

This is the Author's Post-print version of the following article: *L.A. Torres, V. Ibarra-Junquera, P. Escalante-Minakata, H.C. Rosu, High-gain nonlinear observer for simple genetic regulation process, Physica A: Statistical Mechanics and its Applications, Volume 380, 2007, Pages 235-240*, which has been published in final form at <https://doi.org/10.1016/j.physa.2007.02.105> This article may be used for non-commercial purposes in accordance with Terms and Conditions for Self-Archiving

High-Gain Nonlinear Observer for Simple Genetic Regulation Process

L.A. Torres^{1*}, V. Ibarra-Junquera^{2†}, P. Escalante-Minakata^{1‡}, and H.C. Rosu^{1§}

¹ Instituto Potosino de Investigación Científica y Tecnológica
A.P. 3-74 Tangamanga, San Luis Potosí, SLP 78216, México

² Faculty of Chemical Sciences, University of Colima
Coquimatlán, Col 28400, Mexico

published: Physica A 380 (1 July 2007) 235-240

doi: 10.1016/j.physa.2007.02.105

arXiv:q-bio/0409036v2 [q-bio.QM]

Abstract

High-gain nonlinear observers occur in the nonlinear automatic control theory and are in standard usage in chemical engineering processes. We apply such a type of analysis in the context of a very simple one-gene regulation circuit. In general, an observer combines an analytical differential-equation-based model with partial measurement of the system in order to estimate the non-measured state variables. We use one of the simplest observers, that of Gauthier et al., which is a copy of the original system plus a correction term which is easy to calculate. For the illustration of this procedure, we employ a biological model, recently adapted from Goodwin's old book by De Jong, in which one plays with the dynamics of the concentrations of the messenger RNA coding for a given protein, the protein itself, and a single metabolite. Using the observer instead of the metabolite, it is possible to rebuild the non-measured concentrations of the mRNA and the protein.

Keywords: High-gain observer; Diffeomorphism; Gene regulation; Gene expression

*E-mail: ltorres@ipicyt.edu.mx

†E-mail: vj@ucol.mx

‡E-mail: minakata@ipicyt.edu.mx

§E-mail: hcr@ipicyt.edu.mx; corresponding author. calis07b.tex

1 Introduction

According to textbooks, gene expression is a very complicated dynamical process which is regulated at a number of its stages during the synthesis of proteins [1]. Similar to many big cities, with heavy traffic, biological cells host complicated traffic of biochemical signals at all levels. At the nanometer scale, clusters of molecules in the form of proteins drive the dynamics of the cellular network that schematically can be divided into four regulated parts: the DNA or genes, the transcribed RNAs, the set of interacting proteins, and the metabolites. Genes can only affect other genes through specific proteins, as well as through some metabolic pathways that are regulated by proteins themselves. They act to catalyze the information stored in DNA, all the way from the fundamental processes of transcription and translation to the final quantities of produced proteins.

For the purpose of modeling, it is essential to generate simple models that help to understand elementary dynamical components of these complex regulatory networks as molecular tools that participate in an important way in the machinery of cellular decisions, that is to say, in the behavior and genetic program of cells. The central importance of control theory in Biology can be assessed through the recent problem of identifying control motifs (or modules), which are patterns that occur in a gene network far more often than in randomized networks of biological regulators [2]. This hot issue has been first pinpointed in a breakthrough paper of Doyle and collaborators [3] in which the regulation of bacterial chemotaxis was interpreted in terms of the simple integral control ‘adaptive module’ introduced by Barkai and Leibler [4]. Since gene regulation appears to occur only at some definite states of the whole process, which in general are not well known, we are from the point of view of control engineering in the case of the reconstruction of those specific states under the condition of limited information.

It is quite clear that the availability of all state variables to direct measurement is an extremely rare occasion for gene expression phenomena or when it is possible it could be too expensive. For this particular task, but in completely different technological areas, the engineers have developed software sensors (state observers) that accurately reconstruct the state variables of various technological processes [5]. The basic concept of state of a system or process could have many different empirical meanings in biology. For the particular case of gene expression, the meaning of a state is essentially that of a concentration. The typical problem in control engineering that appears to be tremendously useful in biology is the reconstruction of some specific regulated states under conditions of limited information.

