

Accelerating Reactions at the DNA Can Slow Down Transient Gene Expression

Pavol Bokes³, Julia Klein¹, and Tatjana Petrov^{1,2}(✉)

¹ Department of Computer and Information Sciences, University of Konstanz,
Konstanz, Germany

tatjana.petrov@gmail.com

² Centre for the Advanced Study of Collective Behaviour, University of Konstanz,
Konstanz, Germany

³ Department of Applied Mathematics and Statistics, Comenius University,
Bratislava, Slovakia

Abstract. The expression of a gene is characterised by the upstream transcription factors and the biochemical reactions at the DNA processing them. Transient profile of gene expression then depends on the amount of involved transcription factors, and the scale of kinetic rates of regulatory reactions at the DNA. Due to the combinatorial explosion of the number of possible DNA configurations and uncertainty about the rates, a detailed mechanistic model is often difficult to analyse and even to write down. For this reason, modelling practice often abstracts away details such as the relative speed of rates of different reactions at the DNA, and how these reactions connect to one another. In this paper, we investigate how the transient gene expression depends on the topology and scale of the rates of reactions involving the DNA. We consider a generic example where a single protein is regulated through a number of arbitrarily connected DNA configurations, without feedback. In our first result, we analytically show that, if all switching rates are uniformly speeded up, then, as expected, the protein transient is faster and the noise is smaller. Our second result finds that, counter-intuitively, if all rates are fast but some more than others (two orders of magnitude vs. one order of magnitude), the opposite effect may emerge: time to equilibration is slower and protein noise increases. In particular, focusing on the case of a mechanism with four DNA states, we first illustrate the phenomenon numerically over concrete parameter instances. Then, we use singular perturbation analysis to systematically show that, in general, the fast chain with some rates even faster, reduces to a slow-switching chain. Our analysis has wide implications for quantitative modelling of

TP's research is supported by the Ministry of Science, Research and the Arts of the state of Baden-Württemberg, and the DFG Centre of Excellence 2117 'Centre for the Advanced Study of Collective Behaviour' (ID: 422037984), JK's research is supported by Committee on Research of Univ. of Konstanz (AFF), 2020/2021. PB is supported by the Slovak Research and Development Agency under the contract No. APVV-18-0308 and by the VEGA grant 1/0347/18. All authors would like to acknowledge Jacob Davidson and Stefano Tognazzi for useful discussions and feedback.

gene regulation: it emphasises the importance of accounting for the network topology of regulation among DNA states, and the importance of accounting for different magnitudes of respective reaction rates. We conclude the paper by discussing the results in context of modelling general collective behaviour.

1 Introduction

Gene regulation is one of the most fundamental processes in living systems. The experimental systems of *lac* operon in bacteria *E. coli* and the genetic switch of bacteriophage lambda virus allowed to unravel the basic molecular mechanisms of how a gene is turned on and off. These were followed by a molecular-level explanation of stochastic switching between lysis and lysogeny of phage [22], all the way to more complex logic gate formalisms that attempt to abstract more complex biological behaviour [6, 12, 21]. To date, synthetic biology has demonstrated remarkable success in engineering simple genetic circuits that are encoded in DNA and perform their function *in vivo*. However, significant conceptual challenges remain, related to the still unsatisfactory quantitative but also qualitative understanding of the underlying processes [19, 30]. Partly, this is due to the unknown or unspecified interactions in experiments *in vivo* (crosstalk, host-circuit interactions, loading effect). Another major challenge towards rational and rigorous design of synthetic circuits is computational modelling: gene regulation has a combinatorial number of functional entities, it is inherently stochastic, exhibits multiple time-scales, and experimentally measuring kinetic parameters/rates is often difficult, imprecise or impossible. In such context, predicting the transient profile - how gene expression *E* in a population of cells evolves over time - becomes a computationally expensive task. However, predicting how the transient phenotype emerges from the mechanistic, molecular interactions, is crucial both for engineering purposes of synthetic biology (e.g. when composing synthetic systems), as well as for addressing fundamental biological and evolutionary questions (e.g. for understanding whether the cell aims to create variability by modulating timing).

Mechanistically, the transcription of a single gene is initiated whenever a subunit of RNA polymerase binds to that gene's promoter region at the DNA [35]. While such binding can occur spontaneously, it is typically promoted or inhibited through other species involved in regulation, such as proteins and transcription factors (TFs). Consequently, the number of possible molecular configurations of the DNA grows combinatorially with the number of operator sites regulating the gene in question. For instance, one hypothesised mechanism in lambda-phage, containing only three left and three right operators, leads to 1200 different DNA configurations [31]. The combinatorial explosion of the number of possible configurations makes the model tedious to even write down, let alone execute and make predictions about it. The induced stochastic process enumerates states which couple the configuration of the DNA, with the copy number of the protein, and possibly other species involved in regulation, such as mRNA and transcription factors. In order to faithfully predict the stochastic evolution of the

gene product (protein) over time, the modelled system can be solved numerically, by integrating the Master equation of the stochastic process. This is often prohibitive in practice, due to large dimensionality and a combinatorial number of reachable states. For this reason, modelling practice often abstracts away details and adds assumptions. One popular approach is simulating the system by Gillespie simulation [7] and statistically inferring the protein expression profile, hence trading off accuracy and precision. Other approaches are based on mean-field approximations (e.g. deterministic limit [17] and linear-noise approximation [5]), significantly reducing the computational effort. However, mean-field models do not capture the inherent stochasticity, which is especially prominent in gene regulation. Further model reduction ideas exploit multi-scaleness of the system: fast subsystems are identified (possibly dynamically), and assumed to be reaching an equilibrium fast, relative to the observable dynamics [2, 13, 29, 34]. A special class of reductions based on steady-state assumption is the experimentalists' favourite approach of *statistical thermodynamics limit*. This widely and successfully used method (e.g. [3, 24, 33]) estimates the probability of being in any of the possible DNA binding configurations from their relative binding energies (Boltzmann weights) and the protein concentrations, both of which can often be experimentally accessed. The statistical thermodynamics limit model is rooted in the argument that, when the switching rates among DNA configurations are fast, the probability distribution over the configurations is rapidly arriving at its stationary distribution. While this model takes into account the stochasticity inherent to the DNA binding configurations, it neglects the transient probabilities in the DNA switching, before the equilibrium is reached. It abstracts away the relative speed of rates of different reactions at the DNA, and how they connect to one another. The question arises: how does the transient gene expression - its shape and duration - depend on the topology and scale of the rates of reactions involving the DNA? Is it justifiable, in this context, to consider sufficiently fast propensities as an argument for applying a (quasi)-steady-state assumption?

