Adaptive Filtering for Unknown Genetic Regulatory Networks with Disturbance Attenuation

Yu Wenwu1,2, Lü Jinhua, Chen Guanrong1
1. Department of Electronic Engineering, City University of Hong Kong, Hong Kong, P.R. China
E-mail: wenwuyu@gmail.com, gchen@ee.cityu.edu.hk
2. Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, P.R. China
E-mail: jhlu@iss.ac.cn

Abstract: In this paper, the adaptive filtering problem for unknown genetic regulatory networks with disturbance attenuation is investigated based on Lyapunov method and adaptive techniques. A new regulatory network and several adaptive laws are designed to ensure the stochastic stability of the error states between the unknown and the estimated networks. The designed adaptive laws are independent of the unknown system states and parameters, where the only required information is the structure of the network. A simulation example is given to verify the theoretical results.

Key Words: Genetic regulatory network, Lyapunov method, Adaptive filtering, Disturbance attenuation

1 INTRODUCTION

Regulatory networks constitute an important new research subject in biological and biomedical sciences and received great attention over the past few years1-7. Genetic regulatory networks (GRNs) provide a mechanism to regulate the expressions of genes in production. Each gene contains some regulatory sequences, called cis elements. Transcription factors and their cofactors as well as other proteins can bind to such elements, and increase or reduce the gene expression levels. This results in the change of the corresponding protein levels, which in turn affect other genes' expression levels.

Basically, there are two types of genetic network models, i.e., the Boolean model8-10 and the differential equation model1-3. In the Boolean model, the activity of each gene is expressed by two states, ON or OFF, which are determined by a Boolean function of its own and other related states. On the other hand, the differential equation model describes the concentrations of gene products, such as mRNA and proteins, which are used as the state variables in genetic regulatory networks. Here, a genetic network model is considered, where each transcription factor acts additively to regulate a gene, i.e., the regulation function sums over all the inputs, which is called the SUM logic14,15. In [2-4, 7], the regulation function with the SUM logic in a genetic regulatory network has been studied in some detail.

When modelling genetic regulatory networks, molecular noise has been shown to play important roles in biological functions since noise is unavoidable in reactions of transcription, translation, and translocation processes, and also due to the external fluctuations. Recently, some studies on the genetic regulatory network models with stochastic disturbance have been reported5,6. In practice, some concentrations of products, such as mRNA, protein, and network parameters may not be fully known, therefore, they need to be estimated from the observed data. The combination of noise and uncertainties makes the analysis of such network models very difficult in general.

In a real genetic regulatory network, it is still not fully understood today as how the genes are expressed in the right time and right place, and at the right amount, throughout the development of the organism. Studying living organisms is a fairly complex process. It requires significant work on observing data and collecting data. Sometimes, it is even impossible for biologists to get all the information due to various technical difficulties. Thus, the study of system biology comes into place where it attempts to estimate the missing information from the available data. This gives rise to the following filtering problem: given a genetic regulatory network with unknown parameters and random disturbance, how to estimate the unknown products and parameters by using the observed data from the network?

In this paper, the adaptive filtering technique is applied to estimate the unknown information such as mRNA degradation rate, transcription rate, protein, and so on, from the known mRNA data. More precisely, the adaptive filtering problem for unknown genetic regulatory networks with disturbance attenuation is considered here. In [8-10, 12] the adaptive synchronization of unknown dynamical systems without random disturbance was already studied. In [11, 17], the synchronization of nonlinear systems with random disturbance was investigated, where the noise was assumed to exist only in the response system and vanishes on the synchronization manifold, which is not realistic. In this paper, the case when noise exists in the master system is considered, that is, an uncertain genetic regulatory network is discussed, where noise does not vanish on the synchronization manifold. Note also that in [9, 11-13], all the states of the networks are supposed to be known which is not always the case in real biological systems. In contrast, in this paper, some states and system parameters are allowed to be unknown in the concerned genetic regulatory networks with random disturbances.
The rest of this paper is organized as follows: In Section II, model formulation and preliminaries of the genetic regulatory network model are briefly outlined, and a genetic regulatory network with unknown parameter and random disturbance is described. An adaptive filter is then designed to estimate the unknown states and parameters of network in Section III. In section IV, a numerical example is provided to show the effectiveness of the proposed adaptive filtering technique. Conclusions are finally drawn in Section V.

