Analytical results for a stochastic model of gene expression with arbitrary partitioning of proteins

Hugo Tschirhart and Thierry Platini

Original citation & hyperlink:
Tschirhart, Hugo, and Thierry Platini. "Analytical results for a stochastic model of gene expression with arbitrary partitioning of proteins." *Journal of Physics A: Mathematical and Theoretical* 51.19 (2018): 195601.
https://dx.doi.org/10.1088/1751-8121/aab968

DOI 10.1088/1751-8121/aab968
ISSN 1751-8113

Publisher: IOP Publishing

Analytical results for a stochastic model of gene expression with arbitrary partitioning of proteins

Hugo Tschirhart[1,2,3] **& Thierry Platini**[4]

[1] University of Luxembourg, Physics and Materials Science Research Unit, Avenue de la Faïencerie 162a, L-1511 Luxembourg, Luxembourg

[2] Groupe de Physique Statistique, Institut Jean Lamour (CNRS UMR 7198), Université de Lorraine Nancy, B.P. 70239, F54506 Vandoeuvre-lès-Nancy Cedex, France,

[3] Doctoral College for the Statistical Physics of Complex Systems, Leipzig-Lorraine-Lviv-Coventry ($\mathbb{L}^4$)

[4] Applied Mathematics Research Center, Coventry University, Coventry, CV1 5FB, England,

E-mail: `thierry.platini@coventry.ac.uk`

Abstract. In biophysics, the search for analytical solutions of stochastic models of cellular processes is often a challenging task. In recent work on models of gene expression, it was shown that a mapping based on partitioning of Poisson arrivals (PPA-mapping) can lead to exact solutions for previously unsolved problems. While the approach can be used in general when the model involves Poisson processes corresponding to creation or degradation, current applications of the method and new results derived using it have been limited to date. In this paper, we present the exact solution of a variation of the two-stage model of gene expression (with time dependent transition rates) describing the arbitrary partitioning of proteins. The methodology proposed makes full use of the the PPA-mapping by transforming the original problem into a new process describing the evolution of three biological switches. Based on a succession of transformations, the method leads to a hierarchy of reduced models. We give an integral expression of the time dependent generating function as well as explicit results for the mean, variance, and correlation function. Finally, we discuss how results for time dependent parameters can be extended to the three-stage model and used to make inferences about models with parameter fluctuations induced by hidden stochastic variables.

PACS numbers: 87.10.Ca, 87.10.Mn, 87.18.Cf, 87.18.Tt

1. Introduction

Gene expression is the biological process by which information from a gene is used to synthesize RNA macromolecules and proteins. With a few exceptions, until the 1990s, this process was commonly understood from "a deterministic viewpoint" [1, 2]. Since then, the combination of experimental and theoretical approaches has clarified that gene expression is often stochastic in nature (see [3, 4, 5, 6, 7] for review articles). The effect of fluctuations (noise) is usually limited when we are dealing with large numbers of molecules [8]. In cells however, wherein genes and mRNAs are often present in low numbers, stochasticity has an important role on cellular functions. The importance of fluctuations [9, 10, 11, 12, 13, 14, 15] can be illustrated by the observation that, amongst a genetically identical population in a homogenous environment, cell-to-cell variations in gene expression can result in phenotypic heterogeneity.

There exists various mechanisms, some more complex than others, allowing cells to tame and exploit randomness [16]. In order to unveil those processes, research efforts are directed on both experimental [17, 18, 19, 20] and theoretical fronts [21, 22, 23, 24, 25, 26, 27, 28, 29]. Collaborations between biologists, physicists and mathematicians aim to reveal the conditions under which transcriptional noise may, or may not, cascade to affect downstream genetic products.

The two stage and three stage models [30, 31, 32, 33, 34] give a minimalist description of the simplest yet non-trivial biological processes leading to gene expression. The two-stage model includes only transcription and translation processes, while the three-stage model also incorporates free and repressed states of the DNA promoter region. Analytical techniques and results [35, 36, 37] for the previously mentioned processes are the cornerstone for further theoretical developments. These models are the elementary bricks allowing for the construction of more complex reaction networks including non-exponential waiting times, transcriptional burst, feedback loops *et cetera* [38, 39, 40, 41, 42]. Even when it is possible to derive the exact mean and variance of protein and mRNA numbers, obtaining an exact closed-form expression for the generating function is often a challenging problem. The two-stage model is a perfect example. It has been the subject of numerous studies since the paper of Thattai and Oudenaarden [40] in 2001. The model presents linear propensities so that all moments can be derived exactly. Such problems (like the one considered in this paper) are said to be "exactly solvable". Nevertheless, the exact generating function for the two stage model [37] was obtained only after ten years of extensive theoretical and experimental studies.

The search for exact solutions is often challenging because a small variation of a model's definition can make analytical results unattainable. Typically methods aiming for a full characterisation of a given process (beyond results for the mean and variance), focus on the master equation approach and its partner equation for the generating function. Once the generating function is obtained, all moments are in principle known: given by successive derivatives. This approach can provide insights into different limiting cases and into the behaviour of the distribution in different regions of parameter space. It is important to mention that analytical results (for the probability distribution) have been obtained for a class of models such as monomolecular reactions systems [61] or deficiency zero networks [62, 63]. Unfortunately, once outside these classes there exists no systematic analytical recipe applicable independently of a model's structure.

Research efforts are naturally turning towards numerical simulations which though powerful "bring no intuition into the underlying [...] interactions" [40]. To reach a better understanding one needs to investigate the joint distribution of mRNA and proteins, as well as temporal data, beyond the two-time autocorrelation function. As research progresses, emphasis is given to real time measurements with the hope to "expose the true cell dynamics buried in the average" [23]. Nowadays experimental advances allow for the count of individual molecules over time [43, 44, 45, 46] highlighting the need for both time-dependent and steady-state theoretical results .

In recent work, the partitioning of Poisson arrivals [47] was invoked to map Poisson processes to simple biological switches. This method is based on the separation of creation events (mRNA creation or protein creation) into independent groups. When applicable, this procedure leads to a mapping between creation/degradation process and a simple two-states biological switch. Applied to the two-stage model, this method led to the time dependent protein distributions [47] using already known results [21, 18, 36] for mRNA distributions in models with promoter-based regulation. The PPA-mapping needs, however, to be applied with care. It is important to warn the reader that, in a given model, not all creation/degradation process can be mapped onto a biological switch. For the PPA-mapping to apply, one needs to be able to partition a given creation event into independent processes. And for a given model, this will depend on upstream regulation of the creation/degradation process under consideration. This restriction is strong and appears as a serious limitation of the mapping applicability. We therefore need a more systematic way to use the PPA-mapping. As it is, the PPA mapping, can only be applied on models presenting a mixture of zero and first order reactions. It is unclear as if and how this method can be used or adapted to study models in the presence of feedback. In the simplest model describing bursty mRNA production, a variation of PPA mapping leads to an alternative derivation of the mRNA generating function [48]. But so far, this method has not been the subject of much attention and few models have been solved using this approach.

In direct connection with the applicability question of the PPA mapping is the inverse problem: Assuming the arbitrary partition of a creation event into two 'types' (type 1 and type 2), the latter process being itself regulated by an upstream mechanism, what correlation (between 1 and 2) should we expect? Is the correlation bounded? Does it vanish under particular conditions? Also, it is of experimental interest, to search for ways to infer the protein levels of a given type using measurement data on the other.

The model and methodology proposed in this paper were designed to (1) study correlations induced by the arbitrary partition of proteins arrival and (2) obtain the generating function making full use of the idea implicit in the PPA-mapping. The process we consider appears to be a simplified version of mechanisms involved in alternating splicing processes allowing a single gene to code for multiple proteins. With alternating splicing, a particular pre-mRNA can lead to different messenger RNAs, each being responsible for the production of isoform proteins (differing in their amino acid sequence). Results recently published in [49] focused on both bursty and constitutive pre-mRNA creation. One should mention that alternating splicing is far from being rare. Many genes have multiple splicing patterns [50, 51, 52] and numerous examples confirm that alternating splicing contributes to the development of cancer (see [53] and [54] for review articles). Our goal is to obtain the time dependent solution of the proposed model in term of the generating function. Our method is based on the construction of different mappings. Each transformation aims to reduce the study of a given model to the analysis of a simpler one. After a succession of transformations the problem is condensed to the study of two-state biological switches. The nesting between models and reduced models is reflected in a set of relations between generating functions. This hierarchy allows us to derive relations between mean numbers and higher order moments. We show that the PPA mapping allows us to consider arbitrary time dependent transition rates. Other studies such as [58, 59] and [60] have considered explicit time dependent parameters to investigate the effect of upstream hidden dynamics on downstream populations. Here, we obtain the time dependent generating function without solving any complicated differential equation but rather a simple first order equation (with time dependent coefficients). Accessing analytical results for time dependent coefficient provides a way to tackle models with noisy transition rates induced by hidden stochastic variables. In this paper, we show how results for the mean and correlations for time dependent model can be used to access the solution for the three-stage model.

The paper is organised as follows. In section (2.1), we start with the presentation of the model under consideration. We give the master equation governing the evolution of the probability distribution and present the solution of the first order moments. In section (2.2), we define the generating function, and outline the three different steps defining our method. Step 1 and 3 both describe the transformation under the PPA-mapping at different stages of the derivation. The intermediate step 2 defines the decomposition of a given process over all possible histories. Each step is detailed in sections (2.4), (2.5) and (2.6). The succession of transformations takes us relatively far away from the solution of the original problem. To proceed further, we give in section (2.7), the probability associated to each relevant histories. Finally, the time dependent generating function is presented in section (3), where our result is generalised to arbitrary partition numbers. We show how for some particular cases our results match the known solution for the two-stage model (see reference [37] and [47]). Finally we discuss how to use results for time dependent parameters to extend this work to the three-stage model and others processes including additional random variables.

2. Theory

2.1. The model

The model under consideration describes the stochastic evolution of protein numbers in a variation of the two-stage model, for which proteins are arbitrarily separated into two groups ($\mathcal{P}_1$ and $\mathcal{P}_2$). Note that our results will be easily generalized to an arbitrary

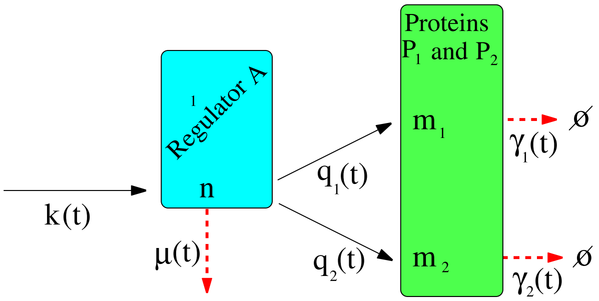

Figure 1. The original model (model-0): Two types of proteins ($\mathcal{P}_1$ and $\mathcal{P}_2$) are regulated by an upstream molecule $\mathcal{A}$. Transition rates for protein production and degradation are respectively written $q_j(t)$ and $\gamma_j(t)$ (with $j = 1, 2$), while $k(t)$ and $\mu(t)$ denote production and degradation rates for regulator $\mathcal{A}$.

number of protein types. We denote by $\mathcal{A}$ the upstream molecule regulating proteins production (see Figure (1)). Time dependent transition rates for proteins production and degradation are respectively written $q_j(t)$ and $\gamma_j(t)$ (with $j = 1, 2$). The level of regulator $\mathcal{A}$ is itself governed by the transition rates $k(t)$ and $\mu(t)$, respectively associated to creation and degradation (see table (1)). Because the method presented in this paper invokes mappings to other processes, it is convenient to refer to the original model as model-0.