In this paper, we investigate how the transient gene expression depends on the topology and scale of the rates of reactions involving the DNA. In Sect. 2, we introduce reaction networks, a stochastic process assigned to it, and the equations for the transient dynamics. In Sect. 3, we describe a generic example where a single protein is regulated through a number of arbitrarily connected DNA configurations, without feedback. This means that any transition between two states of the network is possible. In our first result, we analytically show that, if all switching rates are uniformly speeded up, then, as expected, the protein transient is faster and the noise is smaller. In Sect. 3.1, we introduce concrete parameter instances to illustrate the phenomenon numerically. Then, in Sect. 4, we present our main result: counter-intuitively, if all rates are fast but some more than others (two orders of magnitude vs. one order of magnitude), the opposite effect may emerge: time to equilibration is slower and protein noise increases. We use singular perturbation analysis to systematically show that, in general, the fast chain with some rates even faster, exactly reduces to a slow-switching

chain. We conclude the paper by discussing the implications of our results in Sect. 5.

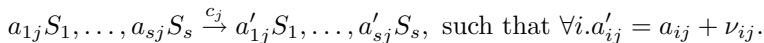
1.1 Related Works

Timing aspects of gene regulation are gaining increasing attention, such as explicitly modelling delays in gene expression [28], showcasing dramatic phenotypic consequences of small delays in the arrival of different TFs [11], resolving the temporal dynamics of gene regulatory networks from time-series data [10], as well as the study of transient hysteresis and inherent stochasticity in gene regulatory networks [27]. Following the early works on examining the relation between topology and relaxation to steady states of reaction networks [9], stochastic gene expression from a promoter model has been studied for multiple states [14]. Singular perturbation analysis has been used for lumping states of Markov chains arising in biological applications [4, 34]. To the best of our knowledge, none of these works showcases the phenomenon of obtaining slower dynamics through faster rates, or, more specifically, slowing down gene expression by speeding up the reactions at the DNA.

2 Preliminaries

The default rate of gene expression, also referred to as the basal rate, can be modified by the presence of transcriptional activators and repressors. *Activators* are transcription factors (TFs) that bind to specific locations on the DNA, or to other TFs, and enhance the expression of a gene by promoting the binding of RNAP. *Repressors* reduce the expression of gene g , by directly blocking the binding of RNAP, or indirectly, by inhibiting the activators, or promoting direct repressors. The mechanism of how and at which rates the molecular species are interacting is transparently written in a list of reactions. Reactions are equipped with the stochastic semantics which is valid under mild assumptions [7]. In the following, we will model gene regulatory mechanisms with the standard Chemical Reaction Network formalism (CRN).

Definition 1. A reaction system is a pair (S, R) , such that $S = \{S_1, \dots, S_s\}$ is a finite set of species, and $R = \{r_1, \dots, r_r\}$ is a finite set of reactions. The state of a system can be represented as a multi-set of species, denoted by $\mathbf{x} = (x_1, \dots, x_s) \in \mathbb{N}^s$. Each reaction is a triple $r_j \equiv (\mathbf{a}_j, \boldsymbol{\nu}_j, c_j) \in \mathbb{N}^s \times \mathbb{N}^s \times \mathbb{R}_{\geq 0}$, written down in the following form:



The vectors \mathbf{a}_j and \mathbf{a}'_j are often called respectively the *consumption* and *production* vectors due to j th reaction, and c_j is the respective *kinetic rate*. If the j th reaction occurs, after being in state \mathbf{x} , the next state will be $\mathbf{x}' = \mathbf{x} + \boldsymbol{\nu}_j$. This will be possible only if $x_i \geq a_{ij}$ for $i = 1, \dots, s$.

Stochastic Semantics. The species' multiplicities follow a continuous-time Markov chain (CTMC) $\{X(t)\}_{t \geq 0}$, defined over the state space $S = \{\mathbf{x} \mid \mathbf{x}$ is reachable from \mathbf{x}_0 by a finite sequence of reactions from $\{r_1, \dots, r_r\}\}$. In other words, the probability of moving to the state $\mathbf{x} + \boldsymbol{\nu}_j$ from \mathbf{x} after time Δ is

$$P(X(t + \Delta) = \mathbf{x} + \boldsymbol{\nu}_j \mid X(t) = \mathbf{x}) = \lambda_j(\mathbf{x})\Delta + o(\Delta),$$

with λ_j the propensity of j th reaction, assumed to follow the principle of mass-action: $\lambda_j(\mathbf{x}) = c_j \prod_{i=1}^s \binom{x_i}{a_{ij}}$. The binomial coefficient $\binom{x_i}{a_{ij}}$ reflects the probability of choosing a_{ij} molecules of species S_i out of x_i available ones.

Computing the Transient. Using the vector notation $\mathbf{X}(t) \in \mathbb{N}^n$ for the marginal of process $\{X(t)\}_{t \geq 0}$ at time t , we can compute this transient distribution by integrating the *chemical master equation* (CME). Denoting by $p_{\mathbf{x}}(t) := P(\mathbf{X}(t) = \mathbf{x})$, the CME for state $\mathbf{x} \in \mathbb{N}^s$ reads

$$\frac{d}{dt} p_{\mathbf{x}}(t) = \sum_{j=1, \mathbf{x} - \boldsymbol{\nu}_j \in S}^r \lambda_j(\mathbf{x} - \boldsymbol{\nu}_j) p_{(\mathbf{x} - \boldsymbol{\nu}_j)}(t) - \sum_{j=1}^r \lambda_j(\mathbf{x}) p_{\mathbf{x}}(t). \quad (1)$$

The solution may be obtained by solving the system of differential equations, but, due to its high (possibly infinite) dimensionality, it is often statistically estimated by simulating the traces of $\{X_t\}$, known as the stochastic simulation algorithm (SSA) in chemical literature [7]. As the statistical estimation often remains computationally expensive for desired accuracy, for the case when the deterministic model is unsatisfactory due to the low multiplicities of many molecular species [18], different further approximation methods have been proposed, major challenge to which remains the quantification of approximation accuracy (see [32] and references therein for a thorough review on the subject).

3 Moment Calculations

We consider a generic example with m different DNA states regulating a single protein, without feedback. The configurations of the DNA are indexed by $1, 2 \dots m$, and we denote the transition rates between them (reaction propensities) by q_{ij} (we additionally define $q_{ii} = -\sum_{j=1}^m q_{ij}$). We assume that the gene chain is irreducible, justified by the reversibility of all reactions at the DNA. The dynamics of the protein copy number is modelled as usually by a birth-death process (with gene-state-dependent birth rate k_i and linear death rate δ per protein). The respective reaction system is schemed in Table 1, left.

