Preprint, 11th IFAC Symposium on Dynamics and Control of Process Systems, including Biosystems
June 6-8, 2016. NTNU, Trondheim, Norway

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.

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.

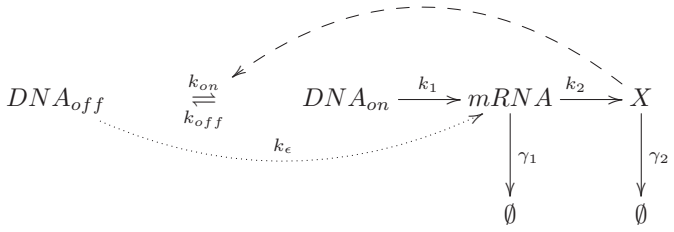

Fig. 1. Schematic representation of the transcription-translation mechanism under study. The promoter associated with the gene of interest is assumed to switch between active (DNA_{on}) and inactive (DNA_{off}) states, with rate constants k_{on} and k_{off} per unit time, respectively. In this study, 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):

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

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).

The stationary solution for (3) (identified as that satisfying $\partial p/\partial \tau = 0$) that we denote by P can be written as (Pájaro et al., 2015):

$$P(x) = C(\rho(x))^{\frac{a(1-\varepsilon)}{H}} x^{-(1-a\varepsilon)} e^{\frac{-x}{b}}, \tag{5}$$

where C is an integration constant that normalizes the corresponding probability distribution function so that $\int_0^\infty P(x)\,\mathrm{d}x = 1$.

The genetic regulatory circuit we have just described supports two possible phenotypes which correspond with the following types of stationary distributions:

1 A bimodal distribution having two maxima, which characterizes transitions between two positive protein levels.
2 A binary distribution having a (positive) minimum and a maximum, which characterizes transitions between a positive and a zero protein level, with the latter denoting the absence of protein.

In order to identify such distributions, conditions for extremal points in (5) are studied. The first derivative of P can be written as:

$$\frac{dP}{dx}(x) = \pi(x)P(x)\,, \tag{6}$$

where the function π takes the form:

$$\pi(x) = \frac{a(1-\varepsilon)}{x}\left[r(x) - \rho(x)\right]\,, \tag{7}$$

Extremal points are found as those which make zero expression (6). Since $P(x)$ has to be positive for positive protein levels, these must coincide with the roots of the following equation:

$$r(x) - \rho(x) = 0, \tag{8}$$

with r being a linear function of the form:

$$r(x) = \frac{-x}{ab(1-\varepsilon)} + \frac{(a-1)}{a(1-\varepsilon)}\,. \tag{9}$$

2.2 A macroscopic approximation

A deterministic (or macroscopic) approximation of the system depicted in Fig. 1 consists of a mass action law based kinetic description, which results into the following ODE set:

$$\frac{\mathrm{d}m}{\mathrm{d}\tau} = a\bar{c}(x) - \frac{\gamma_1}{\gamma_2}m \tag{10}$$

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = b\frac{\gamma_1}{\gamma_2}m - x \tag{11}$$

where m and x represent the concentrations of $mRNA$ and protein, respectively, and $\tau = \gamma_2 t$ is the dimensionless time associated with the scale of protein degradation. From the steady state solution (i.e. that which satisfies $\mathrm{d}m/\mathrm{d}\tau = \mathrm{d}x/\mathrm{d}\tau = 0$), we have that:

$$m(x) = a\frac{\gamma_2}{\gamma_1}\bar{c}(x). \tag{12}$$

Replacing the above expression in (11), using (1) and reordering terms, we finally obtain the equilibrium states for the system, which correspond with the roots of the equation:

$$r_d(x) - \rho(x) = 0, \tag{13}$$

where:

$$r_d(x) := \frac{-x}{ab(1-\varepsilon)} + \frac{1}{1-\varepsilon}\,. \tag{14}$$

Bistability will occur whenever the above expression (13) has exactly three roots, i. e., two stable points and one unstable. The stationary level of protein x_s can be computed by replacing (12) in (11), so to get:

$$x_s = ab\bar{c}(x_s). \tag{15}$$

Note that, as it can be deduced from the above expression, stationary protein levels are related to the burst frequency and size, so the number of protein units increases proportionally to the product ab.