The state of the system is, at any time, characterised by the numbers n, m_1 and m_2, of molecules $\mathcal{A}$ and proteins $\mathcal{P}_1$ and $\mathcal{P}_2$ respectively. We write $P_{n,m_1,m_2}(t)$ the probability distribution of state (n, m_1, m_2). We should keep in mind that the latter quantity is conditional on the initial state. In particular, we will consider $P_{0,0,0}(t = 0) = 1$. Further down in this paper we will explain how the latter initial condition is imposed by the initial state of the reduced model resulting from successive mappings. In order to consider other initial states, an extension of the proposed method is needed. This procedure would add an extra layer of complexity to the work presented here and is not discussed further in this paper. The probability distribution is governed by the master equation

$$\begin{aligned}\frac{d}{dt}P_{n,m_1,m_2} &= k(t)[P_{n-1,m_1,m_2} - P_{n,m_1,m_2}] \\ &+ \mu(t)[(n+1)P_{n+1,m_1,m_2} - nP_{n,m_1,m_2}] \\ &+ q_1(t)n[P_{n,m_1-1,m_2} - P_{n,m_1,m_2}] \\ &+ q_2(t)n[P_{n,m_1,m_2-1} - P_{n,m_1,m_2}] \\ &+ \gamma_1(t)[(m_1+1)P_{n,m_1+1,m_2} - m_1P_{n,m_1,m_2}] \\ &+ \gamma_2(t)[(m_2+1)P_{n,m_1,m_2+1} - m_2P_{n,m_1,m_2}].\end{aligned} \tag{1}$$

At this point, it is premature to solve the full master equation, and we start by deriving equations for mean population numbers. Let us denote $\langle\mathcal{O}\rangle$ the average of the observable

$\mathcal{O}$, given by $\sum_{n,m_1,m_2}\mathcal{O}(n,m_1,m_2)P_{n,m_1,m_2}$. For mean numbers we derive the equations:

$$\frac{d\langle n\rangle}{dt} = k(t)-\mu(t)\langle n\rangle, \tag{2}$$

$$\frac{d\langle m_j\rangle}{dt} = q_j(t)\langle n\rangle-\gamma_j(t)\langle m_j\rangle, \quad j=1,2. \tag{3}$$

If we restrict ourself to the well known case of constant reaction rates, with $\langle n\rangle(t=0)=\langle m_j\rangle(t=0)=0$, the solutions of these equations are:

$$\frac{\langle n\rangle(t)}{\langle n\rangle^*} = \left(1-e^{-\mu t}\right), \tag{4}$$

and

$$\frac{\langle m_j\rangle(t)}{\langle m_j\rangle^*} = \left\{\begin{array}{l} 1-\frac{1}{\gamma_j-\mu}(\gamma_j e^{-\mu t}-\mu e^{-\gamma_j t}), \quad \gamma_j/\mu\neq 1 \\ 1-(1+\mu t)e^{-\mu t},\ \gamma_j/\mu=1, \end{array}\right. \tag{5}$$

with the stationary values $\langle n\rangle^*=k/\mu$ and $\langle m_j\rangle^*=kq_j/\mu\gamma_j$. At this stage, it is not hard to extend these results to time-dependent coefficients. With a bit of work, one can show that the solution of equation (2) is

$$\langle n\rangle(t) = \int_0^t \mathrm{d}s\ k(s)e^{W_\mu(s)-W_\mu(t)}, \tag{6}$$

where we define $W_\mu(t)=\int_0^t \mathrm{d}\lambda\ \mu(\lambda)$. For arbitrary functions $k(t)$ and $\mu(t)$ it is however impossible to comment on the existence of a stationary state unless we assume the existence of the limits $k(t\to\infty)=k^*$ and $\mu(t\to\infty)=\mu^*$. It is here important to note that the solution of equation (3) can be expressed as a double integral. To proceed we use equation (6) with the substitution $k(t)\to q_j(t)\langle n\rangle(t)$ and $\mu(t)\to\gamma_j(t)$. A few lines of calculation leads to

$$\langle m_j\rangle(t) = \int_0^t \mathrm{d}s\int_s^t \mathrm{d}s'\ \mathcal{K}_j(s,s',t), \tag{7}$$

with kernel

$$\mathcal{K}_j(s,s',t) = k(s)q_j(s')e^{W_\mu(s)-W_\mu(s')+W_{\gamma_j}(s')-W_{\gamma_j}(t)}, \tag{8}$$

where $W_{\gamma_j}(t) = \int_0^t \mathrm{d}\lambda\ \gamma_j(\lambda)$. Interestingly, it is this quantity $\mathcal{K}_j(s,s',t)$ which will reappear explicitely in the final expression for the generating function. Along the same lines, it is possible to push further, writing equations for second order moments such as $\langle n^2\rangle$

$$\frac{d\langle n^2\rangle}{dt} = k(t)+(2k(t)+\mu(t))\langle n\rangle-2\mu(t)\langle n^2\rangle. \tag{9}$$

Once again, the solution for constant coefficients is easy to derive and can be expressed as a function of $\langle n\rangle(s)$:

$$\langle n^2\rangle(t) = \int_0^t \mathrm{d}s\ [k(s)+(2k(s)+\mu(s))\langle n\rangle(s)]e^{2W_\mu(s)-2W_\mu(t)}. \tag{10}$$

To evaluate correlations of the form $\langle m_1m_2\rangle(t)$ and $\langle nm_1\rangle(t)$ we write

$$\frac{d\langle nm_j\rangle}{dt} = k(t)\langle m_j\rangle(t)+q_j(t)\langle n^2\rangle-(\mu(t)+q_j(t))\langle nm_j\rangle, \tag{11}$$

$$\frac{d\langle m_1m_2\rangle}{dt} = q_1(t)\langle nm_2\rangle+q_2(t)\langle nm_1\rangle-(\gamma_1(t)+\gamma_2(t))\langle m_1m_2\rangle. \tag{12}$$

Together with $d\langle m_j\rangle/dt$ ($j = 1, 2$), $d\langle n\rangle/dt$ and $d\langle n^2\rangle/dt$, Eq. (11) and (12) define a system of seven equations. Importantly, equations governing the evolution of correlators do not involve higher order moments. As a consequence, correlations at any order can be obtained by solving a finite set of equations [55, 56, 57]. Even if one considers constant reaction rates, the generalisation of the solution to three or more protein types is not trivial. It requires the solution of a new and bigger set of equations. One possible avenue is to pursue with approximation of the "mean field" type, which consists in assuming $\langle m_1 m_2\rangle \simeq \langle m_1\rangle\langle m_2\rangle$. A priori, the later approximation holds for weakly correlated systems only. Hence we need to quantify correlation numbers in order to select the appropriate analytical methods.

Event	Update	Transition rates
$\mathcal{A}$-production	$n \to n+1$	$k(t)$
$\mathcal{A}$-degradation	$n \to n-1$	$n\mu(t)$
$\mathcal{P}_j$-production	$m_j \to m_j + 1$	$nq_j(t)$
$\mathcal{P}_j$-degradation	$m_j \to m_j - 1$	$m_j\gamma_j(t)$

Table 1. Transitions and associated rates for the original model (model-0).

2.2. The generating function

Let us start by defining the generating function of the original model

$$G^{(0)}(x, z_1, z_2, t) = \sum_{n,m_1,m_2} x^n z_1^{m_1} z_2^{m_2} P_{n,m_1,m_2}(t), \tag{13}$$

which obeys the differential equation

$$\begin{aligned}\frac{dG^{(0)}}{dt} &= (x-1)(k(t) - \mu(t)\partial_x)G^{(0)} \\ &+ (z_1 - 1)(q_1(t)x\partial_x - \gamma_1(t)\partial_{z_1})G^{(0)} + (z_2 - 1)(q_2(t)x\partial_x - \gamma_2(t)\partial_{z_2})G^{(0)}.\end{aligned} \tag{14}$$

Focusing our attention on the numbers of proteins only, we define the marginal probability

$$P_{m_1,m_2}(t) = \sum_{n=0}^{\infty} P_{n,m_1,m_2}(t), \tag{15}$$

for which the generating function is $G^{(0)}(z_1, z_2, t) = G^{(0)}(1, z_1, z_2, t)$. In order to attain an analytical expression we will successively reduce the original model into simpler ones.

To avoid confusion we choose to denote as model-1, model-2 and model-3, the processes which will be emerging from these successive mappings. We write $G^{(1)}$, $G^{(2)}$ and $G^{(3)}$ the generating functions for each model respectively. The following gives a short description of the steps taken in this paper, while each of them is further developed in sections (2.4), (2.5) and (2.6).

Step 1: The PPA-mapping [47] is based on the partitioning of Poisson processes (see Figure (2)). It allows for simplification of the original problem to N independent processes all identical to model-1 (see Figure (3) and section (2.4)). In the reduced model, the production of protein is regulated by a biological switch taking values $\theta = 0$ (OFF) and $\theta = 1$ (ON). Since N appears as a parameter of the reduced model, we write $G_N^{(1)}(z_1, z_2, t)$ the generating function of model-1. The latter is related to the original generating function via:

$$G^{(0)}(z_1, z_2, t) = \lim_{N\to\infty} \left[G_N^{(1)}(z_1, z_2, t)\right]^N. \tag{16}$$

Step 2: Denoting by Θ a particular history (or path) generated by the time evolution of the variable θ, we define $\Psi_N(\Theta)$ to be the probability of a given path. Model-2 is defined for one particular history as if frozen (Figure (4)). We write $G_\Theta^{(2)}$ as the associated generating function and express $G_N^{(1)}$ as an average over all possible histories (see section (2.5))

$$G_N^{(1)}(z_1, z_2, t) = \sum_\Theta \Psi_N(\Theta) G_\Theta^{(2)}(z_1, z_2, t). \tag{17}$$

Once the differential equation for $G_\Theta^{(2)}$ has been derived, we will be able to show that protein numbers are uncorrelated in model-2. It follows that $G_\Theta^{(2)}$ can be expressed as the product of two functions, each associated to a given protein type:

$$G_\Theta^{(2)}(z_1, z_2, t) = \prod_{j=1,2} G_{j|\Theta}^{(2)}(z_j, t). \tag{18}$$

Step 3: To access the solution of model-2, we exploit the PPA-mapping one more time. Splitting the creation process into M independent processes, it ultimately reduces to the study of biological switches (see model-3 in Figure (5)). Writing $G_{M;j|\Theta}^{(3)}(z_j, t)$ as the generating function of the switch j ($j = 1, 2$), we show the relation (see section (2.6))

$$G_{j|\Theta}^{(2)} = \lim_{M\to\infty} \left[G_{M;j|\Theta}^{(3)}\right]^M. \tag{19}$$

Finally, nesting all steps together, the original generating function is given by

$$G^{(0)}(z_1, z_2, t) = \lim_{N\to\infty} \underbrace{\left[\sum_\Theta \Psi_N(\Theta) \underbrace{\prod_{j=1,2} \underbrace{\lim_{M\to\infty} \left[G_{M;j|\Theta}^{(3)}(z_j, t)\right]^M}_{G_{j|\Theta}^{(2)}(z_j,t)}}_{G_\Theta^{(2)}(z_1,z_2,t)} \right]}_{G_N^{(1)}(z_1,z_2,t)}^N. \tag{20}$$

2.3. Consequences: hierarchy in mean and correlation numbers

Before entering the heart of the subject with the application of the PPA-mapping, one can investigate consequences of these successive transformations. The nesting of generating functions allows us to derive direct relations between mean numbers in the different models. We write $\langle m_j\rangle^{(1)}_N$, $\langle m_j\rangle^{(2)}_\Theta$ and $\langle m_j\rangle^{(3)}_{M|\Theta}$ the mean numbers of j-proteins in model-1, 2 and 3 respectively. For simplicity, we choose to omit the superscript 0 so that $\langle m_j\rangle$ denotes the average number of proteins in the original model. To ease the notations further we choose not to make the time dependance explicit, since the relations derived bellow are true for all time t. Equations (16), (17) and (19) bring us to

$$\langle m_j\rangle \quad = \lim_{N\to\infty} N\langle m_j\rangle^{(1)}_N, \tag{21}$$

$$\langle m_j\rangle^{(1)}_N = \sum_\Theta \Psi_N(\Theta)\langle m_j\rangle^{(2)}_\Theta, \tag{22}$$

$$\langle m_j\rangle^{(2)}_\Theta = \lim_{M\to\infty} M\langle m_j\rangle^{(3)}_{M|\Theta}. \tag{23}$$

The calculation of $\langle m_j\rangle^{(3)}$ is a pretty simple affair. Each protein being reduced to a biological switch, $m_j^{(3)}$ is restricted to the value 0 and 1. We give here, the expression of $\langle m_j\rangle^{(3)}$, for which the derivation is presented in section (2.6):

$$\langle m_j\rangle^{(3)}_{M|\Theta} = \frac{1}{M}\int_0^t \mathrm{d}\lambda\ \Theta(\lambda)q_j(\lambda)e^{W_{\gamma_j}(\lambda)-W_{\gamma_j}(t)}. \tag{24}$$

To continue further, eq. (22) requires knowledge of the probability $\Psi_N(\Theta)$ for a given path. This is not particularly difficult as one only needs to consider paths with probability up to the order $1/N$ (see section (2.7)). Without further knowledge of the generating function, once $\Psi_N(\Theta)$ and $\langle m_j\rangle^{(3)}$ given, the reader can derive the time evolution for mean number of proteins using equations (21), (22) and (23). Practically, those steps give a convoluted way to reach the result already presented in (5). It however reflects on the strategy adopted here to access the generating function.
Considering the correlation function, with the help of Eq. (16), we can show

$$C_{1,2} = \langle m_1 m_2\rangle - \langle m_1\rangle\langle m_2\rangle = \lim_{N\to\infty} N\langle m_1 m_2\rangle^{(1)}_N. \tag{25}$$

The protein number being uncorrelated in model-2 we have $\langle m_1 m_2\rangle^{(2)}_\Theta = \langle m_1\rangle^{(2)}_\Theta\langle m_2\rangle^{(2)}_\Theta$, which leads us to

$$\langle m_1 m_2\rangle^{(1)}_N = \sum_\Theta \Psi_N(\Theta)\langle m_1\rangle^{(2)}_\Theta\langle m_2\rangle^{(2)}_\Theta. \tag{26}$$

From the latter two equations, we conclude that $C_{1,2} > 0$ unless at least one of $\langle m_j\rangle = 0$ ($j = 1,2$). Hence, there is no non-trivial point in parameter space such that the correlation between protein number vanishes. As a consequence, there is no region of the parameter space in which the mean field approach is valid. In [49], the authors focus on alternative splicing mechanism, investigating the stationary state of a slightly different model from the one presented here. This study considers the transition from a pre-mRNA to two different mature mRNAs. For constitutive expression (no bursty pre-mRNA creation), they show that (in the stationary state) the mRNA numbers (of type 1 and 2) are independent. They however observe, for bursty pre-mRNA production, the emergence of correlations between the two mature mRNA types.