The underlying stochastic process $\{X(t)\}$ takes values in the state space $S \subseteq \mathbb{N}^{m+1}$, such that the first m components represent the DNA states, and the last one is the protein count.

In the following, we will use notation $\mathbf{X}_{1:m}(t) \in \{0, 1\}^m$, to denote the projection of the marginal process at time t , to the DNA-regulatory elements, and, for better readability, we introduce $N(t) := \mathbf{X}_{m+1}(t)$ to denote the protein count at time t .

In total, since there is exactly one copy of the DNA, any state in S can be seen as a gene state coupled with the protein copy number, i.e. $S \cong \{1, 2, \dots, m\} \times \mathbb{N}$. We introduce short-hand notation $s_{(i,n)}$ for state $\mathbf{x} = \underbrace{(0, \dots, 1, \dots, 0)}_i, n \in S$.

Allowable transitions and their rates are summarised in Table 1, right.

Table 1. Two equivalent formulations of a multi-state gene expression model. *Left:* a reaction system with $m + 1$ reaction species S_1, \dots, S_m (gene states) and $S_{m+1} = P$ (protein) with copy numbers $X_1(t), \dots, X_m(t)$ and $X_{m+1}(t) = N(t)$, whereby $X_1(t) + \dots + X_m(t) = 1$ holds initially (and throughout time). *Right:* a two-component Markov chain in which the first component indexes the gene state and the second component gives the protein copy number.

Reaction	Rate	Reset map	Transition	Rate
$S_i \rightarrow S_j$	$q_{ij}X_i$	$X_i \rightarrow X_i - 1$ $X_j \rightarrow X_j + 1$	$(i, n) \rightarrow (j, n)$	q_{ij}
$S_i \rightarrow S_i + P$	k_iX_i	$N \rightarrow N + 1$	$(i, n) \rightarrow (i, n + 1)$	k_i
$P \rightarrow \emptyset$	δN	$N \rightarrow N - 1$	$(i, n) \rightarrow (i, n - 1)$	δn

We arrange the probabilities $p_{n,i}(t) := P(\mathbf{X}(t) = s_{(i,n)})$ of being in gene state i and having n protein into a column vector

$$\mathbf{p}_n(t) = (p_{n,1}(t), \dots, p_{n,m}(t))^T.$$

The probability vector satisfies a system of difference–differential equations

$$\frac{d\mathbf{p}_n}{dt} = \mathbf{A}\mathbf{p}_n + \mathbf{A}_k(\mathbf{p}_{n-1} - \mathbf{p}_n) + \delta((n+1)\mathbf{p}_{n+1} - n\mathbf{p}_n), \quad (2)$$

where $\mathbf{A} = \mathbf{Q}^\top$ is the Markovian generator matrix and \mathbf{A}_k is a diagonal square matrix with the elements of the vector $\mathbf{k} = (k_1, \dots, k_m)^T$ placed on the main diagonal. We study (2) subject to the initial condition

$$\mathbf{p}_n(0) = \delta_{n,n_0} \mathbf{e}_{j_0} \quad (3)$$

in which n_0 is the initial protein copy number, j_0 is the initial gene state, δ_{n,n_0} represents the Kronecker delta, \mathbf{e}_{j_0} is the j_0 -th element of the standard basis in the m -dimensional Euclidean space.

Let us introduce the variables

$$\mathbf{p}(t) = \sum_{n=0}^{\infty} \mathbf{p}_n(t), \quad \langle N(t) \rangle = \sum_{n=0}^{\infty} n \mathbf{1}^T \mathbf{p}_n(t), \quad (4)$$

$$\mathbf{f}(t) = \left(\sum_{n=0}^{\infty} n \mathbf{p}_n(t) \right) - \langle N(t) \rangle \mathbf{p}(t), \quad (5)$$

$$\sigma^2(t) = \left(\sum_{n=0}^{\infty} n^2 \mathbf{1}^T \mathbf{p}_n(t) \right) - \langle N(t) \rangle^2. \quad (6)$$

Note that $\mathbf{1}^T \mathbf{p}_n(t) = p_{n,1}(t) + \dots + p_{n,m}(t)$, where $\mathbf{1}^T = (1, \dots, 1)$ is the m -dimensional row vector of ones, gives the marginal protein probability mass function. It is instructive to interpret the variables $\mathbf{p}(t)$ and $\mathbf{f}(t)$ from the standpoint of the reaction-network formulation of the model (Table 1, left). The elements of the copy-number vector $\mathbf{X}_{1:m}(t)^\top$ of gene states can be zero or one, with exactly one of them being equal to one; the gene-state copy-number statistics can be expressed in terms of (4)–(6) as

$$\begin{aligned} \langle \mathbf{X}_{1:m}(t) \rangle &= \mathbf{p}(t), \\ \langle N(t) \mathbf{X}_{1:m}(t) \rangle - \langle N(t) \rangle \langle \mathbf{X}_{1:m}(t) \rangle &= \mathbf{f}(t), \\ \langle \mathbf{X}_{1:m}(t) \mathbf{X}_{1:m}(t)^\top \rangle - \langle \mathbf{X}_{1:m}(t) \rangle \langle \mathbf{X}_{1:m}(t) \rangle^\top &= \boldsymbol{\Sigma}(t) = \mathbf{A}_{\mathbf{p}(t)} - \mathbf{p}(t) \mathbf{p}(t)^\top. \end{aligned}$$

The variables (4)–(6) thus fully describe the mean and covariance of the reaction system in Tabl 1, right. In particular, $\mathbf{f}(t)$ is the covariance between the gene state and protein copy number; $\boldsymbol{\Sigma}(t)$ is the covariance matrix of the gene state (with itself).

The variables (4)–(6) satisfy a system of differential equations (see Appendix B for derivation)

$$\frac{d\mathbf{p}}{dt} = \mathbf{A}\mathbf{p}, \quad \frac{d\langle N(t) \rangle}{dt} = \mathbf{k}^T \mathbf{p} - \delta \langle N(t) \rangle, \quad (7)$$

$$\frac{d\mathbf{f}}{dt} = \mathbf{A}\mathbf{f} - \delta \mathbf{f} + \boldsymbol{\Sigma}\mathbf{k}, \quad \text{where } \boldsymbol{\Sigma} = \mathbf{A}_{\mathbf{p}} - \mathbf{p}\mathbf{p}^T, \quad (8)$$

$$\frac{d\sigma^2}{dt} = 2\mathbf{k}^T \mathbf{f} + \mathbf{k}^T \mathbf{p} + \delta \langle N(t) \rangle - 2\delta\sigma^2 \quad (9)$$

subject to initial conditions

$$\mathbf{p}(0) = \mathbf{e}_{j_0}, \quad \langle N(0) \rangle = n_0, \quad \mathbf{f}(0) = \mathbf{0}, \quad \sigma^2(0) = 0, \quad (10)$$

where \mathbf{e}_{j_0} is the j_0 -th element of the standard basis in m -dimensional Euclidean space.

