$

Defining

$$r_R = k_R \frac{R_{\mathrm{tot}}}{K_R + T^I}, \qquad r_B = k_B \frac{B^*_{\mathrm{tot}}}{K_B + T^A},$$

which represent the effective rates of the methylation and demethylation reactions, we finally obtain that

$$\frac{d}{dt}T_i = r_R(1-\alpha_{i-1}(L))T_{i-1} + r_B\alpha_{i+1}(L)T_{i+1} - r_R(1-\alpha_i(L))T_i - r_B\alpha_i(L)T_i,$$

where the first and second terms represent transitions into this state via methylation or demethylation of neighboring states (see Figure 5.16) and the last two terms represent transitions out of the current state by methylation and demethylation, respectively. Note that the equations for T_0 and T_4 are slightly different since the demethylation and methylation reactions are not present, respectively.

Finally, we write the phosphotransfer and dephosphorylation reactions among CheA, CheB, and CheY, and the binding of CheY^* to the motor complex M:

$$\begin{aligned}
&\mathrm{A}^* + \mathrm{Y} \xrightarrow{k_1} \mathrm{A} + \mathrm{Y}^*, \quad \mathrm{Y}^* \xrightarrow{k_2} \mathrm{Y},\\
&\mathrm{A}^* + \mathrm{B} \xrightarrow{k_3} \mathrm{A} + \mathrm{B}^*, \quad \mathrm{B}^* \xrightarrow{k_4} \mathrm{B},\\
&\mathrm{T}_i^{\mathrm{A}}{:}\mathrm{B}^* \xrightarrow{k_4} \mathrm{T}_i^{\mathrm{A}}{:}\mathrm{B}, \; i \in \{1,2,3,4\},\\
&\mathrm{Y}^* + \mathrm{M} \underset{d}{\overset{a}{\rightleftharpoons}} \mathrm{Y}^*{:}\mathrm{M},
\end{aligned}$$

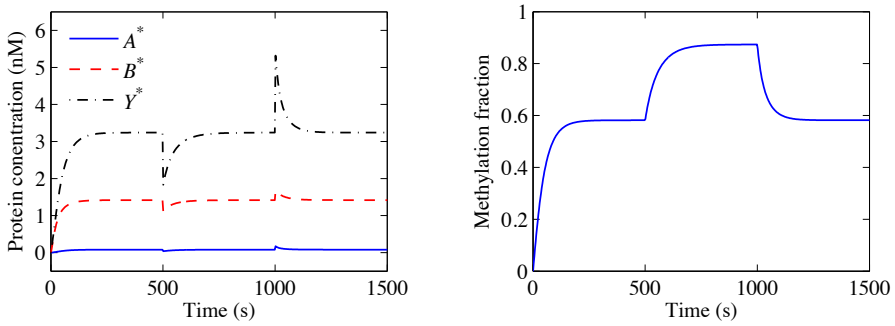
in which the first reaction is the phosphotransfer from CheA^* to CheY^*, the second reaction is the dephosphorylation of CheY^* by CheZ, the third reaction is the phosphotransfer from CheA^* to CheB, the fourth and fifth reactions are the dephosphorylation of CheB^*, which can be dephosphorylated also when bound to the receptor, and the last reaction is the binding of CheY^* with the motor complex M to form the complex $\mathrm{Y}^* : \mathrm{M}$, which increases the probability that the motor will rotate clockwise. The resulting ODE model is given by

$$\begin{aligned}
\frac{d}{dt}A^* &= k_0 T^A A - k_1 A^* Y - k_3 A^* B,\\
\frac{d}{dt}Y^* &= k_1 A^* Y - k_2 Y^* - aMY^* + d[\mathrm{M{:}Y^*}],\\
\frac{d}{dt}B^*_{\mathrm{tot}} &= k_3 A^* B - k_4 B^*_{\mathrm{tot}},\\
\frac{d}{dt}[\mathrm{M{:}Y^*}] &= aMY^* - d[\mathrm{M{:}Y^*}],
\end{aligned}$$

with conservation laws

$$\begin{aligned}
A + A^* &= A_{\mathrm{tot}}, \qquad B + B^*_{\mathrm{tot}} = B_{\mathrm{tot}},\\
Y + Y^* + [\mathrm{M{:}Y^*}] &= Y_{\mathrm{tot}}, \qquad M + [\mathrm{M{:}Y^*}] = M_{\mathrm{tot}}.
\end{aligned}$$