In general, an observer is expected to provide a good estimate $\hat{X}(t)$ of the natural state $X(t)$ of the original system. For this, one usually can think that

some distance $d(\hat{X}(t), X(t))$ (in the sense of a norm $\|\cdot\|$ in a vectorial space) goes to zero as $t \rightarrow \infty$. Such softwares can be constructed using the mathematical model of the process to obtain an estimate \hat{X} of the true state X . This estimate can then be used as a substitute for the unknown state X . The usage of state observers has proven useful in process monitoring and for many other tasks. The concept of observer is used herein in the sense of control theory, defining an algorithm capable of giving a reasonable estimation of the unmeasured variables of a process. In the case of gene expression processes the description is made very concrete in the following by looking at quite simple mathematical models that refer to single gene cases and which in principle can be extended to some *operons* that are single gene clusters.

In this paper we will examine in detail a particularly simple observer due to Gauthier and collaborators [6] possessing arbitrary exponential decay and linear error dynamics for the case of a three-state genetic regulation process. We were led to consider this observer because of its simplicity and its *high gain* property. The gain is defined as the amount of increase in error in the dynamics of the observer. This amount is directly related to the velocity with which the observer recovers the unknown signal. For the observer of Gauthier *et al* the amount of increase in error is constant and usually of high values leading to a fast recover of the unmeasurable states.

2 Mathematical Model for a Simple Gene Regulation Process

A kinetic model of a simple genetic regulation process was first developed by Goodwin as long ago as 1963 [7]. It has been further generalized by Tyson and Othmer [8] and clearly explained by De Jong in his recent review [9]. We consider here the most simple version of this kinetic model. For three concentrations X_1, X_2, X_3 , corresponding to the messenger RNA (mRNA) that codes for the unstable enzyme, the enzyme, and the metabolite, respectively, we write Tyson's model in the form

$$\Gamma_{\text{biology}} : \begin{cases} \dot{X}_1 = K_1 H(X, \vartheta) - \gamma_1 X_1 \\ \dot{X}_2 = K_2 X_1 - \gamma_2 X_2 \\ \dot{X}_3 = K_3 X_2 - \gamma_3 X_3 . \end{cases} \quad (1)$$

The parameters K_1, K_2, K_3 are all strictly positive and represent production constants, whereas $\gamma_1, \gamma_2, \gamma_3$ are also strictly positive degradation constants. These rate equations express a balance between the number of molecules appearing and disappearing per unit time. Notice that the model assumes that the concentration X_2 increases linearly with X_1 and the concentration X_3 linearly with X_2 , which are natural assumptions. In the case of X_1 , the first term is the production term involving a nonlinear nondissipative *regulation function* H

that we take of the m -steepen Hill form ($m > 0$ is the steepness parameter) in common use

$$\begin{aligned}
 H^+(X, \vartheta) &= \frac{X_3^m}{X_3^m + \vartheta^m}, \\
 H^-(X, \vartheta) &= 1 - \frac{X_3^m}{X_3^m + \vartheta^m},
 \end{aligned}
 \tag{2}$$

for the activation and inhibition cases, respectively. The parameter θ gives the threshold for the regulatory influence of the concentration of the metabolite on the target gene, whereas the steepness parameter m is a measure of the collective effect of groups of metabolite molecules and also defines the shape of the Hill curve. This nonlinear parametrization describes the ‘biological regulation process’ that includes the production of the mRNA by transcription of its structural gene, its possible intranuclear processing by cleavage, its enzymatic degradation within the nucleus, and its migration to the cytoplasm by some form of diffusion or biological transport. Once in the cytoplasm, the mRNA is both translated into the unstable enzyme and enzymatically degraded.

System Γ_{biology} and its trivial chain generalization in the linear part is considered to be a good model for the simplest type of allosteric regulation in biochemistry, i.e., the inhibition or activation of an enzyme or protein by a small regulatory molecule that interacts with the enzyme at a site (allosteric site) other than the active site at which catalytic activity occurs. The interaction changes the shape of the enzyme, thus affecting the active site of the standard catalysis. This change of shape of the enzyme is sufficient to change its ability to catalyze a reaction in either negative or positive way and enables a cell to regulate needed metabolites. The allosteric regulation has the typical features of a feedback loop in control theory if the regulatory protein acts on the enzyme in the pathway of its own synthesis.