Interestingly $r_d(x)$ and $r(x)$ share the same slope, but their values at zero are different. Since for every x, $(a-1)/a < 1$, this results into $r(x) < r_d(x)$. We will make use of (9) and (14) to obtain the reliability of the deterministic model versus the stochastic one.

3. CHARACTERIZING BISTABILITY AND BIMODAL OR BINARY RESPONSE

In order to characterize the regions in the a, b parameter space which sustain bimodality and the corresponding bistability for the deterministic counterpart, we proceed to study the roots of (8) and (13), respectively.

We first note that for both cases, the roots coincide with the number of intersections between (2) and the straight line (9) or (14). In addition, bimodality or bistability, require multiple intersections, what as depicted in Fig. 2, occurs whenever (9) - or equivalently (14)- lie within the band delimited by the straight lines $r_1(x)$ and $r_2(x)$. This fact leads to four possible scenarios to be considered:

1. A bistable but unimodal system, $(r(x) < r_2(x) < r_d(x) < r_1(x))$.
2. A bistable and bimodal or binary system, $(r_2(x) < r(x) < r_d(x) < r_1(x))$.
3. A monostable but bimodal or binary system, $(r_2(x) < r(x) < r_1(x) < r_d(x))$.
4. A monostable and unimodal system, $(r(x) < r_d(x) < r_2(x)$ or $r_1(x) < r(x) < r_d(x)\ )$.

In order to construct the regions in the parameter space that sustain bimodal or binary behaviour we make use of the algorithm proposed in (Pájaro et al., 2015).

Essentially, the algorithm estimates the boundaries of the regions which formally can be characterized with the following set inequalities:

$$\mathrm{h}(\bar{x}_2; b) < a < \mathrm{h}(\bar{x}_1; b) \quad \text{with} \quad \mathrm{h}(x; b) = \frac{b+x}{b\bar{c}(x)}. \tag{16}$$

These inequalities come from the necessary and sufficient condition for binary or bimodal distributions $r_2(x) < r(x) < r_1(x)$. Where $\bar{c}(x)$ in the above expression is given by (1) whereas $\bar{x}_1, \bar{x}_2$ are the values which determine the slope of $r(x)$, which in turn coincide with $r_1(x)$ and $r_2(x)$.

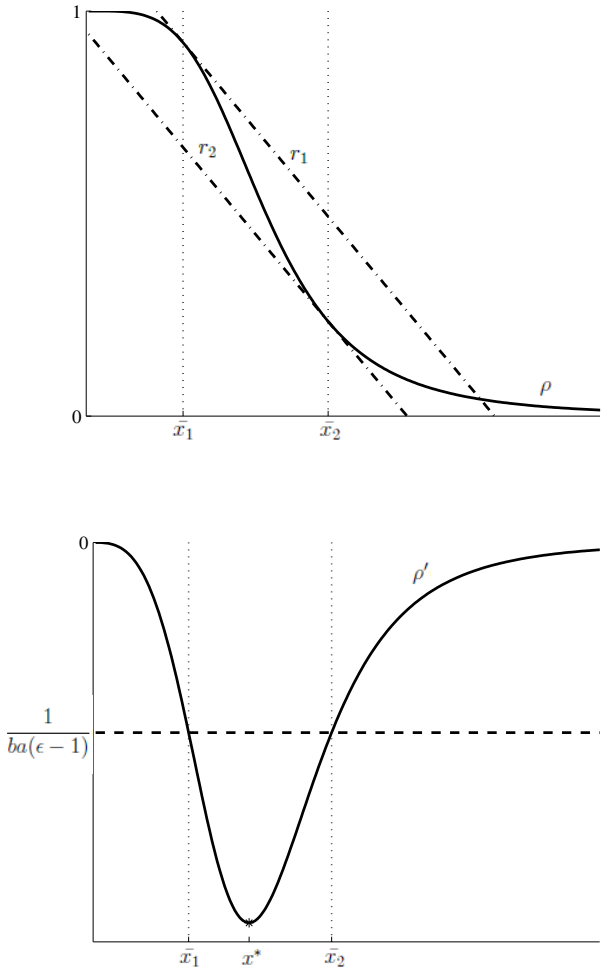

Fig. 2. The upper plot depicts function $\rho(x)$ as given in (2) for $H < -1$ and two possible functions $r_1(x)$ and $r_2(x)$ tangent to $\rho(x)$ at points ($\bar{x}_1$ and $\bar{x}_2$), respectively ($r_1(x) = \rho'(\bar{x}_1)(x - \bar{x}_1) + \rho(\bar{x}_1)$ and $r_2(x) = \rho'(\bar{x}_2)(x - \bar{x}_2) + \rho(\bar{x}_2)$). The lower plot represents the first derivative $\rho'(x)$ and a possible value of $r'(x)$ represented by the horizontal dashed line.

