

Stochastic Models and Numerical Algorithms for a Class of Regulatory Gene Networks

Thomas Fournier^a, Jean-Pierre Gabriel^a, Jérôme Pasquier^a,
Christian Mazza^{a,*}, José Galbete^b, Nicolas Mermod^b

^aDepartment of Mathematics, University of Fribourg, Chemin du Musée 23, 1700 Fribourg, Switzerland

^bInstitute of Biotechnology, University of Lausanne, 1015 Lausanne, Switzerland

Received: 7 February 2008 / Accepted: 10 February 2009 / Published online: 22 April 2009
© Society for Mathematical Biology 2009

Abstract Regulatory gene networks contain generic modules, like those involving feedback loops, which are essential for the regulation of many biological functions (Guido et al. in *Nature* 439:856–860, 2006). We consider a class of self-regulated genes which are the building blocks of many regulatory gene networks, and study the steady-state distribution of the associated Gillespie algorithm by providing efficient numerical algorithms. We also study a regulatory gene network of interest in gene therapy, using mean-field models with time delays. Convergence of the related time-nonhomogeneous Markov chain is established for a class of linear catalytic networks with feedback loops.

Keywords Gillespie algorithm · Gene network · Self-promoter · Quasi-equilibrium · Dimerization · Mean field · Time delay

1. Introduction

Modeling of the regulation of all genes in a given cell is a tantalizing problem in biology and medicine (see, e.g. Guido et al., 2006). Recent developments allow rapid experimental determination of the expression of nearly all genes in a given biological setting, to an extent that in-depth analysis and proper mathematical understanding of these vast arrays of data have become limiting. Qualitative models of regulatory networks, where particular genes code for proteins that activate or repress other genes, are being assembled, but models taking the stochastic and quantitative nature of gene regulation remain scarce, and they often rely on assumptions or simplifications that rest untested experimentally. Thus, it would be useful to build validated mathematical models of particular regulatory modules, as a first step towards constructing models of genome-wide gene expression.

*Corresponding author.

E-mail addresses: christian.mazza@unifr.ch (Christian Mazza), nicolas.mermod@unil.ch (Nicolas Mermod).

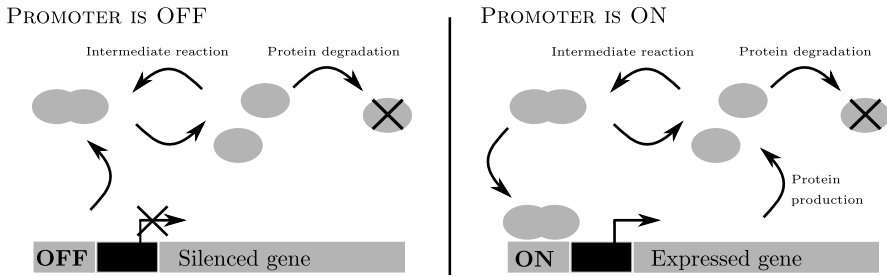


Fig. 1 The self-regulated gene.

Here, we consider a class of self-regulated genes, as depicted in Fig. 1. This auto-regulated module is a very common building block of many gene networks, as it may form the basis of stochastic gene switches that contribute to biological decisions such as cell differentiation, and has been studied extensively in the literature in some special settings, as in Hornos et al. (2005), Kepler and Elston (2001) or Peccoud and Ycart (1995). The belief that stochastic simulation is the only way to handle such problems is widely spread among researchers, and exact methods are only known in some very particular cases, see Peccoud and Ycart (1995) or Hornos et al. (2005). In a previous work (Fournier et al., 2007), we provided the exact steady-state distribution of the stochastic expression level of the autoregulated gene solving the chemical master equation in a general setting. We will present a direct version of the method and study more deeply this stationary distribution by providing efficient numerical algorithms. We will also consider a synthetic regulatory network acting as a genetic switch that was studied in living cells (Imhof et al., 2000).

1.1. A self-regulated gene

The system is composed of a promoter and a gene, as schematized in Fig. 1. As stated previously, one source of molecular noise is the random nature of the states taken by the promoter (on/off). Figure 1 shows protein monomers produced by the RNA polymerase during the transcription and translation processes. Protein monomers react quickly to form dimers: we assume a quasi-equilibrium where fast reactions equilibrate instantaneously. For a global amount of n proteins, the proportion of dimers at quasi-equilibrium is a well defined function of n . Dimers can bind to some sites near the promoter, and therefore enhance transcription, corresponding to a positive feedback loop. These binding events can be assumed to be fast with respect to events like protein formation. They are however included in some chain of events which ends with a state where the right positioning of the RNA polymerase is possible. This will correspond to the on state \mathcal{O}_1 . When these conditions are not satisfied, the promoter is off \mathcal{O}_0 . The rates of transitions between these two states are functions of the proportion of dimers, and therefore of n when the cell contains n proteins. These random events are usually modeled by supposing that the probability that the promoter switches from the off to on state in a small time interval of length $h \approx 0$ is of order $g(n)h$ for n proteins, where the function g can be chosen according to the specificity of the setting. To be as general as possible, and to eventually allow negative feedback loops, we also assume that the probability of transition of the reverse reaction is given by some function $\kappa(n)$. Basal activity is introduced by supposing that $g(0)$ is

positive, so that the required conditions for an eventual transcription event can be realized without protein dimers. The remaining involved chemical reactions are essentially protein monomers production and degradation, which are summarized in Fig. 1. Transcription is stopped when the promoter is off, so that we assume that the probability $\mu_0 h$ that a protein is created during a small time interval of length h vanishes, with $\mu_0 = 0$. When the promoter is on, transcription is possible, and the probability that a transcription event occurs is of order $\mu_1 h$. Degradation of protein dimers is summarized by the rate $\nu(n)$, for some function ν , which is usually linear as a function of n . The time evolution of the state of this self-regulated gene is described by a pair of time-continuous stochastic process $N(t)$ and $Y(t)$, where $N(t)$ gives the number of proteins present in the cell at time t and where $Y(t)$ takes the values 0 and 1 corresponding to the off and on states of the promoter. The usual way of simulating $N(t)$ and $Y(t)$ proceeds by running the Gillespie algorithm (see e.g. Cao et al., 2005 and Gillespie, 2001). The mean steady-state expression level is thus obtained through Monte Carlo experiments.

1.2. A regulatory network for efficient control of transgene expression

A more elaborate gene network consists of three genes. A first gene encodes a transcriptional repressor. Because this gene is expressed from an unregulated promoter, it mediates a stable number of repressor. This repressor binds to and inhibits the promoters of the two other genes, coding for a transactivator protein and for a quantifiable or a therapeutic protein, respectively (Fig. 2A). The activity of the repressor is inhibited by doxycycline, a small antibiotic molecule that acts as a ligand of the repressor and thereby controls its activity. Addition of the antibiotic will inhibit the repressor and relieve repression, allowing low levels of expression of the regulated genes and synthesis of some transactivator protein. This, in turn, allows further activation of the two regulated genes, in a positive feedback loop (Fig. 2B). When introduced in mammalian cells, this behaves as a signal amplifier and as a potent genetic switch, where the expression of a therapeutic gene can be controlled to vary from almost undetectable to very high levels in response to the addition of the antibiotic to the cells (Fournier et al., 2007; Imhof et al., 2000).

1.3. Results

Section 2 considers the Gillespie algorithm for simulating the time evolution of the number of proteins $N(t)$ and of the state of the related promoter $Y(t)$, by focusing on the associated steady-state distribution π . In a previous work (Fournier et al., 2007), we gave an explicit formula for the steady state associated with self-regulated genes, based on the embedded jump chain of the time-continuous Markov process. Here we introduce a direct version dealing with the time-continuous process. For concrete computation, we have to use a bounded state space with a total number of proteins that cannot exceed a fixed but arbitrary integer Λ . Algorithm 1 gives an efficient way of computing π , while Theorem 2 gives the convergence of the sequence of steady-state distributions measures indexed by Λ to the unique invariant distribution of the process defined on the unbounded state space $\mathbb{N} \times \{0, 1\}$. We also provide information on the variance of the gene product at steady state using generating functions and differential equations.

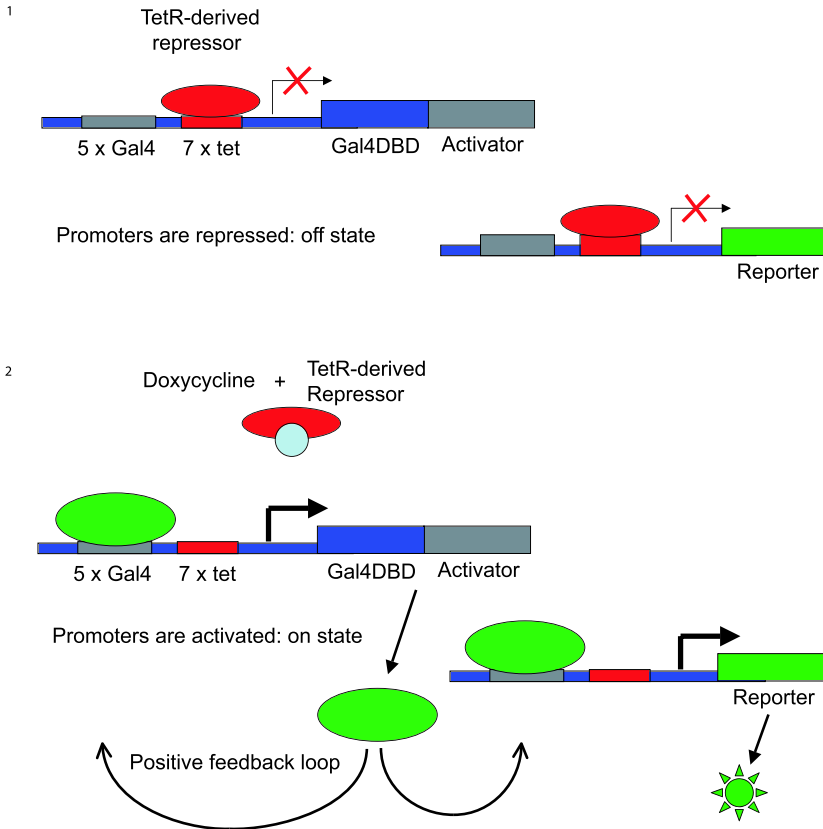


Fig. 2 The regulatory network.

Section 3 proposes a mean-field model with time delays, generalizing a model considered recently in this setting by Goutsias and Kim (2006), by including stochastic signals related to promoters. The feedback rates $g(N(t))$ and $\kappa(N(t))$ are replaced by $g(\mathbb{E}(N(t - \theta)))$ and $\kappa(\mathbb{E}(N(t - \theta)))$. In Fournier et al. (2007), we studied the regulatory gene network in living cells; the obtained experimental results were in good concordance with the model's predictions. The related functions $E(t) = \mathbb{E}(N(t))$ and $G(t) = P(Y(t) = 1)$ satisfy the time delayed differential system

$$\frac{dE}{dt} = \mu G(t) - \nu E(t),$$

$$\frac{dG}{dt} = g(E(t - \theta))(1 - G(t)) - \kappa(E(t - \theta))G(t),$$

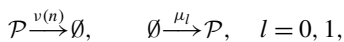
which can be deduced from the chemical master equation, see Section 5. As it is well known, this kind of differential systems can possess oscillating or periodic solutions, see e.g. Bratsun et al. (2005). We show that there is a globally asymptotically stable equilibrium point when g is such that $g(n)/n$ is decreasing as function of n and $\kappa(n) \equiv \kappa$. We

next provide Algorithm 2 for computing the steady-state variance. Section 4 deals with two time scales stochastic simulations and processes evolving at quasi-equilibrium. We also consider a generic dimerization process which occurs in most biochemical reaction networks, and provide an efficient Algorithm 3 for computing the first two moments of the related steady-state distribution, which are then used when dealing with systems evolving at quasi-equilibrium. Finally, Section 5 focus on the regulatory network; we model extrinsic and intrinsic noise using a mean-field model, which permits to study the fluctuations of the variance of the number of therapeutic proteins as function of the number of doxycycline molecules.

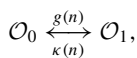
2. Mathematical models related to the self-regulated gene

In what follows, we consider the time-continuous Markov chain known as *the Gillespie algorithm* for simulating the self-regulated gene with arbitrary feedback mechanisms, and give precise formulas for the related steady-state distribution π . This defines a generic class of gene networks, which are the building block of most of the existing regulatory networks. We will then use our method to study a family of mean-field models in Sections 3 and 5.4. The related steady-state distribution is obtained as the product of steady-state distributions of sub modules corresponding to self-regulated genes. A complete understanding of basic modules like the self-regulated gene is thus fundamental for understanding the global network, see e.g. Guido et al. (2006) and Section 5.4. For more details on the model, see Fournier et al. (2007).

The module is composed of a promoter and a gene. Its time evolution is given by the following set of chemical reactions:



represents degradation of gene product when they are n molecules, here proteins (\mathcal{P}), and protein production, where $l = 0$ means that the promoter is off: no transcription factor (a complex composed of gene product) is bound to some operator sites near the promoter, so that the RNA polymerase cannot bind well in the neighborhood of the basal promoter. We assume here that the transcription rate μ_0 is such that $\mu_0 \approx 0$. When $l = 1$, meaning that the promoter is on, transcription occurs at a rate $\mu_1 = \mu$. The fluctuations of the state of the promoter are described by the following reaction



where \mathcal{O}_l , $l \in \{0, 1\}$, indicates the state of the promoter. The transitions from the on to off states $\mathcal{O}_1 \xrightarrow{\kappa(n)} \mathcal{O}_0$ occur at rate $\kappa(n)$, and the reverse reactions at rate $g(n)$. Here g and κ are two functions of the number of proteins modeling positive and negative feedback loops.