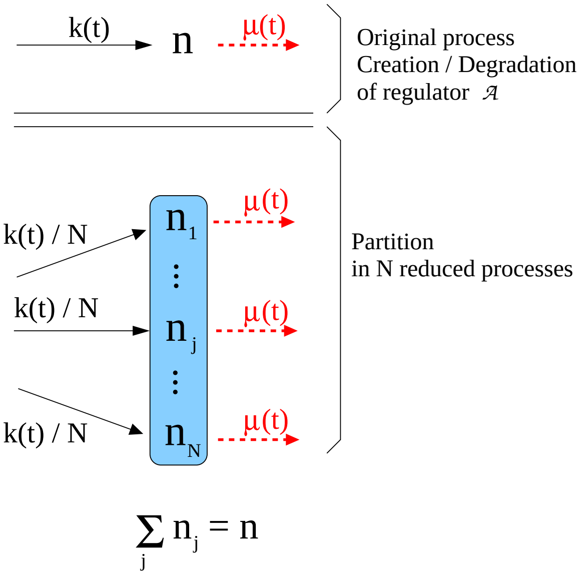

Figure 2. We partition each creation event into N 'types'. The creation rate associated to a particular type is given by $k(t)/N$. The sum of molecules numbers n_j over each type is equal to the total number of molecules n in the original model.

2.4. First transformation: from model-0 to model-1

The PPA-mapping is based on the partitioning property of Poisson processes. Without entering into technical details, the mapping can be understood as follow:

(i) Consider the creation/degradation process of regulator $\mathcal{A}$ (with rates $k(t)$ and $\mu(t)$).

(ii) Partition every creation events into N 'types' (Figure (2)). The partition is homogeneous so that each $\mathcal{A}$ molecule is equally likely to be assigned to a given type. It follows that the creation rate associated to a particular type is given by $k(t)/N$.

(iii) Take the limit $N \gg 1$. As a consequence, the probability of observing more than one $\mathcal{A}$ molecule of a particular type can be neglected. It follows that the random variable describing the number of molecules $\mathcal{A}$ (of a given type) is restricted to the value 0 or 1.

Model-1 as defined under this procedure is illustrated on Figure (3). Note that N appears as a parameter in the reduced model. Along the lines presented in [47] we write $G^{(0)} = [G_N^{(1)}]^N$. Equation (14) shows that $G_N^{(1)}$ obeys the same differential equation under the transformation $k(t) \to k(t)/N$. As a consequence the probability of observing (in the reduced model) more than one $\mathcal{A}$ molecule is of order $1/N^2$ and can be neglected

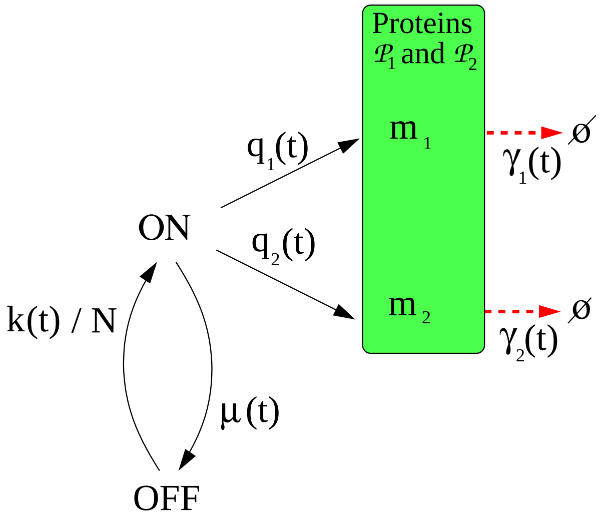

Figure 3. Model-1: The creation and degradation of two types of proteins ($\mathcal{P}_1$ and $\mathcal{P}_2$) is regulated by an upstream switch. Transition rates for proteins production and degradation are respectively written $q_j(t)$ and $\gamma_j(t)$ (with $j = 1, 2$), while $k(t)/N$ and $\mu(t)$ denote the probabilities of transition from $OFF \to ON$ and $ON \to OFF$.

as $N \to \infty$. While the previous logical argument shows how model-1 is emerging from model-0, an alternative derivation, based on the probability distribution instead of the generating function, allows for the reversed construction: building model-0 starting with N independent model-1. This derivation, not presented in the literature so far, is presented in an appendix.

2.5. Second transformation: from model-1 to model-2

Let us remind the reader that θ is the new stochastic variable (taking value in $\{0, 1\}$) emerging in model-1. The decomposition over all possible histories, generated by the variable θ, emerges from the use of conditional probabilities. To be more explicit we write Θ as a particular history associated to the variable θ. For a given path, we write $\Theta(t)$ as the value taken by the random variable θ at time t. We continue further by writing $\varphi_{N;(\Theta,a,b)}(t)$ as the probability associated to a particular history Θ and protein numbers a and b. The generating function $G_N^{(1)}$ can be rewritten as

$$G_N^{(1)}(z_1, z_2, t) = \sum_{\Theta,a,b} z_1^a z_2^b \varphi_{N;(\Theta,a,b)}(t). \tag{27}$$

Defining $\psi_{a,b|\Theta}(t)$ as the conditional probability on Θ while $\Psi_N(\Theta)$ is the probability of a given history, the equality $\varphi_{N;(\Theta,a,b)}(t) = \Psi_N(\Theta)\psi_{a,b|\Theta}(t)$ leads to

$$G_N^{(1)}(z_1, z_2, t) = \sum_{\Theta} \Psi_N(\Theta) G_\Theta^{(2)}(z_1, z_2, t), \tag{28}$$

with

$$G_\Theta^{(2)}(z_1, z_2, t) = \sum_{a,b} z_1^a z_2^b \psi_{a,b|\Theta}(t). \tag{29}$$

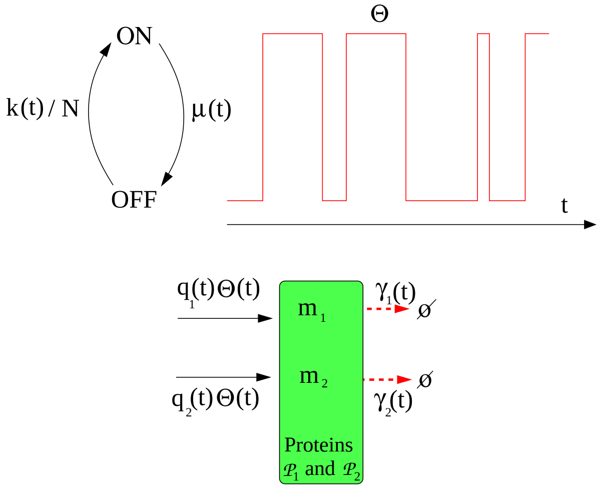

Figure 4. Model-2: The creation and degradation of two types of proteins ($\mathcal{P}_1$ and $\mathcal{P}_2$) for a fixed history Θ. Transition rates for proteins production and degradation are respectively written $q_j(t)\Theta(t)$ and $\gamma_j(t)$ (with $j = 1, 2$).

For a known history $\Theta(t)$, we have

$$\begin{aligned}\frac{dG^{(2)}_{\Theta}}{dt} &= (z_1 - 1)(q_1(t)\Theta(t) - \gamma_1(t)\partial_{z_1})G^{(2)}_{\Theta} \\ &\quad + (z_2 - 1)(q_2(t)\Theta(t) - \gamma_2(t)\partial_{z_2})G^{(2)}_{\Theta}.\end{aligned} \tag{30}$$

Note that in the last equation $G^{(2)}_{\Theta}$ is clearly independent of N and so is the conditional probability $\psi_{a,b|\Theta}$. The dependence in N is now carried by the probability Ψ_N. At this point we see that $G^{(2)}_{\Theta}$ can be written as

$$G^{(2)}_{\Theta}(z_1, z_2, t) = \prod_{j=1,2} G^{(2)}_{j|\Theta}(z_j, t), \tag{31}$$

where each generating function is governed by

$$\frac{dG^{(2)}_{j|\Theta}}{dt} = (z_j - 1)(q_j(t)\Theta(t) - \gamma_j(t)\partial_{z_j})G^{(2)}_{j|\Theta}. \tag{32}$$

Thereupon the two protein numbers are uncorrelated in model-2.

2.6. Third transformation: from model-2 to model-3

To reach the expression of $G^{(2)}_{\Theta}$, one applies the PPA-mapping one more time. This will reduce the original model to the study of biological switches (Figure (5)). For each protein type $\mathcal{P}_j$ ($j = 1, 2$), we once again, choose to partition every creation event into M groups. The partition being homogeneous, each protein is equally likely to be assigned to a given group. The creation rate for a particular group is given by $q_j(t)/M$. Taking the limit $M \gg 1$ allows us to neglect the creation of more than one protein in each group. To be explicit one writes $G^{(2)}_{j|\Theta} = [G^{(3)}_{M;j|\Theta}]^M$ in equation (32). This procedure leads to the same differential equation with the transformation $q_j(t)\Theta(t) \to q_j(t)\Theta(t)/M$. Hence, in the limit $M \to \infty$, the number of proteins of type j are restricted to 0 and 1. It follows that

$$G^{(2)}_{\Theta} = \prod_{j=1,2} \lim_{M\to\infty} \left[G^{(3)}_{M;j|\Theta}\right]^M. \tag{33}$$

The function $G^{(3)}_{M;j|\Theta}$ describes the dynamics of a two-state model and can be written has

$$G^{(3)}_{M;j|\Theta}(z,t) = 1 + (z-1) f_{M;j|\Theta}, \tag{34}$$

where $f_{M;j|\Theta}$ is the probability to find the switch j in the ON-state, knowing the history Θ. The latter is the solution of the following equation

$$\frac{df_{M;j|\Theta}}{dt} = q_j(t)\Theta(t)/M - \left[\gamma_j(t) + q_j(t)\Theta(t)/M\right] f_{M;j|\Theta}. \tag{35}$$

We now have reached the point where one needs to define the initial state. We choose to consider $f_{M;j|\Theta}(t=0) = 0$. Let us remind the reader that, in model-3, the total number of switches j in the ON-state equals the number of proteins $\mathcal{P}_j$ in model-2. The hierarchy builds up to the number of proteins in model-0. As we look at equation (21), (22) and (23), we see that choosing (at time $t=0$) all switches ($j=1,2$) in the OFF-state, imposes the following initial state on to the original model

$$m_1(t=0) = m_2(t=0) = 0. \tag{36}$$

A simple calculation gives

$$f_{M;j|\Theta}(t) = \int_0^t \mathrm{d}\lambda\ \frac{q_j(\lambda)}{M}\Theta(\lambda) \exp\left[-\int_\lambda^t \mathrm{d}s\ \left\{\gamma_j(s) + \frac{q_j(s)}{M}\Theta(s)\right\}\right], \tag{37}$$

which, to the first order in $1/M$, simplifies to

$$f_{M;j|\Theta}(t) = \Lambda_{j|\Theta}(t)/M, \tag{38}$$

with

$$\Lambda_{j|\Theta}(t) = \int_0^t \mathrm{d}\lambda\ \Theta(\lambda) q_j(\lambda) e^{W_{\gamma_j}(\lambda) - W_{\gamma_j}(t)}. \tag{39}$$

Nesting equation (38) into (34) leads to:

$$G^{(3)}_{M;j|\Theta}(z,t) = 1 + (z-1)\Lambda_{j|\Theta}(t)/M. \tag{40}$$

With Eq. (33) the latter result allows us to write

$$G^{(2)}_{\Theta}(z_1,z_2,t) = \exp\left[\sum_{j=1,2}(z_j-1)\Lambda_{j|\Theta}(t)\right]. \tag{41}$$

2.7. Summing over all histories

In order to derive the generating function $G^{(1)}_N$ from $G^{(2)}_\Theta$, using equation (17), we focus on the expression of the probability $\Psi_N(\Theta)$ for all relevant histories Θ. As mentioned earlier, we simply need to evaluate $\Psi_N(\Theta)$ up to the order $1/N$. We choose to consider

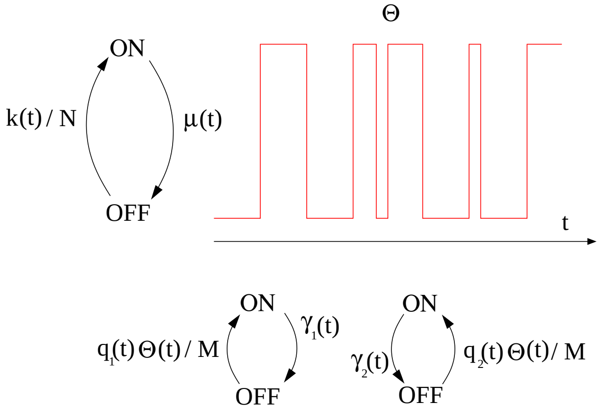

Figure 5. Model-3 describes, for a fixed history Θ, two biological switches for which transition rates are respectively written $q_j(t)\Theta(t)/M$ and $\gamma_j(t)$ (with $j = 1, 2$).

