Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-7 (2016) 206–211

On the applicability of deterministic approximations to model genetic circuits ⋆

M. Pájaro * **A. A. Alonso** **

* *Process Engineering Group, IIM-CSIC. Spanish Council for Scientific Research. Eduardo Cabello 6, 36208 Vigo, Spain (e-mail: mpajaro@iim.csic.es).*

** *Process Engineering Group, IIM-CSIC. Spanish Council for Scientific Research. Eduardo Cabello 6, 36208 Vigo, Spain (e-mail: antonio@iim.csic.es)*

Abstract: Theoretical results and simulations support the idea that deterministic models provide an acceptable description only for large numbers of molecules. In the context of GRN, which usually involve a small number of molecules, such arguments might lead to disregard deterministic models as unsuitable representations.
We found, however, strong evidences that justify their use to model self-regulatory genetic circuits, even for small number of molecules. In fact, we show that under some conditions, a stochastic system showing a switching-like behaviour (manifested on a bimodal distribution) nearly coincides with a deterministic counterpart exhibiting bistability. Moreover, and contrary to what it might be expected, we find situations involving large numbers of molecules where the deterministic model results into a poor approximation. The analysis and methods presented are expected to help selecting the most adequate system's representation.

Keywords: Gene circuits, mathematical models, biochemical switches, bistability, stochastic systems, model reduction.

1. INTRODUCTION

Essentially, gene regulatory networks (GRN) (considered as the software-hardware architecture of the cell) execute the program written in the genome to adapt the physiological state of the cell in response to environmental signals. Such networks usually comprise a large number of biochemical reactions which can be conceptually described as the assembly of simple biochemical structures, conceived as efficient abstractions of the central (transcription-translation) dogma (Sherman and Cohen, 2014). Regulatory functions produced by negative or positive feedback are among the most common mechanisms (Paulsson and Ehrenberg, 2000; Friedman et al., 2006; Shahrezaei and Swain, 2008; Sherman and Cohen, 2014).

The underlying biochemical machinery typically involves a few number of molecules, what makes its behavior inherently stochastic. In describing GRN dynamics, a number of microscopic (stochastic) and deterministic modelling approximations has been attempted with mixed results (Kepler and Elston, 2001; Gillespie, 2007; Rosenfeld et al., 2002; Mackey et al., 2011). Microscopic descriptions revolve around the chemical master equation (CME), with different approximations such as moment methods (Engblom, 2006), finite state projection (Munsky and Khammash, 2006), hybrid models (Jahnke, 2011), or direct stochastic simulation algorithms (SSA) (Gillespie, 2007), oriented to reduce complexity. Deterministic models, on the other hand, are based on classical biochemical kinetics and can be formally represented by sets of ordinary differential equations (ODE). They have been used over the past recent years to get qualitative insights on GRN dynamics Mackey et al. (2011).

Theoretical results and simulations (Van Kampen, 2007; Gillespie, 2009; Wallace et al., 2012) support the idea that deterministic models provide an acceptable description of systems with large numbers of molecules, whereas the quality of the approximation deteriorates as that number reduces (Shmulevich and Aitchison, 2009). In the context of GRN, which usually involve small number of molecules, such arguments might lead to disregard deterministic models as unsuitable representations.

However, for a general class of self-regulatory genetic circuits we found out strong evidences that justify their use even under a small number of molecules condition. The class comprises those GRN where proteins are produced in bursts (e.g. Shahrezaei and Swain, 2008; Dar et al., 2012), what seems to be often the case, both in prokaryotic and eukaryotic cell types (Dar et al., 2012).

We show that under some conditions, a stochastic system showing a switching-like behaviour (manifested on a bimodal distribution) nearly coincides with a deterministic counterpart exhibiting bistability, what confirms the validity of the deterministic approximation for small number of molecules. Note however that bimodality and bistability are not completely interchangeable, as one can find many

⋆ This work has been partially supported by Ministerio de Economia y Competitividad AGL2012-39951-C02-01. M. Pájaro acknowledge support from grant BES-2013-063112.

10.1016/j.ifacol.2016.07.251

other instances in which the bimodal/binary stationary distribution associated to the stochastic system does correspond with a monostable deterministic counterpart.

In order to identify parameter regions, where deterministic approximations capture the essential features of the stochastic dynamics (average protein levels or coexistence of multiple stationary states), we adapt the method develop in Pájaro et al. (2015) to cope with bistability. We show that the quality of the deterministic approximation is at a large extent conditioned by the average number of bursts, and it improves as the value of this parameter increases. On the other hand, and contrary to what it might be expected, we find situations involving a high number of molecules where the deterministic model results into a poor approximation.

Hopefully, the analysis and methods presented can be of help for selecting the most adequate representation of system dynamics or to decide which one is preferable depending on the network structure and parameters.