Figure 5.18a shows the concentration of the phosphorylated proteins based on a simulation of the model. Initially, all species are started in their unphosphory-



(a) Protein concentrations (b) Methylation

Figure 5.18: Simulation and analysis of reduced-order chemotaxis model. The parameters are taken from Rao et al. [83] and given by $k_0 = 50\ \text{s}^{-1}\text{nM}^{-1}$, $k_1 = 100\ \text{s}^{-1}\text{nM}^{-1}$, $k_3 = 30\ \text{s}^{-1}\text{nM}^{-1}$, $k_2 = 30\ \text{s}^{-1}\text{nM}^{-1}$, $a = 50\ \text{s}^{-1}\text{nM}^{-1}$, $d = 19\ \text{s}^{-1}$, and $k_4 = 1\ \text{s}^{-1}$. The concentrations are $R_{\text{tot}} = 0.2$ nM, $\sum_{i=0}^{4} T_i = 5$ nM, $A_{\text{tot}} = 5$ nM, $B_{\text{tot}} = 2$ nM, and $Y_{\text{tot}} = 17.9$ nM. Also, we have $k_B = 0.5\ \text{s}^{-1}$, $K_B = 5.5$ nM, $k_R = 0.255\ \text{s}^{-1}$, and $K_R = 0.251$ nM.

lated and demethylated states. At time 500 s the ligand concentration is increased to $L = 10$ μM and at time 1000 s it is returned to zero. We see that immediately after the ligand is added, the CheY* concentration drops, allowing longer runs between tumble motions. After a short period, however, the CheY* concentration adapts to the higher concentration and the nominal run versus tumble behavior is restored. Similarly, after the ligand concentration is decreased the concentration of CheY* increases, causing a larger fraction of tumbles (and subsequent changes in direction). Again, adaptation over a longer time scale returns the CheY concentration to its nominal value. The chemotaxis circuit pathway from T to CheY (to the motor) explains the sudden drop of CheY* when the ligand L is added. The slower methylation of the receptor complex catalyzed by CheR and removed by CheB slowly takes the value of CheY* to the pre-stimulus level.

Figure 5.18b helps explain the adaptation response. We see that the average amount of methylation of the receptor proteins increases when the ligand concentration is high, which decreases the activity of CheA (and hence decreases the phosphorylation of CheY).

Integral action

The perfect adaptation mechanism in the chemotaxis control circuitry has the same function as the use of integral action in control system design: by including a feedback on the integral of the error, it is possible to provide exact cancellation to constant disturbances. In this section we demonstrate that a simplified version of the dynamics can indeed be regarded as integral action of an appropriate signal. This interpretation was first pointed out by Yi et al. [102].

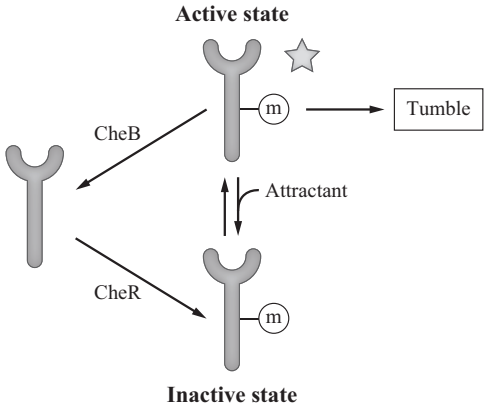


Figure 5.19: Reduced-order model of receptor activity. Star indicates activated complex and "m" indicates methylated complex. Figure adapted from [4], Figure 7.9.

We begin by formulating an even simpler model for the system dynamics that captures the basic features required to understand the integral action. We consider the receptor complex T and the kinase CheA as a single entity denoted by X, which can be either methylated or not. We ignore the number of methylated sites and simply group all the methylated forms into a lumped stated called X_m. Also, we assume that only the methylated state can be activated and the activity is determined by the ligand L concentration (through the functions $\alpha_i(L)$). We let X_m^* represent this active state and ignore the additional phosphorylation dynamics of CheY, so that we take the concentration X_m^* as our measure of overall activity.