the initial state $\Theta(t = 0) = 0$. Amongst N identical models, the number of switches in the ON-states defines the number n of molecules $\mathcal{A}$. It follows that the initial state must satisfy:

$$n(t = 0) = 0. \tag{42}$$

Together, equations (36) and (42) fully specify the initial state, so that $P_{0,0,0}(t = 0) = 1$. We remind the reader that the probability of the transition form $ON \rightarrow OFF$ (between time t and $t + \delta t$) is given by $\mu(t)\delta t$. In addition, the probability of observing the transition $OFF \rightarrow ON$ is given by $k(t)\delta t/N$. At the first order in $1/N$, three different types of histories are relevant. They are symbolically represented by $\longrightarrow$, $\lrcorner\!\rightarrow$ and $\lrcorner\sqcap\llcorner\!\rightarrow$ and detailed in table (2). The probability associated to each path is

$$\Psi_N\left(\longrightarrow\right) \simeq 1 - \frac{1}{N}\int_0^t \mathrm{d}s\ k(s), \tag{43}$$

$$\Psi_N\left(\lrcorner\!\rightarrow\right) \simeq \frac{k(s)}{N} e^{W_\mu(s) - W_\mu(t)}, \tag{44}$$

$$\Psi_N\left(\lrcorner\sqcap\llcorner\!\rightarrow\right) \simeq \frac{k(s)}{N}\mu(s')e^{W_\mu(s) - W_\mu(s')}. \tag{45}$$

It is particularly useful to rewrite the last equation as

$$\Psi_N\left(\lrcorner\sqcap\llcorner\!\rightarrow\right) \simeq -\frac{k(s)}{N} e^{W_\mu(s)}\left(\partial_{s'} e^{-W_\mu(s')}\right), \tag{46}$$

which can be used to verify the conservation of probability:

$$\Psi_N\left(\longrightarrow\right) + \int_0^t \mathrm{d}s\ \ \Psi_N\left(\lrcorner\!\rightarrow\right) + \int_0^t \mathrm{d}s \int_s^t \mathrm{d}s'\ \Psi_N\left(\lrcorner\sqcap\llcorner\!\rightarrow\right) = 1. \tag{47}$$

The latter relation confirms that all relevant paths have been taken into consideration. The expression of $G_N^{(1)}$ is symbolically given by

$$G_N^{(1)} = \Psi_N\left(\longrightarrow\right) G^{(2)}_{\rightarrow} + \int_0^t \mathrm{d}s\ \Psi_N\left(\lrcorner\!\rightarrow\right) G^{(2)}_{\lrcorner\rightarrow} + \int_0^t \mathrm{d}s \int_s^t \mathrm{d}s'\ \Psi_N\left(\lrcorner\sqcap\llcorner\!\rightarrow\right) G^{(2)}_{\lrcorner\sqcap\llcorner\rightarrow}. \tag{48}$$

The explicit calculation (using $G^{(2)}_{\longrightarrow} = 1$) points us to

$$G^{(1)}_N(z_1, z_2, t) \simeq 1 + \frac{A(z_1, z_2, t)}{N}, \tag{49}$$

with

$$A(z_1, z_2, t) = \int_0^t \mathrm{d}s\ k(s) e^{W_\mu(s)} \int_s^t \mathrm{d}s'\ e^{-W_\mu(s')} \partial_{s'} G^{(2)}_{\_\!\sqcap\!\!\rightarrow}. \tag{50}$$

Representation	Description
$\longrightarrow$	θ is constantly in the OFF-state $\Theta(\tau) = 0, \forall \tau \in [0, t[$
$\_\!\ulcorner\!\!\rightarrow$	θ is switching state at time s $\Theta(\tau) = 1$, if $\tau \in [s, t[$ and $\Theta(\tau) = 0$ otherwise
$\_\!\sqcap\!\!\rightarrow$	θ is switching state at time s and s' $\Theta(\tau) = 1$, if $\tau \in [s, s'[$ and $\Theta(\tau) = 0$ otherwise

Table 2. Three different types of histories need to be considered. Each path starts with $\theta = 0$, and transits no more than once from OFF to ON.

3. Result: final expression of $G^{(0)}(z_1, z_2, t)$

The methodology presented in the previous section leads us to the following generating function (obtained by taking the limit $N \to \infty$ in equation (49)):

$$G^{(0)}(z_1, z_2, t) = \exp\left[\sum_{j=1,2} (z_j - 1) A_j(z_1, z_2, t)\right], \tag{51}$$

where

$$A_j = \int_0^t \mathrm{d}s \int_s^t \mathrm{d}s'\ \mathcal{K}_j(s, s', t) G^{(2)}_{\_\!\sqcap\!\!\rightarrow}, \tag{52}$$

with the same kernel $\mathcal{K}_j$ defined in eq. (8)).Finally, $G^{(2)}_\Theta$ for the path $\_\!\sqcap\!\!\rightarrow$ is explicitly given by

$$G^{(2)}_{\_\!\sqcap\!\!\rightarrow} = \exp\left[\sum_{i=1,2} (z_i - 1) \int_s^{s'} \mathrm{d}\lambda\ q_i(\lambda) e^{W_{\gamma_i}(\lambda) - W_{\gamma_i}(t)}\right]. \tag{53}$$

To keep notation as compact as possible, we will simply write $\mathcal{K}_j$, omitting the variables s, s' and t. The relation between $G^{(0)}$ and $G^{(2)}_{\Theta}$ is particularly interesting and allows the kernel $\mathcal{K}_j$ to play a key role in a new set of relations between model-2 and the original model. In model-0, the mean $\langle m_j \rangle$ and correlation $\langle m_j(m_j-1)\rangle$ are obtained using $\langle m_j\rangle(t) = \partial_{z_j} G^{(0)}(z_1,z_2,t)|_{z_1,z_2\to 1}$ and $\langle m_j(m_j-1)\rangle(t) = \partial^2_{z_j} G^{(0)}(z_1,z_2,t)|_{z_1,z_2\to 1}$. For the mean, a simple calculation leads to $\langle m_j\rangle(t) = \int_0^t \mathrm{d}s \int_s^t \mathrm{d}s' \; \mathcal{K}_j(s,s',t)$ (identical to eq. (7)). For the variance, defined by $\mathrm{Var}[m_j](t) = \langle m_j^2\rangle(t) - [\langle m_j\rangle(t)]^2$, writing $\langle m_j\rangle^{(2)}_{\sqcap\!\!\rightarrow} = \partial_{z_j} G^{(2)}_{\sqcap\!\!\rightarrow}|_{z_1,z_2\to 1}$, we obtain :

$$\mathrm{Var}[m_j](t) = \langle m_j\rangle(t) + 2\int_0^t \mathrm{d}s \int_s^t \mathrm{d}s' \; \mathcal{K}_j \langle m_j\rangle^{(2)}_{\sqcap\!\!\rightarrow}, \tag{54}$$

while the correlation function $C_{i,j}(t) = \langle m_i m_j\rangle(t) - \langle m_i\rangle(t)\langle m_j\rangle(t)$ becomes

$$C_{i,j}(t) = \int_0^t \mathrm{d}s \int_s^t \mathrm{d}s' \; \left[\mathcal{K}_i \langle m_j\rangle^{(2)}_{\sqcap\!\!\rightarrow}(t) + \mathcal{K}_j \langle m_i\rangle^{(2)}_{\sqcap\!\!\rightarrow}(t)\right]. \tag{55}$$

With the mean and correlation numbers in hand, the variance can be reached easily using $\mathrm{Var}[m_j](t) = \langle m_j\rangle(t) + C_{j,j}(t)$. It is clear that the generating function can be generalized to an arbitrary number J of proteins by replacing $\sum_{j=1,2} \to \sum_{j=1}^{J}$ in equation (51) and (53). In this situation, the generating function depends of J variables: z_1, z_2, ..., z_J.

3.1. Constant reaction rates

Let us first focus on simplifications occurring when all transition rates are constant. We have $W_\mu(t) = \mu t$, $W_{\gamma_j}(t) = \gamma_j t$ and $\mathcal{K}_j(s,s',t) = kq_j e^{\mu(s-s')+\gamma_j(s'-t)}$. It is then convenient to define $u = e^{\gamma_j(s-t)}$ and $v = e^{\gamma_j(s'-t)}$ so that A_j takes the form

$$A_j = \frac{kq_j}{\gamma_j^2} \int_{e^{-\gamma_j t}}^{1} \mathrm{d}u \int_u^1 \mathrm{d}v \; \Omega_j(u,v), \tag{56}$$

with

$$\Omega_j(u,v) = \frac{u^{\mu/\gamma_j - 1}}{v^{\mu/\gamma_j}} \exp\left[\sum_i (z_i - 1)\frac{q_i}{\gamma_i}\left(v^{\gamma_i/\gamma_j} - u^{\gamma_i/\gamma_j}\right)\right]. \tag{57}$$

At this stage it is not hard to show that the mean protein number is given by (5). Figure (6) confirms the validity of our results. For correlation numbers, it is useful to define the normalised function $\tilde{C}_{i,j}(t) = C_{i,j}(t)/\langle m_i\rangle^*\langle m_j\rangle^*$. For $\gamma_i + \gamma_j \neq \mu$, $\gamma_i \neq \mu$ and $\gamma_j \neq \mu$ the latter quantity is given by

$$\begin{aligned} \frac{k}{\mu}\tilde{C}_{i,j} = {} & \frac{\gamma_i+\gamma_j}{\gamma_i+\gamma_j-\mu}\left(1-e^{-\mu t}\right) \\ & - \frac{\gamma_i}{\gamma_i-\mu}\frac{\mu}{\gamma_j+\mu}\left[1-e^{-(\gamma_j+\mu)t}\right] - \frac{\gamma_j}{\gamma_j-\mu}\frac{\mu}{\gamma_i+\mu}\left[1-e^{-(\gamma_i+\mu)t}\right] \\ & + \frac{\mu}{\gamma_i+\gamma_j}\frac{\gamma_i}{\gamma_i-\mu}\frac{\gamma_j}{\gamma_j-\mu}\frac{\gamma_i+\gamma_j-2\mu}{\gamma_i+\gamma_j-\mu}\left[1-e^{-(\gamma_i+\gamma_j)t}\right]. \end{aligned} \tag{58}$$

The case $\gamma_i = \mu$ (or $\gamma_j = \mu$) has to be treated separately. To proceed one can (1) set $\gamma_i = \mu$ in the kernel K_i, or alternatively (2) write $\gamma_i = \mu + \epsilon$ and take the limit $\epsilon \to 0$.

The case $\gamma_i + \gamma_j = \mu$ (or $2\gamma_i = \mu$ when considering the variance) has to be treated similarly. Those limits lead to relatively more compact expressions, for example when $\gamma_i = \gamma_j = \mu$, we have:

$$\frac{k}{\mu}\tilde{C}_{i,j} = \frac{1}{2} - 2e^{-\mu t} + e^{-2\mu t}(3/2 + \mu t). \tag{59}$$

The agreement (for all time t) between analytical expressions and numerical simulations can be seen in Figure (7). In the limit $t \to \infty$ all expressions of $\tilde{C}_{i,j}$ converge to a single form. In the stationary state, the correlation function has a unique expression

$$\tilde{C}^*_{i,j} = \frac{\mu}{k}\frac{\eta_i}{\eta_i + 1}\frac{\eta_j}{\eta_j + 1}\frac{\eta_i + \eta_j + 2}{\eta_i + \eta_j}, \tag{60}$$

with $\eta_i = \gamma_i/\mu$ and $\eta_j = \gamma_j/\mu$. So that

$$C^*_{i,j} = \frac{k}{\mu}\frac{q_i}{\gamma_i + \mu}\frac{q_j}{\gamma_j + \mu}\frac{\gamma_i + \gamma_j + 2\mu}{\gamma_i + \gamma_j}. \tag{61}$$

For homogeneous degradation rates ($\gamma_i = \gamma_j = \gamma$), the correlation is invariant under the exchange $\gamma \leftrightarrow \mu$. The last equation clearly shows that the correlation function does not vanish (unless one out of $\langle m_i \rangle^*$ and $\langle m_j \rangle^*$ vanishes). We note that the correlation $C^*_{i,j}$ is strictly monotonic (decreasing) in terms of γ_i and γ_j (keeping all other parameters constant). As a consequence, if one can estimate lower and upper bounds of both γ_i and γ_j it is, in principle, possible to restrain the range of correlation values to an interval: $[C^*_{min}, C^*_{max}]$. In addition, we observe, for a fixed value of η_j, that the correlation $\tilde{C}^*$ presents a maximum at $(\eta_i)_{max} = \eta_j + \sqrt{2\eta_j(\eta_j + 1)}$. If C^* is strictly monotonic, it is when varying γ_i while keeping $\langle m_i \rangle^*$ constant that a non monotonic behaviour is observed. In this case, $C^*_{i,j}$ can be rewritten as

$$C^*_{i,j} = \langle m_i \rangle^* \frac{q_j/\mu}{\eta_j + 1}\frac{\eta_i}{\eta_i + 1}\frac{\eta_i + \eta_j + 2}{\eta_i + \eta_j}, \tag{62}$$

and presents a maximum in $(\eta_i)_{max}$. Keeping both protein levels $\langle m_i \rangle^*$ and $\langle m_j \rangle^*$ constant, the correlation function becomes $C^*_{i,j} = \langle m_i \rangle^* \langle m_j \rangle^* \tilde{C}^*_{i,j}$. Looking for an upper bound into the (η_i, η_j)-plane, one needs to solve $\partial_{\eta_i} C^* = 0$ and $\partial_{\eta_j} C^* = 0$ simultaneously. However there are no strictly positive solutions to the latter system of equations. Hence, under this constrain, $C^*_{i,j}$ does not present a maximum when varying both γ_i and γ_j.