3 The Nonlinear Observer

Many attempts have been made to develop nonlinear observer design methods. One could mention the industrially popular extended Kalman filter, whose design is based on a local linearization of the system around a reference trajectory, restricting the validity of the approach within a small region in the state space [5],[10]. The first systematic approach for the development of a theory of nonlinear observers was proposed some time ago by Krener and Isidori [11]. In further works, nonlinear transformations of the coordinates have also been employed to put the considered nonlinear system in a suitable ‘observer canonical form’, in which the observer design problem may be easily solved [6],[12],[13]. The main idea in this case is to find a state transformation to represent the system as a linear differential equation plus a nonlinear term, which is a function of the measured state.

In this section, we present the design of a nonlinear software sensor in which the metabolite concentration is the naturally measured state (the most easy to measure) and corresponds to the mathematical state X_3 in the model introduced in the previous section. Therefore, it seems logical to take X_3 as the output of the system

$$y = h(X) = X_3 . \quad (3)$$

We now apply the technique of high-gain observers that works for many nonlinear systems and guarantees that the output feedback controller recovers the performance of the state feedback controller when the observer's gain is sufficiently high. The model given by the aforementioned system Γ_{biology} has the form

$$\Gamma_y : \begin{cases} \dot{X} = f(X) \\ y = h(X) , \end{cases} \quad (4)$$

in which $X \in \mathbb{R}^3$, and moreover there is a "physical subset" $\Omega \subset \mathbb{R}^3$ where the system lies. To make this mathematically precise we must introduce some further mathematical terminology. Let us construct the j th time derivative of the output. This can be expressed using Lie differentiation of the function h by the vector field f , $L_f^j(h)(X(t))$. $L_f^j(h)(X(t))$ is the j th Lie derivative of h by f and a function of X defined inductively as follows

$$\begin{aligned} L_f^0(h)(X) &= h(X) \\ L_f^j(h)(X) &= \frac{\partial}{\partial X} (L_f^{j-1}(h)(X)) f(X) . \end{aligned} \quad (5)$$

When Γ_y is observable, the map $\Phi : X \rightarrow \Phi(X)$ is a diffeomorphism where

$$\xi = \Phi(X) = \begin{pmatrix} L_f^0(h)(X) \\ L_f^1(h)(X) \\ L_f^2(h)(X) \end{pmatrix} = \begin{bmatrix} X_3 \\ K_3 X_2 - \gamma_3 X_3 \\ K_3(K_2 X_1 - \gamma_2 X_2) - \gamma_3(K_3 X_2 - \gamma_3 X_3) \end{bmatrix} . \quad (6)$$

For $\Phi(X)$ to be a local diffeomorphism in a region Ω , it is necessary and sufficient that the Jacobian $d\Phi(X)$ should be nonsingular on Ω and moreover that $\Phi(X)$ is one-to-one from Ω to $\Phi(\Omega)$, see [14]. Notice, that no matter if we choose $H^+(X, \vartheta)$ or $H^-(X, \vartheta)$, the coordinate transformation is the same. This means that the structure of the observer will be the same for both cases: gene activation or inhibition.

When the system is observable on Ω , it can be rewritten in the global coordinate system defined by $\Phi(X)$ in the following matrix form:

$$\Gamma'_y : \begin{cases} \dot{\xi} = F'(\xi) = \begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\xi}_3 \end{bmatrix} = \begin{bmatrix} \xi_2 \\ \xi_3 \\ \varphi(\xi) \end{bmatrix} \\ y = C \xi = [1 \ 0 \ 0] \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} \end{cases}, \quad (7)$$

where, moreover, φ can be extended from Ω to the entire \mathbb{R}^3 by a C^∞ function globally Lipschitz on \mathbb{R}^3 . The latter form allows us to make use of the following result proven by Gauthier and collaborators [6]:

Consider the system

$$\Gamma_G : \dot{\hat{\xi}} = F'(\hat{\xi}) - S^{-1}C^T (C\hat{\xi} - y), \quad (8)$$

where $S(\theta)$ is the solution of the matrix equation

$$\theta S - A^T S - SA + C^T C = 0, \quad (9)$$

for θ large enough, with A a matrix of Brunovsky form ($A = \delta_{i,j+1}$; δ_{ij} is the Kronecker symbol), which plays the role of a shift operator on \mathbb{R}^n . Then, Eq. (8) defines an observer for Γ'_y , with

$$\|\hat{\xi} - \xi\| \leq M \exp\left(-\frac{\theta}{3} t\right) \|\hat{\xi}_0 - \xi_0\|. \quad (10)$$

In our case, an observer is a dynamical system as given by Eq. (8) that Hill track the trajectory of the original system (here Γ'_y). Notice that both systems are identical unless an additional term that compensate the error in the observer, where the error is given by the difference $\|\hat{\xi} - \xi\|$, which is seen to be exponentially decreasing in time. The Gauthier observer is particularly simple since it appears to be only a copy of Γ'_y , together with a correction term that depends only on the dimension of the state space and not on the system Γ' itself. In others words, the structure of the observer does not depend on the Hill steepness parameter m (Eq. (1)).

