International Journal of Computational Intelligence in Control
Vol. 12 No. 2 (July-December, 2020)

AN ALGORITHM FOR ESTIMATING THE EQUILIBRIUM POINTS IN GENETIC REGULATORY NETWORKS WITH POLYTOPIC UNCERTAINTIES

Graziano Chesi

Department of Electrical and Electronic Engineering, University of Hong Kong
Pokfulam Road, Hong Kong, *E-mail: chesi@eee.hku.hk*

ABSTRACT: This paper addresses the estimation of the equilibrium points in uncertain genetic regulatory networks with regulation functions of various type. The uncertainty is represented as an unknown vector constrained in a polytope and affects the coefficients of the mathematical model of the genetic regulatory network via affine functions. An algorithm is hence proposed for estimating the equilibrium points, which progressively splits the concentrations space into smaller sets discarding those that do not contain equilibrium points of the considered uncertain model. Some numerical examples illustrate the proposed algorithm.

Keywords: Genetic Network; Uncertainty; Equilibrium Point; Robustness.

INTRODUCTION

It is well-known that genetic regulatory networks play a key role in systems biology, as they explain the fundamental interactions between genes and proteins in living organisms, see e.g. [1-11]. An important problem in genetic regulatory networks consists of determining the equilibrium points. In fact, the knowledge of the equilibrium points provides qualitative and quantitative information about the temporal evolution of mRNA and protein concentrations, see e.g. [12] which considers the problem of establishing stability of equilibrium points of genetic regulatory networks. Determining the equilibrium points amounts to solving a system of nonlinear equations since the temporal derivative of the mRNA and protein concentrations is a nonlinear functions of these concentrations. This operation is non-trivial since there do not exist techniques that guarantee to find all solutions of a generic nonlinear system, see e.g. [13-16].

The problem, however, is even more difficult in practice. In fact, mathematical models of genetic regulatory networks are never exactly known. This is due to various reasons, in particular to the fact that the experimental data used to identify the coefficients of the model are unavoidably affected by noise and measurement errors. This means that mathematical models of genetic regulatory networks contain uncertain parameters as proposed e.g. in [17]. These uncertain parameters affect the equilibrium points that, as a result, are uncertain as well.

This paper addresses the estimation of the equilibrium points in uncertain genetic regulatory networks. Specifically, regulation functions of various type are considered through a generalized model, and the uncertainty is represented as an unknown vector constrained in a polytope, which affects the mathematical model of the genetic regulatory network through affine functions. An algorithm is hence proposed for estimating the equilibrium points, which progressively splits the concentrations space into smaller sets discarding those that do not contain equilibrium points of the considered uncertain model. The proposed algorithm is illustrated through some numerical examples.

The paper is organized as follows. The next section introduces some preliminaries on genetic regulatory networks. Then, we describe the proposed algorithm and its properties. Some numerical examples are hence provided in order

to illustrate the proposed algorithm. Lastly, we conclude the paper with some final remarks.

PRELIMINARIES

In this section we provide some preliminaries about genetic regulatory networks and we state the problem formulation. Before proceeding, let us introduce the notation used throughout the paper:

- R: space of real numbers;
- R_+: space of non-negative real numbers, i.e. $\{x \in R : x \geq 0\}$;
- 0_n: null vector of size $n \times 1$;
- ver(Z) : set of vertices of polytope Z;
- TF: transcriptor factor.

We consider genetic regulatory networks described by differential equation models, in particular according to

$$\dot{m}(t) = -Am(t) + b(p(t))$$

where m, $p \in R_+^n$ are vectors containing the concentrations of mRNA and protein; A, C and D are diagonal positive definite matrices; and b is a nonlinear function such that each entry of $b(p)$ is bounded and monotonic with respect to each entry of p.