3.1.1. The 2-stage model: $J = 1$ In the case $J = 1$, the model with constant transition rates, reduces to the conventional two-stage model. To pursue, we define $r = \mu/\gamma$ and $\delta(z) = q(z-1)/\gamma$. We can show that our result leads to the solution first presented in [37] and later in [47]:

$$G^*(z) = \exp\left(\frac{k}{\mu}\int_0^{\delta(z)} \mathrm{d}s\ {}_1F_1\left[1, r+1, s\right]\right), \tag{63}$$

where ${}_1F_1$ is the confluent hypergeometric function. As it is, the identity between equation (51) (for one protein type only) and equation (63) is not obvious. To proceed, we use the Taylor expansion of $e^{\delta v}$ and $e^{-\delta u}$ and write $\epsilon = e^{-\gamma t}$ which we assume small compared to one. Considering $r = \mu/\gamma \neq 1$ and keeping the lowest order in ϵ (see appendix) we show that

$$G^{(0)}(z,t) \underset{t\gg 1}{\simeq} G^*(z)H(z,t), \tag{64}$$

with $G^*(z)$ given by equation (63) and

$$\ln(H(z,t)) = \left\{ \begin{array}{cc} \frac{k}{\gamma-\mu} e^{-\gamma t}\delta(z) & \gamma < \mu \\ -\frac{k}{\mu} e^{-\mu t} \sum_{m=0}^{\infty} \frac{(\delta(z))^{m+1}}{m!(m+1-r)} & \gamma > \mu, \end{array} \right. \tag{65}$$

such that $\lim_{z\to 1} H(z,t) = \lim_{t\to\infty} H(z,t) = 1$. The latter approximation leads to:

$$\frac{\langle m\rangle(t)}{\langle m\rangle^*} \underset{t\gg 1}{\simeq} 1 + \frac{1}{\gamma-\mu} \left\{ \begin{array}{cc} \mu e^{-\gamma t} & \gamma < \mu \\ (-1)\gamma e^{-\mu t} & \gamma > \mu, \end{array} \right. \tag{66}$$

in agreement with equation (5). The case $\gamma = \mu$ is treated separately in appendix.

3.1.2. Homogeneous degradation rates: $\gamma_j = \gamma$, $\forall j$ When dealing with J protein types ($J > 1$) and homogeneous degradation rates ($\gamma_j = \gamma\ \forall j$), the generating function reduces to a form close to the one previously obtained for the two-stage model. Defining $\Delta(\{z_j\}) = \sum_j q_j(z_j - 1)/\gamma$, we can show that $G^{(0)}(\{z_j\},t)$ is given by equation (64) under the substitution $\delta(\{z_j\}) \to \Delta(\{z_j\})$. It follows that

$$\lim_{t\to\infty} G^{(0)}(\{z_j\},t) = \exp\left(\frac{k}{\mu}\int_0^{\Delta(\{z_j\})} \mathrm{d}s\ {}_1F_1\left[1, r+1, s\right]\right). \tag{67}$$

The generating function $\mathcal{G}$, associated to the total number of proteins ($M = \sum_j m_j$), is defined by $\mathcal{G}(z) = \sum_M P_M z^M$, with

$$P_M = \sum_{m_1,m_2,\ldots,m_J} P_{m_1,m_2,\ldots,m_J}\delta\left(\sum_j m_j - M\right). \tag{68}$$

We see that $\mathcal{G}$ is given by $\mathcal{G}(z,t) = G^{(0)}(\{z_j = z\},t)$:

$$\mathcal{G}^*(z) = \exp\left(\frac{k}{\mu}\int_0^{J\bar{q}(z-1)/\gamma} \mathrm{d}s\ {}_1F_1\left[1, r+1, s\right]\right), \tag{69}$$

with the average creation rate defined by $J\bar{q} = \sum_j q_j$. The mean of total protein number ($M = \sum_j m_j$) satisfies $\langle M\rangle/J = k\bar{q}/(\mu\gamma)$.

3.2. Time dependent transition rates: a bridge towards other models

Results for time dependent parameters allow for the study of fluctuations (induced by hidden variables) in production and/or degradation rates. In a recent paper Dattani and Barahona [60] proposed a framework to model gene expression with stochastic or deterministic transcription and degradation rates. Along the same lines, let us start by defining random variables x_ϕ, for all parameters ϕ of the model ($\phi = k, \mu, q_1, q_2, \ldots, \gamma_1, \gamma_2, \ldots$). We choose to write $\phi(t) = \phi_0 + \phi_1 x_\phi(t)$ with $\phi_0, \phi_1 \in$

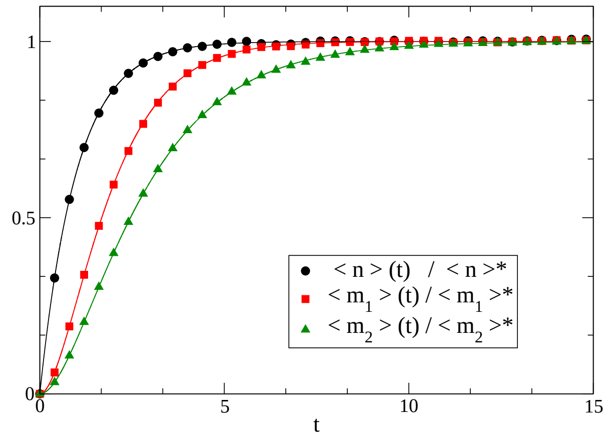

Figure 6. Time evolution of the ratios $\langle n\rangle(t)/\langle n\rangle^*$, $\langle m_1\rangle(t)/\langle m_1\rangle^*$ and $\langle m_2\rangle(t)/\langle m_2\rangle^*$ for the following set of constant parameters: $k = 10$, $\mu = 1$, $q_1 = 3$, $\gamma_1 = 1$, $q_2 = 5$ and $\gamma_2 = 1/2$. We observe an excellent agreement between simulation results (circles, squares and triangles) and the analytical expressions (lines). Simulation data, obtained using the Gillespie algorithm, are the result of an average over 10^4 sampled histories.

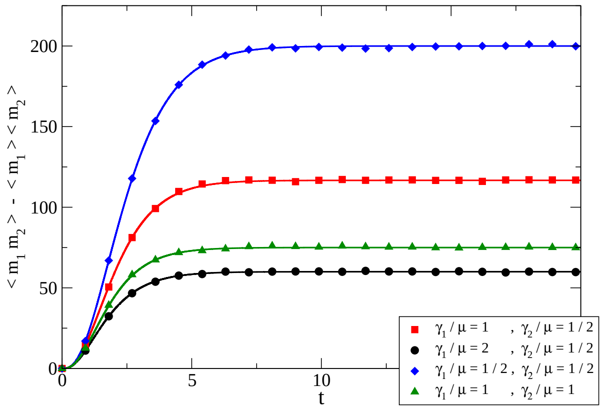

Figure 7. Time evolution of the correlation function $C(t) = \langle m_1 m_2\rangle - \langle m_1\rangle\langle m_2\rangle$ for constant parameter values $k = 10$, $\mu = 1$, $q_1 = 3$, $q_2 = 5$, considering four possible scenarios (1) $(\gamma_1 + \gamma_2)/\mu \neq 1$ and $\gamma_j/\mu \neq 1$ (black cirlces), (2) $(\gamma_1 + \gamma_2)/\mu = 1$ (blue diamonds), (3) $\gamma_1/\mu = 1$ and $\gamma_2/\mu \neq 1$ (red squares) and (4) $\gamma_1/\mu = \gamma_2/\mu = 1$ (green triangles). Analytical results (lines) are compared to numerical results (symbols). Simulation data, obtained using the Gillespie algorithm, are the result of an average over 10^5 sampled histories.

$\mathbb{R}$, but other functional forms could be chosen. More explicitly, we have for $k(t)$: $k(t) = k_0 + k_1 x_k(t)$. A particular time history of the random variable $x_k(t)$ is written $X_k = (x_k(t)|\forall t)$. Identically, we write $X_\phi = (x_\phi(t)|\forall t)$ the history for random variable x_ϕ. Finally, we define $\mathbb{X}$ as the set of histories $\mathbb{X} = \{X_k, X_\mu, X_{q_1}, ...X_{\gamma_1}, ...\}$ so that the generating function is now explicitly dependent on $\mathbb{X}$: we write $G^{(0)}_{\mathbb{X}}$. Writing $\mathcal{P}(\mathbb{X})$ as the probability of the set of histories, the new generating function is

$$\langle G^{(0)}\rangle = \sum_{\mathbb{X}} \mathcal{P}(\mathbb{X}) G^{(0)}_{\mathbb{X}}, \tag{70}$$

where $\sum_{\mathbb{X}}$ symbolically represents the sum over all possible histories of all parameters. The mean protein number is given by

$$\langle\langle m_j \rangle\rangle = \sum_{\mathbb{X}} \mathcal{P}(\mathbb{X}) \langle m_j \rangle_{\mathbb{X}}, \tag{71}$$

with $\langle m_j \rangle_{\mathbb{X}} = \partial_{z_j} G^{(0)}_{\mathbb{X}} |_{\forall z_i = 1}$. To give a concrete example, we will restrict ourself to time dependent production rate $k(t)$ while all other transition rates are constant. We show how our results bridge towards the three-stage model, allowing us to access the exact mean and correlation functions. In this situation $\mathcal{K}_j(s, s', t) = k(s) q_j e^{\mu(s-s') + \gamma_j (s'-t)}$. Choosing the appropriate function $k(t)$ can give information on the behaviour induced by state fluctuation of the DNA operational site (Figure (8)). We will write $\langle m_j \rangle_{X_k}$ the mean protein numbers, for a particular history X_k. For $\gamma_j \neq \mu$ and without restriction on $k(t)$, we show that

$$\langle m_j \rangle_{X_k}(t) = \frac{q_j}{\gamma_j - \mu} k \star \left[e^-_{\mu} - e^-_{\gamma_j} \right], \tag{72}$$

with the convolution product

$$(k \star e^{\pm}_a)(t) = \int_0^t \mathrm{d}s \; k(s) e^{\pm a(t-s)}. \tag{73}$$

For $\gamma_j = \mu$ we write $\gamma_j = \mu + \epsilon$ in (72) together with the limit $\epsilon \to 0$. ‡ The previous equation becomes $\langle m_j \rangle_{X_k}(t) = q_j \left(-\frac{\partial}{\partial \mu} \right) (k \star e^-_{\mu})(t)$. For all values of γ_j and μ, the Laplace transform of the mean number of protein $L[\langle m_j \rangle]$ simplifies to a single expression:

$$L[\langle m_j \rangle_{X_k}](s) = \frac{q_j L[k](s)}{(s + \mu)(s + \gamma_j)}, \tag{74}$$

with $L[k]$ as the Laplace transform of $k(t)$. Assuming the limit $k(t \to \infty) = k^*$ exists, the final value theorem leads to $\langle m_j \rangle^* = k^* q_j / (\gamma_j \mu)$. With the three-stage model in mind, we set $k_0 = 0$ and write $k(t) = k_1 x_k(t)$, where the random variable $x_k(t)$ takes value in $\{0, 1\}$. It describes the possible states, active ($x_k = 1$) or inactive ($x_k = 0$), of the promoter region. Governed by a simple two state dynamics (with transition rates $W_{0\to1} = \alpha$ and $W_{1\to0} = \beta$) the variable $x_k(t)$ "oscillates" between those states (see illustration (8)). This motivates in [59] the choice of a sinusoidal function: $k(t) = c_1 \sin(\omega t + \phi) + c_2$. However, the time evolution of the variable $x_k(t)$ is stochastic and, starting from initial condition $\langle x_k \rangle (t = 0) = 0$, it satisfies

$$\frac{\langle x_k \rangle (t)}{\langle x_k \rangle^*} = \chi(t) = 1 - e^{-(\alpha+\beta)t}, \tag{75}$$

with stationnary state $\langle x_k \rangle^* = \alpha / (\alpha + \beta)$. It follows that the mean for the three stage model is given by $\langle\langle m_j \rangle\rangle$ representing the average over the history of the variable x_k. For $\gamma_j \neq \mu$ we have to evaluate

$$\frac{\langle\langle m_j \rangle\rangle (t)}{\langle\langle m_j \rangle\rangle^*} = \frac{\mu \gamma_j}{\gamma_j - \mu} \chi \star \left[e^-_{\mu} - e^-_{\gamma_j} \right] \tag{76}$$