2 MODEL FORMULATION AND PRELIMINARIES

The activities of a gene is regulated by other genes through the interactions among them, i.e., by transcription and translation. The following genetic regulatory network was proposed in [2-4]:

\[
\begin{align*}
\frac{dm_i(t)}{dt} &= -a_i m_i(t) + G_i(p_1(t), p_2(t), \ldots, p_n(t)) \\
\frac{dp_i(t)}{dt} &= -c_ip_i(t) + d_i m_i(t)
\end{align*}
\]

(1)

where \( m_i(t), p_i(t) \in R \) are concentrations of mRNA and protein of the \( i \)the node at time \( t \), and \( a_i > 0 \) and \( c_i > 0 \) are the degradation rates of the mRNA and protein, respectively, \( d_i \) is the translation rate, and the functions \( G_i \) represent the feedback regulation of the protein on the transcription, which is generally a nonlinear monotonically increasing function\(^{[13]}\).

The gene activities are being controlled in a cell, and the gene regulation function \( G_i \) plays a key role in the dynamical behavior of the network. Generally, the term of \( G_i \) can be very complex. In [3, 4], it is assumed that each transcription factor acts additively to regulate the \( i \)th gene, and the regulatory function is of the form \( G_i = \sum_{j=1}^{n} G_{ij}(p_j(t)) \), which is also called the SUM logic\(^{[14, 15]}\), and each \( G_{ij} \) is a monotonic function in the Hill form\(^{[6]}\). If the transcription factor \( j \) is an activator of gene \( i \), then

\[
G_{ij}(p_j(t)) = \alpha_{ij} \frac{(p_j(t))^\beta_j}{1 + (p_j(t))^\beta_j} \]

(2)

if transcription factor \( j \) is a repressor of gene \( i \), then

\[
G_{ij}(p_j(t)) = \alpha_{ij} \frac{1}{1 + (p_j(t))^\beta_j} \]

(3)

where \( \beta_j \) are the Hill coefficients, \( \beta_j \) are positive constants, and \( \alpha_{ij} \) are the dimensionless transcriptional rate of transcription factor \( j \) to \( i \). System (1) can be rewritten as

\[
\begin{align*}
\frac{dm_i(t)}{dt} &= -a_i m_i(t) + \sum_{j=1}^{n} w_{ij} f_j(p_j(t)) + L_i \\
\frac{dp_i(t)}{dt} &= -c_ip_i(t) + d_i m_i(t)
\end{align*}
\]

(4)

where \( f_j(x) = (x/\beta_j)^{\beta_j}/(1+(x/\beta_j)^{\beta_j}) \), \( L_i = \sum_{j \in I_i} \alpha_{ij} \) in which \( I_i \) is the set of all the repressors of gene \( i \), and \( W = (w_{ij}) \in R^{n \times n} \) is defined as follows:

\[
w_{ij} = \begin{cases} 
\alpha_{ij} & \text{if transcription factor } j \text{ is an activator of gene } i \\
0 & \text{if there is no link from node } j \text{ to } i \\
-\alpha_{ij} & \text{if transcription factor } j \text{ is a repressor of gene } i
\end{cases}
\]

System (4) can be written into a compact matrix form, as

\[
\begin{align*}
\frac{dm(t)}{dt} &= -Am(t) + Wf(p(t)) + L \\
\frac{dp(t)}{dt} &= -Cp(t) + Dm(t)
\end{align*}
\]

(5)

where \( m(t) = [m_1(t), m_2(t), \ldots, m_n(t)]^T \), \( p(t) = [p_1(t), p_2(t), \ldots, p_n(t)]^T \), \( f(p(t)) = [f_1(p_1(t)), f_2(p_2(t)), \ldots, f_n(p_n(t))]^T \), \( m(t) = [m_1(t), m_2(t), \ldots, m_n(t)]^T \), \( \hat{L}(t) = (L_1, L_2, \ldots, L_n)^T \), \( A = \text{diag}(a_1, a_2, \ldots, a_n) \), \( C = \text{diag}(c_1, c_2, \ldots, c_n) \), \( D = \text{diag}(d_1, d_2, \ldots, d_n) \), and \( W = (w_{ij})_{n \times n} \).