This is graphically described in the lower plot of Figure 2, which represents $\rho'(x)$ and the intersection with the slope of $r(x)$. The set of possible values $\bar{x}_1, \bar{x}_2$ associated to the whole range of derivatives $\rho'(x)$ from its minimum at x^* is computed by solving the following equation:

$$\rho'(x) - \sigma\rho'(x^*) = 0. \tag{17}$$

for the parameter σ within the interval (0, 1).

Two types of regions have been identified depending on the crossing of bounds $h(\overline{x}_2; b)$ and $h(\overline{x}_1; b)$ in (21). The distance between bounds can be computed as:

$$\triangle \mathrm{h}(\sigma, b) = \frac{\bar{c}(\bar{x}_2) - \bar{c}(\bar{x}_1)}{\bar{c}(\bar{x}_1)\bar{c}(\bar{x}_2)} \left(1 - \frac{b^*(\sigma)}{b}\right), \tag{18}$$

where $b^*(\sigma)$ is given by

$$\begin{aligned} b^*(\sigma) &= \frac{\bar{x}_2\bar{c}(\bar{x}_1) - \bar{x}_1\bar{c}(\bar{x}_2)}{\bar{c}(\bar{x}_2) - \bar{c}(\bar{x}_1)} \\ &\equiv \frac{\bar{x}_2 - \bar{x}_1 - (1-\varepsilon)(\rho(\bar{x}_1)\bar{x}_2 - \rho(\bar{x}_2)\bar{x}_1)}{(1-\varepsilon)(\rho(\bar{x}_1) - \rho(\bar{x}_2))} \end{aligned} \tag{19}$$

From the above expressions it follows that whenever $b^*(\sigma) \leq 0$ the bimodal region has a strip-like shape since then $\triangle \mathrm{h}(\sigma, b) > 0$. On the other hand, if $b^*(\sigma) > 0$ we have that $\triangle \mathrm{h}(\sigma, b) > 0$ only for $b > b^*(\sigma)$, what leads to a horn-like region. As it has been demonstrated in Pájaro et al. (2015), such conditions are related to a critical leakage ε^* that depends on the cooperativity level H as:

$$\varepsilon^* = \left(\frac{H+1}{H-1}\right)^2, \tag{20}$$

and determines whether the region is strip-like ($\varepsilon \leq \varepsilon^*$) or horn-like ($\varepsilon > \varepsilon^*$). One example of each type of region can be seen in Figure 3.

The algorithm we just have discussed can be employed to construct the regions in the parameter space associated to the deterministic approximation, which sustain bistability ($r_2(x) < r_d(x) < r_1(x)$). In this case, region boundaries become:

$$\mathrm{h}_d(\bar{x}_2; b) < a < \mathrm{h}_d(\bar{x}_1; b) \quad \text{with} \quad \mathrm{h}_d(x; b) = \frac{x}{b\bar{c}(x)}, \tag{21}$$

and the distance between bounds computed as:

$$\triangle \mathrm{h}_d(\sigma, b) = \frac{1}{\bar{c}(\bar{x}_1)\bar{c}(\bar{x}_2)} \left(-\frac{b_d^*(\sigma)}{b}\right), \tag{22}$$

with function $b_d^*(\sigma)$ in (22) being of the form:

$$\begin{aligned} b_d^*(\sigma) &= \bar{x}_2\bar{c}(\bar{x}_1) - \bar{x}_1\bar{c}(\bar{x}_2) \\ &\equiv \bar{x}_2 - \bar{x}_1 - (1-\varepsilon)\left(\rho(\bar{x}_1)\bar{x}_2 - \rho(\bar{x}_2)\bar{x}_1\right) \end{aligned} \tag{23}$$

For $b_d^*(\sigma) < 0$ we have that $\triangle \mathrm{h}_d(\sigma, b) > 0$ what leads to a strip-like region of bistability. As in the previous situation there also exists a critical leakage ε^* which makes $b_d^*(\sigma) = 0$. However, because of (22) no crossing of the boundaries is possible. Thus for any leakage above a critical value (i.e. $\varepsilon \geq \varepsilon^*$) no bistability is possible. In other words, bistability never coexists with bimodality above such values.