Basal activity is introduced at the level of the reaction $\mathcal{O}_0 \xrightarrow{g(n)} \mathcal{O}_1$, by supposing that $g > 0$. The Gillespie algorithm for simulating the above chemical reactions is a bivariate Markov process $\eta(t) = (N(t), Y(t))$, with $N(t) \leq \Lambda$ and $Y(t) = 0, 1$, where $N(t)$ denotes the number of proteins present in the cell at time t , and $Y(t)$ represents the state of the

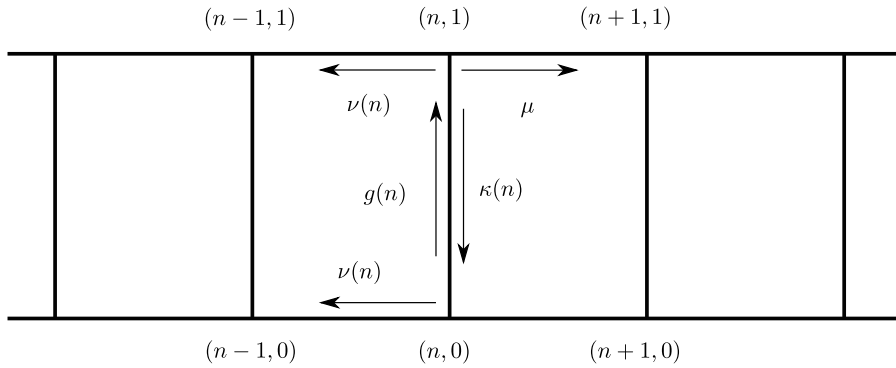


Fig. 3 Visualization of the state space as a strip. The possible transitions are represented by the arrows with corresponding rates.

operator. The time evolution of $Y(t)$ is coupled to that of $N(t)$ to model autoregulation, using functions g and κ . For small time interval $(t, t + h)$, $h \approx 0$, the probability that the operator switches from the off to the on state is of order $g(N(t))h$.

Let $p_n^0(t) = P(N(t) = n, Y(t) = 0)$ and $p_n^1(t) = P(N(t) = n, Y(t) = 1)$ give the probability of having n proteins at time t when the states of the promoter are \mathcal{O}_0 and \mathcal{O}_1 , respectively. We assume here that $0 \leq n \leq \Lambda$ for some fixed but arbitrary integer Λ . The related Gillespie algorithm is given as a time-continuous Markov chain $\eta(t) = (N(t), Y(t))$ (see e.g. Gillespie (1977)), where $N(t) \in \{0, 1, \dots, \Lambda\}$ and $Y(t) \in \{0, 1\}$, with transition rates given by

$$\begin{aligned}
 P((n, y), (n + 1, y)) &= \mu_y, & P((n, y), (n - 1, y)) &= \nu(n), \\
 P((n, y), (n, 1 - y)) &= \kappa(n) \quad \text{when } y = 1, \quad \text{and,} \\
 P((n, y), (n, 1 - y)) &= g(n) \quad \text{when } y = 0.
 \end{aligned}$$

Figure 3 provides a visualization of the state space and the possible transitions with corresponding rates. The *chemical master equation* associated with the reaction scheme is then given by

$$\begin{aligned}
 \frac{dp_n^s(t)}{dt} &= \mu_s (p_{n-1}^s(t) - p_n^s(t)) + \nu(n + 1)p_{n+1}^s(t) - \nu(n)p_n^s(t) \\
 &\quad + (-1)^s (\kappa(n)p_n^1(t) - g(n)p_n^0(t)),
 \end{aligned} \tag{1}$$

where $s \in \{0, 1\}$, see e.g. Kepler and Elston (2001).

The steady-state distribution π associated with (1) is obtained by letting $t \rightarrow \infty$: π is defined as

$$\pi_n(0) = \lim_{t \rightarrow \infty} p_n^0(t) \quad \text{and} \quad \pi_n(1) = \lim_{t \rightarrow \infty} p_n^1(t),$$

and solves the linear system obtained from (1) by imposing $dp_n^s/dt = 0$:

$$\begin{aligned}
 0 &= \mu_s (\pi_{n-1}(s) - \pi_n(s)) + \nu(n + 1)\pi_{n+1}(s) - \nu(n)\pi_n(s) \\
 &\quad + (-1)^s (\kappa(n)\pi_n(1) - g(n)\pi_n(0)), \quad s = 0, 1.
 \end{aligned}$$

$\pi_n(0)$ is the probability to find n proteins and that the promoter is off; $\pi_n(1)$ is defined similarly but for the on state. The probability of observing n proteins at equilibrium is just $\pi_n(0) + \pi_n(1)$.

In what follows, we derive the steady-state distribution π . This probability measure is used in the mean-field delayed model of Section 3 and in Section 5 for the study of the network.

2.1. Computing the steady state

In this section, we assume that $\mu_0 = 0$, and that $\mu_1 = \mu$, where an explicit formula for the steady state is available. The degradation propensity function $v(n)$ and the feedbacks $\kappa(n)$ and $g(n)$ are arbitrary positive functions.

The method of generating functions can be used in some particular special cases to compute the invariant measure, see Peccoud and Ycart (1995) for the simple case without feedback with $v(n) = v \cdot n$, $\kappa(n) \equiv \kappa$ and $g(n) \equiv g$, or Hornos et al. (2005) for the case with linear negative feedback $v(n) = v \cdot n$, $\kappa(n) \equiv \kappa \cdot n$ and $g(n) \equiv g$, or the recent work of Visco et al. (2008a, 2008b) where interesting phase diagrams are provided. Although this method provides a powerful tool for analytic description, the method of generating functions is very particular in the sense that a little change in the form of one of the feedback propensity function can induce major changes in the generating function, and for each particular propensity function one has to derive the whole set of equations anew. Furthermore, an explicit form for the generating function can only be found when the feedback propensity functions are simple, either constant or linear in the protein numbers. In practice, the feedback propensity functions are related to the number of sites in the promoter on which the proteins bind, either directly in monomer form or in more complicated bound forms like dimers or higher order polymers, see Dill and Bromberg (2003).

In Fournier et al. (2007), we presented a general formula for the steady state for arbitrary degradation and feedback propensity functions. The method relies on the asymptotic behavior of the jump matrix of the embedded discrete jump chain. We provide here a direct version of this method allowing to compute the invariant distribution of the time-continuous Markov process directly. We recall that we consider a bounded total number of protein $N(t) \leq \Lambda$, a restriction that is biologically meaningful due to the finite volume of a cell but that is mainly supposed for technical reasons since the formula is recursive and we have to find a starting point $(\pi_\Lambda(0), \pi_\Lambda(1))$ to begin with. However, the condition is not restrictive and we show in Theorem 2 that the sequence of steady-state distributions indexed by the boundary Λ converges weakly to the unique invariant distribution of the unbounded process on $\mathbb{N} \times \{0, 1\}$.

Consider the transfer matrices

$$\alpha_n = \frac{v(n+1)}{\mu} \begin{bmatrix} \frac{\kappa(n)+\mu}{g(n)+v(n)} & 1 \\ \frac{\kappa(n)}{g(n)+v(n)} & 1 \end{bmatrix}, \quad 0 < n < \Lambda,$$

$$\alpha_0 = \frac{v(1)}{\mu} \begin{bmatrix} \frac{\kappa(0)+\mu}{g(0)} & 1 \\ \frac{\kappa(0)}{g(0)} & 1 \end{bmatrix},$$

and the vector $w_\Lambda = (\kappa(\Lambda), g(\Lambda) + v(\Lambda))$.

Theorem 1. For $0 \leq n \leq \Lambda$, the invariant distribution

$$\pi_n = (\pi_n(0), \pi_n(1)) = \lim_{t \rightarrow \infty} (P((N(t), Y(t)) = (n, 0)), P((N(t), Y(t)) = (n, 1)))$$

of the time-continuous Markov process $\{\eta(t)\}_{t>0}$ on the strip $\{0, 1, \dots, \Lambda\} \times \{0, 1\}$ is given by

$$\pi_n = \frac{w_\Lambda \alpha_{\Lambda-1} \alpha_{\Lambda-2} \cdots \alpha_n}{Z_\Lambda} \quad \text{for } 0 \leq n < \Lambda, \quad \text{and} \quad \pi_\Lambda = \frac{w_\Lambda}{Z_\Lambda},$$

with the normalization constant

$$Z_\Lambda = w_\Lambda \cdot (1, 1)^T + \sum_{j=0}^{\Lambda-1} w_\Lambda \alpha_{\Lambda-1} \alpha_{\Lambda-2} \cdots \alpha_j \cdot (1, 1)^T.$$

Proof: At equilibrium, (1) reads

$$\begin{aligned} 0 &= \pi_0 R_0 + \pi_1 D_1, \\ 0 &= \pi_{n-1} U + \pi_n R_n + \pi_{n+1} D_{n+1}, \quad 0 < n < \Lambda, \\ 0 &= \pi_{\Lambda-1} U + \pi_\Lambda R_\Lambda, \end{aligned}$$

with the 2×2 matrices $U = \begin{bmatrix} 0 & 0 \\ 0 & \mu \end{bmatrix}$, $D_n = \begin{bmatrix} v(n) & 0 \\ 0 & v(n) \end{bmatrix}$ and

$$R_n = \begin{bmatrix} -(g(n) + v(n)) & g(n) \\ \kappa(n) & -(\kappa(n) + v(n) + \mu) \end{bmatrix}$$

for $1 \leq n \leq \Lambda - 1$, and the boundaries $R_0 = \begin{bmatrix} -g(0) & g(0) \\ \kappa(0) & -(\kappa(0) + \mu) \end{bmatrix}$,

$$D_\Lambda = \begin{bmatrix} v(\Lambda) & 0 \\ 0 & v(\Lambda) \end{bmatrix} \quad \text{and} \quad R_\Lambda = \begin{bmatrix} -(g(\Lambda) + v(\Lambda)) & g(\Lambda) \\ \kappa(\Lambda) & -(\kappa(\Lambda) + v(\Lambda)) \end{bmatrix}.$$

Simple linear algebra shows that the above defined matrices α_n , $0 \leq n < \Lambda$, satisfy the relation $\pi_n = \pi_{n+1} \alpha_n$. Indeed one only has to check that for $0 < n < \Lambda$, the 2×2 matrices $\alpha_{n-1} U + R_n$ are invertible and that the matrices α_n solve the matrix continuous fraction

$$\begin{aligned} \alpha_0 &= -D_1 R_0^{-1}, \\ \alpha_n &= -D_{n+1} (\alpha_{n-1} U + R_n)^{-1}, \quad 0 < n < \Lambda. \end{aligned} \quad \square$$

The formula given in Theorem 1 must be used with care numerically since, when Λ is large, both the numerator and denominator rapidly diverge. It can be improved with the following normalization algorithm, that is exactly the same as the one used for the embedded jump chain in Fournier et al. (2007). Let $\mathbb{R}_{\geq 0}^2 := \{w = (w_1, w_2), w_1, w_2 \in \mathbb{R}_{\geq 0}\}$ with the 1-norm $\|w\| := w \cdot (1, 1)^T$.

Algorithm 1.

(STEP 1): Define \tilde{v}_n for $n = \Lambda - 1$ to 0 as

$$\tilde{v}_\Lambda := \frac{w_\Lambda}{\|w_\Lambda\|}, \quad \text{and} \quad \tilde{v}_n := \frac{\tilde{v}_{n+1}\alpha_n}{\|\tilde{v}_{n+1}\alpha_n\|}.$$

(STEP 2): Given the \tilde{v}_n , define $v_0 = \tilde{v}_0$ and, for $n = 1$ to Λ , set

$$v_n := \frac{\tilde{v}_n}{\|\tilde{v}_n\alpha_{n-1}\| \cdot \|\tilde{v}_{n-1}\alpha_{n-2}\| \cdots \|\tilde{v}_1\alpha_0\|}.$$

(STEP 3): Compute the steady-state distribution as

$$\pi_n = \frac{v_n}{V_\Lambda}, \quad \text{where } V_\Lambda := \sum_{i=0}^\Lambda v_i \cdot \mathbf{1}.$$

It immediately results from their definition that the \tilde{v}_n and v_n satisfy $\|\tilde{v}_n\| = 1$,

$$v_n = \frac{\tilde{v}_\Lambda \alpha_{\Lambda-1} \alpha_{\Lambda-2} \cdots \alpha_n}{\|\tilde{v}_\Lambda \alpha_{\Lambda-1}\| \cdot \|\tilde{v}_{\Lambda-1} \alpha_{\Lambda-2}\| \cdots \|\tilde{v}_{n+1} \alpha_n\| \cdot \|\tilde{v}_n \alpha_{n-1}\| \cdots \|\tilde{v}_1 \alpha_0\|},$$

the denominator of the above expression is independent of n and \tilde{v}_Λ is proportional to w_Λ . Hence v_n is proportional to the invariant measure π_n , and (STEP 3) of the algorithm effectively compute the actual steady-state distribution.

Proposition 1 below provides conditions under which the normalization constant V_Λ remains bounded as Λ is large. The function $v(n)$ gives the monomer degradation rates for n proteins, and is assumed to be increasing with $v(0) = 0$, and strictly positive for $n \geq 1$. Usually, $v(n)$ is taken to be a constant times n , here we assume the less restrictive condition that $\inf_{n \geq 1} v(n)/n$ is strictly positive to allow situations where for example proteins that are present as chemical complexes (dimer, trimer, ...) cannot be degraded, or situations where $v(n)/n \rightarrow \infty$ as $n \rightarrow \infty$.

