Revisiting the reduction of stochastic models of genetic feedback loops with fast promoter switching

J. Holehouse, R. Grima

April 9, 2019

Abstract

Propensity functions of the Hill-type are commonly used to model transcriptional regulation in stochastic models of gene expression. This leads to an effective reduced master equation for the mRNA and protein dynamics only. Based on deterministic considerations, it is often stated or tacitly assumed that such models are valid in the limit of rapid promoter switching. Here, starting from the chemical master equation describing promoter-protein interactions, mRNA transcription, protein translation and decay, we prove that in the limit of fast promoter switching, the probability distribution of protein numbers is different than that given by standard stochastic models with Hill-type propensities. We show that the differences are pronounced whenever the protein-DNA binding rate is much larger than the unbinding rate, a special case of fast promoter switching. Furthermore we show using both theory and simulations that use of the standard stochastic models leads to drastically incorrect predictions for the switching properties of positive feedback loops and that these differences decrease with increasing mean protein burst size. Our results confirm that commonly used stochastic models of gene regulatory networks are only accurate in a subset of the parameter space consistent with rapid promoter switching.

1 Introduction

Many biochemical systems have one or more species with low molecule numbers which implies that the dynamics can be highly noisy and consequently a deterministic description may not be accurate [1–4]. Rather a more appropriate mathematical description is stochastic and given by the chemical master equation [5]. When the system is made up of zero and first order reactions only, exact solutions at both steady state and in time are occasionally possible [6]. However many systems have at least one bimolecular reaction and in such cases only a few exact solutions of the CME are known (see for example [7, 8]). A common example of such systems are auto-regulatory feedback loops, whereby a protein produced by a gene binds to its own promoter region to activate or suppress its own production [9–11]. In the absence of exact solutions, we become either reliant: (i) on the stochastic simulation algorithm (SSA) [12] or (ii) on approximations of the original network so that analytic results become tractable [13–15]. Generally, it is a challenge to utilise approximations to simplify the CME such that the resulting reduced equation is representative of the true system dynamics.

A common set of approximation methods are based on timescale separation. At the microscopic level, there are several different scenarios which can lead to timescale separation conditions. Depending on their firing rate, reactions can be classified as either slow or fast. Depending on the reaction system, it is possible that fast and slow reactions do not involve the same species but more commonly fast and slow reactions share some species and hence it is generally unclear what should be considered a fast or a slow species. Methods in the literature differ according to the definition of what is a slow or fast species. Zeron and Santillan [16] assume that fast species are only involved in fast reactions whereas slow species can participate in both slow and fast reactions. In contrast, Cao et al. [17] define slow species as those involved in slow reactions only and fast species as those participating in at least one fast reaction and any number of slow reactions. These two approaches lead to a reduced CME description in the slow species only. Other approaches due to Haseltine and Rawlings [18] and Goutsias [19] model the state of

the system using extents of reaction as opposed to molecules of species. A singular perturbation theory based method has also recently been used to obtain a reduced stochastic description [20].

Despite the wide breadth of rigorous approaches, by far the most popular approach in the literature to obtain a reduced master equation is heuristic. The key idea is to use the results of timescale separation for deterministic kinetics. Under the quasi-steady state or fast equilibrium approximations, the mean concentration of a subset of species (the fast species) reaches steady-state on a much shorter time scale than the rest of the species (the slow species). Using the deterministic rate equations it is then possible to express the concentration of the fast species in terms of the concentration of the slow species. This leads to a reduced chemical system composed of effective reactions with non-mass action kinetics describing the dynamics of the slow species. The reduced chemical master equation is then obtained by writing effective propensities analogous to the non-mass action reaction rates obtained from the deterministic analysis. For example, Hill-type effective protein production rates in the deterministic rate equations result if the gene equilibrates on a much faster timescale than mRNA and protein, i.e the fast promoter switching limit (see for example [21] for experimental evidence of this limit), and hence by analogy, Hill-type propensities for the protein production rates are commonly used in stochastic simulations of gene regulatory networks [22–29]. All of these studies and many others assume that such effective propensities are justified in the limit of fast promoter switching.