‡ Along the same line, when considering time dependent production rate $q_j(t)$ while keeping all other parameters constant, we get an equation similar to (72): $\langle m_j \rangle_{X_{q_j}}(t) = (k/\mu) \left[q_j - q_j e^-_{\mu} \right] \star e^-_{\gamma_j}$.

with $\langle\langle m_j\rangle\rangle^* = \langle x_k\rangle^* k_1 q_j/(\mu\gamma_j)$. A simple calculation leads to the exact expression:

$$\begin{aligned}\frac{\langle\langle m_j\rangle\rangle(t)}{\langle\langle m_j\rangle\rangle^*} &= 1 + \frac{\mu}{\gamma_j-\mu}\frac{\alpha+\beta}{\alpha+\beta-\gamma_j}e^{-\gamma_j t} + \frac{\gamma_j}{\mu-\gamma_j}\frac{\alpha+\beta}{\alpha+\beta-\mu}e^{-\mu t} \\ &\quad - \frac{\mu}{\alpha+\beta-\mu}\frac{\gamma_j}{\alpha+\beta-\gamma_j}e^{-(\alpha+\beta)t}, \end{aligned} \tag{77}$$

as long as $\gamma_j \neq \mu$, $\mu \neq \alpha+\beta$ and $\gamma_j \neq \alpha+\beta$. Once again, the Laplace transform gives one single expression valid in all parameter space:

$$L\left[\frac{\langle\langle m_j\rangle\rangle}{\langle\langle m_j\rangle\rangle^*}\right](s) = \frac{1}{s}\frac{\alpha+\beta}{s+\alpha+\beta}\frac{\mu}{s+\mu}\frac{\gamma_j}{s+\gamma_j}. \tag{78}$$

This result is not new and could have alternatively been obtained by writing $d\langle\langle m_j\rangle\rangle/dt = q_j\langle\langle n\rangle\rangle - \gamma_j\langle\langle m_j\rangle\rangle$, which solution is $\langle\langle m_j\rangle\rangle = q\int_0^t \mathrm{d}s\ \langle\langle n\rangle\rangle(s)e^{-\gamma_j(t-s)}$ and using the time evolution of mRNA level $\langle\langle n\rangle\rangle$ (presented in [21]). Figure (9) shows agreement between analytical predictions and numerical simulations. If it is mathematically convenient to consider the time evolution starting from an "empty" initial state (all stochastic variable to zero), this situation does not seem to be biologically relevant. One could however, consider the similar scenario starting from the state $x_k = 0$, with initial numbers n, m_1, m_2 of A macromolecules and proteins. Solving this new problem requires a different approach based on an variation of the PPA mapping, which is not considered in this paper.

Finally, let us discuss how to infer on correlation numbers between protein types. First, with the help of

$$\langle m_j\rangle^{(2)}_{\sqcap\!\!\rightarrow}(t) = \frac{q_j}{\gamma_j}\left(e^{\gamma_j(s'-t)} - e^{\gamma_j(s-t)}\right), \tag{79}$$

we can use equation (55) to express the correlation function $C_{i,j|X_k} = \langle m_i m_j\rangle_{X_k}(t) - \langle m_i\rangle_{X_k}(t)\langle m_j\rangle_{X_k}(t)$:

$$\begin{aligned}\frac{C_{i,j|X_k}}{(q_i/\gamma_i)(q_j/\gamma_j)} &= k \star \left[\frac{\gamma_i+\gamma_j}{\gamma_i+\gamma_j-\mu}e^{-}_{\mu} - \frac{\gamma_i}{\gamma_i-\mu}e^{-}_{\gamma_j+\mu} - \frac{\gamma_j}{\gamma_j-\mu}e^{-}_{\gamma_i+\mu}\right. \\ &\quad \left. + \ \frac{\gamma_i}{\gamma_i-\mu}\frac{\gamma_j}{\gamma_j-\mu}\frac{\gamma_i+\gamma_j-2\mu}{\gamma_i+\gamma_j-\mu}e^{-}_{\gamma_i+\gamma_j}\right]. \end{aligned} \tag{80}$$

One should note that the variance is given by $C_{j,j|X_k}(t)+\langle m_j\rangle_{X_k}(t)$ and can be evaluated using equations (72) and (80). For singular cases $\gamma_i = \mu$, $\gamma_j = \mu$ or $\gamma_i+\gamma_j = \mu$ a similar expression can be derived from the previous equation taking the limit appropriately. To continue further one has to proceed more carefully. In fact the correlations in the model presented in figure 6 are defined by $\mathcal{C}_{i,j} = \langle\langle m_i m_j\rangle\rangle - \langle\langle m_i\rangle\rangle\langle\langle m_j\rangle\rangle$, which can be expressed using $\langle C_{i,j}\rangle$ (the average of $C_{i,j|X_k}$ over the history X_k):

$$\mathcal{C}_{i,j} = \langle C_{i,j}\rangle + \langle\langle m_i\rangle\langle m_j\rangle\rangle - \langle\langle m_i\rangle\rangle\langle\langle m_j\rangle\rangle. \tag{81}$$

One can notice that $\langle C_{i,j}\rangle$ (just like $\langle\langle m_i\rangle\rangle$ and $\langle\langle m_j\rangle\rangle$) is a functional of the mean $\langle x_k\rangle(t)$, and can be evaluated easily. The challenge comes from the term $\langle\langle m_i\rangle\langle m_j\rangle\rangle$ as

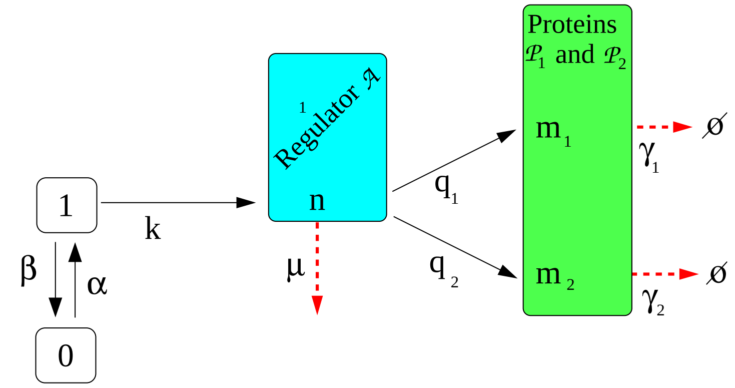

Figure 8. The three-stage model of gene expression, with arbitrary partition of proteins.

it requires knowledge of correlators $\langle x_k(s)x_k(s')\rangle$:

$$\langle\langle m_i\rangle\langle m_j\rangle\rangle = \frac{k_1^2 q_i q_j}{(\gamma_i-\mu)(\gamma_j-\mu)} \times \tag{82}$$

$$\int_0^t \mathrm{d}s \int_0^t \mathrm{d}s' \, \langle x_k(s)x_k(s')\rangle (e^{-\mu(t-s)} - e^{-\gamma_i(t-s)})(e^{-\mu(t-s')} - e^{-\gamma_j(t-s')}).$$

In fact the x_k-correlation can be calculated exactly. For $s < s'$, it is given by

$$\langle x_k(s)x_k(s')\rangle = \left(\frac{\alpha}{\alpha+\beta}\right)^2 \left(1 - e^{-(\alpha+\beta)s}\right)\left(1 + \frac{\beta}{\alpha} e^{-(\alpha+\beta)(s'-s)}\right). \tag{83}$$

This last expression, together with the help of (80) and (77) allow for the evaluation of $\mathcal{C}_{i,j}$. To compare correlations $\mathcal{C}_{i,j}$ for the three-stage model with $C_{i,j}$ (61) for the two-stage model, we impose the equality $\alpha k_1 = (\alpha+\beta)k$ which assures, in both models, identical regulator, and proteins levels. It follows that $\mathcal{C}^*_{i,j}/C^*_{i,j} = 1+R$, where R has a cumbersome expression, dependent on all parameters but q_i and q_j. It vanishes for $\beta = 0$ (as well as $k = 0$) and satisfies $R > 0$ for all other finite parameter values. Hence we conclude that, in the stationary state, for identical regulator and protein levels, correlations between protein numbers are higher in the model with promoter-based regulation: $\mathcal{C}^*_{i,j} > C^*_{i,j}$. Restraining ourself to homogeneous degradation rates ($\gamma_i = \gamma_j = \gamma$) the expression for R is more manageable :

$$R = \frac{\beta}{\alpha+\beta}\frac{k_1(\alpha+\beta+\mu+\gamma)}{(\alpha+\beta+\mu)(\alpha+\beta+\gamma)}. \tag{84}$$

Once again, we note that R and $\mathcal{C}^*_{i,j}$ are invariant under the exchange $\mu \leftrightarrow \gamma$. Figure (9) shows a comparison between the time evolution of protein number in the two and three stage model. Our data validate the equality $\mathcal{C}^*_{i,j}/C^*_{i,j} = 1+R$ and seem to indicate that $\mathcal{C}_{i,j}(t)/C_{i,j}(t) \simeq 1+R$ is an acceptable approximation, at least for the set of parameter selected.

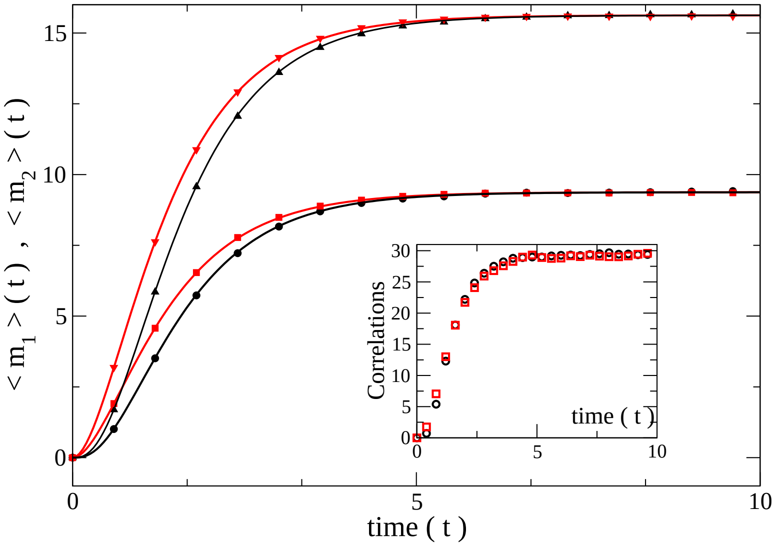

Figure 9. Comparison of the time evolution of the mean numbers $\langle m_1 \rangle(t)$, $\langle m_2 \rangle(t)$ for the two stage model (in red - for parameter values $k = 6.25$, $\mu = 1$, $q_1 = 3$, $q_2 = 5$, $\gamma_1 = \gamma_2 = 2$) and three state model (in black - for parameter values $\alpha = 2.5$, $\beta = 1.5$, $k_1 = 10$, $\mu = 1$, $q_1 = 3$, $q_2 = 5$, $\gamma_1 = \gamma_2 = 2$). Analytical results (lines) are compared to numerical results (symbols). In the insert, we compare correlations between protein numbers for the two stage model (in red) and three stage model (in black). For the same parameter values, we plot the time evolution of $(1+R)C_{1,2}(t)$ (red) together with $\mathcal{C}_{1,2}(t)$ (black).

4. Conclusion

In this paper, we present a variation on the two-stage model including the arbitrary partition of protein and arbitrary time dependent parameters. The mechanism considered is similar to the one involved in alternative splicing [49]. Our aim was to develop an analytical approach leading to the time dependent generating function, based on the PPA-mapping, and which did not require knowledge of already known results. Constructed on a succession of transformations, our work simplifies the original problem to the analysis of two-states biological switches. As a consequence, a series of different reduced models emerges, with clear relations linking their respective generating functions. In particular we show how the kernel $\mathcal{K}_j$ plays a important role in the final expression of the generating function. We show that the emerging hierarchy allows us to connect mean numbers and higher order moments between models. This leads to an explicit relation between correlation $C_{i,j}$ in the original model and mean protein number in model-2. For constant parameter values, we derived the exact time dependent expression of correlation $C_{i,j}$. We note that particular cases such as $\gamma_j = \mu$, $\gamma_j = \mu/2$ or $\gamma_1 + \gamma_2 = \mu$ have to be treated separately. However, in the stationary state this distinction vanishes as each function converges towards a common asymptotic expression. Considering constant parameters, for $J = 1$ or for homogeneous degradation rates ($\gamma_j = \gamma\ \forall j$), we show how our results reduce to the solution of two stage model.

Finally, we show how results for arbitrary time dependent transition rates can be used to study models presenting parameter fluctuations induced by hidden random variable. We give an explicit example by extending our results for the mean and correlation to the three state model. In particular we show that promoter based regulation leads to higher protein-protein correlations $\mathcal{C}^*_{i,j} > C^*_{i,j}$. The method presented here is applicable in other scenarios (with zero and first order reactions) and may be use as a guide in the study of other biological systems. We hope this methodology will contribute to the development of new analytical avenues for future research.

Acknowledgements

The authors would like to thanks the AMRC in Coventry and the Stat. Phys. Group in Nancy for their support and constant efforts to protect curiosity driven research. In particular, T. Platini extends his acknowledgements to K. Aujogue, R. Low, S. Mukherjee, S. Vantieghem and R. Kulkarni.