Equating the derivatives in (7)–(9) to zero, we obtain steady state protein mean and Fano factor in the form

$$\langle N(t) \rangle = \frac{\mathbf{k}^T \bar{\mathbf{p}}}{\delta}, \quad \frac{\sigma^2}{\langle N(t) \rangle} = 1 + \frac{\mathbf{k}^T \bar{\mathbf{f}}}{\mathbf{k}^T \bar{\mathbf{p}}}, \quad (11)$$

where $\bar{\mathbf{p}}$ and $\bar{\mathbf{f}}$ satisfy algebraic equations

$$\mathbf{A}\bar{\mathbf{p}} = 0, \quad (\mathbf{A} - \delta)\bar{\mathbf{f}} + \boldsymbol{\Sigma}\mathbf{k} = 0. \quad (12)$$

We note that the solution $\bar{\mathbf{f}}$ to (12) can be represented as

$$\bar{\mathbf{f}} = \int_0^\infty e^{-\delta t} (e^{t\mathbf{A}} - \bar{\mathbf{p}}\mathbf{1}^T) dt \boldsymbol{\Sigma}\mathbf{k}. \quad (13)$$

Equation (13) connects the magnitude of $\bar{\mathbf{f}}$ to the equilibration timescale of the gene-state Markov chain (note that $\bar{\mathbf{p}}\mathbf{1}^\top = \lim_{t \rightarrow \infty} e^{t\mathbf{A}}$). Specifically, if $\mathbf{A} = \tilde{\mathbf{A}}/\varepsilon$,

where $\varepsilon \ll 1$, i.e. the gene transition rates are $O(1/\varepsilon)$ large, then substituting $t = \varepsilon s$ into (13) implies that

$$\bar{\mathbf{f}} = \varepsilon \int_0^\infty \left(e^{s\tilde{\mathbf{A}}} - \bar{\mathbf{p}}\mathbf{1}^T \right) ds \Sigma \mathbf{k} + O(\varepsilon^2),$$

which is $O(\varepsilon)$ small. Correspondingly, the (steady-state) protein Fano factor (11) will differ from the Poissonian value of 1 by an $O(\varepsilon)$ quantity. This concludes the argument that, in agreement with intuition, if all rates are faster by an order of magnitude (ε^{-1}), then, as expected, the magnitude of equilibration time-scale of the whole chain scales down with the same factor. The fast fluctuations of the gene chain are thereby averaged out at the downstream level of the protein.

However, as will be demonstrated in the next section with a specific example of a four state chain, the largeness of transition rates does not guarantee, on its own, fast equilibration, and the connectivity of the chain can play a crucial role. Indeed, we will show that one can slow down equilibration (and hence increase protein noise) by increasing some of the transition rates.

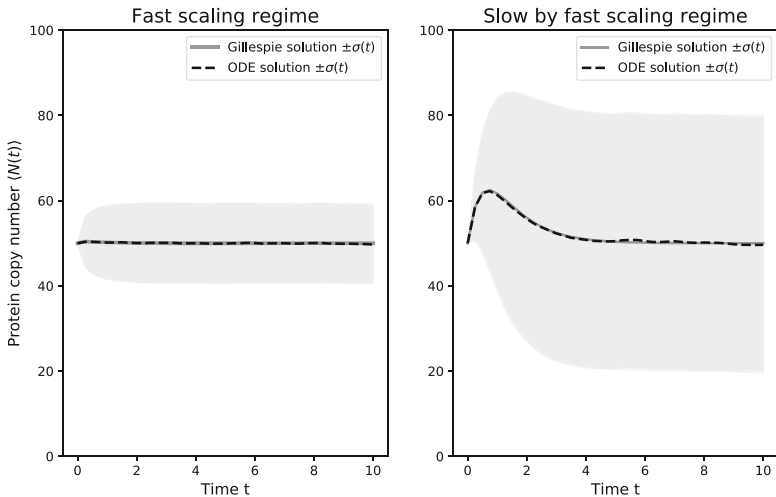


Fig. 1. Dependence of $\langle N(t) \rangle \pm \sigma(t)$ on t for a value of $\varepsilon = 0.01$, using two regimes. Left: “fast” scaling regime where all transition rates are $O(1/\varepsilon)$. Right: “slow by fast” scaling regime, where the backward rates are speeded up to $O(1/\varepsilon^2)$. ODE results (dashed line) are cross-validated by Gillespie simulations (solid line). The chain is initially at state $j_0 = 2$, and the amount of protein is set to $n_0 = 50$. The transition matrix parameters are set to $\tilde{a}_g = \tilde{b}_g = \tilde{a}_r = \tilde{b}_r = \tilde{a}_b = \tilde{b}_b = 1$.

3.1 Four-State Chain

We specifically focus on a case with four gene states, with transition matrix

$$\mathbf{A} = \begin{pmatrix} -a_r & a_b & 0 & 0 \\ a_r & -a_b - a_g & b_g & 0 \\ 0 & a_g & -b_g - b_b & b_r \\ 0 & 0 & b_b & -b_r \end{pmatrix} \quad (14)$$

(recall that the matrix \mathbf{A} is shown as a transpose of the graph of connections, the respective graph is depicted in Fig. 2, left). We investigate two alternative, different scaling regimes with respect to a small dimensionless parameter ε :

Fast. We assume that all transition rates are $O(1/\varepsilon)$, i.e.

$$a_g = \frac{\tilde{a}_g}{\varepsilon}, \quad a_b = \frac{\tilde{a}_b}{\varepsilon}, \quad a_r = \frac{\tilde{a}_r}{\varepsilon}, \quad b_g = \frac{\tilde{b}_g}{\varepsilon}, \quad b_b = \frac{\tilde{b}_b}{\varepsilon}, \quad b_r = \frac{\tilde{b}_r}{\varepsilon},$$

where $\tilde{a}_g, \tilde{a}_b, \tilde{a}_r, \tilde{b}_g, \tilde{b}_b,$ and \tilde{b}_r are $O(1)$.