For instance, in genetic regulatory networks with SUM form, the i-th entry of $b(p)$ is expressed as a linear combination of functions of a single variable, i.e.

$$b_i(p) = \alpha_{i,1} b_{i,1}(p_1) + \ldots + \alpha_{i,n} b_{i,n}(p_n)$$

where $\alpha_{i,1},\ldots, \alpha_{i,n} \in R$ and $b_{i,1}(p_1),\ldots, b_{i,n}(p_n)$ are nonlinear, bounded and monotonic. In genetic regulatory networks with PROD form, the function b(p) is expressed as product of the functions, i.e.

$$b_i(p) = \alpha_i b_{i,1}(p_1)\ldots b_{i,n}(p_n)$$

where $\alpha_i \in R$. Each function $b_{i,j}(p_j)$ is typically expressed as

$b_{i,j}(p_j) = f(p_j)$ if TF j is an activator of gene i

$b_{i,j}(p_j) = 1 - f(p_j)$ if TF j is a repressor of gene i

$b_{i,j}(p_j) = \gamma$ otherwise

where $\gamma \in R$ is a constant depending on the model, in particular $\gamma = 0$ for SUM form and $\gamma = 1$ for PROD form. The function $f(x)$ is a saturation function, i.e. a function satisfying the following properties:

- $f : R_+ \to [0,1]$
- $f(0) = 0$
- $f(x) \to 1$ as $x \to \infty$
- $f(x_2) \geq f(x_1)$ for all x_1, $x_2 : x_1 \leq x_2$.

In the case of Hill functions, $f(x)$ has the form

$$f(x) = x^h / (\beta^h + x^h)$$

where $\beta \in R$ and h is an integer known as Hill coefficient.

As explained in the introduction, mathematical models of genetic regulatory networks always contain uncertainties on their coefficients. This fact can be expressed by introducing uncertain parameters in the previous model, according to

where z is a vector containing the uncertain parameters constrained according to

$$z \in Z$$

where Z is a polytope that we describe as

$Z = \{z \in R^r : z = a_1 z^{(1)} + \ldots + a_w z^{(w)}, a_i \geq 0, a_1 + \ldots + a_w = 1\}$

for some vectors $z^{(1)},\ldots, z^{(w)}$. The functions $A(z)$, $C(z)$, $D(z)$ and $b(p, z)$ are affine in z and describe an admissible genetic regulatory network for all $z \in Z$.

The problem addressed in this paper consists of estimating the set of possible equilibrium points of the uncertain genetic regulatory network, i.e.

$S = \{(m,p) \in R_+^{2n} :$ and $\dot{p} = 0$ for some $z \in Z\}$.

ESTIMATING THE EQUILIBRIUM POINT LOCATIONS

This section describes the proposed strategy. In particular, we first explain how estimates of the equilibrium point locations can be obtained

through an iterative strategy based on worst-case evaluations of some appropriate functions. Then, we explain how these worst-case evaluations can be performed based on the type of functions and number of variables involved. First of all, let us observe that S can be written as

$$S = \cup_{z \in Z} S_z$$

where S_z is the set of equilibrium points for the considered value of z, i.e.

$$S_z = \{(m, p) \in R_+^{2n} : \dot{m} = 0 \text{ and } \dot{p} = 0\}$$

Let us also observe that (m, p) belongs to S_z if and only if the following system of nonlinear equations is satisfied:

$$-A(z)m + b(p, z) = 0_n$$
$$-C(z)p + D(z)m = 0_n$$
$$m,p \in R_+^n$$

From the second equation one can determine m as a function of p since $D(z)$ is nonsingular. This implies that the system can be equivalently rewritten as

$$A(z)D(z)^{-1}C(z)p + b(p, z) = 0_n$$
$$m = -D(z)^{-1}C(z)p$$
$$p \in R_+^n$$

Therefore, in the sequel we will focus on the computation of the vectors p fulfilling this system, which represent the p-part of S_z. We indicate the set of such vectors as

$$P_z = \{p \in R_+^n : A(z)D(z)^{-1}C(z)p + b(p,z) = 0_n\}.$$

Similarly, we define the p-part of S as

$$P = \{p \in R_+^n : A(z)D(z)^{-1}C(z)p + b(p, z) = 0_n \text{ for some } z \in Z\}.$$