The advantage of this heuristic approach is its simplicity and ease of use and this is the main reason for its widespread use. However clearly the use of a reduced master equation obtained from deterministic considerations is doubtful. This has led to a number of studies evaluating the accuracy of these reduced master equations. Thomas et al. in a series of papers [30–32] showed using Langevin approximation theory that in the limit of large molecule numbers and for parameters consistent with the quasi-steady state approximation, the mean number of molecules of slow species predicted by the reduced master equation agrees with that predicted from the master equation of the full system but the variance of molecule number fluctuations does not. Similar results have been shown using stochastic simulations by Kim et al. (in particular see Fig. 1 of [33]). In contrast, Bundschuh et al. [34] have shown that the SSA corresponding to the heuristic reduced master equation of a negative feedback loop, whereby the DNA-protein binding reactions are assumed fast compared to the rest of the reactions and hence are eliminated, is in very good agreement with the SSA of the full system for parameter values specific to the phage λ system (this case is referred to as "Michaelis-Menten system" in their paper). At first sight the results of Thomas et al. and Bundschuh et al. may appear contradictory but in reality they are not: while the results of Thomas et al. prove that the heuristic approach of obtaining reducing master equations cannot be considered equivalent to the stochastic version of the quasi-steady state approximation (see also [35]), nevertheless it is possible that the error in the predictions of the heuristic approach are small for specific parameter values which would be consistent with the results of Bundschuh et al. What is clear from these studies is that more work is needed to identify the precise regions of parameter space where the heuristic reduced master equation of gene regulatory networks can be safely used – the study reported in this article identifies such regions and hence fills a gap in the literature.

The structure of this article is as follows. In Section 2 we obtain the steady-state solution of the heuristic reduced master equation with Hill-type protein production propensities for a non-bursty genetic feedback loop and prove that it is different than the solution of the master equation of the non-bursty genetic feedback loop in the limit of fast promoter switching. It is then shown that the differences between the probability distributions of protein numbers predicted by the two master equations tend to zero when the rate of DNA-protein binding is much smaller than the unbinding rate. In contrast, the differences maximize when the rate of DNA-protein binding is much larger than the unbinding rate. The results are confirmed by stochastic simulations across large swaths of parameter space. In Section 3 we extend the analysis to bursty feedback loops. We finish with Section 4, providing a discussion and then by concluding our results.

2 Model reduction for non-bursty feedback loops

2.1 Deterministic description and reduction

The reaction scheme for a genetic non-bursty feedback loop is given by:

$$G \xrightarrow{\rho_u} G + P,\ P \xrightarrow{1} \emptyset,\ P + G \underset{\sigma_u}{\overset{\sigma_b}{\rightleftharpoons}} G^*,\ G^* \xrightarrow{\rho_b} G^* + P \tag{1}$$

This models the production of proteins, their degradation, DNA-protein binding and unbinding. For simplicity we do not have an mRNA description (though this will be added in Section 3). The gene can be in one of two states: an unbound state G and a bound state G^*. The rate of protein production depends on the gene state. Note that here we have scaled all parameters by the protein degradation rate. The deterministic rate equations for the time evolution of the mean molecule numbers are:

$$\frac{d\langle g(t)\rangle}{dt} = -\sigma_b\langle g(t)\rangle\langle n(t)\rangle + \sigma_u(1 - \langle g(t)\rangle), \tag{2}$$

$$\frac{d\langle n(t)\rangle}{dt} = -\sigma_b\langle g(t)\rangle\langle n(t)\rangle + \sigma_u(1 - \langle g(t)\rangle) + \rho_u\langle g(t)\rangle + \rho_b(1 - \langle g(t)\rangle) - \langle n(t)\rangle, \tag{3}$$

where $\langle n(t)\rangle$ denotes the mean number of molecules of protein P at time t, $\langle g(t)\rangle$ denotes the mean number of molecules of gene G at time t and $\langle g^*(t)\rangle$ denotes the mean number of molecules of gene G^* at time t. Note that we used the relation $\langle g(t)\rangle + \langle g^*(t)\rangle = 1$, i.e. there is one gene copy. Note also that t is non-dimensional time, i.e. actual time multiplied by the protein degradation rate.