Lemma 1. *If $\inf_{n \geq 1} v(n)/n$ is strictly positive, there exists a constant $k > 0$ depending only on μ, ν, κ, g (and not on Λ) such that for all $n \geq 1, \|\tilde{v}_n \alpha_{n-1}\| \geq nk$.*

Proof: Each \tilde{v}_j lies in the line segment $S \subset \mathbb{R}^2$ between the points $(0, 1)$ and $(1, 0)$ and depends on Λ . To break this dependence, we prove the results for an arbitrary vector $v = (t, 1 - t) \in S, t \in [0, 1]$.

$$v\alpha_n = \frac{v(n)}{\mu}(\star, t + (1 - t)) = \frac{v(n)}{\mu}(\star, 1),$$

with $\star > 0$ for all n . Hence, uniformly in S ,

$$\frac{\|v\alpha_n\|}{n} \geq \frac{v(n)}{n\mu} \geq \inf_{n \geq 1} \frac{v(n)}{n\mu} =: k > 0. \quad \square$$

With Lemma 1 we can give bounds uniformly in Λ :

Proposition 1. *Assume that $\inf_{n \geq 1} v(n)/n$ is strictly positive. There exists $M > 0$, depending only on μ, v, κ, g (and not on Λ), such that*

$$1 \leq V_\Lambda = \sum_{n=0}^\Lambda \|v_n\| \leq M.$$

Proof: Notice that $\|v_0\| = \|\tilde{v}_0\| = 1$. The 1-norm of v_n is

$$\|v_n\| = \frac{\|\tilde{v}_n\|}{\|\tilde{v}_n \alpha_{n-1}\| \cdots \|\tilde{v}_1 \alpha_0\|} = \frac{1}{\|\tilde{v}_n \alpha_{n-1}\| \cdots \|\tilde{v}_1 \alpha_0\|}$$

and with Lemma 1

$$V_\Lambda = 1 + \sum_{n=1}^\Lambda \|v_n\| \leq 1 + \sum_{n=1}^\infty \frac{k^{-n}}{n!} = e^{1/k} =: M.$$

Lemma 1 shows that the preceding algorithm is efficient. But this lemma can also be used to demonstrate that the steady-state distribution of the continuous time process $\eta(t)$ converges when $\Lambda \rightarrow \infty$. We can also show that this limiting distribution is the invariant distribution of the process on the unbounded strip. We will write $\pi_n^{(\Lambda)}$ and $v_n^{(\Lambda)}$ instead of π_n and v_n . The previous results deal with a finite state space, where the number of proteins is bounded by Λ . Even if this model is realistic (an organism cannot contain an infinite number of proteins), it is relevant to check what happens when Λ is large. In other words, we want to show that under a sufficient condition, the invariant measure converges in Λ . We thus define $\pi^{(\Lambda)} = (\pi_n^{(\Lambda)})_{0 \leq n \leq \Lambda}$, and consider the family of probability measures $\Pi = (\pi^{(\Lambda)})_{\Lambda \in \mathbb{N}}$ embedded in $\mathbb{N} \times \{0, 1\}$. The proof of the following Theorem is given in [Appendix A](#). □

Theorem 2. *The sequence of invariant measures $\pi^{(\Lambda)}$ converges weakly as $\Lambda \rightarrow \infty$ to the invariant distribution of the process defined on the unbounded strip.*

2.2. The method of generating functions for the mean and variance

We consider the problem of computing the mean and variance of the gene product $N(t)$ at steady state, that is when t is large, using generating functions. As discussed in Section 2.1, generating functions allows in some cases to compute the steady-state distribution, see e.g. Hornos et al. (2005) or Peccoud and Ycart (1995), with simple feedback functions. See also the more recent works of Visco et al. (2008a, 2008b) where phase diagrams are obtained using generating functions for linear feedback. Here we show that even when the feedbacks are arbitrary, the method can be used to gain insight in the relations between variance, mean and probability to be ON. To avoid boundary conditions, we suppose here that the number of protein is arbitrary ($\Lambda = \infty$), and the only assumption concerning the propensity functions is that $v(n) = v n$, i.e. degradation is directly proportional to the number of proteins, while $\kappa(n)$ and $g(n)$ are arbitrary positive functions and μ_0 is not necessarily 0.

Let $\pi_n(y) = \lim_{t \rightarrow \infty} P(N(t) = n, Y(t) = y)$ and

$$\alpha(z) = \sum_{n \geq 0} \pi_n(1)z^n, \quad \beta(z) = \sum_{n \geq 0} \pi_n(0)z^n,$$

be the partial generating functions related to the steady state, and

$$R(z) := \sum_{n \geq 0} (\pi_n(1)\kappa(n) - \pi_n(0)g(n))z^n.$$

From the master equation (1) at equilibrium, we deduce

$$0 = -R(z) + \mu_1(z-1)\alpha(z) + \nu(1-z)\frac{d\alpha(z)}{dz}, \quad \text{and}$$

$$0 = +R(z) + \mu_0(z-1)\beta(z) + \nu(1-z)\frac{d\beta(z)}{dz}.$$

Adding these two relations and assuming that $z \neq 1$ gives

$$\frac{d\alpha(z)}{dz} + \frac{d\beta(z)}{dz} = \frac{\mu_1\alpha(z) + \mu_0\beta(z)}{\nu}.$$

This shows that the following general relations hold:

$$\mathbb{E}(N(\infty)) = \frac{\mu_1}{\nu} P(Y(\infty) = 1) + \frac{\mu_0}{\nu} P(Y(\infty) = 0),$$

$$\text{Var}(N(\infty)) = \frac{\mu_1}{\nu} \frac{d\alpha(z)}{dz} \Big|_{z=1} + \frac{\mu_0}{\nu} \frac{d\beta(z)}{dz} \Big|_{z=1} + \mathbb{E}(N(\infty))(1 - \mathbb{E}(N(\infty))),$$

where we recall that

$$\alpha(1) = P(Y(\infty) = 1),$$

$$\frac{d\alpha(z)}{dz} \Big|_{z=1} = \mathbb{E}(N(\infty)Y(\infty)) \quad \text{and}$$

$$\frac{d\beta(z)}{dz} \Big|_{z=1} = \mathbb{E}(N(\infty)(1 - Y(\infty))).$$

These formulas make sense since, when the promoter is on (resp. off), the process evolves as a birth and death process with birth rate μ_1 (resp. μ_0) and death rate νn , and has a Poisson distribution of parameter μ_1/ν (resp. μ_0/ν) as a stationary distribution, of mean and the variance given by μ_1/ν (resp. μ_0/ν). The last term is related to promoter fluctuations, see the following example.

Example 1. Assume that $g(n) \equiv g$ and $\kappa(n) \equiv \kappa$. Let $F(t) = \sum_{n \geq 0} nP(N(t) = n, Y(t) = 1)$. The master equation yields

$$\frac{dF(t)}{dt} = gE(t) + \mu_1G(t) - F(t)(g + \kappa + \nu),$$

$$\frac{dE(t)}{dt} = \mu_1 G(t) + \mu_0(1 - G(t)) - \nu E(t),$$

where $E(t) = \mathbb{E}(N(t))$. Using the differential relation $dG(t)/dt = g(1 - G(t)) - \kappa G$, one gets that

$$G(\infty) = \frac{g}{g + \kappa}.$$

Notice that $G(\infty)$ is related to the stationary law of the 0/1 Markov chain given by the transition rates $p(0, 1) = g$ and $p(1, 0) = \kappa$, which models the fluctuations of the state of the promoter. Then

$$\begin{aligned} \text{Var}(N(\infty)) &= \left. \frac{d^2\alpha(z)}{d^2z} \right|_{z=1} + \left. \frac{d^2\beta(z)}{d^2z} \right|_{z=1} + E(\infty) - E(\infty)^2 \\ &= \frac{\mu_1}{\nu} F(\infty) + \frac{\mu_0}{\nu} H(\infty) + E(\infty) - E(\infty)^2, \end{aligned}$$

where we set $H(t) = E(t) - F(t)$. It follows that

$$\text{Var}(N(\infty)) = \frac{\mu_1}{\nu} G(\infty) + \frac{\mu_0}{\nu} (1 - G(\infty)) + \frac{\mu_1}{\nu} F(\infty) + \frac{\mu_0}{\nu} H(\infty) - E(\infty)^2.$$

We finally obtain, after some algebra,

$$\text{Var}(N(\infty)) = \frac{\mu_1}{\nu} G(\infty) + \frac{\mu_0}{\nu} (1 - G(\infty)) + \frac{\tau_2}{\tau_1 + \tau_2} \frac{(\mu_1 - \mu_0)^2}{\nu^2} \text{Var}(Y(\infty)),$$

where the characteristic times τ_1 and τ_2 are defined by

$$\tau_1 = \frac{1}{\nu} \quad \text{and} \quad \tau_2 = \frac{1}{g + \kappa}.$$

The interpretation of this formula is obtained by observing that, when the promoter is on with probability $G(\infty)$, the process evolves as a birth and death process with steady-state distribution given by a Poisson distribution of parameter μ_1/ν . The interpretation of the second term is similar. The third term corresponds to the variance of a Bernoulli random variable (on/off) multiplied by a factor accounting for characteristic times related to protein degradation and promoter fluctuation. When $\mu_0 = 0$, the coefficient of variation can be then given as

$$CV_N^2 = \frac{\text{Var}(N(\infty))}{\mathbb{E}(N(\infty))^2} = \frac{1}{\mathbb{E}(N(\infty))} + \frac{\tau_2}{\tau_1 + \tau_2} \frac{\text{Var}(Y(\infty))}{\mathbb{E}(Y(\infty))^2},$$

as given in Paulsson (2005). The above relations yield moreover that

$$CV_N^2 = \frac{g + \kappa}{\rho g} + \frac{\nu \kappa}{g(g + \nu + \kappa)},$$

where $\rho = \mu_1/\nu$, and it follows that CV_N^2 is decreasing as a function of g and increasing as a function of κ .

3. Mean-field models

Most mathematical works on gene networks, like Gadgil et al. (2005), consider networks with linear transition rates in which the state space of each chemical species equals \mathbb{N} . Results on networks involving catalytic transitions rates are very scarce. Lipan and Wong (2005) also focus on such models but allow time dependent transition rates. In this situation, one gets interesting linear differential equations for the first and second moments, and for covariance functions. When some state space is finite, boundary effects transform the equations which become more involved.

The model for the self-regulated gene defined in Section 2 is similar to a mathematical model for an epidemic of schistosomiasis provided by Nasell and Hirsch (1972, 1973). In their model, the authors consider a similar Markov chain, where they replace every external random variables in the transitions probabilities by functions of their expected values. This means for example that the transition rate $P((n, y), (n + 1, y)) = y\mu$ is replaced by $P((n, y), (n + 1, y)) = \mathbb{E}(Y(t))\mu$, and $P((n, y), (n, 1 - y))$ by $\kappa(\mathbb{E}(N(t)))y + g(\mathbb{E}(N(t)))(1 - y)$, since for this last transition, the external random variable corresponding to this transition is $N(t)$. One gets a time-nonhomogeneous Markov chain. One can show that the pair $(\mathbb{E}(N(t)), \mathbb{E}(Y(t)))$ converges to a limit $(\mathbb{E}(N(\infty)), \mathbb{E}(Y(\infty)))$ (see e.g. Nasell and Hirsch, 1972). This model is then asymptotically equivalent to the model of the self-regulated gene given in Example 1 where $\mu_1 = \mu_0 = \mathbb{E}(Y(\infty))\mu$: it is easy to check that the stationary distribution of $N(\infty)$ is Poisson of parameter μ_1/ν when $\nu(n) \equiv \nu n$. The mean and the variance are then equal to μ_1/ν . The behavior of the propagation of noise in gene networks can be counter-intuitive, as shown for example by Pedraza and van Oudenaarden (2005), where the mean gene expression at steady state is increasing as function of some inducer, but where the variance exhibits a peak. The same phenomenon occurs with the therapeutic network of Section 5. This shows that this model cannot predict this qualitative behavior. We shall adopt a different point of view below by conserving the external variable $Y(t)$ and taking only the average of $N(t)$.

Models in which one considers the average of $N(t)$ in transition rates, but not involving promoters and therefore $Y(t)$, have been considered more recently in the gene regulation setting by Goutsias and Kim (2006), where the authors introduce biologically meaningful delays in feedback interactions. They replace occurrences in transition rates of $\kappa(N(t))$ and $g(N(t))$ by expressions involving their expected values, that is by $\kappa(\mathbb{E}(N(t - \theta)))$ and $g(\mathbb{E}(N(t - \theta)))$ for some delay θ . Time delays are biologically very meaningful since, in fact, proteins move around at random and the delay θ might represent the average time a protein takes to move back in the neighborhood of the promoter. As stated in the Introduction, their models however do not involve $Y(t) \in \{0, 1\}$, and therefore promoters. For the self-regulated gene, assuming linear degradation transition rates of the form $\nu(n) = \nu n$, the limiting steady state is again Poisson, so that this model is not completely satisfactory for predicting the propagation of noise in gene expression levels.

In a previous work Fournier et al. (2007), we proposed a mean-field model, which includes promoter states and time delays, extending a model of Bratsun et al. (2005). The regulatory network was studied in living cells, and the experimental data were in good agreement with the model's predictions. We also provided a rationale for introducing mean-field interactions: The many steps and relatively slow transitions between states of chromatin in mammalian cells between the permissive and the non-permissive states of chromatin may dampen the noise that stems from the stochastic binding of a low number

of activator proteins to the promoter and from noise amplification resulting from the gene auto-activation feedback. In this setting chromatin may act as a noise-filtering device that allows graded response from stochastic events. This new Markov chain $\eta(t)$ evolves in the same state space, but has transition rates given by (we assume that $\mu_0 = 0$ and $\mu_1 = \mu > 0$)

$$\begin{aligned}
 q(N(t+h) = n+1, Y(t+h) = y | N(t) = n, Y(t) = y) &= y \mu h + o(h), \\
 q(N(t+h) = n, Y(t+h) = 1-y | N(t) = n, Y(t) = y) \\
 &= \kappa(\mathbb{E}(N(t-\theta))) h + o(h) \quad \text{when } y = 1, \\
 q(N(t+h) = n, Y(t+h) = 1-y | N(t) = n, Y(t) = y) \\
 &= g(\mathbb{E}(N(t-\theta))) h + o(h) \quad \text{when } y = 0,
 \end{aligned}$$

and

$$q(N(t+h) = n-1, Y(t+h) = y | N(t) = n, Y(t) = y) = \nu n h + o(h).$$

The main difference with the basic model is that transition rates like $p((n, y), (n, 1-y)) = \kappa(n)$ are replaced by time nonhomogeneous rates $q_t((n, y), (n, 1-y)) = \kappa(\mathbb{E}(N(t-\theta)))$, so that the related Markov chain is time nonhomogeneous. Let us denote by Q_t the related transition matrix at time t , of instantaneous steady-state distribution π^t , with $\pi^t Q_t = 0$. In what follows, we shall use the family of transition matrices $Q_{(b,c)}$ given by

$$\begin{aligned}
 q_{(b,c)}((n, y), (n+1, y)) &= y \mu, & q_{(b,c)}((n, y), (n-1, y)) &= \nu n, \\
 q_{(b,c)}((n, y), (n, 1-y)) &= b \quad \text{when } y = 1, & \text{and} \\
 q_{(b,c)}((n, y), (n, 1-y)) &= c \quad \text{when } y = 0,
 \end{aligned}$$

of steady-state distribution $\pi^{(b,c)}$. Then $Q_t = Q_{(b(t),c(t))}$, where $b(t) = \kappa(\mathbb{E}(N(t-\theta)))$ and $c(t) = g(\mathbb{E}(N(t-\theta)))$.

When dealing with time nonhomogeneous Markov chains, the main problem is that the law of the stochastic process $P(\eta(t) = (n, y))$ does not necessarily converges toward the limiting steady-state distribution (when it exists) $\lim_{t \rightarrow \infty} \pi^t$, and can lead to oscillations, as provided for example in Bratsun et al. (2005), or in Goutsias and Kim (2006). The first thing we can do is to check the asymptotic behavior of the functions $b(t)$ and $c(t)$. Suppose that these functions converge toward positive numbers $b(\infty)$ and $c(\infty)$. Then one ask if the following holds

$$\lim_{t \rightarrow \infty} P(\eta(t) = (n, y)) = \lim_{t \rightarrow \infty} \pi^t(n, y) = \pi^{(b(\infty),c(\infty))}(n, y) ? \tag{2}$$

Assume that this is true: Then one gets that the steady-state behavior of the self-regulated gene is given by computing the steady-state and the basic statistical descriptors related to the Markov chain of transition kernel $Q_{(b(\infty),c(\infty))}$, which is much simpler. We will see that in such a situation, one can get exact formulas for the mean and for the variance of the number of proteins (see also Example 1).