The basic idea of the proposed strategy consists of understanding whether a hyperrectangle may contain points of P. To this, let us denote a generic hyperrectangle as

$$H = \{p \in R_+^n : p_i \in [p_{i-}, p_{i+}]\}$$

for some p_{i-}, $p_{i+} \in R_+$, and let us denote the vertices of H as

$$\text{ver}(H) = \{p^{(1)}, \ldots, p^{(l)}\}.$$

The first step of the proposed strategy consists of defining the function

$$A(H) = \{p \in R_+^n : p_i \in [q_{i-}, q_{i+}]\}$$

where

$$q_{i-} = \min_{j=1,\ldots,l} q_{ij-}$$
$$q_{i+} = \max_{j=1,\ldots,l} q_{ij+}$$

and

$$q_{ij-} \le \min_{z \in Z} u_i(p^{(1)}, z)$$
$$q_{ij+} \ge \max_{z \in Z} u_i(p^{(1)}, z)$$

where $u_i(p, z)$ is the i-th entry of the vector function $u(p,z)$ defined as

$$u(p, z) = -C(z)^{-1}D(z)A(z)^{-1}b(p, z).$$

We have that

$$p \in H \cap P \Rightarrow p \in A(H).$$

Moreover,

$$H \cap A(H) = \text{empty-set} \Rightarrow H \cap P = \text{empty-set}.$$

Evaluating the function $A(H)$ requires the computation of the quantities q_{i-} and q_{i+} which are the minimum and the maximum of finite sequences. The construction of these sequences amounts to finding bounds of the function $u_i(p,z)$ for z variable over Z. In the sequel we will explain how this step can be addressed depending on the dependence of $u(p,z)$ on z and on the dimension of z.

From the function $A(H)$ we define the function $B(H)$ according to the following rules:

- (Step B1) set $H^{(0)} = H$ and $k = 0$;
- (Step B2) set $B_1 = H^{(k)} \cap A(H^{(k)})$;
- (Step B3) if B_1 is empty, set $B(H) = B_1$ and exit;
- (Step B4) if $B_1 = H^{(k)}$, set $B(H) = B_1$ and exit;
- (Step B5) set $k = k+1$, $H^{(k)} = B_1$ and go to Step B2.

The function B(H) returns either the empty-set, a point, or a hyperrectangle. Moreover:

$$B(H) \text{ is included in } H$$

and

$$p \in H \cap P \Rightarrow p \in B(H).$$

The function B(H) transforms a given hyperrectangle via a sequence of applications of the function A(·), and returns a set which can

be either the empty set, a point, or a hyperrectangle. By exploiting the function B(H) we derive the algorithm for the computation of the sought set P as follows.

Let H be a hyperrectangle and let us define the function C(H,k) in the following way:

- (Step C1) if k = 0, set C(H, k) = B(H) and exit;
- (Step C2) if B(H) is either the empty set or a point, set C(H,k) = B(H) and exit;
- (Step C3) divide the hyperrectangle B(H) in hyperrectangles $H_1,\ldots, H_s$ such that

$$H_i \cap H_j = \text{empty-set for all } i \neq j$$

$$\cup_{i=1,\ldots,s} H_i = H$$

- (Step C4) set

$$C(H, k) = \cup_{i=1,\ldots,s} C(H_i,k\text{-}1)$$

and exit.

The function $C(H,k)$ satisfies the following properties:

$$C(H,k+1) \text{ is included in } C(H,k)$$

and

$$P \text{ is included in } C(R_+^n, k).$$

The proposed strategy for estimating the set of possible equilibrium points amounts to evaluating the function $C(R_+^n,k)$ for some integer $k \geq 0$. As explained above, the output of this function is a set that does not increase with k, and possibly decreases. Moreover, for any k, the output is guaranteed to contain all the possible equilibrium points, i.e. the set P.