Slow by fast. We speed up the backward rates by making them $O(1/\varepsilon^2)$, i.e.

$$a_g = \frac{\tilde{a}_g}{\varepsilon}, \quad a_b = \frac{\tilde{a}_b}{\varepsilon^2}, \quad a_r = \frac{\tilde{a}_r}{\varepsilon}, \quad b_g = \frac{\tilde{b}_g}{\varepsilon}, \quad b_b = \frac{\tilde{b}_b}{\varepsilon^2}, \quad b_r = \frac{\tilde{b}_r}{\varepsilon}. \quad (15)$$

where $\tilde{a}_g, \tilde{a}_b, \tilde{a}_r, \tilde{b}_g, \tilde{b}_b,$ and \tilde{b}_r are $O(1)$.

We first numerically analyse the transient protein dynamics for these two scaling scenarios. In Fig. 1, we plot the average protein count and the standard deviation. Increasing the speed of rates from states 2 (resp. 3) to state 1 (resp. 4) not only does not increase the scale and decrease the protein noise, but significantly slows down the protein dynamics and increases noise. In the next section, we systematically derive that the case ‘slow by fast’ regime is approximated by a slow-switching 2-state chain (shown in Fig. 2, right).

4 Singular-Perturbation Analysis of the Slow-by-fast Regime

The probability dynamics generated by the transition matrix (14) in the slow-by-fast scaling regime (15) is given by a system of four differential equations

$$\varepsilon^2 \frac{dp_1}{dt} = \tilde{a}_b p_2 - \varepsilon \tilde{a}_r p_1, \quad (16)$$

$$\varepsilon^2 \frac{dp_2}{dt} = \varepsilon \tilde{a}_r p_1 - \varepsilon \tilde{a}_g p_2 - \tilde{a}_b p_2 + \varepsilon \tilde{b}_g p_3, \quad (17)$$

$$\varepsilon^2 \frac{dp_3}{dt} = \varepsilon \tilde{a}_g p_2 - \varepsilon \tilde{b}_g p_3 - \tilde{b}_b p_3 + \varepsilon \tilde{b}_r p_4, \quad (18)$$

$$\varepsilon^2 \frac{dp_4}{dt} = \tilde{b}_b p_3 - \varepsilon \tilde{b}_r p_4. \quad (19)$$

Equations such as (16)–(19) whose right-hand sides depend on a small parameter ε are referred to as perturbation problems. Additionally, problems in which, like in (16)–(19), the small parameter multiplies one or more derivatives on the left-hand side, are classified as singularly perturbed [16, 23]. We study solutions to system (16)–(19) that satisfy an initial condition

$$p_i(0) = p_i^{\text{init}}, \quad i = 1, 2, 3, 4, \quad (20)$$

where the right-hand side of (20) is a prescribed probability distribution. The aim of what follows is to characterise the behaviour as $\varepsilon \rightarrow 0$ of the solution to (16)–(19) subject to (20).

We look for a solution to (16)–(19) in the form of a regular power series

$$p_i(t; \varepsilon) = p_i^{(0)}(t) + \varepsilon p_i^{(1)}(t) + O(\varepsilon^2), \quad i = 1, 2, 3, 4. \quad (21)$$

Inserting (21) into (16) and (19) and collecting terms of same order yields

$$O(1) : \quad p_2^{(0)} = p_3^{(0)} = 0, \quad (22)$$

$$O(\varepsilon) : \quad \tilde{a}_b p_2^{(1)} = \tilde{a}_r p_1^{(0)}, \quad (23)$$

$$\tilde{b}_b p_3^{(1)} = \tilde{b}_r p_4^{(0)}. \quad (24)$$

Equations (22) imply that the probability of states 2 or 3 is $O(\varepsilon)$ -small. Equation (23) means that, at the leading order, the probability of state 2 is proportional to that of state 1; equation (24) establishes the analogous for states 3 and 4. Adding (16) to (17), and (18) to (19), yield

$$\varepsilon \frac{d}{dt} (p_1 + p_2) = -\tilde{a}_g p_2 + \tilde{b}_g p_3, \quad (25)$$

$$\varepsilon \frac{d}{dt} (p_3 + p_4) = \tilde{a}_g p_2 - \tilde{b}_g p_3. \quad (26)$$

Inserting (21) into (25)–(26) and collecting $O(\varepsilon)$ terms gives

$$\frac{d}{dt} (p_1^{(0)} + p_2^{(0)}) = -\tilde{a}_g p_2^{(1)} + \tilde{b}_g p_3^{(1)}, \quad (27)$$

$$\frac{d}{dt} (p_3^{(0)} + p_4^{(0)}) = \tilde{a}_g p_2^{(1)} - \tilde{b}_g p_3^{(1)}. \quad (28)$$

Inserting (22)–(24) into (27)–(28) yields

$$\frac{dp_1^{(0)}}{dt} = -\tilde{a}_g \tilde{a}_r \tilde{a}_b^{-1} p_1^{(0)} + \tilde{b}_g \tilde{b}_r \tilde{b}_b^{-1} p_4^{(0)}, \quad (29)$$

$$\frac{dp_4^{(0)}}{dt} = \tilde{a}_g \tilde{a}_r \tilde{a}_b^{-1} p_1^{(0)} - \tilde{b}_g \tilde{b}_r \tilde{b}_b^{-1} p_4^{(0)}. \quad (30)$$

Equations (29)–(30) describe the probability dynamics of a two-state (or random-telegraph) chain with states 1 and 4 and transition rates $\tilde{a}_g \tilde{a}_r / \tilde{a}_b$ and $\tilde{b}_g \tilde{b}_r / \tilde{b}_b$ between them. Intriguingly, the emergent dynamics of (29)–(30) occurs on the $t = O(1)$ scale although the original system (16)–(19) featured only $O(1/\varepsilon)$ rates (or faster).

4.1 Inner Solution and Matching

In singular-perturbation studies, the leading-order term of a regular solution (21) is referred to as the outer solution [15, 23]. As is typical in singularly perturbed problems, the outer solution satisfies a system, here (29)–(30), that is lower-dimensional than the original system (16)–(19); the remaining components of the outer solution are trivially given by (22). Therefore, the initial condition (20) cannot be immediately imposed on the outer solution. In order to formulate an appropriate initial condition for (29)–(30), we need to study (16)–(19) on the fast timescale, find the so-called inner solution, and use an asymptotic matching principle [15, 23] to connect the two asymptotic solutions together.

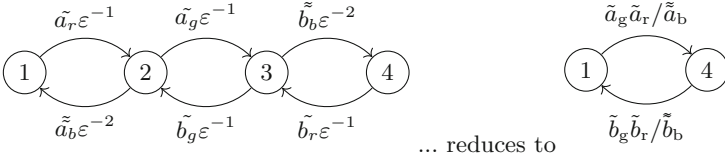


Fig. 2. (left) Four-state chain at the DNA: all rates are faster than $O(1)$: some by order or magnitude ε^{-1} , and some even faster, by order of magnitude ε^{-2} . (left) The emergent dynamics is approximated by a two-state chain. Intriguingly, the emergent dynamics of occurs on the $t = O(1)$ scale although the original system featured only $O(1/\varepsilon)$ rates and faster.