Equation (2) holds under fairly general assumptions. Theorem 6 of Appendix A gives that

$$\lim_{t \rightarrow \infty} P(\eta(t) = (n, y)) = \pi^{(b(\infty), c(\infty))}(n, y),$$

when the limiting process of transition kernel $Q_{(b(\infty), c(\infty))}$ is ergodic,

$$\begin{aligned} \int_0^\infty \left(\sqrt{b(t)} - \sqrt{b(\infty)} \right)^2 dt < +\infty, \quad \text{and} \\ \int_0^\infty \left(\sqrt{c(t)} - \sqrt{c(\infty)} \right)^2 dt < +\infty, \end{aligned} \quad (3)$$

and under an additional hypothesis which is automatically satisfied in our model (see the Appendix A).

The chemical master equation yields differential equations for $G(t) = P(Y(t) = 1)$ and $E(t) := \mathbb{E}(N(t))$, given by

$$\begin{aligned} \frac{dE}{dt} &= \mu G(t) - \nu E(t), \\ \frac{dG}{dt} &= c(t)(1 - G(t)) - b(t)G(t). \end{aligned} \quad (4)$$

Remark 1. Bratsun et al. (2005), and Goutsias and Kim (2006), consider delayed differential systems similar to the system given by (4) with solutions oscillating toward limit cycles. The steady state exists however only when the solutions of this system converge as $t \rightarrow \infty$. We will see that this is the case for a linear positive feedback.

Algorithm 2.

(STEP 1): Check the convergence of the orbits of the system given by (4), for a given initial condition $G(0)$, $E(t)$, $-\theta \leq t \leq 0$. When convergence holds, proceed to the next step

(STEP 2): Let $e_\infty = E(\infty)$, with $e_\infty = \mu/\nu G(\infty)$. Solve the equation

$$g(e_\infty) \left(1 - \frac{\nu}{\mu} e_\infty \right) - \kappa(e_\infty) \frac{\nu}{\mu} e_\infty = 0.$$

(STEP 3): Let

$$\tau_2 := \left(g(e_\infty) + \kappa(e_\infty) \right)^{-1}.$$

Compute the coefficient of variation as

$$CV_N^2 = \frac{\text{Var}(N(\infty))}{\mathbb{E}(N(\infty))^2} = \frac{1}{e_\infty} + \frac{\tau_2}{\tau_1 + \tau_2} \frac{(1 - G(\infty))}{G(\infty)},$$

where $\tau_1 = 1/\nu$.

For more insight in these formulas, see the remarks in Example 1.

3.1. Convergence for linear positive feedbacks

Recall that $c(t) = g(\mathbb{E}(N(t - \theta)))$ and assume that $\kappa(n) \equiv \kappa$. We generalize the linear case by assuming that $g(x)/x$ is decreasing. Notice that even when $g(x)$ is affine in n , this does not mean that the stochastic system is linear: for example, assuming fast promoters or a quasi-equilibrium, one gets a time-nonhomogeneous birth and death process with birth rate $\mu c(t)/(c(t) + \kappa)$ and death rate νn . We prove below that the above dynamical system is such that there is a globally asymptotically stable critical point $(E(\infty), G(\infty))$ with

$$G(\infty) > 0 \quad \text{and} \quad c(\infty) > 0$$

(see Gabriel et al., 1981, for more general mathematical results). In this case, the mean and the variance of the number of proteins are obtained by studying the transition kernel $Q_{(\kappa, c(\infty))}$.

In the following, we focus on the system (4) that reads in our setting

$$\begin{aligned} \frac{dE}{dt} &= \mu G(t) - \nu E(t), \\ \frac{dG}{dt} &= g(\mathbb{E}(N(t - \theta)))(1 - G(t)) - \kappa G(t), \end{aligned} \tag{5}$$

where $g(\cdot)$ is continuously differentiable and increasing over \mathbb{R}_+ , $g(0) > 0$, $\frac{g(E)}{E}$ is decreasing, the initial condition $E(t)$ is continuous and non-negative over $[-\theta, 0]$ and $0 \leq G(0) \leq 1$. We use the notation \dot{f} for the derivative df/dt . We proceed step by step to show that condition (3) holds.

In Lemma 2, we prove that the evolution equations defining the system are well defined, providing a unique solution, then we show in Lemma 3 that the system converges to the unique biologically meaningful critical point of the system and in Lemma 4 that the speed of convergence is exponential. The methods used are adapted from Gabriel et al. (1981). Finally, using our hypothesis on the function g , it is easy to conclude that the condition (3) holds, and the main result is stated in Theorem 3.

The theory of delayed differential equations is very different from the usual theory of differential equations, here the initial condition is no more a point in the finite dimensional space $\mathbb{R}_+ \times [0, 1]$ but a continuous non-negative function $E(t)$ over the interval $[-\theta, 0]$ and a value $G(0) \in [0, 1]$. To solve the system (5), we have to first integrate the second equation over the interval $[0, \theta]$, then plug the solution in the first equation and integrate using the variation of constant over the interval $[0, \theta]$ and begin the whole procedure anew over the interval $[\theta, 2\theta]$ with initial condition given by $E(t)$ over the interval $[0, \theta]$, and so on.

Lemma 2 (Existence and uniqueness). *For any initial condition $E(t)$ non-negative and continuous over $[-\theta, 0]$ and $0 \leq G(0) \leq 1$, there exists a unique solution of the system (5) defined over $[0, +\infty)$. Furthermore,*

$$0 < E(t) \leq \max\{E(0), \mu/\nu\}, \quad t \geq 0,$$

and

$$0 < G(t) < 1, \quad t > 0.$$

Proof: For any initial condition $0 \leq G(0) \leq 1$ and $E(t)$ non-negative and continuous over $[-\theta, 0]$, (5) admits obviously a unique solution over $[0, \theta]$. If $G(0) > 0$, then by continuity G remains strictly positive over some open interval to the right of 0. If $G(0) = 0$, then according to the second equation of (5), $\dot{G}(0) > 0$ and the same conclusion holds. The same reasoning shows that $G < 1$ over some open interval to the right of 0. Clearly, if they exist, $t_0 = \inf\{t \in (0, \theta], G(t) = 0\}$ and $t_1 = \inf\{t \in (0, \theta], G(t) = 1\}$ are both strictly positive. By definition, $\dot{G}(t_0) \leq 0$ and by continuity, $G(t_0) = 0$. The second equation of (5) entails $\dot{G}(t_0) > 0$. We have a similar contradiction for t_1 , thus $0 < G(t) < 1$ over $(0, \theta]$. The variation of constant formula entails

$$0 < E(t) \leq \max\{E(0), \mu/\nu\} \quad \text{over } (0, \theta].$$

Iterating the procedure provides existence and uniqueness of a solution defined over $[0, +\infty)$ and the preceding inequalities are preserved. \square

We are interested in the possible equilibria of (5) in $\mathbb{R}_+ \times [0, 1]$, i.e. the solutions (E_0, G_0) in $\mathbb{R}_+ \times [0, 1]$ of

$$0 = \mu G_0 - \nu E_0, \quad 0 = g(E_0)(1 - G_0) - \kappa G_0.$$

Clearly $G_0 = 0$ and $G_0 = 1$ lead to contradictions. We thus have $0 < G_0 < 1$ and consequently $E_0 > 0$. Plugging $G_0 = \frac{\nu}{\mu} E_0$ in the second equation yields

$$\frac{g(E_0)}{E_0} \left(\frac{\mu}{\nu} - E_0 \right) = \kappa.$$

If $g(0) > 0$ and $\frac{g(E)}{E}$ is decreasing over $(0, +\infty)$, then $\frac{g(E)}{E} \left(\frac{\mu}{\nu} - E \right)$ is strictly decreasing. Since it starts at $+\infty$ and becomes ultimately negative, we conclude to the existence of a unique solution $(E_0, G_0) \in \mathbb{R}^2 \times [0, 1]$.

We will use the fluctuation Lemma 6 given in Appendix A to prove the convergence to the critical point (E_0, G_0) .

Lemma 3 (Convergence). *For any initial condition $E(t)$ non-negative and continuous over $[-\theta, 0]$ and $0 \leq G(0) \leq 1$, the unique solution $(E(t), G(t))$ converges to (E_0, G_0) as $t \rightarrow \infty$.*

Proof: The fluctuation Lemma 6 and the monotonicity of g imply that

$$\begin{aligned} 0 &\geq \mu \underline{G} - \nu \underline{E}, & 0 &\geq g(\underline{E})(1 - \underline{G}) - \kappa \underline{G} \\ 0 &\leq \mu \overline{G} - \nu \overline{E}, & 0 &\leq g(\overline{E})(1 - \overline{G}) - \kappa \overline{G}. \end{aligned} \tag{6}$$

We prove the last inequality to exemplify the method. We choose $t_n \uparrow +\infty$ so that $G(t_n) \rightarrow \overline{G}$ and $\dot{G}(t_n) \rightarrow 0$ as $n \rightarrow +\infty$. Since the sequence $\{E(t_n)\}_{n \geq 1}$ is bounded, there exists a subsequence $(t_{n_k})_{k \geq 1}$ so that $E(t_{n_k})$ converges as $k \rightarrow \infty$ to a certain value that we call E_∞ . Evaluating the equation for G over the subsequence $(t_{n_k})_{k \geq 1}$ and letting $k \rightarrow +\infty$, we get

$$0 = g(E_\infty)(1 - \overline{G}) - \kappa \overline{G} \leq g(\overline{E})(1 - \overline{G}) - \kappa \overline{G}$$

since g is increasing and $\overline{G} \leq 1$. The proof of the other inequalities in (6) is similar.

We already know that $0 < \underline{G}, \underline{E} > 0$ and $\overline{G} < 1$, and (6) entails $\underline{G} \leq \frac{\nu}{\mu} \underline{E}$ and $\overline{G} \geq \frac{\nu}{\mu} \overline{E}$, and in particular $\overline{E} \leq \frac{\mu}{\nu} \overline{G} < \frac{\mu}{\nu}$. Consequently

$$\frac{\kappa \nu}{\mu} \overline{E} \leq \kappa \overline{G} \leq g(\overline{E})(1 - \overline{G}) \leq g(\overline{E}) \left(1 - \frac{\nu}{\mu} \overline{E}\right),$$

hence

$$\kappa \leq \frac{g(\overline{E})}{\overline{E}} \left(\frac{\mu}{\nu} - \overline{E}\right).$$

Repeating the same argument for \underline{E} , one gets

$$\kappa \geq \frac{g(\underline{E})}{\underline{E}} \left(\frac{\mu}{\nu} - \underline{E}\right).$$

By assumption, $\frac{g(E)}{E}(\frac{\mu}{\nu} - E)$ is decreasing, so that the two last equations then give that $\overline{E} \leq E_0$ and $\underline{E} \geq E_0$. Clearly we have $\underline{E} = \overline{E} = E_0$, so that $E(t)$ converges as $t \rightarrow +\infty$. According to Lemma 3.1 in Coppel (1965), $E(t)$ and its first two derivatives being bounded on $[\theta, +\infty)$, we have $\lim_{t \rightarrow +\infty} \dot{E}(t) = 0$ and the relation $\dot{E} = \mu G - \nu E$ entail the convergence of $G(t)$ as $t \rightarrow +\infty$. \square

From this Lemma, we deduce that the propensity function

$$a_3(t, y) = g(\mathbb{E}(N(t - \theta))) y$$

converges to $g_\infty y = g(E_0) y$ as $t \rightarrow \infty$. To show that the convergence speed is exponential, we use Theorem 7 cited in Appendix A.

Lemma 4 (Exponential convergence). *The convergence of $(E(t), G(t))$ to (E_0, G_0) is exponential.*

Proof: Near a critical point, the asymptotic behavior of the system is determined by the asymptotic behavior of the linearized system

$$\begin{bmatrix} \dot{E}(t) \\ \dot{G}(t) \end{bmatrix} = A \cdot \begin{bmatrix} E(t) \\ G(t) \end{bmatrix} + B \cdot \begin{bmatrix} E(t - \theta) \\ G(t - \theta) \end{bmatrix},$$

where A and B are the matrices

$$A = \begin{bmatrix} -\nu & \mu \\ 0 & -(g(E_0) + \kappa) \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ g'(E_0)(1 - G_0) & 0 \end{bmatrix}.$$

We show that all roots λ of the characteristic equation $\det(A + e^{-\lambda\theta} B - \lambda I) = 0$ have negative real parts. The characteristic equation is here

$$\lambda^2 + \lambda(\nu + g(E_0) + \kappa) + \nu(g(E_0) + \kappa) - \mu g'(E_0)(1 - G_0)e^{-\lambda\theta} = 0,$$

and all roots λ of this equation have negative real part if and only if all roots z of

$$H(z) := (z^2 + pz + q)e^z + r = 0$$

have negative real parts, with

$$p := (v + g(E_0) + \kappa)\theta, \quad q := v(g(E_0) + \kappa)\theta^2, \quad r := -\mu g'(E_0)(1 - G_0)\theta^2,$$

and the change of variable $z := \lambda\theta$. According to Theorem 7 given in Appendix A, since $r < 0$ and

$$p^2 = 2q + v^2\theta^2 + (g(E_0) + \kappa)^2\theta^2 \geq 2q,$$

we have to check that $-q < r < 0$ and $r \sin(a_2)/(pa_2) \leq 1$, where a_2 is the unique root of the equation $\cot(a) = (a^2 - p)/q$ which lies in the interval $(2\pi, 3\pi)$. The second inequality is clear since $r/p < 0$ and $\sin(x)/x \geq 0$ on $(2\pi, 3\pi)$. For the inequality $-r < q$, notice that $g'(x) \leq g(x)/x$ since $g(x)/x$ is decreasing, and using the equilibrium equation $g(E_0)(1 - G_0) = \kappa G_0 = \frac{v\kappa}{\mu}E_0$, we have

$$-r = \mu g'(E_0)(1 - G_0)\theta^2 \leq \mu \frac{g(E_0)}{E_0}(1 - G_0)\theta^2 = v\kappa\theta^2 < v(\kappa + g(E_0))\theta^2 = q.$$

Hence all roots of the characteristic equation have negative real parts and the system is asymptotically stable. Since our system is autonomous, asymptotic stability implies uniform asymptotic stability. According to Theorem 4.6 in Halanay (1966), the convergence is exponential. □

Using the exponential convergence of $\mathbb{E}(N(t - \theta))$ and the hypothesis on g , it is now easy to show that condition (3) is satisfied.