In the following, consider an uncertain genetic regulatory network with random disturbance:

\[
\begin{align*}
\frac{dm(t)}{dt} &= [-Am(t) + Wf(p(t)) + L]dt + v(t)dw \\
\frac{dp(t)}{dt} &= [-Cp(t) + Dm(t)]dt + \sigma(t)d\omega \\
y &= C_0x(t)
\end{align*}
\]

(6)

where \( y \) is the output, \( v(t) \) and \( \sigma(t) \) are external noise intensity functions, \( \omega(t) \) and \( \nu(t) \) are two independent one-dimensional Brownian motions satisfying \( E[\omega(t)] = 0, E[\omega(t)^2] = 1 \), \( E[\nu(t)] = 1 \), \( E[\nu(t)^2] = 1 \) (the mathematical expectation), \( x(t) = [m^T(t), p^T(t)]^T \), \( C_0 = (I_n 0_n) \), \( I_n \) and \( 0_n \) denote the \( n \)-dimension identity and zero matrices, respectively. The initial conditions of (6) are given by \( x(t) = x_0 \).

Sometimes, the concentrations of mRNA can be obtained. However, the regulation process in mRNA, i.e, the degradation rates \( a_i \), transcriptional factors \( w_{ij} \), repressor factors \( L_i \), and concentrations of protein \( p(t) \), are not known in many cases. Thus, assume that \( A, W, L \) are uncertain matrices and \( p(t) \) is unknown. Assume also that in system (6), the output \( y = m(t) \) can be observed. The objective is to design a filter to estimate the concentrations of mRNA. The nonlinear filter is of the form:

\[
\begin{align*}
\frac{d\tilde{m}(t)}{dt} &= -\hat{A}(t)\tilde{m}(t) + \tilde{W}(t)f(\tilde{p}(t)) + \tilde{L}(t) \\
\frac{d\tilde{p}(t)}{dt} &= -\hat{C}\tilde{p}(t) + D\tilde{m}(t) + \sigma(t)d\omega \\
y &= C_0\tilde{x}(t)
\end{align*}
\]

(7)

where \( \tilde{m}(t) = [\tilde{m}_1(t), \tilde{m}_2(t), \ldots, \tilde{m}_n(t)]^T \), \( \tilde{p}(t) = [\tilde{p}_1(t), \tilde{p}_2(t), \ldots, \tilde{p}_n(t)]^T \), \( \hat{K}(t) = \text{diag}(k_1(t), k_2(t), \ldots, k_n(t)) \) is the adaptive feedback matrix, \( \tilde{W}(t) = (\overline{w}_{ij}(t))_{n \times n} \), \( \tilde{A}(t) = \text{diag}(\tilde{a}_1(t), \tilde{a}_2(t), \ldots, \tilde{a}_n(t)) \), and \( \tilde{L}(t) = (\overline{L}_1(t), \overline{L}_2(t), \ldots, \overline{L}_n(t))^T \) are matrix and vector functions of time \( t \). Systems (6) and (7) can be considered as master and slave systems as in [9, 12].

Subtracting (6) from (7) yields the following error dynamical system:

\[
\begin{align*}
d\epsilon_m(t) &= [-\hat{A}(t)\tilde{m}(t) + \tilde{W}(t)f(\tilde{p}(t)) + \tilde{L}(t) \\
&\quad + Am(t) - Wf(p(t)) - L - K_0m(t)]dt - v(t)d\nu \\
d\epsilon_p(t) &= [-C\epsilon_p(t) + D\epsilon_m(t)]dt - \sigma(t)d\omega \\
y(t) &= C_0x(t)
\end{align*}
\]

(8)
Note that in system (4), the function \( f_i : \mathbb{R} \to \mathbb{R} \) is monotonically increasing and satisfies the Lipschitz condition

\[
|f_i(u) - f_i(v)| \leq h_i|u - v| \quad \forall u, v \in \mathbb{R}
\]

with the Lipschitz constant \( h_i > 0 \) for all \( i = 1, 2, \ldots, n \).

**Definition 1** [6] The two networks (6) and (7) are said to be stochastically synchronous with disturbance attenuation \( \gamma > 0 \) if

i) network (8) with \( v(t) = 0 \) and \( \sigma(t) = 0 \) is asymptotically stable;

ii) under zero initial conditions, there exists a scalar \( \gamma > 0 \) such that

\[
\mathbb{E} \int_0^\infty (\|e_m(s)\|^2 + \|e_p(s)\|^2)ds \leq \gamma^2 \mathbb{E} \int_0^\infty (\|v(s)\|^2 + \|\sigma(s)\|^2)ds
\]

for all nonzero \( v, \sigma \in L_2[0, \infty) \).