Let us observe that $C(R_+^n,k)$ is typically a family of hyperrectangles, and that in special cases this family can also contain isolated points or be the empty set. Specifically, isolated points may appear for example when $A(z)D(z)^{-1}C(z)p + b(p,z) = 0_n$ admits solutions for p that are independent on z. Then, $C(R_+^n,k)$ may be the empty set when P is empty.

Now, let us address the construction of the quantities q_{ij-} and q_{ij+} which are a lower bound and an upper bound of the function $u_i(p^{(j)},z)$ for z variable over Z. We will discuss how these bounds can be found in some cases of interest, depending on the dependence of the function $u(p,z)$ on z and on the dimension of z.

First of all, let us observe that we aim to determine bounds that are as tight as possible, since the tighter q_{ij-} and q_{ij+} are, the less conservative the estimate of P found is. In fact, tighter bounds q_{ij-} and q_{ij+} yield smaller hyperrectangles A(H), and consequently smaller estimates B(H) and $C(H,k)$. We indicate the tightest values for q_{ij-} and q_{ij+} with q_{ij-}^* and q_{ij+}^* where

$$q_{ij-}^* = \min_{z \in Z} u_i(p^{(j)},z)$$

$$q_{ij+}^* = \max_{z \in Z} u_i(p^{(j)},z).$$

Let us consider the case where the function $u_i(p^{(j)},z)$ is monotonic with respect to each entry of z. For instance, this is the case when $u_i(p^{(j)},z)$ is linear in z. In such a case, q_{ij-}^* and q_{ij+}^* can be simply found as

$$q_{ij-}^* = \min_{k=1,\ldots,w} u_i(p^{(j)},z^{(k)})$$

$$q_{ij+}^* = \max_{k=1,\ldots,w} u_i(p^{(j)},z^{(k)}).$$

Another case of interest is when z is a scalar, e.g. $r = 1$. In fact, in such a case, $u_i(p^{(j)},z)$ is a function of one (scalar) variable only, and Z is an interval. This means that the computation of q_{ij-}^* and q_{ij+}^* amounts to finding the roots of an univariate function. Indeed:

$$q_{ij-}^* = \min_{z \in W} u_i(p^{(j)},z)$$

$$q_{ij+}^* = \max_{z \in W} u_i(p^{(j)},z)$$

where W is a finite set given by

$$W = W_1 \cup W_2$$

$$W_1 = \text{frontier of } Z$$

$$W_2 = \{z \in Z : du_i(p^{(j)},z)/dz = 0\}.$$

Observe that, since $u_i(p^{(j)},z)$ is rational in z, determining W_2 (and hence W) amounts to computing the roots of an univariate polynomial, operation that can be easily done.

Lastly, for the case where $u_i(p^{(j)},z)$ is a generic rational function in z, bound q_{ij-} and q_{ij+} (and possibly their tight values q_{ij-}^* and q_{ij+}^*) can be found by solving convex optimization problems with linear matrix inequality constraints. See [18,19] and references therein for details.

ILLUSTRATIVE EXAMPLES

This section presents some illustrative examples where the proposed strategy is adopted to estimate the equilibrium point locations of uncertain genetic regulatory networks. The computations are performed with Matlab 7 running under Windows 7 on a standard personal computer (Intel Core 2, 3 GHz, 4 GB RAM).

Example 1

Let us start by considering a simple example with an uncertain genetic regulatory network described by

$$\dot{m}_1(t) = -m_1(t) + 1 - f(p_2(t))$$

where $f(x)$ is a Hill function according to

$$f(x) = x^2/(1+x^2)$$

and where the uncertain parameter z is constrained according to

$$z \in [0, 1].$$

This genetic regulatory network is characterized by the fact that the TF 1 is a regressor of the gene 2, and the TF 2 is a regressor of the gene 1. The mathematical model contains an uncertain parameter which affects the linear part of the regulatory function, in particular the dependence of the temporal derivative of the protein concentration p_1 on the mRNA concentration m_1.