Theorem 3. *Assume that $g(n) = g_0 + g_1n$ with $g_0 > 0$ and that $\kappa(n) \equiv \kappa$. The limiting distribution of the time-nonhomogeneous process $(N(t), Y(t))$ is such that*

$$\lim_{t \rightarrow \infty} P(N(t) = n, Y(t) = y) = \pi_n^{(\kappa, g(E_0))}(y),$$

where $\pi^{(g(e_\infty), \kappa)}$ is the steady-state distribution given by Theorem 1 for a self-regulated gene with the simpler transitions

$$\begin{aligned} P((n, y), (n + 1, y)) &= y \mu, & P((n, y), (n - 1, y)) &= vn, \\ P((n, y), (n, 1 - y)) &= \kappa \quad \text{when } y = 1, \\ P((n, y), (n, 1 - y)) &= g(E_0) \quad \text{when } y = 0. \end{aligned}$$

The steady-state coefficient of variation of $N(\infty)$ is given by

$$CV_N^2 = \frac{1}{E_0} + \frac{\tau_1}{\tau_1 + \tau_2} \frac{\kappa}{g(E_0)},$$

where

$$\tau_2 = \frac{1}{g(E_0) + \kappa}, \quad \tau_1 = \frac{1}{\nu}.$$

Proof: According to Theorem 6 in Appendix A, we only have to show that condition (3) holds. Using the positiveness and boundedness of $g(E(t - \theta))$, $0 < g(0) \leq g(E(t - \theta)) \leq g(\mu/\nu)$, and the expansion

$$\sqrt{g(E(t - \theta))} - \sqrt{g(E_0)} = \frac{g(E(t - \theta)) - g(E_0)}{\sqrt{g(E(t - \theta))} + \sqrt{g(E_0)}},$$

we have

$$\frac{|g(E(t - \theta)) - g(E_0)|}{2\sqrt{g(\mu/\nu)}} \leq \left| \sqrt{g(E(t - \theta))} - \sqrt{g(E_0)} \right| \leq \frac{|g(E(t - \theta)) - g(E_0)|}{2\sqrt{g(0)}}$$

and condition (3) is in our case equivalent to

$$\int_0^\infty (g(E(t - \theta)) - g(E_0))^2 dt < \infty.$$

Let ε be positive, $\varepsilon < E_0$ and T_ε be such that $|E(t - \theta) - E_0| < \varepsilon$ for all $t \geq T_\varepsilon$. Using the mean value theorem, for all $t \geq T_\varepsilon$, there exists a ξ_t in the interval delimited by $E(t - \theta)$ and E_0 such that

$$|g(E(t - \theta)) - g(E_0)| = g'(\xi_t)|E(t - \theta) - E_0|.$$

Since for all $t \geq T_\varepsilon$, $E_0 - \varepsilon < \min(E(t - \theta), E_0)$, and furthermore $0 \leq g'(x) \leq g(x)/x$ and $g(x)/x$ is decreasing,

$$\begin{aligned} |g(E(t - \theta)) - g(E_0)| &= g'(\xi_t)|E(t - \theta) - E_0| \leq \frac{g(\xi_t)}{\xi_t}|E(t - \theta) - E_0| \\ &\leq \frac{g(E_0 - \varepsilon)}{E_0 - \varepsilon}|E(t - \theta) - E_0| =: L_\varepsilon|E(t - \theta) - E_0|, \end{aligned}$$

and finally with the exponential convergence of $E(t - \theta)$ to E_0 , condition (3) holds

$$\begin{aligned} \int_0^\infty (g(E(t - \theta)) - g(E_0))^2 dt &\leq \int_0^{T_\varepsilon} (g(E(t - \theta)) - g(E_0))^2 dt \\ &\quad + L_\varepsilon \int_{T_\varepsilon}^\infty (E(t - \theta) - E_0)^2 dt < \infty. \quad \square \end{aligned}$$

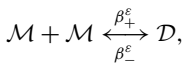
Remark 2. When the positive feedback rate $g(\mathbb{E}(N(t - \theta)))$ is such that $\frac{g(x)}{x}$ is increasing, for example when g is a polynomial of degree ≥ 2 , there can possibly exist several biologically meaningful equilibrium points and it cannot be excluded that for some initial conditions the solutions of (5) oscillate endlessly.

When g is constant but the negative feedback $\kappa(\mathbb{E}(N(t - \theta)))$ is an increasing function of $\mathbb{E}(N(t - \theta))$, the biologically meaningful equilibrium point is unique but similar application of the fluctuation lemma as in the proof of Lemma 3 yields the trivial observation that $\underline{E} \leq E_0 \leq \bar{E}$, and oscillating solutions cannot be excluded in this case either.

4. Two-time-scale stochastic simulations

The self-regulated gene and the network presented in the Introduction involve slow and fast species, like therapeutic proteins and activator dimers. We recall existing known probabilistic results concerning quasi-equilibrium. Let $\varepsilon > 0$ be a small parameter, which will be useful for describing fast species. In what follows, $\eta^\varepsilon(t)$ is a random vector describing the number of molecules of each species present in the cell at time t . For example, considering the self-regulated gene, $N^\varepsilon(t)$ gives the number of protein molecules at time t , and $Y^\varepsilon(t) = 0, 1$ gives the state of the promoter. The pair $\eta_s^\varepsilon(t) = (N^\varepsilon(t), Y^\varepsilon(t))$ stands for the slow process. $\eta_f^\varepsilon(t)$ models the fast process, and the global process is $\eta^\varepsilon(t) = (\eta_s^\varepsilon(t), \eta_f^\varepsilon(t))$.

A generic example of fast reaction is dimerization, as given by the chemical reaction



where \mathcal{M} represents protein monomers and \mathcal{D} protein dimers. Protein dimers form a fast species, while protein (involved in dimers or monomers) is a slow species. The rates of these reactions are fast when for example the rate constants β_-^ε and β_+^ε are such that $\beta_-^\varepsilon = c_-/\varepsilon$ and $\beta_+^\varepsilon = c_+/\varepsilon$, for positive constants c_- and c_+ , when $\varepsilon \approx 0$. In this setting, the number of protein monomers is then given by $N^\varepsilon(t) - 2D^\varepsilon(t)$, where $D^\varepsilon(t)$ gives the number of protein dimers present in the cell at time t . Here

$$\eta_s^\varepsilon(t) = (N^\varepsilon(t), Y^\varepsilon(t)) \quad \text{and} \quad \eta_f^\varepsilon(t) = D^\varepsilon(t).$$

When $\varepsilon \approx 0$, a quasi-equilibrium is attained, meaning that for given $\eta_s^\varepsilon(t) = k$, one can assume a local steady state for the number of dimers. More generally, we assume that the slow process evolves in some finite space $E_s = \{1, \dots, L\}$, and that, given $k \in E_s$, $\eta_f^\varepsilon(t) \in \mathbb{N}$ is described by a Markov transition kernel $A^k(t)/\varepsilon$ (see below). We follow essentially (Zhang and Yin, 1997). We assume that the generator $Q(t) = Q^\varepsilon(t)$ of the Gillespie algorithm $\eta^\varepsilon(t)$ can be decomposed as

$$Q^\varepsilon(t) = \frac{1}{\varepsilon}A(t) + B(t),$$

where $A(t)$ and $B(t)$ are matrix valued functions. Following Zhang and Yin (1997), assume that $A(t)$ has the block diagonal form

$$A(t) = \begin{pmatrix} A^1(t) & 0 & 0 & \dots & 0 \\ 0 & A^2(t) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A^L(t) \end{pmatrix},$$

where each block $A^k(t)/\varepsilon$ is a transition matrix representing the transitions rates of the fast variables given $\eta_s^\varepsilon = k$. The generator $B(t)$ gives the slow transition rates and in particular transitions of the form $((k, u), (j, v))$, where $u, v \in \mathbb{N}$ and $k, j \in E_s$. Following Zhang and Yin (1997), we partition the state space E as

$$E = E_1 \cup E_2 \cup \dots \cup E_L,$$

where each E_k contains m_k elements, with

$$E_k = \{e_{k1}, e_{k2}, \dots, e_{km_k}\}, k \in E_s.$$

Each E_k corresponds to some subset of $\{k\} \times \mathbb{N}$, with $k \in E_s$.

Hypothesis. We suppose that the process is time-homogeneous, that is that both $A(t)$ and $B(t)$ do not depend on t , and that each generator A^k is irreducible with a unique invariant probability measure $\sigma^k = (\sigma^k(e_{k1}), \dots, \sigma^k(e_{km_k}))$, such that $\sigma^k A^k = 0$.

Following Zhang and Yin (1997), each E_k can be aggregated, and represented by a single state k , corresponding to a particular slow state; The Markov process $\eta^\varepsilon(t)$ of transition kernel Q^ε is then approximated by an *aggregated process* $\tilde{\eta}^\varepsilon(t)$ defined by

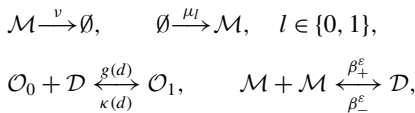
$$\tilde{\eta}^\varepsilon(t) = k \quad \text{if } \eta^\varepsilon(t) \in E_k, \quad k = 1, \dots, L.$$

This process converges in distribution as $\varepsilon \rightarrow 0$ toward a Markov process $\eta(t)$ generated by the kernel $Q = (\gamma_{kj})_{k,j \in E_s}$, with

$$\gamma_{kj} = \sum_{u=1}^{m_k} \sum_{v=1}^{m_j} \sigma^k(e_{ku}) B(e_{ku}, e_{jv}), \quad k \neq j.$$

4.1. Transcription with fast dimerization

The model is similar to that given in Section 2, with dimerization as a fast component, see e.g. Burrage et al. (2004), Cao et al. (2005), or Goutsias (2005). It is described by the following set of chemical reactions:



where \mathcal{D} represent dimers, g is function of the number of dimers, and the rates β_-^ε and β_+^ε involve a small number $\varepsilon > 0$ modeling the speed of dimerization, see below. The number $N^\varepsilon(t)$ of proteins \mathcal{P} present at time t is related to the number of dimers as $0 \leq 2D^\varepsilon(t) \leq N^\varepsilon(t)$, and the number of free monomers is such that $\mathcal{M} + 2\mathcal{D} = \mathcal{P}$.

The running process is a Markov process $\eta^\varepsilon(t) = (N^\varepsilon(t), Y^\varepsilon(t), D^\varepsilon(t))$, $0 \leq 2D^\varepsilon(t) \leq N^\varepsilon(t)$, $t \geq 0$. The dimerization process $D^\varepsilon(t)$ is given by the transition rates

$$P(\eta_s^\varepsilon(t+h) = \eta_s^\varepsilon(t), D^\varepsilon(t+h) = D^\varepsilon(t) + 1)$$

$$\begin{aligned}
&= \beta_+^\varepsilon (N^\varepsilon(t) - 2D^\varepsilon(t))(N^\varepsilon(t) - 2D^\varepsilon(t) - 1)h + o(h), \\
P(\eta_s^\varepsilon(t+h) = \eta_s^\varepsilon(t), D^\varepsilon(t+h) = D^\varepsilon(t) - 1) &= \beta_-^\varepsilon D^\varepsilon(t)h + o(h),
\end{aligned}$$

where the slow process is $\eta_s^\varepsilon(t) = (N^\varepsilon(t), Y^\varepsilon(t))$.

The transition rates of $D^\varepsilon(t)$ depend on $N^\varepsilon(t)$ but are independent of the state of the promoter. We can fit the setup of this section by setting

$$\beta_-^\varepsilon = \frac{c_-}{\varepsilon} \quad \text{and} \quad \beta_+^\varepsilon = \frac{c_+}{\varepsilon},$$

for positive constants c_- and c_+ . Then, the result of Section 4 yield that the slow process at quasi-equilibrium ($\varepsilon \approx 0$) is well described in the above discussion: For a given slow state $k = (n, y)$, one gets

$$g(n) = \sum_{0 \leq d \leq \lfloor n/2 \rfloor} \sigma^{(n,y)}(d)g(d), \quad \sigma^{(n,y)} = \mu^n,$$

where the quasi-equilibrium stationary measure $\sigma^{(n,y)}$ corresponds to the stationary measure μ^n of the dimerization process (see below). A typical example is given by $g(D^\varepsilon(t)) = \lambda D^\varepsilon(t) + g(0)$, that is depends linearly on the number of dimers at time t . Then, at quasi-equilibrium one gets

$$g(n) = \sum_{0 \leq d \leq \lfloor n/2 \rfloor} \mu^n(d)(\lambda d + g(0)) = \lambda \mathbb{E}_n + g(0),$$

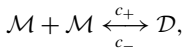
$$\text{where we set } \mathbb{E}_n := \sum_{0 \leq d \leq \lfloor n/2 \rfloor} d \mu^n(d),$$

where $[\dots]$ denotes the integer part. The algorithms developed in Section 2 can be applied efficiently if one can compute the rates $g(n)$. The next section develops efficient algorithms for computing $g(n)$ when $g(d)$ is linear, quadratic or a higher polynomial in the number d of dimers. The polynomial $g(d)$ is proportional to the polynomial known in biochemistry as the binding polynomial, see for example (Dill and Bromberg, 2003), Chapter 28, or Section 5.1. Its degree is given by the number of sites to which dimers bind to on the promoter and to model high cooperativity, one usually uses a Hill model with $g(d)$ proportional to $1 + d^{k_r}$, with k_r the number of sites on the promoter. The next section considers the problem of computing various moments like \mathbb{E}_n at quasi-equilibrium. This is a generic problem, and the usual way of dealing with moment problems consists in providing various kind of (Gaussian) approximations, see e.g. Cao et al. (2004, 2005) or Goutsias (2005). We here provide exact algorithms based on iterative procedure, which are computationally efficient, and which permit to avoid adding more approximative steps when dealing with stochastic simulations at quasi-equilibrium.

4.2. Dimerization

Dimerization appears in most biochemical processes, and is usually considered as a fast reaction. The aim of this section is to give mathematical statements relevant for computational purposes (see also Cao et al., 2005, Darvey et al., 1966 or Kepler and Elston, 2001).