For the sake of concreteness we will construct the observer only for the activation case. However, one should notice that only the function $f(\hat{X})$ will change for the inhibition case.

The Gauthier observer in Eq. (8) in the original coordinates is given by

$$\dot{\hat{X}} = f(\hat{X}) + \Upsilon(\hat{X}) S^{-1}C^T (h(X) - h(\hat{X})), \quad (11)$$

where

$$\Upsilon(\hat{X}) = \frac{\partial \Phi^{-1}}{\partial \hat{\xi}} \Big|_{\hat{\xi}=\Phi(\hat{\xi})}. \quad (12)$$

For the particular three-dimensional state space of Γ_{biology} we get

$$\Upsilon(\hat{X}) = \begin{bmatrix} \frac{\gamma_2\gamma_3}{K_2K_3} & \frac{\gamma_2+\gamma_3}{K_2K_3} & \frac{1}{K_2K_3} \\ \frac{\gamma_3}{K_3} & \frac{1}{K_3} & 0 \\ 1 & 0 & 0 \end{bmatrix}. \quad (13)$$

The matrix $S(\theta)$ in the three dimensional case can be easily computed by means of Eq. (9) given in the Gauthier theorem and its inverse $S^{-1}(\theta)$ appears to be

$$S^{-1}(\theta) = \begin{bmatrix} 3\theta & 3\theta^2 & \theta^3 \\ 3\theta^2 & 5\theta^3 & 2\theta^4 \\ \theta^3 & 2\theta^4 & \theta^5 \end{bmatrix}. \quad (14)$$

Plugging the matrices (13) and (14) in Eq. (11), we get the following equation for the observer introduced by Gauthier and collaborators as applied to our biological case

$$\dot{\hat{X}}_{\text{biology}} = f(\hat{X}) + \begin{bmatrix} 3\frac{\gamma_2\gamma_3\theta}{K_2K_3} + 3\frac{(\gamma_2+\gamma_3)\theta^2}{K_2K_3} + \frac{\theta^3}{K_2K_3} \\ 3\frac{\gamma_3\theta}{K_3} + 3\frac{\theta^2}{K_3} \\ 3\theta \end{bmatrix} (X_3 - \hat{X}_3). \quad (15)$$

We use this form of the Gauthier observer to estimate the states X_1 and X_2 of the dynamical system Γ_{biology} . We work with $\theta = 1$ and the values of the parameters given in Table 1 that are not necessarily the experimental values but are consistent with the requirements of the model. Figure (1) shows the results of a numerical simulation, where the solid lines represent the true states and the dotted lines stand for the estimates, respectively. In addition, for the real system we have taken $m = 2$ whereas for the observer $m = 1$ in order to show the robustness of this type of nonlinear observer with respect to the steepness parameter.

Table 1: Variables and parameters

Symbol	Meaning	Value (arb. units)
K_1	Production constant of mRNA	0.001
K_2	Production constant of protein	1.0
K_3	Production constant of metabolite	1.0
γ_1	Degradation constant of mRNA	0.1
γ_2	Degradation constant of protein	1.0
γ_3	Degradation constant of metabolite	1.0
ϑ	Hill's threshold parameter	1.0