5. Appendix

5.1. From model-0 to model-1: a detailed derivation

In the following we aim to construct the original problem (model-0) starting with N independent models, all identical to model-1. The derivation is presented for constant parameter, but could easily be extended to time dependent parameter. We will start with the definition of $\phi_{N;(\theta,a,b)}(t)$: the probability of finding, in model-1, the biological switch in the state θ together with a and b proteins of type $\mathcal{P}_1$ and $\mathcal{P}_2$ respectively. We continue by defining $\Phi_N(n, m_1, m_2)$ as the probability of finding amongst N independent identical models (each labeled with subscript ν) the total numbers of active switches $n = \sum_\nu \theta_\nu$ and proteins $m_1 = \sum_\nu a_\nu$, $m_2 = \sum_\nu b_\nu$. Finally when, taking the limit $N \to \infty$, we will be able to show that $\lim_{N\to\infty} \Phi_N(t)$ obeys the master equation (1). In other words $P_{n,m_1,m_2} = \lim_{N\to\infty} \Phi_{N;(n,m_1,m_2)}$.

To start let us try to keep notation as compact as possible by defining the flip operator $\hat{F}$ as the operator acting on the triplet (θ, a, b) and such that $\hat{F}(\theta, a, b) = (1-\theta, a, b)$. In addition, we write $\hat{A}^\pm$ and $\hat{B}^\pm$ as the operators defined by

$$\hat{A}^\pm(\theta, a, b) = (\theta, a \pm 1, b), \tag{85}$$

$$\hat{B}^\pm(\theta, a, b) = (\theta, a, b \pm 1). \tag{86}$$

Let $\phi_{N;(\theta,a,b)}(t)$ be the probability distribution associated to model-1. It obeys the following equation

$$\begin{aligned}\frac{d\phi_{N;(\theta,a,b)}}{dt} =& \ [\theta(k/N) + (1-\theta)\mu]\, \phi_{N;\hat{F}(\theta,a,b)} \\ &+ \theta q_1 \phi_{N;\hat{A}^-(\theta,a,b)} + \theta q_2 \phi_{N;\hat{B}^-(\theta,a,b)} \\ &+ (a+1)\gamma_1 \phi_{N;\hat{A}^+(\theta,a,b)} + (b+1)\gamma_2 \phi_{N;\hat{B}^+(\theta,a,b)} \\ &- \ [(1-\theta)(k/N) + \theta\mu + \theta q_1 + \theta q_2 + a\gamma_1 + b\gamma_2]\, \phi_{N;(\theta,a,b)}.\end{aligned} \tag{87}$$

Considering N independent but identical models, we define $\mathbb{S}$ as the set of triplets: $\mathbb{S} = \{(\theta_\nu, a_\nu, b_\nu); \nu = 1, 2, \ldots, N\}$. Note that $\mathbb{S}$ gives a full description of the state of all N independent reduced models. In addition, we will write $\theta_\nu(\mathbb{S})$, $a_\nu(\mathbb{S})$ and $b_\nu(\mathbb{S})$ the variable θ, a and b in the ν^{th} triplet of the set $\mathbb{S}$. The definitions of the operator $\hat{F}$, $\hat{A}^\pm$ and $\hat{B}^\pm$ are extended onto the set $\mathbb{S}$. So that $\hat{F}_\nu$, $\hat{A}^\pm_\nu$ and $\hat{B}^\pm_\nu$ act on the ν^{th} triplet of the set $\mathbb{S}$, leaving all others unchanged. We can now define the overall probability $\phi_{N,\mathbb{S}} = \prod_{\nu=1}^N \phi_{N,(\theta_\nu,a_\nu,b_\nu)}$, which evolution is governed by

$$\frac{d}{dt}\phi_{N,\mathbb{S}} = \sum_{i=1}^N \prod_{\nu \neq i} \phi_{N,(\theta_\nu,a_\nu,b_\nu)} \frac{d}{dt}\phi_{N,(\theta_i,a_i,b_i)}. \tag{88}$$

With a little bit of effort, the latter equation leads to

$$\begin{aligned} \frac{d}{dt}\phi_{N,\mathbb{S}} = k/N \sum_i &\left\{\theta_i(\mathbb{S})\phi_{N,\hat{F}_i\mathbb{S}} - [1-\theta_i(\mathbb{S})]\phi_{N,\mathbb{S}}\right\} \\ + \mu \sum_i &\left\{[1-\theta_i(\mathbb{S})]\phi_{N,\hat{F}_i\mathbb{S}} - \theta_i(\mathbb{S})\phi_{N,\mathbb{S}}\right\} \\ + q_1 \sum_i &\theta_i(\mathbb{S})\left\{\phi_{N,\hat{A}_i^-\mathbb{S}} - \phi_{N,\mathbb{S}}\right\} \\ + q_2 \sum_i &\theta_i(\mathbb{S})\left\{\phi_{N,\hat{B}_i^-\mathbb{S}} - \phi_{N,\mathbb{S}}\right\} \\ + \gamma_1 \sum_i &\left\{[a_i(\mathbb{S})+1]\phi_{N,\hat{A}_i^+\mathbb{S}} - a_i(\mathbb{S})\phi_{N,\mathbb{S}}\right\} \\ + \gamma_2 \sum_i &\left\{[b_i(\mathbb{S})+1]\phi_{N,\hat{B}_i^+\mathbb{S}} - b_i(\mathbb{S})\phi_{N,\mathbb{S}}\right\}. \end{aligned} \tag{89}$$

Moving forward, we define the probability distribution Φ_N as

$$\Phi_{N;(n,m_1,m_2)} = \sum_{\mathbb{S}} \phi_{N;\mathbb{S}} \Delta_n^{\{\theta_\nu\}} \Delta_{m_1}^{\{a_\nu\}} \Delta_{m_2}^{\{b_\nu\}}, \tag{90}$$

where the constraint Δ is defined by

$$\Delta_y^{\{x_\nu\}} = \delta\left(\sum_\nu x_\nu, y\right), \tag{91}$$

where δ is the Kronecker symbol. To keep notations as compact as possible we write

$\Delta^{(3)} = \Delta_n^{\{\theta_\nu\}}\Delta_{m_1}^{\{a_\nu\}}\Delta_{m_2}^{\{b_\nu\}}$. The master equation for the distribution Φ_N is

$$\begin{aligned}\frac{d}{dt}\Phi_{N,(n,m_1,m_2)} &= k/N\sum_{\mathbb{S}}\Delta^{(3)}\sum_i\left\{\theta_i(\mathbb{S})\phi_{N,\hat{F}_i\mathbb{S}} - [1-\theta_i(\mathbb{S})]\phi_{N,\mathbb{S}}\right\} \\ &+ \mu\sum_{\mathbb{S}}\Delta^{(3)}\sum_i\left\{[1-\theta_i(\mathbb{S})]\phi_{N,\hat{F}_i\mathbb{S}} - \theta_i(\mathbb{S})\phi_{N,\mathbb{S}}\right\} \\ &+ q_1\sum_{\mathbb{S}}\Delta^{(3)}\sum_i\left\{\theta_i(\mathbb{S})\phi_{N,\hat{A}_i^-\mathbb{S}} - \theta_i(\mathbb{S})\phi_{N,\mathbb{S}}\right\} \\ &+ q_2\sum_{\mathbb{S}}\Delta^{(3)}\sum_i\left\{\theta_i(\mathbb{S})\phi_{N,\hat{B}_i^-\mathbb{S}} - \theta_i(\mathbb{S})\phi_{N,\mathbb{S}}\right\} \\ &+ \gamma_1\sum_{\mathbb{S}}\Delta^{(3)}\sum_i\left\{[a_i(\mathbb{S})+1]\phi_{N,\hat{A}_i^+\mathbb{S}} - a_i(\mathbb{S})\phi_{N,\mathbb{S}}\right\} \\ &+ \gamma_2\sum_{\mathbb{S}}\Delta^{(3)}\sum_i\left\{[b_i(\mathbb{S})+1]\phi_{N,\hat{B}_i^+\mathbb{S}} - b_i(\mathbb{S})\phi_{N,\mathbb{S}}\right\}.\end{aligned} \tag{92}$$

Every sum, for which $\phi_{N;\mathbb{S}}$ appears explicitly can be easily evaluated. As an example we give here the second term of the first line in the previous equation. Using the constraint $\Delta^{\{\theta_\nu\}}$, which impose $\sum_\nu \theta_\nu = n$, we have

$$\begin{aligned}\sum_{\mathbb{S}}\Delta_n^{\{\theta_\nu\}}\Delta_{m_1}^{\{a_\nu\}}\Delta_{m_2}^{\{b_\nu\}}\sum_i[1-\theta_i(\mathbb{S})]\phi_{N,\mathbb{S}} &= \sum_{\mathbb{S}}\Delta_n^{\{\theta_\nu\}}\Delta_{m_1}^{\{a_\nu\}}\Delta_{m_2}^{\{b_\nu\}}[N-n]\phi_{N,\mathbb{S}} \\ &= (N-n)\Phi_{N,(n,m_1,m_2)}.\end{aligned} \tag{93}$$

When $\phi_{N,\mathbb{S}}$ does not appear explicitly we need to re-labeled the sum over $\mathbb{S}$. As an example we present details to the calculation of the first term of the first line in which we have $\phi_{N,\hat{F}_i\mathbb{S}}$. Defining $\tilde{\mathbb{S}} = \hat{F}_i\mathbb{S}$ we have $\hat{F}_i\tilde{\mathbb{S}} = \mathbb{S}$ so that

$$\theta_j(\mathbb{S}) = \theta_j(\hat{F}_i\tilde{\mathbb{S}}) = \begin{cases} \theta_\nu(\tilde{\mathbb{S}}), & \text{if} \quad i \neq \nu \\ 1-\theta_i(\tilde{\mathbb{S}}), & \text{if} \quad i = \nu. \end{cases} \tag{94}$$

In relabelling $\mathbb{S}$ to $\tilde{\mathbb{S}}$ the expression of $\Delta^{\{\theta_\nu\}}$ has changed. To keep track of this change we will replace $\Delta^{\{\theta_\nu\}} \to \tilde{\Delta}^{\{\theta_\nu\}}$ where

$$\Delta_n^{\{\theta_\nu\}} = \delta\left(\sum_\nu \theta_\nu(\mathbb{S}), n\right) = \delta\left(\sum_\nu \theta_\nu(\tilde{\mathbb{S}}) + 1 - 2\theta_i(\tilde{\mathbb{S}}), n\right) = \tilde{\Delta}_n^{\{\theta_\nu\}} \tag{95}$$

It follows that

$$\sum_{i=1}^N\sum_{\mathbb{S}}\Delta_n^{\{\theta_\nu\}}\Delta_{m_1}^{\{a_\nu\}}\Delta_{m_2}^{\{b_\nu\}}\theta_i(\mathbb{S})\phi_{N,\hat{F}_i\mathbb{S}} = \sum_{i=1}^N\sum_{\tilde{\mathbb{S}}}\tilde{\Delta}_n^{\{\theta_\nu\}}\Delta_{m_1}^{\{a_\nu\}}\Delta_{m_2}^{\{b_\nu\}}[1-\theta_i(\tilde{\mathbb{S}})]\phi_{N,\tilde{\mathbb{S}}}. \tag{96}$$

Note that the only elements which will contribute are such that $\theta_i(\tilde{\mathbb{S}}) = 0$, which allows us to write

$$\tilde{\Delta}_n^{\{\theta_j\}} = \delta\left(\sum_\nu \theta_\nu(\tilde{\mathbb{S}}), (n-1)\right). \tag{97}$$

Finally, we are able to express the first summation in term of $\Phi_{N,(n,m_1,m_2)}$:

$$\sum_{i=1}^{N}\sum_{\mathbb{S}} \Delta_n^{\{\theta_\nu\}}\Delta_{m_1}^{\{a_\nu\}}\Delta_{m_2}^{\{b_\nu\}}\theta_i(\mathbb{S})\phi_{N,\hat{F}_i\mathbb{S}} = [N-(n-1)]\phi_{N,(n-1,m_1,m_2)}. \tag{98}$$

Proceeding along the same line for every summation symbol we have

$$\begin{aligned}\frac{d}{dt}\Phi_{N,(n,m_1,m_2)} &= k\left[1-\frac{(n-1)}{N}\right]\Phi_{N,(n-1,m_1,m_2)} - k\left[1-\frac{n}{N}\right]\Phi_{N,(n,m_1,m_2)}\\ &+ \mu(n+1)\Phi_{N,(n+1,m_1,m_2)} - \mu n\Phi_{N,(n,m_1,m_2)} \qquad (99)\\ &+ q_1 n\Phi_{N,(n,m_1-1,m_2)} - q_1 n\Phi_{N,(n,m_1,m_2)}\\ &+ q_1 n\Phi_{N,(n,m_1,m_2-1)} - q_2 n\Phi_{N,(n,m_1,m_2)}\\ &+ \gamma_1[m_1+1]\Phi_{N,(n,m_1+1,m_2)} - \gamma_1 m_1\Phi_{N,(n,m_1,m_2)}\\ &+ \gamma_2[m_2+1]\Phi_{N,(n,m_1,m_2+1)} - \gamma_2 m_2\Phi_{N,(n,m_1,m_2)}.\end{aligned}$$

Taking the limit $N\to\infty$, we see that the latter equation converges towards the master equation (1). In other words $P_{n,m_1,m_2} = \lim_{N\to\infty}\Phi_{N;(n,m_1,m_2)}$, from which it naturally follow $G^{(0)} = \lim_{N\to\infty}\left(G_N^{(1)}\right)^N$.