Given a fixed number of proteins n , the dimerization process is given by the reaction



where we recall that \mathcal{M} and \mathcal{D} represent protein monomers and dimers. The infinitesimal transitions probabilities are such that

$$\begin{aligned} P(D(t+h) = i+1 | D(t) = i) &= c_+(n-2i)(n-2i-1)h + o(h), \\ P(D(t+h) = i-1 | D(t) = i) &= c_-ih + o(h). \end{aligned}$$

The stationary distribution μ^n of the process is given explicitly by

$$\mu^n(i) = \left(\frac{c_+}{c_-}\right)^i \frac{1}{(n-2i)!i!} \cdot \frac{1}{Z_n}, \quad 0 \leq i \leq n_2, \quad n_2 := \lfloor n/2 \rfloor,$$

where

$$Z_n = \sum_{i=0}^{n_2} \left(\frac{2c_+}{c_-}\right)^i \frac{1}{(n-2i)!i! \cdot 2^i}.$$

It can be shown that the generating function at equilibrium $M(s) = \sum_{i=0}^{n_2} \mu^n(i)s^i$ can be written using confluent hypergeometric functions

$$M(s) = \begin{cases} s^{n_2} \frac{{}_1F_1(-n_2, \frac{3}{2}, -\frac{c_-}{4c_+s})}{{}_1F_1(-n_2, \frac{3}{2}, -\frac{c_-}{4c_+})} & \text{if } n \text{ is odd,} \\ s^{n_2} \frac{{}_1F_1(-n_2, \frac{1}{2}, -\frac{c_-}{4c_+s})}{{}_1F_1(-n_2, \frac{1}{2}, -\frac{c_-}{4c_+})} & \text{if } n \text{ is even.} \end{cases}$$

This gives a theoretical way of computing the invariant measure as

$$\mu^n(i) = \frac{M^{(i)}(0)}{i!}, \quad 0 \leq i \leq n_2,$$

and the mean number of dimers in the stationary regime is given by

$$\mathbb{E}_n = M'(1).$$

Numerical computation based on this last formula is tedious and in the case described in Section 5.3 we have to compute moments repeatedly for each n between 0 and Λ . The recursive method described in the next section provides an alternative adapted to this situation.

4.3. An approach of the invariant measure adapted to numerical computation

We provide a different approach, which will allow efficient computations of the mean and second moment. If the feedback function in the slow process is linear or quadratic

in the number of dimers, the infinitesimal transition rates of the slow process at quasi-equilibrium will only depend on the first two moments. Set $y = c_+/c_-$, so that $Z_n = \sum_{i=0}^{n_2} \frac{y^i}{(n-2i)!i!}$. Using the following polynomial identities:

$$\sum_{i=1}^{n_2} \frac{iy^i}{(n-2i)!i!} = yZ_{n-2} \quad \text{and} \quad \sum_{i=1}^{n_2} \frac{i^2y^i}{(n-2i)!i!} = y^2Z_{n-4} + yZ_{n-2},$$

the mean $\mathbb{E}_n := \sum_{0 \leq i \leq n_2} i\mu^n(i)$ and second moment $\mathbb{E}_n^2 := \sum_{0 \leq i \leq n_2} i^2\mu^n(i)$ are given by

$$\begin{aligned} \mathbb{E}_n &= \frac{1}{Z_n} \sum_{i=1}^{n_2} \frac{iy^i}{(n-2i)!i!} = y \frac{Z_{n-2}}{Z_n} \\ \mathbb{E}_n^2 &= \frac{1}{Z_n} \sum_{i=1}^{n_2} \frac{i^2y^i}{(N-2i)!i!} = y^2 \frac{Z_{n-4}}{Z_n} + y \frac{Z_{n-2}}{Z_n} = \mathbb{E}_n(1 + \mathbb{E}_{n-2}). \end{aligned} \tag{7}$$

In what follows, we give another description of Z_n based on the involutions of the permutation group S_n . This approach will allow to compute the ratios Z_{n-2}/Z_n recursively.

4.3.1. A description of Z_n based on the involutions of S_n

Let I_n denote the involution subgroup of the permutation group S_n , i.e. the set of permutation of n points $\sigma \in S_n$ so that σ^2 is the identity. For $\sigma \in I_n$, $\text{fix}(\sigma)$ denotes the number of fixed points of σ . For any number $n - 2i$ between 0 and n , the number of involutions with $n - 2i$ fixed points is given by

$$\sum_{\substack{\sigma \in I_n \\ \text{fix}(\sigma) = n-2i}} 1 = \binom{n}{2} \cdot \binom{n-2}{2} \cdots \binom{n-2(i-1)}{2} \cdot \frac{1}{i!} = \frac{n!}{(n-2i)!i! \cdot 2^i},$$

so that, setting $q := \sqrt{c_-/2c_+} = (2y)^{-1/2}$, Z_n can be written as

$$Z_n = \sum_{i=0}^{n_2} \left(\frac{2c_+}{c_-}\right)^i \frac{1}{(n-2i)!i! \cdot 2^i} = \frac{1}{n!} \sum_{\sigma \in I_n} \left(\frac{2c_+}{c_-}\right)^{\frac{n-\text{fix}(\sigma)}{2}} = \frac{q^{-n}}{n!} \sum_{\sigma \in I_n} q^{\text{fix}(\sigma)}.$$

Let Q_n denote the polynomial $Q_n(q) := \sum_{\sigma \in I_n} q^{\text{fix}(\sigma)}$, so that the partition function and the mean (7) can be written as

$$Z_n = \frac{q^{-n}}{n!} Q_n(q), \quad \text{and} \quad \mathbb{E}_n = \frac{n(n-1)}{2} \frac{Q_{n-2}(q)}{Q_n(q)}. \tag{8}$$

According to Randrianarivony (1997), one can identify Q_n as the Taylor coefficient of a Stieltjes type continued fraction. Here we proceed in a recursive way, using the following two propositions.

Proposition 2. $Q_n(q)$ satisfies the relation $Q_{n+1}(q) = q Q_n(q) + Q'_n(q)$.

Proof: Each involution $\sigma \in I_n$ induces $1 + \text{fix}(\sigma)$ involutions in I_{n+1} , namely the one that fixes the point $n + 1$ and the $\text{fix}(\sigma)$ ones that interchange a fixed point of σ with $n + 1$. Partitioning I_{n+1} as the set of involutions that fixe $n + 1$, and those that do not, we see that the first set contains exactly the involutions of I_n except that they have one more fixed point, namely $n + 1$. For each $\sigma \in I_n$ that fixes at least one point, the second set contains $\text{fix}(\sigma)$ involutions with one fixed point less, namely the one that is interchanged with $n + 1$. More precisely, the partition of I_{n+1} is given by

$$\begin{aligned} I_{n+1} &= \{\sigma \in I_{n+1}; \sigma \text{ fixes } n + 1\} \cup \{\sigma \in I_{n+1}; \sigma \text{ does not fix } n + 1\} \\ &= I_n \cup \bigcup_{\sigma \in I_n} \bigcup_{\sigma \text{ fixes } k} \{\sigma \circ (k, n + 1)\} \end{aligned}$$

where $(k, n + 1)$ is the permutation of k and $n + 1$. Therefore, we have the recurrence relation

$$\begin{aligned} Q_{n+1}(q) &= \sum_{\sigma \in I_{n+1}} q^{\text{fix}(\sigma)} = \sum_{\sigma \in I_n} q^{\text{fix}(\sigma)+1} + \sum_{\substack{\sigma \in I_n, \\ \text{fix}(\sigma) \geq 1}} \sum_{1 \leq k \leq \text{fix}(\sigma)} q^{\text{fix}(\sigma)-1} \\ &= q \sum_{\sigma \in I_n} q^{\text{fix}(\sigma)} + \sum_{\substack{\sigma \in I_n, \\ \text{fix}(\sigma) \geq 1}} \text{fix}(\sigma) q^{\text{fix}(\sigma)-1} = q Q_n(q) + Q'_n(q). \end{aligned} \quad \square$$

Proposition 3. *The derivative of Q_n is given by $Q'_n(q) = n \cdot Q_{n-1}(q)$, and hence*

$$Q_{n+1}(q) = q Q_n(q) + n \cdot Q_{n-1}(q). \tag{9}$$

Proof: One can easily compute $Q_1(q) = q$ and $Q_2(q) = q^2 + 1$. If $Q'_n(q) = n Q_{n-1}(q)$ for some n , using Proposition 2 for the first and last equality and by the induction hypothesis for the second one, we have

$$\begin{aligned} Q'_{n+1}(q) &= Q_n(q) + q Q'_n(q) + Q''_n(q) \\ &= Q_n(q) + qn Q_{n-1}(q) + n Q'_{n-1}(q) = (n + 1) Q_n(q). \end{aligned} \quad \square$$

Remark 3. Let $h_q(t) := \sum_{n=0}^{\infty} Q_n(q) \frac{t^n}{n!}$. Multiplying both sides of (9) by $\frac{t^n}{n!}$ and summing over all possible n leads to the equation $h'_q(t) = (t + q)h_q(t)$, with initial condition $h_q(0) = Q_0(q) = 1$, which has the unique solution

$$h_q(t) = e^{\frac{t^2}{2} + qt}.$$

This function is the moment generating function of a normal random variable of mean q and variance 1, so that $Q_n(q)$ is the moment of order n of a random variable $X \sim \mathcal{N}(q, 1)$.

Algorithm 3. Due to the fast increase of its coefficients, Q_n cannot be efficiently computed for large n . However, the computation of the mean \mathbb{E}_n only involves the ratio Q_{n-2}/Q_n .

Let $c_n(q) := \frac{Q_{n-1}(q)}{Q_n(q)}$. From $\frac{Q_{n-2}(q)}{Q_n(q)} = c_{n-1}(q) \cdot c_n(q) = \frac{1}{n-1}(1 - q c_n(q))$, one obtains the continued fraction

$$\frac{1}{c_{n+1}(q)} = q + n c_n(q). \tag{10}$$

From $Q_0(q) = 1$ and $Q_1(q) = q$, the first term c_1 is given by $c_1(q) = 1/q$.

The mean and the second moment (8) or (7) can then be computed recursively as

$$\begin{aligned} \mathbb{E}_n &= \frac{n}{2}(1 - q c_n(q)), \\ \mathbb{E}_n^2 &= \frac{n}{2}(1 - q c_n(q)) \left(1 + \frac{n-2}{2}(1 - q c_{n-2}(q)) \right). \end{aligned}$$

Theorem 4. $c_n(x) \rightarrow 0$ as $n \rightarrow \infty$.

Proof: Suppose that the lim sup of the sequence of non-negative numbers $\{c_n(x)\}_{n \geq 1}$ is strictly positive,

$$\limsup_{n \rightarrow \infty} c_n(x) = a > 0.$$

Using relation (10), one gets

$$a = \limsup_{n \rightarrow \infty} c_n(x) = \frac{1}{x + \liminf_{n \rightarrow \infty} n c_n(x)} =: \frac{1}{x + b},$$

where $b := \liminf_{n \rightarrow \infty} n c_n(x)$ has to be finite. Isolating $c_n(x)$ in (10) yields

$$c_n(x) = \frac{1}{n c_{n+1}(x)} - \frac{x}{n}$$

and we get

$$a = \limsup_{n \rightarrow \infty} c_n(x) = \limsup_{n \rightarrow \infty} \left(\frac{1}{n c_{n+1}(x)} - \frac{x}{n} \right) = \frac{1}{\liminf_{n \rightarrow \infty} n c_{n+1}(x) \cdot \frac{n+1}{n+1}} = \frac{1}{b}.$$

Since $x > 0$ and $b < \infty$, this leads to the contradiction $\frac{1}{x+b} = \frac{1}{b}$.

Notice that the limiting proportion of dimers is given $\frac{1}{2}$ for large n , more precisely $\lim_{n \rightarrow \infty} \frac{\mathbb{E}_n}{n} = \frac{1}{2}$, for every set of positive parameters c_+, c_- .

In our concrete Example of Section 4.1, we are mainly interested in computing higher moments for n proteins. We will show below that the computation of higher moments only requires the knowledge of the first moments for a lower number of proteins. More precisely, let $P_{j+1}(i)$ denote the polynomial

$$P_{j+1}(i) := i \cdot (i - 1) \cdots (i - 2) \cdots (i - j) =: i^{j+1} - \sum_{l=1}^j a_{l,j} i^l.$$

With the convention that $\mathbb{E}_i = 0$ for $i < 0$, the higher moments can be computed as combinations of the means for lower total number of proteins.

Lemma 5. $\mathbb{E}_n(P_{j+1}(D)) = \mathbb{E}_{n-2j} \cdot \mathbb{E}_{n-2(j-1)} \cdots \mathbb{E}_{n-2} \cdot \mathbb{E}_n.$

Proof: We show that both terms are equal to $y^{j+1} \frac{Z_{n-2(j+1)}}{Z_n}.$

$$\begin{aligned} Z_n \cdot \mathbb{E}_n(P_{j+1}(D)) &= \sum_{i=1}^{n_2} \frac{P_{j+1}y^i}{(n-2i)! i!} = \sum_{i=1}^{n_2} \frac{i \cdot (i-1) \cdots (i-2) \cdots (i-j) y^i}{(n-2i)! i!} \\ &= \sum_{i=1}^{n_2} \frac{y^i}{(n-2i)!(i-j-1)!} \cdot \mathbf{1}_{\{i>j\}} \\ &= \sum_{i=j+1}^{n_2} \frac{y^i}{(n-2(i-j-1)-2(j+1))! (i-j-1)!} \\ &= \sum_{i=0}^{n_2-(j+1)} \frac{y^{i+j+1}}{(n-2(j+1)-2i)! i!} \\ &= y^{j+1} Z_{n-2(j+1)}, \end{aligned}$$

and with (8), we have

$$\begin{aligned} \mathbb{E}_{n-2j} \cdot \mathbb{E}_{n-2(j-1)} \cdots \mathbb{E}_{n-2} \cdot \mathbb{E}_n &= y \frac{Z_{n-2(j+1)}}{Z_{n-2j}} \cdot y \frac{Z_{n-2j}}{Z_{n-2(j-1)}} \cdots y \frac{Z_{n-2}}{Z_n} \\ &= y^{j+1} \frac{Z_{n-2(j+1)}}{Z_n}. \end{aligned}$$

□

From the preceding lemma, we can give a formula for arbitrary moments:

Theorem 5. *The moment of order (j + 1) of D is given by*

$$\mathbb{E}_n^{j+1} = \mathbb{E}_{n-2j} \cdot \mathbb{E}_{n-2(j-1)} \cdots \mathbb{E}_{n-2} \cdot \mathbb{E}_n + \sum_{l=1}^j a_{l,j} \mathbb{E}_n^l.$$

Proof: From the definition of the coefficients $a_{l,j},$

$$\mathbb{E}_n(P_{j+1}(D)) = \mathbb{E}_n^{j+1} - \sum_{l=1}^j a_{l,j} \mathbb{E}_n^l,$$

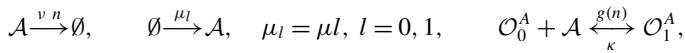
hence with Lemma 5 the statement holds.

Theorem 5 provides then a way of performing efficiently the computation of $g(n)$ (see Section 4.1) for any polynomial function of the number of dimers.