In order to construct the inner solution, we focus on the fast dynamics of system (16)–(19) by means of a transformation

$$t = \varepsilon^2 T, \quad p_i(t) = P_i(T). \quad (31)$$

Inserting (31) into (16)–(19) yields a time-rescaled system

$$\frac{dP_1}{dT} = \tilde{a}_b P_2 - \varepsilon \tilde{a}_r P_1, \quad (32)$$

$$\frac{dP_2}{dT} = \varepsilon \tilde{a}_r P_1 - \varepsilon \tilde{a}_g P_2 - \tilde{a}_b P_2 + \varepsilon \tilde{b}_g P_3, \quad (33)$$

$$\frac{dP_3}{dT} = \varepsilon \tilde{a}_g P_2 - \varepsilon \tilde{b}_g P_3 - \tilde{b}_b P_3 + \varepsilon \tilde{b}_r P_4, \quad (34)$$

$$\frac{dP_4}{dT} = \tilde{b}_b P_3 - \varepsilon \tilde{b}_r P_4. \quad (35)$$

Note that in the time-rescaled system (32)–(35), the time derivative is no longer multiplied by a small parameter.

$$P_i(T; \varepsilon) = P_i^{(0)}(T) + \varepsilon P_i^{(1)}(T) + O(\varepsilon^2), \quad i = 1, 2, 3, 4. \quad (36)$$

into (32)–(35) and collecting the $O(1)$ terms yield

$$\frac{dP_1^{(0)}}{dt} = -\frac{dP_2^{(0)}}{dt} = \tilde{a}_b P_2^{(0)}, \quad \frac{dP_3^{(0)}}{dt} = -\frac{dP_4^{(0)}}{dt} = -\tilde{b}_b P_3^{(0)}. \quad (37)$$

Since the reduced problem (37) retains the dimensionality of the original problem (32)–(35), we can solve it subject to the same initial condition

$$P_i^{(0)}(0) = p_i^{\text{init}}, \quad (38)$$

which yields

$$P_1^{(0)}(T) = p_1^{\text{init}} + p_2^{\text{init}}(1 - e^{-\tilde{a}_b T}), \quad P_2^{(0)}(T) = p_2^{\text{init}} e^{-\tilde{a}_b T}, \quad (39)$$

$$P_3^{(0)}(T) = p_3^{\text{init}} e^{-\tilde{b}_b T}, \quad P_4^{(0)}(T) = p_4^{\text{init}} + p_3^{\text{init}}(1 - e^{-\tilde{b}_b T}). \quad (40)$$

Thus, on the inner timescale, there occurs a fast transfer of probability mass from the states 2 and 3 into the states 1 and 4, respectively.

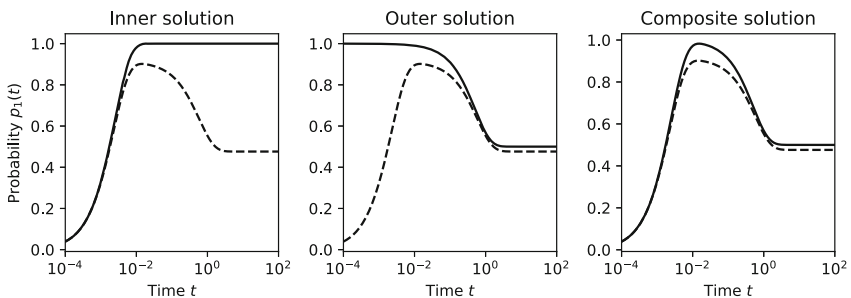


Fig. 3. The inner, outer, and composite approximations (solid curves) to the first component of the exact solution (dashed curve) to (16)–(20) (time is shown at logarithmic scale). The timescale separation parameter is set to $\varepsilon = 0.01$. The chain is initially at state 2, i.e. $p_2^{\text{init}} = 1$, $p_1^{\text{init}} = p_3^{\text{init}} = p_4^{\text{init}} = 0$. The transition matrix parameters are set to $\tilde{a}_g = \tilde{b}_g = \tilde{a}_r = \tilde{b}_r = \tilde{a}_b = \tilde{b}_b = 1$.

According to the asymptotic matching principle [15, 23], the large- T behaviour of the inner and the small- t behaviour of the outer solution overlap, i.e.

$$p_1^{(0)}(0) = P_1^{(0)}(\infty) = p_1^{\text{init}} + p_2^{\text{init}}, \quad (41)$$

$$p_4^{(0)}(0) = P_4^{(0)}(\infty) = p_3^{\text{init}} + p_4^{\text{init}}. \quad (42)$$

Equations (41)–(42) establish the relationship between the original initial condition (20) and the initial condition that needs to be imposed for the outer solution; solving (29)–(30) subject to (41)–(42) yields

$$p_1^{(0)}(t) = \frac{\frac{\tilde{b}_g \tilde{b}_r}{\tilde{b}_b}}{\frac{\tilde{a}_g \tilde{a}_r}{\tilde{a}_b} + \frac{\tilde{b}_g \tilde{b}_r}{\tilde{b}_b}} + \left(p_1^{\text{init}} + p_2^{\text{init}} - \frac{\tilde{b}_g \tilde{b}_r}{\tilde{b}_b} \right) e^{-\left(\frac{\tilde{a}_g \tilde{a}_r}{\tilde{a}_b} + \frac{\tilde{b}_g \tilde{b}_r}{\tilde{b}_b} \right) t}, \quad (43)$$

$$p_4^{(0)}(t) = \frac{\frac{\tilde{a}_g \tilde{a}_r}{a_b}}{\frac{\tilde{a}_g \tilde{a}_r}{a_b} + \frac{\tilde{b}_g \tilde{b}_r}{b_b}} + \left(p_3^{\text{init}} + p_4^{\text{init}} - \frac{\frac{\tilde{a}_g \tilde{a}_r}{a_b}}{\frac{\tilde{a}_g \tilde{a}_r}{a_b} + \frac{\tilde{b}_g \tilde{b}_r}{b_b}} \right) e^{-\left(\frac{\tilde{a}_g \tilde{a}_r}{a_b} + \frac{\tilde{b}_g \tilde{b}_r}{b_b} \right) t}. \quad (44)$$