4. RESULTS AND DISCUSSION

The methods described in the previous section are employed to depict the regions in the a, b parameter space sustaining bimodal and bistable behavior. Such regions are presented in Figure 3 for two leakage rates around a critical value (20). Below ε^* (Figure 3 A), the regions that present bimodality (shaded area) and bistability (area delimited by dashed lines) show strip-like shapes that partially overlap. Above ε^*, only a horn-like region that corresponds with bimodal distributions, remains. Outside each of those regions, parameter combinations lead to graded distributions or to monostable systems, if modelled via the ODE system (10)-(11).

A simulation approach based for instance on SSA, could be used to identify the different qualitative behavior regions depicted in Figure 3. However, because it would require a representative number of simulations for each parameter

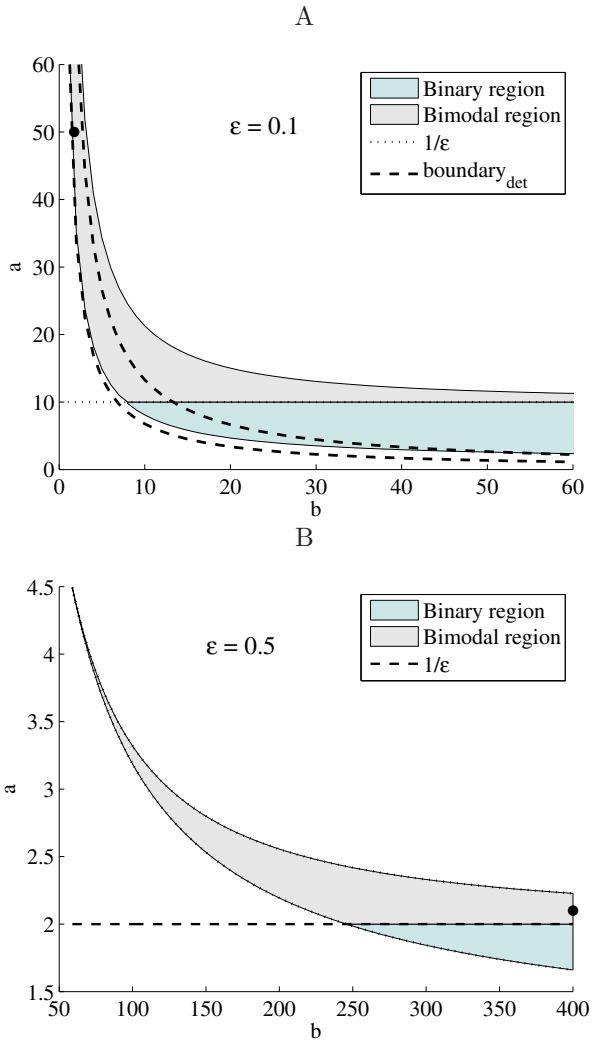

Fig. 3. Regions sustaining bistable, bimodal and binary behaviour in the parameter space (b, a). Plot A shows a strip-like region for $H = -4, K = 40$, and $\varepsilon = 0.1 < \varepsilon^* = 0.36$. The marked point ($b = 1.75$ and $a = 50$) represents a case where bistable behaviour coincides with a bimodal distribution. Plot B shows a horn-like region for $H = -4, K = 140$, and $\varepsilon = 0.5 > \varepsilon^* = 0.36$. The marked point ($b = 400$ and $a = 2.1$) leads to a bimodal distribution which has not a bistable counterpart.

set, the computational cost would be overwhelming even for simple systems involving small number of parameters.