Let us consider the problem of estimating the equilibrium point locations, in particular the set P. To this end, we adopt the proposed strategy, which consists of evaluating the function $C(R_+^2,k)$ for some chosen k, which defines the number of recursive steps and hence the accuracy of the solution.

Specifically, with k = 0 the positive quadrant R_+^2 is shrunk via the function B(H) to the rectangle shown in Figure 1a. With k = 1, this rectangle is divided in four equal rectangles, on which the function B(H) is re-applied. This provides the three rectangles shown in Figure 1b: observe in fact that one rectangle has been shrunk to the empty set. Proceeding in this way, we obtain the estimates of P shown in Figure 1c (with k = 2) and Figure 1d (with k = 3).

Example 2

Let us start by considering a simple example with an uncertain genetic regulatory network described by

$$= -0.2m_1(t) + (2.1 - z_1)(1 - f(p_2(t)))$$

$$= -0.9m_2(t) + (3.7 - 2z_2)(1 - f(p_1(t))) + 0.5f(p_2(t))$$

$$= -0.5p_1(t) + 1.4m_1(t)$$

$$= -0.6p_2(t) + 0.8m_2(t)$$

where $f(x)$ is a Hill function according to

$$f(x) = 1-\exp(-x^2)$$

and z is an uncertain parameter constrained according to

$$z \in [-1, 1]^2.$$

The mathematical model of this genetic regulatory network contains two uncertain parameters which affect the nonlinear part of the regulatory function, in particular the dependence of the temporal derivative of the mRNA concentrations m_1 and m_2 on the protein concentrations p_1 and p_2.

Let us estimate the equilibrium point locations via the function $C(R_+^2, k)$. With $k = 0$ the positive quadrant R_+^2 is shrunk via the function $B(H)$ to the rectangle shown in Figure 2a. With $k = 1$, $k = 2$ and $k = 3$ we obtain the estimates shown in Figures 2b, 2c and 2d, respectively.

Before concluding this example it is worth observing that the saturation function is not a Hill function, in particular $f(pi)$ is irrational, and there do not exist techniques able to find all solutions of a system of nonlinear equations with irrational nonlinearities.

Example 3

Here we consider the repressilator investigated in Escherichia coli [20] where the coefficients

of the mathematical model are not exactly known. Specifically, we consider

$$= -m_i(t) + z_i(1 - f(p_j(t)))$$

$$= -p_i(t) + m_i(t)$$

$$i = \text{lacl,tetR,cl};\ j = \text{cl,lacl,tetR}$$

where $f(x)$ is a Hill function according to

$$f(x) = x^4/(1 + x^4)$$

and z is an uncertain parameter constrained according to

$$z_1 \in [1,3]$$

$$z_2 \in [2,5]$$

$$z_3 \in [3,8].$$

The mathematical model contains three uncertain parameters which affect the nonlinear part of the regulatory function, in particular the dependence of the temporal derivative of the mRNA concentrations m_i on the protein concentrations p_i.

Let us estimate the equilibrium point locations via the function $C(R_+^3,k)$. With $k = 0$ the positive octant R_+^3 is shrunk via the function $B(H)$ to the hyperrectangle shown in Figure 3a. With $k = 1$, this hyperrectangle is divided in eight equal hyperrectangles, on which the function $B(H)$ is re-applied. This provides the hyperrectangle shown in Figure 3b: observe in fact that seven hyperrectangles have been shrunk to the empty set. With $k = 2$ and $k = 3$ we find the family of hyperrectangles shown in Figure 3c–d.

CONCLUSION

We have proposed an algorithm for estimating the equilibrium points in genetic regulatory networks with polytopic uncertainties. The proposed algorithm is based on worst-case evaluations of some appropriate functions of the uncertainty, and provides regions containing all possible equilibrium points via an iterative method that progressively splits the concentrations space into smaller sets discarding those that do not contain equilibrium points of the considered model.

REFERENCES

[1] F. Jacob and J. Monod. Genetic Regulatory Mechanisms in the Synthesis of Proteins. *Journal of Molecular Biology*, 3: 318–356, 1961.