By the fast equilibrium approximation it follows that $\partial_t\langle g(t)\rangle \simeq 0$ (and $\partial_t\langle g^*(t)\rangle \simeq 0$) for all times which implies, from Eq. (2), that $\langle g(t)\rangle = L/(L + \langle n(t)\rangle)$ where $L = \sigma_u/\sigma_b$. Substituting the latter in the right hand side of Eq. (3) and suppressing the time dependence (for notational convenience) we obtain:

$$\frac{d\langle n\rangle}{dt} = \frac{L\rho_u + \rho_b\langle n\rangle}{L + \langle n\rangle} - \langle n\rangle. \tag{4}$$

This is an effective time-evolution equation for the protein numbers. This corresponds to a system with two reactions: an effective zero-order reaction modeling the transcriptional regulation of protein production, and a first-order protein degradation reaction. The rate of protein production is a function of the mean number of proteins and three special cases can be distinguished: (i) if $\rho_u > \rho_b$ then the rate of protein production decreases with increasing $\langle n\rangle$; this is the case of negative feedback. (ii) if $\rho_u < \rho_b$ then the rate of protein production increases with increasing $\langle n\rangle$; this is the case of positive feedback. (iii) if $\rho_u = \rho_b$ then the rate of protein production is independent of $\langle n\rangle$ and effectively there is no feedback.

Intuitively one would expect the solution of Eq. (4) to be an excellent approximation to the time-evolution of the protein in the full model given by Eqs. (2-3), in the limit of fast promoter switching, i.e. $\min(\sigma_u, \sigma_b) \gg \max(1, \rho_u, \rho_b)$. This can be explicitly verified by calculating the ratio of gene and protein timescales, as follows. Replacing σ_u by σ_u/ϵ and σ_b by σ_b/ϵ and taking the limit of $\epsilon \to 0$, it is straightforward to show that to leading order the two eigenvalues of the Jacobian matrix of the rate equations Eqs. (2-3) evaluated at steady-state, are given by:

$$\lambda_1 = -\frac{(\langle g\rangle^2 + L)\sigma_b}{\epsilon\langle g\rangle} + O(\epsilon^0),$$
$$\lambda_2 = -\frac{L + \langle g\rangle^2(\rho_u - \rho_b)}{\langle g\rangle^2 + L} + O(\epsilon), \tag{5}$$

where $\langle g\rangle$ is the steady-state mean gene number given by:

$$\langle g\rangle = \frac{L + \rho_b - \sqrt{(L - \rho_b)^2 + 4L\rho_u}}{2(\rho_b - \rho_u)}. \tag{6}$$

For completeness and since we will use it later, the steady-state mean protein number is given by:

$$\langle n \rangle = \frac{1}{2}\left(\rho_b - L + \sqrt{(L-\rho_b)^2 + 4L\rho_u}\right). \tag{7}$$

Note that $\lambda_{1,2}$ are negative and hence the steady-state of the system is stable to small perturbations. Furthermore Eq. (5) show that as $\epsilon \to 0$, $\lambda_1 \to -\infty$ and λ_2 tends to a constant. Since the timescales of decay of transients in the mean protein and gene numbers are given by the absolute of the inverse of the eigenvalues, it follows that there is clear timescale separation in the limit of fast promoter switching. Note that in the calculation above we assumed that $\rho_u \neq \rho_b$; a similar calculation for the equality case also leads to timescale separation.

Hence to summarize, a deterministic rate equation analysis shows that in the limit of fast promoter switching, the reaction scheme (1) composed of five reactions (four first-order reactions and a bimolecular reaction) reduces to just two reactions: an effective zero-order reaction for the production of proteins with a rate which is a function of the mean number of proteins and a first-order reaction modeling protein degradation.

2.2 Heuristic stochastic model reduction

The most popular stochastic model reduction approach consists of directly writing the chemical master equation for the reduced reaction scheme deduced from the deterministic analysis in the last section. In particular given there are n proteins in the system then we define the effective propensities:

$$\begin{aligned} T^+(n) &= \frac{L\rho_u + \rho_b n}{L+n}, \\ T^-(n) &= n, \end{aligned} \tag{8}$$

where $T^+(n)dt$ is the probability, given n proteins, that a protein production reaction increasing the number of proteins by one, will occur in the time interval $[t, t+dt)$ and $T^-(n)dt$ is the probability, given n proteins, that a protein degradation event reducing the number of proteins by one will occur in the time interval $[t, t+dt)$. These probabilities are deduced directly from the form of the effective rate equation Eq. (4). Essentially the probability per unit time for a particular reaction is taken to be the same as the reaction rate in the effective deterministic rate equation with $\langle n \rangle$ replaced by n. *We emphasize that this is a heuristic rule with no fundamental microscopic basis, except that it guarantees equivalence between the effective equation for the time-evolution of the mean protein numbers of the heuristic master equation and the reduced deterministic rate equation in the limit of small protein number fluctuations*. The chemical master equation for this reduced reaction scheme is then given by:

$$\frac{dP_a(n,t)}{dt} = T^+(n-1)P_a(n-1,t) + T^-(n+1)P_a(n+1,t) - (T^+(n) + T^-(n))P_a(n,t). \tag{9}$$