We take the ligand into account by assuming that the transition between the active form X_m^* and the inactive form X_m depends on the ligand concentration: higher ligand concentration will increase the rate of transition to the inactive state. The activation/deactivation reactions are then written as

$$\texttt{R0}: \quad X_m^* \underset{k^r}{\overset{k^f(L)}{\rightleftharpoons}} X_m \qquad \text{activation/deactivation,}$$

in which the deactivation rate $k^f(L)$ is an increasing function of L. As before, CheR methylates the receptor and $CheB^*$ demethylates it. We simplify the picture by only allowing $CheB^*$ to act on the active state X_m^* and CheR to act on the inactive state. Figure 5.19 shows the transitions between the various forms X.

This model is a considerable simplification from the ligand binding model that is illustrated in Figures 5.15b and 5.16. In the previous models, there is some probability of activity with or without methylation and with or without ligand. In this simplified model, we assume that only three states are of interest: demethylated, methylated/inactive and methylated/active. We also modify the way that ligand binding is captured and instead of keeping track of all of the possibilities in Figure 5.15b, we assume that the ligand transitions us from an active state X_m^* to an

inactive X_m. These states and transitions are roughly consistent with the different energy levels and probabilities in Figure 5.15b, but it is clearly a much coarser model.

Accepting these approximations, the model illustrated in Figure 5.19 results in the methylation and demethylation reactions

$$\begin{aligned} \texttt{R1}: &\quad X + R \rightleftharpoons X{:}R \longrightarrow X_m + R && \text{methylation,} \\ \texttt{R2}: &\quad X_m^* + B^* \rightleftharpoons X_m^*{:}B^* \longrightarrow X + B^* && \text{demethylation.} \end{aligned}$$

For simplicity we take both R and B^* to have constant concentration.

We can further approximate the first and second reactions by their Michaelis-Menten forms, which yield net methylation and demethylation rates (for those reactions)

$$v_+ = k_R R \frac{X}{K_X + X}, \qquad v_- = k_B B^* \frac{X_m^*}{K_{X_m^*} + X_m^*}.$$

If we further assume that $X \gg K_X > 1$, then the methylation rate can be further simplified:

$$v_+ = k_R R \frac{X}{K_X + X} \approx K_R R.$$

Using these approximations, we can write the resulting dynamics for the overall system as

$$\begin{aligned} \frac{d}{dt} X_m &= k_R R + k^f(L) X_m^* - k^r X_m, \\ \frac{d}{dt} X_m^* &= -k_B B^* \frac{X_m^*}{K_{X_m^*} + X_m^*} - k^f(L) X_m^* + k^r X_m. \end{aligned}$$

We wish to use this model to understand how the steady state activity level X_m^* depends on the ligand concentration L (which enters through the deactivation rate $k^f(L)$).

It will be useful to rewrite the dynamics in terms of the activated complex concentration X_m^* and the *total* methylated complex concentration $X_{m,\text{tot}} = X_m + X_m^*$. A simple set of algebraic manipulations yields

$$\begin{aligned} \frac{dX_m^*}{dt} &= k^r (X_{m,\text{tot}} - X_m^*) - k_B B^* \frac{X_m^*}{K_{X_m^*} + X_m^*} - k^f(L) X_m^*, \\ \frac{dX_{m,\text{tot}}}{dt} &= k_R R - k_B B^* \frac{X_m^*}{K_{X_m^*} + X_m^*}. \end{aligned}$$

From the second equation, we see that the the concentration of methylated complex $X_{m,\text{tot}}$ is a balance between the action of the methylation reaction (R1, at rate v_+) and the demethylation reaction (R2, at rate v_-). Since the action of a ligand binding to the receptor complex increases the rate of deactivation of the complex (R0), in the presence of a ligand we will increase the amount of methylated complex

and, (via reaction R1) eventually restore the amount of the activated complex. This represents the adaptation mechanism in this simplified model.

To further explore the effect of adaptation, we compute the equilibrium points for the system. Setting the time derivatives to zero, we obtain

$$X^*_{m,e} = \frac{K_{X^*_m} k_R R}{k_B B^* - k_R R},$$
$$X_{m,\text{tot},e} = \frac{1}{k^r}\left(k^r X^*_m + k_B B^* \frac{X^*_m}{K_{X^*_m} + X^*_m} + k^f(L) X^*_m\right).$$

Note that the solution for the active complex $X^*_{m,e}$ in the first equation does not depend on $k^f(L)$ (or k^r) and hence the steady state solution is independent of the ligand concentration. Thus, in steady state, the concentration of activated complex adapts to the steady state value of the ligand that is present, making it insensitive to the steady state value of this input.