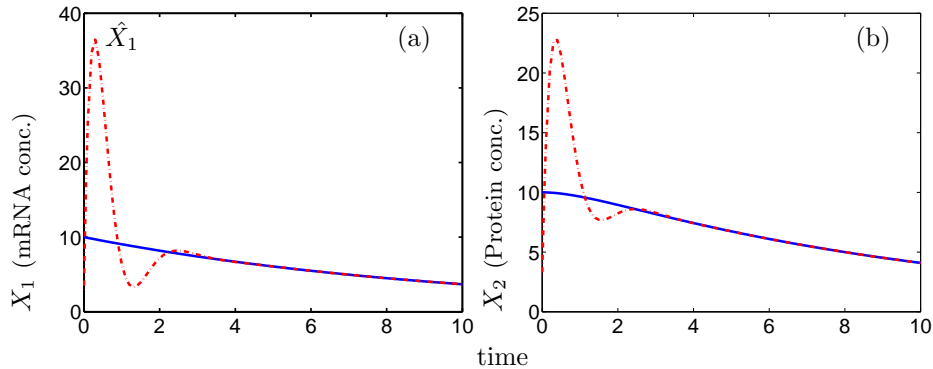


Figure 1: The numerical simulation – solid lines represent the true states and dotted lines represent the Gauthier estimates given by Eq. (15) for an activation case. Plot (a) represents the evolution of mRNA concentration in time and plot (b) the variation of protein (enzyme) concentration in time.

4 Conclusion

We presented here the mathematical exercise of designing a high-gain observer for a simple one-gene regulation dynamic process involving end-product activation (inhibition leads to similar results), which is able to rebuild in an effective way the non-measured concentrations of mRNA and the involved protein. Thus, the limitation of those experiments in which one has available only the metabolite can be overcome by employing this simple observer. In addition, this type of nonlinear observer could be used on line and is robust with respect to m , i.e., it does not need the exact value of the Hill steepness parameter. However, for more complex inputs of more complicated observable dynamical systems, this constant gain observer could have less performance and be overcome by some *adaptive* observers that can change in order to work better or provide more fit for a particular purpose. In the case of more limited information, e.g., for unknown functional form of the regulation function and high noise levels that can spoil the performance of the observer, the completely different mathematical procedure of creating dynamical extensions of the observer system are required [15].

Acknowledgment

This work was partially sponsored by grants from the Mexican Agency *Consejo Nacional de Ciencia y Tecnología* through project 46980.

References

- [1] B. Lewin, *Genes VII*, Oxford University Press, Oxford, 1999.
- [2] S. Mangan and U. Alon, *Structure and function of the feed-forward loop network motif*, Proc. Natl. Acad. Sci. USA 100 (2003) 11980-11985.
- [3] T.M. Yi, Y. Huang, M.I. Simon, and J. Doyle, *Robust perfect adaptation in bacterial chemotaxis through integral feedback control*, Proc. Natl. Acad. Sci. USA 97 (2000) 4649-4653.
- [4] N. Barkai and S. Leibler, *Robustness in simple biochemical networks*, Nature 387 (1997) 913-917.
- [5] G. Stephanopoulos, *Chemical Process Control*, Prentice-Hall, Eaglewood Cliffs, NJ, 1984.
- [6] J.P. Gauthier, H. Hammouri, and S. Othman, *A simple observer for nonlinear systems: applications to bioreactors*, IEEE Trans. Autom. Control 37 (1992) 875-880.
- [7] B.C. Goodwin, *Temporal Oscillations in Cells*, Academic Press, New York, 1963.
- [8] J.J. Tyson and H.G. Othmer, *The dynamics of feedback control circuits in biochemical pathways*, Prog. Theor. Biol. 5 (1978) 1-62.
- [9] H. De Jong, *Modeling and simulation of genetic regulatory systems: a literature review*, J. Comput. Biol. 9 (2002) 67-103.
- [10] W.A. Wolovich, *Automatic Control Systems*, Saunders College Publishing, Fort Worth, 1994.
- [11] A.J. Krener and A. Isidori, *Linearization by output injection and nonlinear observers*, Systems & Control Letters, 3 (1983) 47-52.
- [12] J.P. Gauthier and G. Bornard, *Obsevability for any $u(t)$ of a class of nonlinear systems*, IEEE Trans. on Autom. Control, AC-26 (1981) 922-926.
- [13] J.P. Gauthier and I.A.K. Kupka, *Observability and observers for nonlinear systems*, SIAM J. Control Optimization 32 (1994) 975-994.
- [14] H. Shim, Y.I. Son, and J.H. Seo, *Semi-global observer for multi-output nonlinear systems*, Syst. Control Lett. 41 (2001) 233-244.
- [15] V. Ibarra-Junquera, L.A. Torres, H.C. Rosu, G. Argüello, and J. Collado-Vides, *Nonlinear software sensor for monitoring genetic regulation process with noise and modeling errors*, Phys. Rev. E 72 (2005) 011919 (9 pp).