We note that the second and third components of the outer solution are trivially given by $p_2^{(0)}(t) = p_3^{(0)}(t) = 0$ by (22). The outer solution (43)–(44) provides a close approximation to the original solution for $t = O(1)$ but fails to capture the behaviour of the initial transient; the inner solution (39)–(40) provides a close approximation for $T = O(1)$, i.e. $t = O(\varepsilon^2)$, but disregards the outer dynamics. A uniformly valid composite solution can be constructed by adding the inner and outer solutions up, and subtracting the matched value, i.e.

$$p_1^{\text{comp}}(t) = p_1^{(0)}(t) - p_2^{\text{init}} e^{-\tilde{a}_b t / \varepsilon^2}, \quad p_2^{\text{comp}}(t) = p_2^{\text{init}} e^{-\tilde{a}_b t / \varepsilon^2}, \quad (45)$$

$$p_3^{\text{comp}}(t) = p_3^{\text{init}} e^{-\tilde{b}_b t / \varepsilon^2}, \quad p_4^{\text{comp}}(t) = p_4^{(0)}(t) - p_3^{\text{init}} e^{-\tilde{b}_b t / \varepsilon^2}. \quad (46)$$

Figure 3 shows the exact solution to (16)–(20), the inner solution (39)–(40), the outer solution (43)–(44), and the composite solution (45)–(46).

5 Discussion and Future Work

The key ingredient of our analysis is the separation of temporal scales at the level of the gene state chain. If all gene transition rates are of the same order, say $O(1/\varepsilon)$, then the chain equilibrates on a short $O(\varepsilon)$ timescale (the $O(1)$ timescale is assumed to be that of protein turnover). This situation has been widely considered in literature, e.g. [25, 34]. In the example on which we focused in our analysis, however, some transition rates are of a larger order, $O(1/\varepsilon^2)$. These faster rates generate an $O(\varepsilon^2)$ short timescale in our model. Importantly (and counterintuitively), the acceleration of these rates drives an emergent slow transitioning dynamics on the slow $O(1)$ timescale. This means, in particular, that the transient behaviour, as well as stochastic noise, is not averaged out but retained at the downstream level of protein dynamics. We expect that more general networks of gene states can generate more than two timescales (fast and slow). Results from other works can be used to compute approximations for multiple timescales [9]. In particular, we note that although our example retains some $O(1/\varepsilon)$ transition rates, no distinguished dynamics occurs on the corresponding $O(\varepsilon)$ timescale. We expect, however, the intermediate $O(\varepsilon)$ timescale can play a distinguished role in more complex systems.

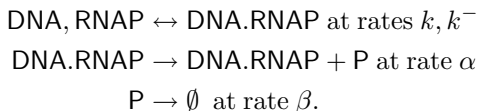
The possibility of realistic GRNs implementing slow gene expression dynamics by accelerating reactions at the DNA, opens up fundamental biological questions related to their regulatory and evolutionary roles. For modelling, the uncertainty about even the magnitude of biochemical reaction rates pressures us to account for the potentially emerging slow-by-fast phenomenon: approximations resting on the argument that all rates are ‘sufficiently fast’, while not accounting for the topology

of interactions at the DNA, can lead to wrong conclusions. The key feature of the 4-state example presented in this paper are very fast rates towards two different states which are poorly connected and consequently hard to leave. This situation will likely be seen in larger, realistic gene regulatory networks, because the rate of forming larger functional complexes typically depends on the order of TF's binding at the DNA. For instance, in a biologically realisable gene regulatory circuit shown in [11], a pair of activators and a pair of repressors compete to bind the DNA, so to rapidly transition to highly stable conformational change at the DNA. One of the interesting directions for future work is automatising the derivation of singular perturbation reduction shown in Sect. 4. Such a procedure would allow us to systematically explore reductions for larger gene regulatory networks. Additionally, we want to examine different topologies and sizes of networks to generalize our results. This could reveal if the backward reactions are always the crucial factor in causing the slow-by-fast phenomenon.

Slow-by-fast phenomena we show here, could appear in application domains beyond gene regulation, i.e., wherever nodes over a weighted network regulate a collective response over time. For instance, in network models used to predict the spread of information or spread of disease, among coupled agents [26, 36], or in network-models for studying the role of communication in wisdom of the crowds (known to be enhanced by interaction, but at the same time hindered by information exchange [1, 20]). Finally, networks of neurons are known to have different intrinsic time-scales, in addition to the time-scales that arise from network connections [8].

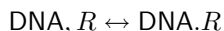
Appendix A: Mechanism of Gene Regulation - Examples

Example 1 (basal gene expression). Basal gene expression with RNAP binding can be modelled with four reactions, where the first reversible reaction models binding between the promoter site at the DNA and the polymerase, and the second two reactions model the protein production and degradation, respectively:



The state space of the underlying CTMC $S \cong \{0, 1\} \times \{0, 1, 2, \dots\}$, such that $s_{(1,x)} \in S$ denotes an active configuration (where the RNAP is bound to the DNA) with $x \in \mathbb{N}$ protein copy number.

Example 2 (adding repression). Repressor blocking the polymerase binding can be modelled by adding a reaction



In this case, there are three possible promoter configurations, that is, $S \cong \{\text{DNA}, \text{DNA.RNAP}, \text{DNA.R}\} \times \{0, 1, 2, \dots\}$ (states DNA and DNA.R are inactive promoter states).

Appendix B: Derivation of Moment Equations

Multiplying the master equation (2) by $n(n-1)\dots(n-j+1)$ and summing over all $n \geq 0$ yields differential equations [37]

$$\frac{d\boldsymbol{\nu}_j}{dt} = \mathbf{A}\boldsymbol{\nu}_j + j(\mathbf{A}_k\boldsymbol{\nu}_{j-1} - \delta\boldsymbol{\nu}_j) \quad (\text{B1})$$

for the factorial moments

$$\boldsymbol{\nu}_j(t) = \sum_{n=0}^{\infty} n(n-1)\dots(n-j+1)\mathbf{p}_n(t). \quad (\text{B2})$$

The quantities (4)–(6) can be expressed in terms of the factorial moments as

$$\mathbf{p} = \boldsymbol{\nu}_0, \quad \langle n \rangle = \mathbf{1}^\top \boldsymbol{\nu}_1, \quad \mathbf{f} = \boldsymbol{\nu}_1 - (\mathbf{1}^\top \boldsymbol{\nu}_1) \boldsymbol{\nu}_0, \quad \sigma^2 = \mathbf{1}^\top \boldsymbol{\nu}_2 + \mathbf{1}^\top \boldsymbol{\nu}_1 - (\mathbf{1}^\top \boldsymbol{\nu}_1)^2. \quad (\text{B3})$$

Differentiating (B3) with respect to t and using (B1), one recovers equations (7)–(9).