In order to demonstrate that after a perturbation due to addition of the ligand the value of X^*_m returns to its equilibrium, we need to prove that the equilibrium point $(X_{m,\text{tot},e}, X^*_{m,e})$ is asymptotically stable. To do so, let $x = X_{m,\text{tot}}$, $y = X^*_m$ and rewrite the system model as

$$\frac{dx}{dt} = k_R R - k_B B^* \frac{y}{K_{X^*_m} + y},$$
$$\frac{dy}{dt} = k^r(x - y) - k_B B^* \frac{y}{K_{X^*_m} + y} - k^f(L) y,$$

which is in the standard integral feedback form introduced in Section 3.2. The stability of the equilibrium point $(x_e, y_e) = (X_{m,\text{tot},e}, X^*_{m,e})$ can be determined by calculating the Jacobian matrix J of this system at the equilibrium. This gives

$$J = \begin{pmatrix} 0 & -k_B B^* \frac{K_{X^*_m}}{(y_e + K_{X^*_m})^2} \\ k^r & -k_B B^* \frac{K_{X^*_m}}{(y_e + K_{X^*_m})^2} - k^f(L) - k^r \end{pmatrix},$$

for which

$$\text{tr}(J) < 0 \qquad \text{and} \qquad \det(J) > 0,$$

implying that the equilibrium point is asymptotically stable.

The dynamics for $X_{m,\text{tot}}$ can be viewed as an integral action: when the concentration of X^*_m matches its reference value (with no ligand present), the quantity of methylated complex $X_{m,\text{tot}}$ remains constant. But if $X_{m,\text{tot}}$ does not match this reference value, then $X_{m,\text{tot}}$ increases at a rate proportional to the methylation "error" (measured here by difference in the nominal reaction rates v_+ and v_-). It can be shown that this type of integral action is necessary to achieve perfect adaptation in a robust manner [102].

Exercises

5.1 Consider the negatively autoregulated system given in equation (5.4). Determine the frequency response with respect to noise input Γ_2 and compare its magnitude to that of the open loop system in equation (5.2).

5.2 Consider the contribution of the negative autoregulation given by the parameter g in equation (5.6). Study how g changes when the value of the dissociation constant K_d is changed.

5.3 Consider the negatively autoregulated system

$$\frac{dA}{dt} = \frac{\beta}{1+(A/K)^n} - \gamma A.$$

Explore through linearization how increasing the Hill coefficient affects the response time of the system. Also, compare the results of the linearization analysis to the behavior of the nonlinear system obtained through simulation.

5.4 Consider the toggle switch model

$$\frac{dA}{dt} = \frac{\beta_A}{1+(B/K)^n} - \gamma A, \qquad \frac{dB}{dt} = \frac{\beta_B}{1+(A/K)^m} - \gamma B.$$

Here, we are going to explore the parameter space that makes the system work as a toggle switch. To do so, answer the following questions:

(i) Consider $m = n = 1$. Determine the number and stability of the equilibria as the values of β_A and β_B are changed.

(ii) Consider $m = 1$ and $n > 1$ and determine the number and stability of the equilibria (as other parameters change).

(iii) Consider $m = n = 2$. Determine parameter conditions on $\beta_A, \beta_B, \gamma, K$ for which the system is bistable, i.e., there are two stable steady states.

5.5 Consider the repressilator model and the parameter space for oscillations provided in Figure 5.5. Determine how this parameter space changes if the value of K in the Hill function is changed.

5.6 Consider the "generalized" model of the repressilator in which we have m repressors (with m an odd number) in the loop. Explore via simulation the fact that when m is increased, the system oscillates for smaller values of the Hill coefficient n.

5.7 Consider the activator-repressor clock model given in equations (5.11). Determine the number and stability of the equilibria as performed in the text for the case in which $m > 1$.

5.8 Consider the feedforward circuit shown in Figure 5.11. Assume that we account for cooperativity such that the model becomes

$$\frac{dA}{dt} = k_0 u - \gamma A, \qquad \frac{dB}{dt} = \frac{k_1 u}{1 + (A/K_d)^n} - \gamma B.$$

Show that the adaptation property still holds under suitable parameter conditions.