Note that we have labelled the solution of this *approximate* heuristic master equation P_a to distinguish it from the solution of the full master equation P, which we discuss in Section 2.3. The exact solution of the above one variable master equation in steady-state conditions can be obtained using standard methods [36] and is given by:

$$P_a(n) = P_a(0)\prod_{y=1}^{n}\frac{T^+(y-1)}{T^-(y)} = \frac{\rho_b^n [LN]_n}{n![L]_n M(LN, L, \rho_b)}, \tag{10}$$

where $N = \rho_u/\rho_b$, $[x]_n = x(x+1)...(x+n-1)$ (the Pochhammer symbol) and M is the Kummer confluent hypergeometric function. To obtain insight into the discrepancies introduced by the heuristic approach, we now study two limiting cases.

2.2.1 The limit of large L

This is the limit in which the rate at which proteins bind DNA is much smaller than the unbinding rate. In this limit, the propensities given by Eq. (8) reduce to the simpler form:

$$\begin{aligned} T^{+}(n) &= \rho_u, \\ T^{-}(n) &= n. \end{aligned} \tag{11}$$

Hence in this limit, the propensity $T^{+}(n)$ is independent of n and the steady-state solution of Eq. (9) is simply a Poisson with mean ρ_u:

$$P_a(n) = \frac{\exp(-\rho_u)\rho_u^n}{n!} \tag{12}$$

An alternative proof of this result starting from the exact solution Eq. (10) can be found in Appendix B.

2.2.2 The limit of small L

This is the limit in which the rate at which proteins bind DNA is much larger than the unbinding rate. In this limit, the propensities given by Eq. (8) reduce to the simpler form:

$$\begin{aligned} T^{+}(n) &= (\rho_u - \rho_b)\delta(n) + \rho_b, \\ T^{-}(n) &= n, \end{aligned} \tag{13}$$

where $\delta(n)$ is the Kronecker delta. Substituting these in the heuristic reduced master equation Eq. (9), multiplying throughout by z^n and taking the sum over n on both sides of this equation we get the corresponding generating function equation:

$$\frac{dG(z,t)}{dt} = ((\rho_u - \rho_b)G(0,t) + \rho_b G(z,t))(z-1) + (1-z)\frac{dG(z,t)}{dz}, \tag{14}$$

where $G(z) = \sum_n z^n P_a(n,t)$. In steady-state, this equation has the solution:

$$G(z) = \frac{\rho_b + \rho_u(\exp(\rho_b z) - 1)}{\rho_b + \rho_u(\exp(\rho_b) - 1)}. \tag{15}$$

Hence the steady-state probability distribution is given by:

$$P_a(n) = \begin{cases} \frac{1}{1+N(\exp(\rho_b)-1)}, & \text{if } n = 0, \\ \frac{\exp(-\rho_b)\rho_b^n}{n!}\left(1 + \frac{N-1}{1+N(\exp(\rho_b)-1)}\right), & \text{if } n \geq 1 \end{cases} \tag{16}$$

An alternative proof of this result starting from the exact solution Eq. (10) can be found in Appendix B. Eq. (16) is clearly not a Poisson when there is positive or negative feedback ($N \neq 1$). Note that when $N = 1$, the solution is a Poisson but this case is biologically unimportant because it implies that the rate of protein production is independent of the state of the promoter (bound or free) and consequently there is effectively no feedback mechanism at play.

2.3 Exact stochastic model reduction

For the non-bursty feedback loop (1), the chemical master equation can be conveniently formulated as a set of two coupled equations:

$$\begin{aligned} \frac{dP_0(n,t)}{dt} &= \rho_u(P_0(n-1,t) - P_0(n,t)) + ((n+1)P_0(n+1,t) - nP_0(n,t)) + \sigma_u P_1(n-1,t) - \sigma_b n P_0(n,t), \\ \frac{dP_1(n,t)}{dt} &= \rho_b(P_1(n-1,t) - P_1(n,t)) + ((n+1)P_1(n+1,t) - nP_1(n,t)) - \sigma_u P_1(n,t) \\ &\quad + \sigma_b(n+1)P_0(n+1,t), \end{aligned} \tag{17}$$