The article is structured as follows: In Section 2, we describe the gene regulatory system and its stochastic representation together with the deterministic (ODE based) approximation. A method to characterise the regions in the parameter space that sustain bimodal or binary response and bistable behaviour is presented in Section 3. Main results are discussed in Section 4. We end up with some conclusions and future work.

2. THE SYSTEM AND ITS REPRESENTATION

The genetic system under study consists of a transcription-translation network involving a single gene that expresses a protein X which regulates its own production. The representative biochemical steps, including protein and $mRNA$ degradation, are depicted in Fig. 1.

As reported in Huang et al. (2015), RNA transcription may occur also at the inactive promoter state, a phenomenon that is known as transcriptional leakage. We assume that the basal transcription level from the inactive promoter takes place at a rate constant k_ε lower than k_1, (Friedman et al., 2006; Ochab-Marcinek and Tabaka, 2015).

Typically, self-regulation is described by a function of the form (Friedman et al., 2006; Ochab-Marcinek and Tabaka, 2015):

$$\bar{c}(x) = [1 - \rho(x)] + \rho(x)\varepsilon\,, \tag{1}$$

with x representing protein level, $\varepsilon = \frac{k_\varepsilon}{k_1} \in (0, 1)$ the transcriptional leakage constant and $\rho(x)$, a Hill-type function (Alon, 2007) that relates x to the fraction of DNA_{off}:

$$\rho(x) = \frac{x^H}{x^H + K^H}. \tag{2}$$

where $K = \frac{k_{off}}{k_{on}}$ is an equilibrium constant, and H a parameter proportional to the number of protein molecules bonded to the promoter. Its values can be positive or negative depending on whether the circuit represses or promotes protein production, thus resulting into a negative or positive feedback, respectively.

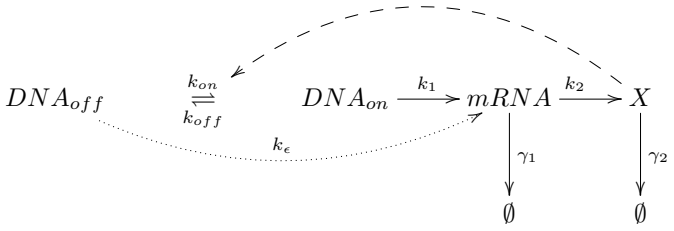

Fig. 1. Schematic representation of the transcription-translation mechanism under study. The promoter associated with the gene of interest is assumed to switch between active (DNA_{on}) and inactive (DNA_{off}) states, with rate constants k_{on} and k_{off} per unit time, respectively. In this study, the transition is assumed to be controlled by a feedback mechanism induced by the binding/unbinding of a given number of X-protein molecules, what makes the network self-regulated. Transcription of messenger RNA ($mRNA$) from the active DNA form, and translation into protein X are assumed to occur at rates (per unit time) k_1 and k_2, respectively. k_ε is the rate constant associated with transcriptional leakage. Both $mRNA$ and X-protein degradation are assumed to occur by first order processes with rate constants γ_1 and γ_2, respectively.

2.1 The microscopic description

In the following we will consider gene regulatory networks where the rate of $mRNA$ degradation is much faster than the corresponding to protein so that $\gamma_1/\gamma_2 \gg 1$. Under such condition, protein will be produced in bursts (e.g. Shahrezaei and Swain, 2008; Dar et al., 2012), what supports a description of the protein probability distribution based on the following partial integro-differential equation firstly proposed by Friedman et al. (2006):

$$\begin{aligned}\frac{\partial p(\tau, x)}{\partial \tau} = &\frac{\partial}{\partial x}(x p(\tau, x)) + \\ &a \int_0^x w(x - x')\bar{c}(x')p(\tau, x')\mathrm{d}x'\,,\end{aligned} \tag{3}$$

where $\tau = \gamma_2 t$ represents a dimensionless time associated with the time scale of protein degradation, and $a = k_1/\gamma_2$ is the dimensionless rate constant for transcription that relates to the mean number of bursts produced per cell cycle (e.g. burst frequency). The first term in the right-hand side of the equation accounts for protein degradation, whereas the integral describes protein production in bursts. Since burst size is assumed to follow an exponential distribution (Elgart et al., 2011), the conditional probability for protein level to jump from a state x' to x after a burst can be expressed as:

$$w(x - x') = (1/b)\exp((x' - x)/b) - \delta(x - x') \tag{4}$$

where parameter $b = k_2/\gamma_1$, is a dimensionless rate constant associated with translation which corresponds with the mean number of proteins produced per burst (i.e. burst size). Finally, the feedback mechanism is modelled by incorporating into the integral term the function $\bar{c}$ previously defined in (1).