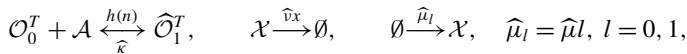
5. Modeling the regulatory gene network

We first recall the basic mathematical steps which lead to the mathematical model studied in Fournier et al. (2007). We shall see that the time evolution of the gene products

involved in the network described in the Introduction can be modeled by the following set of chemical reactions:



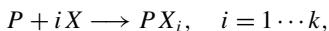
where the symbol \mathcal{A} stands for activator proteins, and \mathcal{O}_l^A , $l = 0, 1$ denotes the state of the promoter related to the activator, and by the chemical reactions related to therapeutic proteins as given by



where \mathcal{A} denotes activator proteins, and \mathcal{O}_l^T is defined in a similar way for the promoter of the therapeutic gene and \mathcal{X} symbolizes therapeutic proteins.

5.1. Equilibrium equations

The modeling of the time evolution of the number of molecules involved in the regulatory network is obtained by assuming that extrinsic noise, here the random fluctuations of the number of repressor and doxycycline molecules attains a chemical equilibrium. We first describe mathematically the effect of this extrinsic noise on the promoters associated with the activator and therapeutic genes. We follow Section 28 of Dill and Bromberg (2003). Consider a multiple binding of a ligand X with $1 \leq i \leq k$ different binding sites on a polymer P ,



with equilibrium constants

$$K_i = \frac{[PX_i]}{[P][X]^i}.$$

The *binding polynomial* is defined by

$$Q(X) = 1 + \sum_{i=1}^k K_i X^i,$$

where in the sequel X will denote the number of ligand molecules. The proportion of P molecules that are in the i th ligand state is $\frac{[PX_i]}{[P]Q}$, and the average number of bound sites is

$$M(X) = \frac{d \ln(Q)}{d \ln(X)} = \frac{\sum_{i=0}^k i K_i X^i}{Q(X)}.$$

Example 2. If the k binding sites are independent, there is no cooperativity, and one has

$$Q(X) = (1 + KX)^k,$$

with

$$M(X) = \frac{kKX}{1 + KX}.$$

Example 3. If a P molecule binds to exactly k ligand molecules at a time, one gets the Hill model $kX + P \rightarrow PX_k$, with equilibrium constant K , and

$$Q(X) = 1 + KX^k,$$

$$M(X) = \frac{kKX^k}{1 + KX^k}.$$

5.2. Transgene expression

5.2.1. Reaction of TetR repressor and doxycycline

The reaction between the doxycycline (Dox) and the repressor (R) is described as $R + \text{Dox} \rightarrow RD$ with some forward rate, and $RD \rightarrow R + D$ with some backward rate; Considering equilibrium of constant K_{RD} , we can write $K_{RD} = [RD]/([R][\text{Dox}]$, where $[\text{Dox}]$ gives the number of molecules of doxycycline. The free proportion of (R) molecules, i.e. not involved in the RD complex, can, when considered as ligand, bind to the k_r sites of the TetR operators (the binding sites where repressor molecules can bind, see e.g. Imhof et al., 2000), inhibiting thus both the transactivator and the synthesis of the transgene product. We next estimate the average fraction $F(R, [\text{Dox}])$ of sites free of repressor. Let k_r denote the number of sites where repressor molecules can bind. Using a Hill model of cooperativity (see Example 3), one gets that the average number of bound sites is then given by

$$M_r([R]) = \frac{k_r K_r [R]^{k_r}}{1 + K_r [R]^{k_r}}.$$

Then,

$$F(R, [\text{Dox}]) = \frac{k_r - M_r([R])}{k_r} = \frac{1}{1 + K_r [R]^{k_r}}.$$

The total number of repressor, denoted by R_{tot} , is such that

$$[R_{\text{tot}}] \approx [R] + K_{RD}[R][\text{Dox}] = [R](1 + K_{RD}[\text{Dox}]),$$

when we neglect the amount of repressor involved in the k_r binding sites. Set $[R_{\text{tot}}] = R_{\text{max}}$. Then

$$F(R, [\text{Dox}]) = \frac{(1 + K_{RD}[\text{Dox}])^{k_r}}{(1 + K_{RD}[\text{Dox}])^{k_r} + K_r R_{\text{max}}^{k_r}}.$$

5.2.2. Transactivator

The transactivator is repressed by the bound repressors, and activated by the positive feedback loop; The above considerations suggest a stochastic model of transactivation with

$$g(d) = VF(R, [\text{Dox}])d^a + g(0),$$

where a denotes the number of binding sites on the activator, d denotes the number of transactivator dimers, and where V is a parameter. $g(0) > 0$ is introduced here to model basal activity for the off to on transitions.

5.3. The regulatory network at quasi-equilibrium

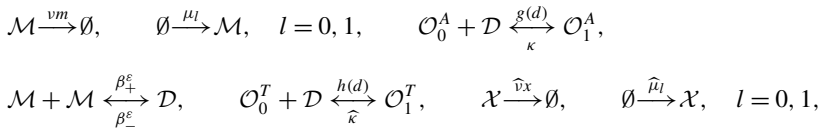
We assume that the promoter switch from the off to on state at rate

$$g(D^\varepsilon(t)) = VF(R, [\text{Dox}]) (D^\varepsilon(t))^a + g(0),$$

where $D^\varepsilon(t)$ is the number of transactivator dimers present at time t , a is the number of activator binding sites and V is a parameter. This models the positive feedback loop. We suppose that degradation occurs at a rate proportional to the number of monomers $N^\varepsilon(t) - 2D^\varepsilon(t)$, with constant of proportion ν . The transactivator process is given by the triplet $(N^\varepsilon(t), Y^\varepsilon(t), D^\varepsilon(t))$, where we assume fast dimerization, as given in the preceding paragraph. The time evolution of the network is described by the random process

$$\eta^\varepsilon(t) = (N^\varepsilon(t), Y^\varepsilon(t), D^\varepsilon(t), X^\varepsilon(t), Z^\varepsilon(t)),$$

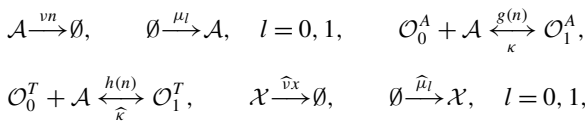
where $X^\varepsilon(t)$ denotes the number of therapeutic proteins (\mathcal{X}) present in the cell at time t , and where $Z^\varepsilon(t) = 0, 1$ denotes the state of its associated promoter (off/on). These chemical reactions are described schematically as



where \mathcal{O}_l^T , $l = 0, 1$ accounts for the state of the promoter related to the therapeutic gene, \mathcal{M} denotes activator proteins (monomers), $\mu_l = \mu l$, $l = 0, 1$, \mathcal{X} denotes therapeutic proteins and $\widehat{\mu}_l = \widehat{\mu} l$, $l = 0, 1$. We again assume a quasi-equilibrium with fast dimerization, to get the limiting process

$$\eta(t) = (N(t), Y(t), X(t), Z(t)),$$

associated with the set of coupled chemical reactions



with quasi-equilibrium transition rates given by (see Section 4)

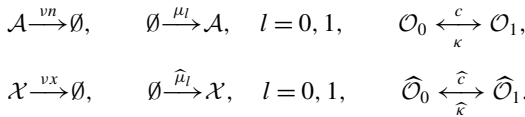
$$\begin{aligned} \nu(n) &= \nu \mathbb{E}_{\mu^n}(n - 2d), \\ g(n) &= VF(R, [\text{Dox}]) \mathbb{E}_{\mu^n}(d^a) + g(0), \\ h(n) &= \widehat{V} F(R, [\text{Dox}]) \mathbb{E}_{\mu^n}(d^a) + h(0). \end{aligned}$$

5.4. A semi-stochastic mean-field model

We consider the time evolution of the network in a semi-stochastic version by supposing that the rates $g(N(t))$ and $h(N(t))$ are replaced by $c(t) = g(\mathbb{E}(N(t - \theta)))$ and $\widehat{c}(t) = h(\mathbb{E}(N(t - \widehat{\theta})))$. The method is similar to what is presented in Section 3. Consider the family of transition kernels $L_{c,\widehat{c}} = (q_{c\widehat{c}}((n, y, x, z)(n', y', x', z')))$ given by

$$\begin{aligned} q_{c\widehat{c}}((n, y, x, z)(n + 1, y, x, z)) &= y\mu, \\ q_{c\widehat{c}}((n, y, x, z)(n - 1, y, x, z)) &= \nu n, \\ q_{c\widehat{c}}((n, y, x, z)(n, 1 - y, x, z)) &= \kappa y + (1 - y)c, \\ q_{c\widehat{c}}((n, y, x, z)(n, y, x + 1, z)) &= z\widehat{\mu}, \\ q_{c\widehat{c}}((n, y, x, z)(n, y, x - 1, z)) &= \widehat{\nu}x, \\ q_{c\widehat{c}}((n, y, x, z)(n, y, x, 1 - z)) &= \widehat{\kappa}z + (1 - z)\widehat{c}. \end{aligned}$$

The nice feature of this kernel is that its steady-state distribution is the product $\pi^{\kappa,c} \otimes \pi^{\widehat{\kappa},\widehat{c}}$ of the stationary distributions associated with the self-regulated genes given by the two sets of chemical reactions



Both measures can be computed efficiently by using either the method of transfer matrices or the exact analytical expressions obtained through generating functions. Coming back to the time evolution of the network under a mean-field model, the method is similar to that given in Section 3 and consists in two basic steps:

- Find the limiting values $c(\infty) = \lim_{t \rightarrow \infty} c(t)$ and $\widehat{c}(\infty) = \lim_{t \rightarrow \infty} \widehat{c}(t)$, when they exist;
- Compute the steady-state distribution $\pi^{\kappa,c(\infty)} \otimes \pi^{\widehat{\kappa},\widehat{c}(\infty)}$, and the related means and variances.

Let $(N(\infty), Y(\infty), X(\infty), Z(\infty))$ be distributed according to the steady-state distribution. Proceeding as in Example 1, the coefficient of variation related to the activator satisfies

$$CV_N^2 = \frac{\text{Var}(N(\infty))}{\mathbb{E}(N(\infty))^2} = \frac{1}{\mathbb{E}(N(\infty))} + \frac{\tau_2}{\tau_1 + \tau_2} \frac{\text{Var}(Y(\infty))}{\mathbb{E}(Y(\infty))^2},$$

where

$$\mathbb{E}(N(\infty)) = \frac{\mu}{\nu} \frac{c(\infty)}{c(\infty) + \kappa}, \quad \tau_1 = \frac{1}{\nu}, \quad \text{and} \quad \tau_2 = \frac{1}{c(\infty) + \kappa}.$$

Similarly the CV of the transgene product is such that

$$CV_X^2 = \frac{\text{Var}(X(\infty))}{\mathbb{E}(X(\infty))^2} = \frac{1}{\mathbb{E}(X(\infty))} + \frac{\widehat{\tau}_2}{\widehat{\tau}_1 + \widehat{\tau}_2} \frac{\text{Var}(Z(\infty))}{\mathbb{E}(Z(\infty))^2},$$

where

$$\mathbb{E}(X(\infty)) = \frac{\hat{\mu}}{\hat{\nu}} \frac{\hat{c}(\infty)}{\hat{c}(\infty) + \hat{\kappa}},$$

$$\hat{\tau}_1 = \frac{1}{\hat{\nu}}, \quad \text{and} \quad \hat{\tau}_2 = \frac{1}{\hat{c}(\infty) + \hat{\kappa}}.$$

In what follows, we consider $c(\infty)$ and $\hat{c}(\infty)$.

5.4.1. $c(\infty)$ and $\hat{c}(\infty)$ for linear feedbacks

In the linear case, $c(t) = g(\mathbb{E}(N(t - \theta))) = g_0 + g_1 \mathbb{E}(N(t - \theta))$ and $\hat{c}(t) = h(\mathbb{E}(N(t - \theta))) = h_0 + h_1 \mathbb{E}(N(t - \hat{\theta}))$. We thus consider the averages

$$E(t) = \mathbb{E}(N(t)), \quad G(t) = \mathbb{E}(Y(t)) \quad \text{and}$$

$$\hat{E}(t) = \mathbb{E}(X(t)), \quad \hat{G}(t) = \mathbb{E}(Z(t)),$$

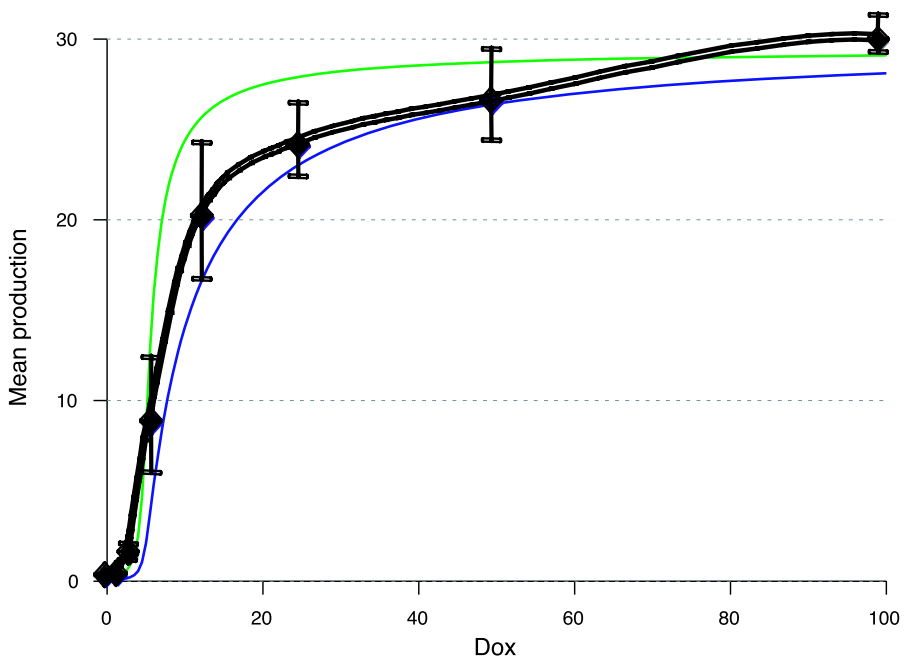


Fig. 4 Assay of the regulation of EGFP expression in living cells. The experimental curve obtained in Imhof et al. (2000) (in *black*) giving the average expression of therapeutic proteins as function of the number [Dox] of doxycycline molecules is compared to the curves obtained from the mean-field model, where the green and blue curves provide the mean expression levels of activator and therapeutic proteins, respectively. Since our approach is mainly qualitative, the scales are arbitrary and represent possible curves predicted by the model in a well chosen parameter regime. (Color figure online.)

which satisfy the system of delayed differential equations

$$\begin{aligned}\frac{dG(t)}{dt} &= (g_0 + g_1 E(t - \theta))(1 - G(t)) - \kappa G(t), & \frac{dE(t)}{dt} &= \mu G(t) - \nu E(t), \\ \frac{d\widehat{G}(t)}{dt} &= (h_0 + h_1 E(t - \widehat{\theta}))(1 - \widehat{G}(t)) - \kappa \widehat{G}(t), & \frac{d\widehat{E}(t)}{dt} &= \widehat{\mu} \widehat{G}(t) - \widehat{\nu} \widehat{E}(t).\end{aligned}$$

The results of Section 3.1 yield that the above delayed differential system has a globally asymptotically stable equilibrium point $(E(\infty), G(\infty), \widehat{E}(\infty), \widehat{G}(\infty))$, with

$$\begin{aligned}E(\infty) &= \frac{\mu}{\nu} G(\infty), & \left(g_0 + g_1 \frac{\mu}{\nu} G(\infty)\right)(1 - G(\infty)) &= \kappa G(\infty), \\ \left(h_0 + h_1 \frac{\mu}{\nu} G(\infty)\right)(1 - \widehat{G}(\infty)) &= \widehat{\kappa} \widehat{G}(\infty) & \text{and} & \widehat{E}(\infty) = \frac{\widehat{\mu}}{\widehat{\nu}} \widehat{G}(\infty).\end{aligned}$$

Finally

$$c(\infty) = g_0 + g_1 E(\infty) \quad \text{and} \quad \widehat{c}(\infty) = h_0 + h_1 \widehat{E}(\infty).$$

A qualitative comparison between the evolution of the mean protein production predicted by the model when varying the doxycycline level and experimental data is shown in Fig. 4.