[2] J. M. Bower and H. Bolouri, Editors. Computational Modeling of Genetic and Biochemical Networks. Computational Molecular Biology. MIT Press, 2001.

[3] L. Chen and K. Aihara. Stability of Genetic Regulatory Networks with Time Delay. *IEEE Trans. on Circuits and Systems I*, 49(5): 602–608, 2002.

[4] H. Toh and K. Horimoto. System for Automatically Inferring a Genetic Network from Expression Profiles. *Journal of Biological Physics*, 28: 449–464, 2002.

[5] T. Akutsu, S. Kuhara, O. Maruyama, and S. Miyano. Identification of Genetic Networks by Strategic Gene Disruptions and Gene Overexpressions under a Boolean Model. *Theoretical Computer Science*, 298: 235–251, 2003.

[6] L. Chen, R. Wang, T. Kobayashi, and K. Aihara. Dynamics of Gene Regulatory Networks with Cell Division Cycle. *Physical Review E*, 70: 011909, 2004.

[7] C. Li, L. Chen, and K. Aihara. Stability of Genetic Networks with Sum Regulatory Logic: Lure System and LMI Approach. *IEEE Trans. on Circuits and Systems I*, 53(11): 2451–2458, 2006.

[8] C. Li, L. Chen, and K. Aihara. A Systems Biology Perspective on Signal Processing in Genetic Network Motifs. *IEEE Signal Processing Magazine*, 221(3): 136–142, 2007.

[9] C. Li, L. Chen, and K. Aihara. Stochastic Stability of Genetic Networks with Disturbance Attenuation. *IEEE Transactions on Circuits and Systems II*, 54(10): 892–896, 2007.

[10] S. R. Taylor, R. Gunawan, L. R. Petzold, and F. J. Doyle. Sensitivity Measures for Oscillating Systems: Application to Mammalian Circadian Gene Network. Joint Special Issue on Systems Biology, IEEE Transactions on Automatic Control & IEEE Transactions on Circuits and Systems I: Regular Papers, 53(1): 177–188, 2008.

[11] W. Yu, J. Lu, G. Chen, Z. Duan, and Q. Zhou. Estimating Uncertain Delayed Genetic Regulatory Networks: An Adaptive Filtering Approach. *IEEE Trans. on Automatic Control*, 54(4): 892–897, 2009.

[12] G. Chesi. Polynomial Relaxation-based Conditions for Global Asymptotic Stability of

Equilibrium Points of Genetic Regulatory Networks. *Int. Journal of Systems Science*, 41(1): 65–72, 2010.

[13] G. Chesi, A. Garulli, A. Tesi, and A. Vicino. Characterizing the Solution Set of Polynomial Systems in Terms of Homogeneous Forms: An LMI Approach. *Int. Journal of Robust and Nonlinear Control*, 13(13): 1239–1257, 2003.

[14] T. Mora. Solving Polynomial Equation Systems II. Cambridge University Press, 2005.

[15] J. Nocedal and S. Wright. Numerical Optimization. Springer, 2006.

[16] G. Chesi. Computing Equilibrium Points of Genetic Regulatory Networks. Trans. on Computational Systems Biology XI, LNBI 5750, pages 268–282, 2009.

[17] G. Chesi and Y. S. Hung. Stability Analysis of Uncertain Genetic SUM Regulatory Networks. Automatica, 44(9): 2298–2305, 2008.

[18] G. Chesi, A. Garulli, A. Tesi, and A. Vicino. Homogeneous Polynomial Forms for Robustness Analysis of Uncertain Systems. Springer, 2009.

[19] G. Chesi. LMI Techniques for Optimization over Polynomials in Control: A Survey. *IEEE Trans. on Automatic Control*, 55(11): 2500–2510, 2010.

[20] M. B. Elowitz and S. Leibler. A Synthetic Oscillatory Network of Transcriptional Regulators. *Nature*, 403:335–338, 2000.