**Lemma 1** For any vectors \( x, y \in \mathbb{R}^n \) and positive definite matrix \( G \in \mathbb{R}^{n \times n} \), the following matrix inequality holds:

\[
2x^Ty \leq x^TGx + y^TG^{-1}y
\]

### 3 ADAPTIVE FILTER DESIGN

In this section, the adaptive laws \( K(t), \tilde{A}(t), \tilde{W}(t) \), and \( \tilde{L}(t) \) are designed to achieve the asymptotical stability of the error dynamical system (8) with disturbance attenuation.

**Theorem 1** The two networks (6) and (7) are stochastically synchronous with disturbance attenuation \( \gamma > 0 \) if

\[
\hat{a}_i = q_ie_{mi} \tilde{m}_i(t), \quad i = 1, 2, \ldots, n,
\]

\[
\hat{w}_{ij} = -r_{ij}e_{mi}f_j(\tilde{p}(t)), \quad i, j = 1, 2, \ldots, n,
\]

\[
\hat{l}_i = -u_ie_{mi}, \quad i = 1, 2, \ldots, n,
\]

\[
\hat{k}_i = \beta_i e_{mi}^2,
\]

and

\[
\gamma > \frac{1}{\sqrt{2(c - \mu)}}
\]

where \( q_i, r_{ij}, u_i, \) and \( \beta_i \) are positive constants, \( c = \min_{1 \leq i \leq n} \{c_i\} \), and \( \mu \) is a positive constant satisfying \( \mu < c \).

**Proof** Consider the Lyapunov functional

\[
V(t) = \frac{1}{2} e_m^T(t)e_m(t) + e_p^T(t)e_p(t) + \sum_{i=1}^n \frac{1}{2h_i}(\tilde{a}_i - a_i)^2 + \sum_{i=1}^n \sum_{j=1}^n \frac{1}{2r_{ij}}(\tilde{w}_{ij} - w_{ij})^2 + \sum_{i=1}^n \frac{1}{2u_i}(\tilde{l}_i - l_i)^2 + \sum_{i=1}^n \frac{1}{2\beta_i}(k_i - \beta_i)^2 \quad \text{(14)}
\]

where \( k_i, \delta_1, \) and \( \delta_2 \) are positive constants.

From Itô formula [10], one obtains the following stochastic differential:

\[
dV(t) = \mathcal{L}V(t)dt - e_m^T(t)e_m(t)dv - e_p^T(t)e_p(t)d\omega \quad \text{(15)}
\]

The weak infinitesimal operator \( \mathcal{L} \) of the stochastic process is given by

\[
\mathcal{L}V(t) = e_m^T(t)(\tilde{A}(t)\tilde{m}(t) + \tilde{W}(t)f(\tilde{p}(t))) + \tilde{L}(t) + Am(t) - Wf(p(t)) - L - K\tilde{e}_m(t)] + e_p^T(t)[-Ce_p(t) + D\tilde{e}_m(t)] + \sum_{i=1}^n (\tilde{a}_i - a_i)e_{mi}
\]

Choosing sufficiently small values \( \eta \) and \( \varepsilon \) so that \( c - \frac{\eta}{2}h^2 - \frac{\varepsilon}{2}\lambda_{max}(DD^T) > 0 \) and let \( \mu = \frac{\eta}{2}h^2 + \frac{\varepsilon}{2}\lambda_{max}(DD^T) \), \( k = -a + \frac{1}{2\varepsilon} + \frac{1}{2\eta}\lambda_{max}(WW^T) + \frac{1}{2\gamma^2} + 1 \). It is then easy...
to see that the condition i) in Definition 1 is satisfied. Next, it is to find the disturbance attenuation $\gamma$ so that condition ii) in Definition 1 is also satisfied. From (13) and (17), it is obvious that $J(t) < 0$. Then definition 1 is satisfied. Therefore, with the designed adaptive laws (9)-(12), the condition ii) in Definition 1 is satisfied. This completes the proof.

**Remark 1** Under adaptive laws (9)-(12), one has $\gamma > \frac{1}{\sqrt{2(e-\mu)}}$. If one chooses a sufficiently small value $\mu$, then one can obtain the low bound of $\gamma = \frac{1}{\sqrt{2e}}$. Thus, by adaptive laws (9)-(12), network synchronous error system (8) is stochastically stable about zero, with disturbance attenuation $\gamma = \frac{1}{\sqrt{2e}}$. In other words, the adaptive filter (7) has been successfully designed.