As it can be seen in the figure, bistability occupies a smaller area than bimodal or binary response and appears at lower burst frequency and size (a and b, respectively). In that range, it almost perfectly overlaps with the region of binary-bimodal response. At the microscopic level, the system with parameters in that region, frequently transits between two protein levels what induces a periodic switching. Such response parallels the bistable behavior exhibited by the deterministic counterpart described by (10)-(11). As an example, Figure 4 (A) presents the resulting stationary distribution as well as the equilibrium states for the deterministic counterpart. Figure 4 (C) compares the stochastic response with the corresponding deterministic states. Note, that despite the low number

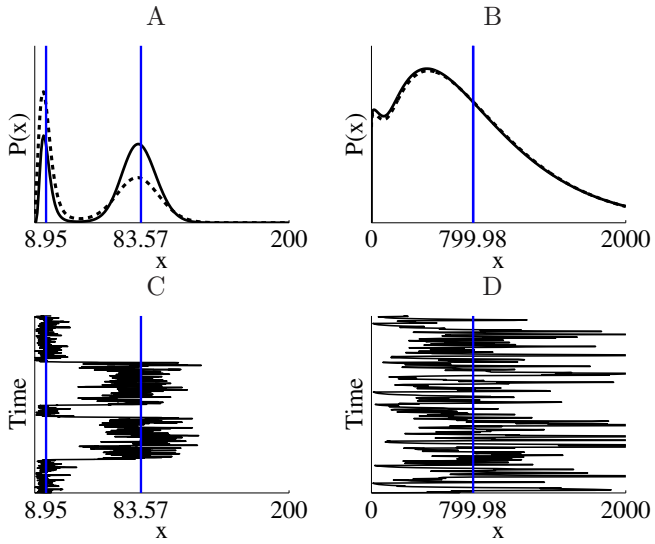

Fig. 4. Bimodal and bistable response for $H = -4$, $K = 40$, $\varepsilon = 0.1$, $b = 1.75$ and $a = 50$ (Plots A and C). Critical leakage $\varepsilon^* = 0.36$ for this example is computed from (20). Bimodal and monostable response for $H = -4$, $K = 140$, $\varepsilon = 0.5$, $b = 400$ and $a = 2.1$ (Plots B and D). A and B show the stationary distributions computed from the CME (dashed line) and the analytical distribution (continuous line). C and D depict the corresponding stochastic dynamics (SSA). Vertical lines represent the deterministic stable points. In Plots A and C we can clearly distinguish a periodic switching between two different protein levels. The scenario depicted in Plots B and C illustrates a very noisy signal, reflected by the long tail of the corresponding distribution. Here, the stable point of the deterministic system is far away of the stochastic behaviour (two peaks).

of molecules involved, the equilibrium points associated to the deterministic approximation are close to the maxima of the distributions, what proves that the deterministic ODE model is able to capture reasonably well (in the form of bistability) the intensity of the stochastic fluctuations.

Essentially, the applicability of the deterministic approximation is indicated by the distance between the mode and the mean of the probability distribution, the latter being related to the deterministic state. On a bimodal distribution the closer each hill will approach to a Gaussian, the closer the two most frequent states (modes) will approach the mean (average) of each hill. This turns out to be the trend as the number of proteins increases (i.e. as the system approaches the deterministic limit). One can show this by noting that according to (15), for any burst size b, stationary protein levels are function of the burst frequency a, so that in the limit:

$$\lim_{a\to\infty} r(0) = \frac{a-1}{a(1-\varepsilon)} = \frac{1}{1-\varepsilon} = r_d(0), \tag{24}$$

and equations (8) and (13) become identical.

Note however, that contrary to what it might be expected, increases in the number of proteins produced by the network does not necessary parallel improvements in the validity of the deterministic approximation. This scenario

is depicted in Figure 4 (plots B and D). Plot B, represents the stationary bimodal probability distribution associated the point marked in Figure 3 (B). Such point belongs to a horn-like region, characteristic of high transcription leakages. In fact, as demonstrated in Pájaro et al. (2015), such regions appear in the range of large burst sizes (and small burst frequencies), whenever leakage is over a critical value, that only depends on the level of cooperativity (see expression (20)).

Plot D in Figure 4 compares a realization of the stochastic dynamics with the corresponding deterministic stationary state. In this situation, the (monostable) state, turns out to be far from the two most frequently visited microscopic states associated to the bimodal distribution. Thus, the deterministic approximation does not properly capture the essential features of the self-regulatory gene network, (namely bimodality or even average state values) despite to have a system with a sufficiently large number of molecules (in the order of 10^3 molecules).

5. CONCLUSION AND FUTURE WORK

In this contribution we made use of a continuous version of the Chemical Master Equation to study the applicability of deterministic approximations to describe the dynamics of self-regulatory genetic circuits.