References

1. Becker, J., Brackbill, D., Centola, D.: Network dynamics of social influence in the wisdom of crowds. *Proc. Natl. Acad. Sci.* **114**(26), E5070–E5076 (2017)
2. Beica, A., Guet, C.C., Petrov, T.: Efficient reduction of kappa models by static inspection of the rule-set. In: Abate, A., Šafránek, D. (eds.) *HSB 2015*. LNCS, vol. 9271, pp. 173–191. Springer, Cham (2015). <https://doi.org/10.1007/978-3-319-26916-0-10>
3. Bintu, L.: Transcriptional regulation by the numbers: applications. *Curr. Opin. Genet. Dev.* **15**(2), 125–135 (2005)
4. Bo, S., Celani, A.: Multiple-scale stochastic processes: decimation, averaging and beyond. *Phys. Rep.* **670**, 1–59 (2017)
5. Cardelli, L., Kwiatkowska, M., Laurenti, L.: Stochastic analysis of chemical reaction networks using linear noise approximation. *Biosystems* **149**, 26–33 (2016)
6. Gardner, T.S., Cantor, C.R., Collins, J.J.: Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**(6767), 339 (2000)
7. Gillespie, D.T.: Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977)
8. Gjorgjieva, J., Drion, G., Marder, E.: Computational implications of biophysical diversity and multiple timescales in neurons and synapses for circuit performance. *Curr. Opin. Neurobiol.* **37**, 44–52 (2016)
9. Goban, A.N., Radulescu, O.: Dynamic and static limitation in multiscale reaction networks, revisited. *Adv. Chem. Eng.* **34**, 103–107 (2008)
10. Greenham, K., McClung, C.R.: Time to build on good design: resolving the temporal dynamics of gene regulatory networks. *Proc. Natl. Acad. Sci.* **115**(25), 6325–6327 (2018)

11. Guet, C., Henzinger, T.A., Igler, C., Petrov, T., Sezgin, A.: Transient memory in gene regulation. In: Bortolussi, L., Sanguinetti, G. (eds.) CMSB 2019. LNCS, vol. 11773, pp. 155–187. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-31304-3_9
12. Guet, C.C., Elowitz, M.B., Hsing, W., Leibler, S.: Combinatorial synthesis of genetic networks. *Science* **296**(5572), 1466–1470 (2002)
13. Gunawardena, J.: Time-scale separation-Michaelis and Menten’s old idea, still bearing fruit. *FEBS J.* **281**(2), 473–488 (2014)
14. da Costa Pereira Innocentini, G., Forger, M., Ramos, A.F., Radulescu, O., Hornos, J.E.M.: Multimodality and flexibility of stochastic gene expression. *Bull. Math. Biol.* **75**(12), 2360–2600 (2013)
15. Kevorkian, J., Cole, J.D.: *Perturbation Methods in Applied Mathematics*. Springer, New York (1981). <https://doi.org/10.1007/978-1-4757-4213-8>
16. Kevorkian, J., Cole, J.D., Nayfeh, A.H.: *Perturbation methods in applied mathematics*. *Bull. Am. Math. Soc.* **7**, 414–420 (1982)
17. Kurtz, T.G.: Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Prob.* **7**(1), 49–58 (1970)
18. Kurtz, T.G.: Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Prob.* **8**(2), 344–356 (1971)
19. Kwok, R.: Five hard truths for synthetic biology. *Nature* **463**(7279), 288–290 (2010)
20. Lorenz, J., Rauhut, H., Schweitzer, F., Helbing, D.: How social influence can undermine the wisdom of crowd effect. *Proc. Natl. Acad. Sci.* **108**(22), 9020–9025 (2011)
21. Marchisio, M.A., Stelling, J.: Automatic design of digital synthetic gene circuits. *PLoS Comput. Biol.* **7**(2), e1001083 (2011)
22. McAdams, H.H., Arkin, A.: It’s a noisy business! genetic regulation at the Nanomolar scale. *Trends Genet.* **15**(2), 65–69 (1999)
23. Murray, J.D.: *Mathematical Biology: I*. Springer, Introduction (2003)
24. Myers, C.J.: *Engineering Genetic Circuits*. CRC Press, Boca Raton (2009)
25. Newby, J., Chapman, J.: Metastable behavior in Markov processes with internal states. *J. Math. Biol.* **69**(4), 941–976 (2013). <https://doi.org/10.1007/s00285-013-0723-1>
26. Pagliara, R., Leonard, N.E.: Adaptive susceptibility and heterogeneity in contagion models on networks. *IEEE Trans. Automatic Control* (2020)
27. Pájaro, M., Otero-Muras, I., Vázquez, C., Alonso, A.A.: Transient hysteresis and inherent stochasticity in gene regulatory networks. *Nat. Commun.* **10**(1), 1–7 (2019)
28. Parmar, K., Blyuss, K.B., Kyrchko, Y.N., Hogan, S.J.: Time-delayed models of gene regulatory networks. In: *Computational and Mathematical Methods in Medicine* (2015)
29. Peleš, S., Munsky, B., Khammash, M.: Reduction and solution of the chemical master equation using time scale separation and finite state projection. *J. Chem. Phys.* **125**(20), 204104 (2006)
30. Rothenberg, E.V.: Causal gene regulatory network modeling and genomics: second-generation challenges. *J. Comput. Biol.* **26**(7), 703–718 (2019)
31. Santillán, M., Mackey, M.C.: Why the lysogenic state of phage λ is so stable: a mathematical modeling approach. *Biophys. J.* **86**(1), 75–84 (2004)
32. Schnoerr, D., Sanguinetti, G., Grima, R.: Approximation and inference methods for stochastic biochemical kinetics—a tutorial review. *J. Phys. A: Math. Theor.* **50**(9), 093001 (2017)
33. Segal, E., Widom, J.: From DNA sequence to transcriptional behaviour: a quantitative approach. *Nat. Rev. Genet.* **10**(7), 443–456 (2009)

34. Srivastava, R., Haseltine, E.L., Mastny, E., Rawlings, J.B.: The stochastic quasi-steady-state assumption: reducing the model but not the noise. *J. Chem. Phys.* **134**(15), 154109 (2011)
35. Trofimenkoff, E.A.M., Roussel, M.R.: Small binding-site clearance delays are not negligible in gene expression modeling. *Math. Biosci.* 108376 (2020)
36. Zhong, Y.D., Leonard, N.E.: A continuous threshold model of cascade dynamics. arXiv preprint [arXiv:1909.11852](https://arxiv.org/abs/1909.11852) (2019)
37. Zhou, T., Liu, T.: Quantitative analysis of gene expression systems. *Quant. Biol.* **3**(4), 168–181 (2015). <https://doi.org/10.1007/s40484-015-0056-8>