**Remark 2** In [11, 17], the synchronization problem of nonlinear systems with stochastic disturbance was also investigated. However, the noise intensity function therein is $\sigma(e(t), e(-\tau))$ in the response system, where $e$ is the error state. This function $\sigma$ can vanish on the synchronization manifold $e(t) = 0$. Furthermore, it is not realistic to assume that noise exists in the response system. In this paper, it is assumed that the noise exists in the master system, i.e., in the genetic regulatory network (6), and it does not vanish on the synchronization manifold.

**Remark 3** In [19], robust Kalman filter is used for discrete linear time-varying uncertain systems with noise and missing measurements, while in this paper the uncertain system parameters are assumed fully unknown and the genetic regulatory network is truly nonlinear.

## 4 SIMULATION EXAMPLE

In this section, a numerical example is given to show the effectiveness of the proposed adaptive filtering technique. The dynamics of the repressilator has been theoretically predicted and experimentally investigated on *Escherichia coli* [3]. Three repressor-protein concentrations $p_i$, and their mRNA concentrations $m_i$ (where $i$ is $lacz$, $tetR$ or $cl$) were treated as continuous dynamical variables. The repressilator is a cyclic negative-feedback loop composing of three genes and their corresponding promoters. The kinetics of the system are determined by six coupled first-order differential equations:

$$
\frac{d m_i(t)}{dt} = -m_i(t) + \frac{\alpha}{1 + p_i^2(t)} + L_i
$$

$$
\frac{d p_i(t)}{dt} = -c_i p_i(t) + d_i m_i(t)
$$

(19)

where $i = lacz, tetR, cl$; $j = cl, lacz, tetR$, and $n$ is a Hill coefficient. In this paper, the following genetic regulatory network is considered:

$$
\frac{d m(t)}{dt} = [-Am(t) + W f(p(t))] + L_1 dt + \nu(t) d\nu
$$

$$
\frac{d p(t)}{dt} = [-Cp(t) + Dm(t)] + \sigma(t) d\omega
$$

$$
y = m(t)
$$

(20)

where $A = \begin{pmatrix} 0.2 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.6 \end{pmatrix}$, $W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

$$
D = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}, L = \begin{pmatrix} 1.5 & 1.5 & 1.5 \end{pmatrix}^T,
$$

$\nu(t) = 0.1 e^{-0.05t}$, $f(z) = e^{-0.1t}$, $f(z) = \frac{z^2}{1 + z^2}$. Assume $w_{13} = -1.5$ and $l_1 = 1.5$ are unknown, and the adaptive filter (7) and (9)-(12) is designed. The error states $e_m$ and $e_p$ in (8), and the adaptive laws $\tilde{w}_{13}$ and $\tilde{l}_1$ in (10) and (11) are shown in Figure 1, where $\tau_{13} = u_{1} = 10$. It is easy to see that the adaptive filter (7) and (9)-(12) works very effectively.

**Remark 4** In [9, 12], the parameters identification problem for dynamical systems is carried out from time series, where a linear independence condition is derived, which is sufficient for parameter identification of general dynamical systems. One can see from Figure 1 that parameters $w_{13} = -1.5$ and $l_1 = 1.5$ could be approximately estimated by $\tilde{w}_{13}$ and $\tilde{l}_1$, as expected.

## 5 CONCLUSION

In this paper, an adaptive filter for unknown genetic regulatory networks with disturbance attenuation has been designed and simulated. Several new adaptive laws have been developed to ensure the stochastic stability of the synchronous error states between the unknown genetic regulatory network and its estimated model with disturbance attenuation. The designed adaptive laws are independent of the unknown system states and parameters. In addition, noise is assumed to exist in the genetic regulatory network and does not vanish on the synchronization manifold. Furthermore, the concentrations of protein and some network parameters are all assumed to be unknown. The success of the proposed adaptive filtering technique shows that it is a powerful and promising approach to complete unknown networks identification, hence should be further studied in the future.

## 6 ACKNOWLEDGEMENT

The authors thank Prof. Zidong Wang, Prof. Jinde Cao, and Dr. Fengli Ren for hot discussions and helpful suggestions.

## REFERENCES


Fig. 1  Error states $e_m$ and $e_p$, and adaptive laws $\tilde{w}_{13}$ and $\tilde{l}_1$.