We have shown 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. Furthermore, and contrary to what it might be expected, there are circumstances for which the deterministic approximation does not properly capture the essential features of self-regulatory gene networks, despite systems with a sufficiently large number of molecules.

The main results of this parametric analysis, could be of help in the context of synthetic biology as a tool to guide the design of artificial engineered genetic circuits with specific properties.

We plan to extend the present methodology to handle more complex GRN, such as those including couple cascades of transcription-translation steps. Hopefully, the analysis and methods here discussed 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.

ACKNOWLEDGEMENTS

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.

REFERENCES

Alon, U. (2007). *An Introduction to Systems Biology. Design Principles of Biological Circuits.* Chapman & Hall/ CRC, London.

Dar, R.D., Razooky, B.S., Singh, A., Trimeloni, T.V., McCollum, J.M., Cox, C.D., Simpson, M.L., and Weinberger, L.S. (2012). Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc. Natl. Acad. Sci. U.S.A.*, 109(43), 17454–17459.

Elgart, V., Jia, T., Fenley, A.T., and Kulkarni, R. (2011). Connecting protein and mRNA burst distributions for stochastic models of gene expression. *Phys. Biol.*, 8, 046001.

Engblom, S. (2006). Computing the moments of high dimensional solutions of the master equation. *Appl. Math. Comput.*, 180(2), 498–515.

Friedman, N., Cai, L., and Xie, X.S. (2006). Linking stochastic dynamics to population distribution: An analytical framework of gene expression. *Phys. Rev. Lett.*, 97(16), 168302.

Gillespie, D.T. (2007). Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58, 35–55.

Gillespie, D.T. (2009). Deterministic limit of stochastic chemical kinetics. *J. Phys. Chem. B*, 113(6), 1640–1644.

Huang, L., Yuan, Z., Liu, P., and Zhou, T. (2015). Effects of promoter leakage on dynamics of gene expression. *BMC Syst. Biol.*, 9(1).

Jahnke, T. (2011). On reduced models for the chemical master equation. *Multiscale Model. Simul.*, 9(4), 1646–1676.

Kepler, T.B. and Elston, T.C. (2001). Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophys. J.*, 81(6), 3116–3136.

Mackey, M.C., Tyran-Kamiska, M., and Yvinec, R. (2011). Molecular distributions in gene regulatory dynamics. *J. Theor. Biol*, 274(1), 84–96.

Munsky, B. and Khammash, M. (2006). The finite state projection algorithm for the solution of the chemical master equation. *J. Chem. Phys.*, 124(4), 1–12.

Ochab-Marcinek, A. and Tabaka, M. (2015). Transcriptional leakage versus noise: A simple mechanism of conversion between binary and graded response in autoregulated genes. *Phys. Rev. E*, 91(1), 012704.

Pájaro, M., Alonso, A.A., and Vázquez, C. (2015). Shaping protein distributions in stochastic self-regulated gene expression networks. *Phys. Rev. E*, 92(3), 032712.

Paulsson, J. and Ehrenberg, M. (2000). Random signal fluctuations can reduce random fluctuations in regulated components of chemical regulatory networks. *Phys. Rev. Lett.*, 84(23), 5447–5450.

Rosenfeld, N., Elowitz, M.B., and Alon, U. (2002). Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.*, 323, 785–793.

Shahrezaei, V. and Swain, P.S. (2008). Analytical distributions for stochastic gene expressions. *Proc. Natl. Acad. Sci. U.S.A.*, 105(45), 17256–17261.

Sherman, M.S. and Cohen, B.A. (2014). A computational framework for analyzing stochasticity in gene expression. *PLoS Comput. Biol.*, 10(5), 1003596.

Shmulevich, I. and Aitchison, J.D. (2009). Deterministic and stochastic models of genetic regulatory networks. *Methods Enzymol.*, 467(C), 335–356.

Van Kampen, N.G. (2007). *Stochastic Processes in Physics and Chemistry.* Elsevier, Netherlands, third edition.

Wallace, E.W.J., Gillespie, D.T., Sanft, K.R., and Petzold, L.R. (2012). Linear noise approximation is valid over limited times for any chemical system that is sufficiently large. *IET Syst. Biol.*, 6(4), 102–115.