6. Conclusion and discussion

In this work, we considered a class of self-regulated genes which are the building blocks of most of the existing gene networks. We provided efficient numerical algorithms for computing the steady-state distribution of the number of produced proteins. These results permit to handle more complex situations, and to understand the effect of positive or negative feedbacks in the network's dynamics. In a recent review paper, Karlebach and Shamir (2008) provides a schematic comparison of regulatory networks models. As compared to other methods, the single molecule level stochastic simulation has the highest level of detail, faithfulness to biological reality and ability to model dynamics, but on the other hand the amount of data needed for modeling is larger, the size of implemented models is smaller, the speed of analysis and the ability to perform inference are lower. Our exact solution to the chemical master equation permits to increase drastically the speed of computation and might also be useful in reverse engineering problems when inferring for example the parameters defining chemical reactions. Next, we considered in Sections 3 and 5 mean-field models with time delays which are of special interest in synthetic biology or in biotechnology, where small engineered regulatory networks are inserted at random in host genomes. Mathematical results in this setting are very scarce, and it is known that such systems can exhibit oscillations (see e.g. Bratsun et al., 2005 or Goutsias and Kim, 2006). Section 3 provides convergence results for mean-field models with time delays, which might open ways for handling more complex gene networks. Experimental results performed in living cells were in good agreement with our predictions. This shows that such models can provide the relevant information concerning complex systems, and that mathematical models can be efficiently used for the design of new regulatory gene networks in synthetic biology or in biotechnology.

Appendix A

A.1 Proof of Theorem 2

A sequence $\{P_n\}$ of probability distributions on a countable and discrete state space E converges to the probability distribution P on E if and only if each of its subsequences $\{P_{n'}\}$ contains a further subsequence $\{P_{n''}\}$ that converges to P . A family Π of probability distributions on E is relatively compact if every sequence of elements of Π contains a convergent subsequence (to a probability distribution on E), and tight if for every positive ϵ there exists a compact set K such that $P(K) > 1 - \epsilon$ for all P in Π . Tightness implies relative compactness, see e.g. Billingsley (1968).

We first show that the family of probability distributions $\Pi = (\pi^{(\Lambda)})_{\Lambda \in \mathbb{N}}$ is tight. Lemma 1 implies that there exists k not depending on Λ such that

$$\|v_n^{(\Lambda)}\| \leq \frac{1}{n!k^n}, \quad n \geq 1.$$

Since $V_\Lambda \geq 1$, we have also

$$\|\pi_n^{(\Lambda)}\| = \frac{\|v_n^{(\Lambda)}\|}{V_\Lambda} \leq \frac{1}{n!k^n}, \quad n \geq 1.$$

Hence, for all $\epsilon > 0$ there exists M_ϵ (not depending on Λ) such that

$$\sum_{j=M_\epsilon}^\infty \|\pi_j^{(\Lambda)}\| \leq \sum_{j=M_\epsilon}^\infty \frac{1}{j!k^j} < \epsilon.$$

Consequently, Π is relatively compact and there exists a convergent subsequence $(\pi^{(\Lambda_k)})_{k \in \mathbb{N}}$ of Π . Define $\pi^{(\infty)} = \lim_{k \rightarrow \infty} \pi^{(\Lambda_k)}$. We check now that $\pi^{(\infty)}$ is the invariant distribution of the continuous time process defined on the unbounded strip. For each $n \in \mathbb{N}$ we have

$$\begin{aligned} \pi_n^{(\infty)} &= \lim_{k \rightarrow \infty} \pi_n^{(\Lambda_k)} = \lim_{k \rightarrow \infty} (\pi_{n+1}^{(\Lambda_k)} D_{n+1} + \pi_n^{(\Lambda_k)} R_n + \pi_{n-1}^{(\Lambda_k)} U) \\ &= \pi_{n+1}^{(\infty)} D_{n+1} + \pi_n^{(\infty)} R_n + \pi_{n-1}^{(\infty)} U. \end{aligned}$$

This shows that $\pi^{(\infty)}$ is indeed an invariant distribution of the limit chain. In fact, what preceded is also valid for any converging subsequence. Besides, the invariant distribution is unique because the process is irreducible. Thus, we conclude that $\pi^{(\Lambda)}$ converges as $\Lambda \rightarrow \infty$ to $\pi^{(\infty)}$, which is the invariant distribution of the process defined on the unbounded strip. □

A.2 Fluctuation lemma

The following result is a slight modification of Lemma 4.2 in Hirsch et al. (1985):

Lemma 6. *Let $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ be bounded and differentiable, \dot{f} denoting its derivative. There exist increasing sequences $t_n \uparrow +\infty$ and $s_n \uparrow +\infty$, such that*

$$f(t_n) \rightarrow \bar{f}, \quad \dot{f}(t_n) \rightarrow 0, \quad \text{and} \quad f(s_n) \rightarrow \underline{f}, \quad \dot{f}(s_n) \rightarrow 0$$

as $n \rightarrow +\infty$, where for a function f we denote

$$\overline{f} := \limsup_{t \rightarrow +\infty} f(t), \quad \underline{f} := \liminf_{t \rightarrow +\infty} f(t).$$

A.3 Convergence of time-nonhomogeneous Markov chains

We consider a nonhomogeneous Markov chain $X(t)$ taking values in \mathbb{N} , of instantaneous transition matrix $Q_t = (q_t(i, j))_{i, j \in \mathbb{N}}$. The following Theorem is proved in Abramov and Liptser (2004).

Theorem 6. Assume that we can find non-negative constants $q(i, j)$ such that

$$\sum_{j \neq i} q(i, j) < +\infty, \quad \int_0^\infty \left(\sqrt{q_t(i, j)} - \sqrt{q(i, j)} \right)^2 dt < +\infty,$$

and

$$\int_{0 \leq s \leq t, q(i, j) > 0} q_s(i, j) ds = \int_0^t q_s(i, j) ds.$$

Let $Q_0 = (q(i, j))_{i, j \in \mathbb{N}}$, and let $X^0(t)$ be the related \mathbb{N} -valued Markov chain. Suppose that Q_0 is ergodic, that is that there is a unique probability measure π such that $\pi Q_0 = 0$ and

$$\lim_{t \rightarrow \infty} P(X^0(t) = j | X^0(s) = i) = \pi_j, \quad \forall s, i, j.$$

Then

$$\lim_{t \rightarrow \infty} P(X(t) = j | X(s) = i) = \pi_j, \quad \forall s, i, j.$$

A.4 Zeros of an exponential polynomial

We consider the exponential polynomial $H(z) = (z^2 + pz + q)e^z + r$, where p is real and positive, q is real and non-negative, and r is real. The following theorem is proved in Bellman and Cooke (1963), p. 449.

Theorem 7. Denote by a_k ($k \geq 0$) the sole root of the equation $\cot(a) = (a^2 - q)/p$ which lies on the interval $(k\pi, k\pi + \pi)$. We define the number w as follows:

1. if $r \geq 0$ and $p^2 \geq 2q$, $w = 1$;
2. if $r \geq 0$ and $p^2 < 2q$, w is the odd k for which a_k lies closest to $\sqrt{q - p^2/2}$;
3. if $r < 0$ and $p^2 \geq 2q$, $w = 2$;
4. if $r < 0$ and $p^2 < 2q$, w is the even k for which a_k lies closest to $\sqrt{q - p^2/2}$.

Then, a necessary and sufficient condition that all roots of $H(z) = 0$ lie to the left of the imaginary axis is that

1. $r \geq 0$ and $r \sin(a_w)/(pa_w) < 1$ or
2. $-q < r < 0$ and $r \sin(a_w)/(pa_w) < 1$.

References

- Abramov, V., Liptser, R., 2004. On existence of limiting distributions for time-nonhomogeneous countable Markov processes. *Queueing Syst.* 46, 353–361.
- Bellman, R., Cooke, K.L., 1963. *Differential-Difference Equations*. Academic Press, New York.
- Billingsley, P., 1968. *Convergence of Probability Measures*, vol. XII. New York, Wiley. 253 p.
- Bratsun, D., Volfson, D., Tsimring, L., Hasty, J., 2005. Delayed-induced stochastic oscillations in gene regulation. *Proc. Natl. Acad. Sci. U.S.A.* 102(41), 14593–14598.
- Burrage, K., Tian, T., Burrage, P., 2004. A multi-scale approach for simulating chemical reaction systems. *Prog. Biophys. Mol. Biol.* 85, 217.
- Cao, Y., Li, H., Petzold, L., 2004. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J. Chem. Phys.* 121, 4059–4067.
- Cao, Y., Gillespie, D., Petzold, L., 2005. The slow scale stochastic simulation algorithm. *J. Chem. Phys.* 122, 014116.
- Coppel, L.W.A., 1965. *Stability and Asymptotic Behavior in Differential Equations*. Heath Mathematical Monographs. Boston.
- Darvey, I.G., Ninham, B.W., Staff, P.J., 1966. Stochastic models for second-order chemical reaction kinetics. The equilibrium state. *J. Chem. Phys.* 45, 2145.
- Dill, K., Bromberg, S., 2003. *Molecular Driving Forces*. Garland, New York.
- Fournier, T., Gabriel, J.-P., Mazza, C., Pasquier, J., Galbete, J.L., Mermod, N., 2007. Steady-state expression of self-regulated genes. *Bioinformatics* 23(23), 3185–3192.
- Gabriel, J.-P., Hanisch, H., Hirsch, W.M., 1981. Dynamic equilibria of helminthic infections? In: Chapman, D.G., Gallucci, V.F. (Eds.), *Quantitative Population Dynamics*, vol. 13, pp. 83–104. International Cooperative Publishing House, Fairland. *Stat. Ecol. Ser.*
- Gadgil, C., Lee, C., Othmer, H., 2005. A stochastic analysis of first-order reaction networks. *Bull. Math. Biol.* 67, 901–946.
- Gillespie, D., 1977. Exact stochastic simulation of coupled chemical reactions. *J. Chem. Phys.* 81, 2340–2361.
- Gillespie, D., 2001. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* 115, 1716–1733.
- Goutsias, J., 2005. Quasi-equilibrium approximation of fast reactions kinetics in stochastic biochemical systems. *J. Chem. Phys.* 122, 184102.
- Goutsias, J., Kim, S., 2006. Stochastic transcriptional regulatory systems with time delays: a mean-field approximation. *J. Comput. Biol.* 13, 1049–1076.
- Guido, N., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C., Elston, T., Collins, J., 2006. A bottom-up approach to gene regulation. *Nature* 439, 856–860.
- Halanay, A., 1966. *Differential Equations: Stability, Oscillations, Time Lags*. Academic Press, New York.
- Hirsch, W.M., Hanisch, H., Gabriel, J.-P., 1985. Differential equation models of some parasitic infections: methods for the study of asymptotic behavior. *Commun. Pure Appl. Math.* 38, 733–753.
- Hornos, J., Schultz, D., Innocentini, G., Wang, J., Walczak, A., Onuchic, J., Wolynes, P., 2005. Self-regulating genes: An exact solution. *Phys. Rev. E* 72, 051907-1–051907-5.
- Imhof, M., Chatellard, P., Mermod, N., 2000. A regulatory network for the efficient control of transgene expression. *J. Gene Med.* 2, 107–116.
- Karlebach, G., Shamir, R., 2008. Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.* 9, 770–780.
- Kepler, T., Elston, T., 2001. Stochasticity in Transcriptional Regulation: Origins, Consequences and Mathematical Representations. *Biophys. J.* 81(1), 3116–3136.
- Lipan, O., Wong, W.H., 2005. The use of oscillatory signals in the study of genetic networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7063–7068.
- Nasell, I., Hirsch, W.M., 1972. A mathematical model of some helminthic infections. *Commun. Pure Appl. Math.* 25, 459–477.
- Nasell, I., Hirsch, W.M., 1973. The transmission dynamics of schistosomiasis. *Commun. Pure Appl. Math.* 26, 395–453.
- Randrianarivony, A., 1997. Fractions continues, q-nombres de Catalan et q-polynômes de Genocchi. *Eur. J. Comb.* 18, 75–92.
- Paulsson, J., 2005. Models of stochastic gene expression. *Phys. Life Rev.* 2, 157–175.
- Peccoud, J., Ycart, B., 1995. Markovian modelling of gene product synthesis. *Theor. Popul. Biol.* 48(2), 222–234.

- Pedraza, J., van Oudenaarden, A., 2005. Noise propagation in gene networks. *Science* 307, 1965–1969.
- Visco, P., Allen, R., Ewans, M., 2008a. Exact Solution of a model DNA-inversion switch with orientational control. *Phys. Rev. Lett.* 101, 118104.
- Visco, P., Allen, R., Ewans, M., 2008b. Statistical physics of a model binary switch with linear feedback. [arXiv:0812.3867](https://arxiv.org/abs/0812.3867).
- Zhang, Q., Yin, G., 1997. Structural properties of Markov chains with weak and strong interactions. *Stoch. Proc. Appl.* 70, 181–197.