5.2. Large time approximation

The large time approximation, for one protein type only, is obtained by writting $r=\mu/\gamma$, $\delta = q(z-1)/\gamma$ and $\epsilon = e^{-\gamma t}$ so that

$$(z-1)A = \delta\frac{k}{\gamma}\int_\epsilon^1 \mathrm{d}u\int_u^1 \mathrm{d}v\ \Omega(u,v), \tag{100}$$

with

$$\Omega(u,v) = \frac{u^{r-1}}{v^r}\exp\left[\delta(z)(v-u)\right]. \tag{101}$$

Using the Taylor expansion leads to

$$(z-1)A = \delta\frac{k}{\gamma}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^m\delta^{n+m}}{m!n!}\int_\epsilon^1 \mathrm{d}u\ u^{m+r-1}\int_u^1 \mathrm{d}v\ v^{n-r}, \tag{102}$$

for which there is no particular problem unless in the last integral we have $n-r=-1$ for some value of n. Avoiding this situation, by choosing $r\notin\mathbb{N}$, leads to

$$\begin{aligned}(z-1)A = I_F + \delta\frac{k}{\gamma}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}&\frac{(-1)^m\delta^{n+m}}{m!n!}\frac{1}{n-r+1}\\ &\times\left[\frac{1}{n+m+1}\epsilon^{n+m+1} - \frac{1}{m+r}\epsilon^{m+r}\right],\end{aligned} \tag{103}$$

with

$$\begin{aligned}I_F &= \delta\frac{k}{\gamma}\sum_{\kappa=0}^{\infty}\frac{\delta^\kappa}{\kappa+1}\sum_{m=0}^{\kappa}\frac{(-1)^m}{m!(\kappa-m)!}\frac{1}{(m+r)} = \delta\frac{k}{\gamma}\sum_{\kappa=0}^{\infty}\frac{\delta^\kappa}{\kappa+1}\frac{(r-1)!}{(r+\kappa)!}\\ &= \frac{k}{\mu}\int_0^\delta \mathrm{d}s\ {}_1F_1[1,r+1,s].\end{aligned} \tag{104}$$

Keeping the lowest order in ϵ we get $(z-1)A \simeq I_F + h$ with

$$h = \begin{cases} \delta \frac{k}{\gamma} \frac{\epsilon}{1-r} & r > 1 \\ -\delta \frac{k}{\gamma} \frac{\epsilon^r}{r} \sum_{n=0}^{\infty} \frac{\delta^n}{n!} \frac{1}{n-r+1} & r < 1. \end{cases} \tag{105}$$

It follows that $G^{(0)}(z,t) \simeq G^*(z)H(z,t)$, with $H = e^h$ as presented in equation (65). Going back to equation (102), we can work with $r = 1$ ($\mu = \gamma$). In this case, when keeping terms of order ϵ and $\epsilon \ln(\epsilon)$ we get :

$$h = -\delta \frac{k}{\mu} e^{-\mu t} \left(\mu t + \sum_{n=1}^{\infty} \frac{\delta^n}{n \times n!} \right). \tag{106}$$

Under the following approximation we get $\langle m \rangle(t)/\langle m \rangle^* \simeq 1 - \mu t e^{-\mu t}$ in agreement with the long time limit of equation (5).

6. Bibliography

[1] Ko M S, Nakauchi H, and Takahashi N 1990 The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates *EMBO J.* **9** 2835-2842

[2] Ko M S 1991 A stochastic model for gene induction *J. Theor. Biol.* **153** 181-194

[3] Raj A and van Oudenaarden A 2008 Nature, nurture, or chance: Stochastic gene expression and its consequences *Cell* **135** 216-226

[4] Larson D R, Singer R H, and Zenklusen D 2009 A single molecule view of gene expression *Trends Cell Biol.* **19** 630

[5] Huang S 2009 Non-genetic heterogeneity of cells in development: more than just noise *Development* **136** 3853-62

[6] Eldar A and Elowitz M B 2010 Functional roles for noise in genetic circuits *Nature* **467** 167

[7] Lionnet T and Singer R H 2012 Transcription goes digital *EMBO reports* **13**(4) 313-21

[8] Van Kampen N G 2007 Stochastic Processes in Physics and Chemistry, 3rd Edition North Holland.

[9] Lobner-Olesen A 1999 Distribution of minichromosomes in individual Escherichia coli cells: implications for replication control *EMBO J.* **18**(6) 1712-21

[10] Becskei A, Seraphin B and Serrano L 2001 Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion *EMBO J.* **20**(10) 2528-35

[11] Elowitz M B, Levine A J, Siggia E D and Swain P S 2002 Stochastic Gene Expression in a Single Cell *Science* **297** 1183

[12] Ozbudak E M, Thattai M, Kurtser I, Grossman A D and van Oudenaarden A 2002 Regulation of noise in the expression of a single gene *Nature Genetics* **31** 69 - 73

[13] Blake W J, Kaern M, Cantor C R, Collins J J 2003 Noise in eukaryotic gene expression *Nature* **422** 633-637

[14] Acar M, Becskei A and van Oudenaarden A 2005 Enhancement of cellular memory by reducing stochastic transitions *Nature* **435** 228-232

[15] Balazsi G, van Oudenaarden A and Collins J J 2011 Cellular decision making and biological noise: from microbes to mammals *Cell* **144**(6) 910-25

[16] Nevozhay D, Adams R M, Murphy K F, Josic K and Balazsi G 2009 Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression *PNAS* **106** (13) 5123-8

[17] Golding I, Paulsson J, Zawilski S M and Cox E C 2005 Real-time kinetics of gene activity in individual bacteria *Cell* **123** (6) 1025-1036

[18] Raj A, Peskin C S, Tranchina D, Vargas D Y and Tyagi S 2006 Stochastic mRNA synthesis in mammalian cells *PLoS Biol* **4**(10) e309

[19] Taniguchi Y, Choi P J, Li G W, Chen H, Babu M, Hearn J, Emili A and Xie X S 2010 Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells *Science* **329** 533

[20] Ferguson M, Le Coq D, Jules M, Aymerich S, Radulescu O, Declerck N and Royer CA 2012 Reconciling molecular regulatory mechanisms with noise patterns of bacterial metabolic promoters in induced and repressed states *PNAS* **109** 155

[21] Peccoud J, and Ycart B 1995 Markovian modeling of gene-product synthesis *Theoret. Popul. Biol.* **48** 222

[22] Hasty J, Pradines J, Dolnik M and Collins J J 2000 Noise-based switches and amplifiers for gene expression *PNAS* **97** (5) 2075-2080

[23] Paulsson J 2004 Summing up the noise in gene networks *Nature (London)* **427** 415

[24] Karmakar R and Bose I, 2004 Graded and binary responses in stochastic gene expression *Physical Biology* **1** (4)

[25] Hornos J E M, Schultz D, Innocentini G C, Wang J, Walczak A M, Onuchic J N and Wolynes P G 2005 Self-regulating gene: an exact solution *Phys. Rev. E* **72** 051907

[26] Friedman N, Cai L and Xie X S 2006 Linking stochastic dynamics to population distribution: an analytical framework of gene expression *Phys. Rev. Lett.* **97** 168302

[27] Okabe Y, Yagi Y and Sasai M 2007 Effects of the DNA state fluctuation on single-cell dynamics of self-regulating gene *J. Chem. Phys.* **127** 105107

[28] Ramos A F, Innocentini G C and Hornos J E 2011 Exact time-dependent solutions for a self-regulating gene *Phys. Rev. E* **83** 062902

[29] Zhang J, Chen L and Zhou T 2012 Analytical distribution and tunability of noise in a model of promoter progress *Biophysical Journal* **102** (6) 1247-1257

[30] McAdams H and Arkin A 1997 Stochastic mechanisms in gene expression *PNAS* **94** 814-819

[31] Coulon A, Gandrillon A, and Beslon G 2010 On the spontaneous stochastic dynamics of a single gene: complexity of the molecular interplay at the promoter *BMC System Biology* **4** 2

[32] Kepler T B and Elston T C 2001 Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations *Biophys J.* **81** (6) 3116-3136

[33] Paulsson J 2005 Models of Stochastic Gene Expression *Phys. Life Rev.* **2** 157-175

[34] Paszek P 2007 Modeling stochasticity in gene regulation: characterization in the terms of the underlying distribution function *Bull Math Biol.* **69** (5) 1567-601

[35] Shahezaei V and Swain P S 2008 Analytical distributions for stochastic gene expression *PNAS* **105** (45) 17256-17261

[36] Iyer-Biswas S, Hayot F and Jayaprakash C 2009 Stochasticity of gene products from transcriptional pulsing *Phys. Rev. E* **79** 031911

[37] Bokes P, King J R, Wood A T and Loose M 2012 Exact and approximate distributions of protein and mRNA levels in the low-copy regime of gene expression *J. Math. Biol.* **64** 829-854

[38] Pedraza J M and Paulsson J 2008 Effects of molecular memory and bursting on fluctuations in gene expression *Science* **319** 339

[39] Stinchcombe A R, Peskin C S and Tranchina D 2012 Population density approach for discrete mRNA distributions in generalized switching models for stochastic gene expression *Phys. Rev. E* **85** 061919

[40] Thattai M and van Oudenaarden A 2001 Intrinsic noise in gene regulatory networks *PNAS* **98** (15) 8614-8619

[41] Xu B L and Tao Y 2006 External noise and feedback regulation: steady-state statistics of auto-regulatory genetic network *J. Theor. Biol.* **243** 214

[42] Kumar N, Platini T and Kulkarni R V 2014 Exact distributions for stochastic gene expression models with bursting and feedback *Phys. Rev. Lett.* **113** 268105

[43] Yu J, Xiao J, Ren X, Lao K and Xie X S 2006 Probing gene expression in live cells, one protein

molecule at a time *Science* **311**

[44] Suter D M, Molina N, Gatfield D, Schneider K, Schibler U and Naef F 2011 Mammalian genes are transcribed with widely different bursting kinetics *Science* **332** 472

[45] Piatkevich K D, Lionnet T, Singer R H and Verkhusha V V 2011 Modern fluorescent proteins and imaging technologies to study gene expression, nuclear localization, and dynamics *Curr. Opin. Cell. Biol.* **23** 310

[46] Lionnet T, Czaplinski K, Darzacq X, Shav-Tal Y, Wells A L, Chao J A, Park H Y, de Turris V, Lopez-Jones M and Singer RH 2011 A transgenic mouse for in vivo detection of endogenous labeled mRNA *Nat. Methods* **8** 165

[47] Pendar H, Platini T and Kulkarni R V 2013 Exact protein distributions for stochastic models of gene expression using partitioning of Poisson processes *Phys. Rev. E* **87** 042720

[48] Kulkarni R 2014 *private communication*

[49] Wang Q and Zhou T 2014 Alternative-splicing-mediated gene expression *Phys. Rev. E* **89** 012713

[50] Modrek B and Lee C 2002 A genomic view of alternative splicing. *Nat Genet.* (1):13-9

[51] Black D L 2000 Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology *Cell* **103** (3) 367-70

[52] Graveley B R 2001 Alternative splicing: increasing diversity in the proteomic world *Trends Genet.* (2): 100-7

[53] Grabowski P J and Black D L 2001 Alternative RNA splicing in the nervous system *Prog. Neurobiol.* **65**(3):289-308

[54] Black D L Mechanisms of alternative pre-messenger RNA splicing 2003 *Annu Rev Biochem.* **72**: 291-336

[55] Lesta I, et al., 2008 *IEEE T Circuits-I* **53** 189-200

[56] Gadgil C, Lee C H and Othmer H G 2005 A stochastic analysis of first-order reaction networks *Bull Math Biol* **67**: 901-946

[57] Singh A and Hespanha J P 2007 A derivative matching approach to moment closure for the stochastic logistic model *Bull Math Biol* **69** 1909-1925

[58] Mugler A, Walczak A M and Wiggins C H 2010 Information-optimal transcriptional response to oscillatory driving *Phys. Rev. Lett* **105** 058101

[59] Jedrak J and Ochab-Marcinek A 2016 Time-dependent solutions for a stochastic model of gene expression with molecule production in the form of a compound Poisson process *Phys. Rev. E* **94** 032401

[60] Dattani J and Barahona M 2017 Stochastic models of gene transcription with upstream drives: exact solution and sample path characterization *J. R. Soc. Interface* **14** (126) 20160833

[61] Jahnke T and Huisinga W 2007 Solving the chemical master equation for monomolecular reactions systems analytically *J of Math Biol* **54** 1 1-26

[62] Anderson D F, Craciun G and Kurtz T G 2010 Product-form stationary distributions for deficiency zero chemical reaction networks *Bull of Math Biol* **72** 8 1947-1970

[63] Anderson D F and Cotter S L 2010 Product-form stationary distributions for deficiency zero networks with non-mass action kinetics *Bull of Math Biol* **78** 12 2390-2407