where $P_0(n,t)$ is the probability that at time t there are n proteins and the gene is in state G while $P_1(n,t)$ is the probability that at time t there are n proteins and the gene is in state G^*. Note that time t is non-dimensional and equal to the actual time multiplied by the protein degradation rate. The probability of n proteins is then given by $P(n,t) = P_0(n,t) + P_1(n,t)$. In the limit of fast promoter switching, i.e., replacing σ_u by σ_u/ϵ and σ_b by σ_b/ϵ and taking the limit of $\epsilon \to 0$, one can show that the steady-state solution of Eqs. (17) (to leading-order in ϵ) is given by:

$$P(n) = \frac{(1+L)N\rho_b^n(n\rho_b + L(L+n+N\rho_b))[1+LN]_n}{AM(1+LN,1+L,\rho_b) + BM(2+LN,2+L,\rho_b)}, \tag{18}$$

where

$$A = (LN+n)n!(1+L)(L+(N-1)\rho_b)[1+L]_n, \tag{19}$$

$$B = (LN+n)n!(1+LN)\rho_b[1+L]_n. \tag{20}$$

See Appendix A for the details of the derivation. Comparing Eq. (18) with the solution of the heuristic reduced master equation Eq. (10), it is immediately obvious that the two are not equal. *Hence it follows the heuristic reduced master equation is generally different than the solution of the full master equation under fast promoter switching conditions, contradicting a major assumption in the literature (as discussed in the Introduction).* It is straightforward to verify that the two agree only if $N = 1$ in which case the protein production rate is the same in state G or G^* and hence there is no effective feedback mechanism.

To understand the nature of the differences between Eq. (10) and Eq. (18), we consider two limiting cases of small and large L. Whilst results in these limits can be obtained directly from consideration of Eq. (18) (see Appendix B) it is both simpler and instructive to consider a different approach which does not need the exact solution of the master equation. The advantage of this approach is that as we shall see later on, it can be easily extended to the analysis of more complex feedback systems.

The limit of fast promoter switching implies that the reversible reaction $P + G \xrightarrow{\sigma_b} G^*, G^* \xrightarrow{\sigma_u} G + P$ in reaction scheme (1) is approximately in equilibrium for all times. From the chemical master equation for this reversible reaction one can derive the following equation for the steady-state moments (see Appendix C of [37]):

$$\frac{d\langle g\rangle}{dt} = -\sigma_b\langle ng\rangle + \sigma_u(1-\langle g\rangle) = 0, \tag{21}$$

where $\langle n\rangle$ and $\langle g\rangle$ are the mean numbers of proteins and gene respectively, and $\langle ng\rangle$ is the mean of the product of the protein and gene numbers. Hence the mean number of proteins conditional on the gene being in state G can be written as:

$$\langle n|G\rangle = \frac{\sum_{n=0}^{\infty} nP_0(n)}{\sum_{n=0}^{\infty} P_0(n)} = \frac{\langle ng\rangle}{\langle g\rangle} = L\frac{(1-\langle g\rangle)}{\langle g\rangle}, \tag{22}$$

from which we obtain after rearrangement:

$$\langle g\rangle = \frac{L}{L+\langle n|G\rangle}. \tag{23}$$

Hence it follows that in the limit of small L, $\langle g\rangle$ is also very small, the gene spends most of its time in state G^* and consequently the principal reactions determining the protein dynamics are $G^* \xrightarrow{\rho_b} G^* + P, P \xrightarrow{1} \emptyset$. Similarly it can be argued that in the limit of large L, $\langle g\rangle \simeq 1$ (the gene spends most of its time in state G) and hence the principal reactions determining the protein dynamics are $G \xrightarrow{\rho_u} G + P, P \xrightarrow{1} \emptyset$. The master equation for both sets of principal reactions is trivial to solve and implies that the steady-state protein number distribution in both limits is a Poisson:

$$P(n) = \frac{\exp(-\rho_b)\rho_b^n}{n!}, \text{ if } L \to 0, \tag{24}$$

$$P(n) = \frac{\exp(-\rho_u)\rho_u^n}{n!}, \text{ if } L \to \infty \tag{25}$$

2.4 Comparison of heuristic and exact reduction for small & large L

Comparing Eqs.(24)-(25) with Eq. (12) and Eq. (16), it is immediately clear that the heuristic method of stochastic model reduction gives the correct answer in the limit of fast promoter switching for large L but the incorrect answer for small L. Note that timescale separation exists in both cases of small and large L (as can be verified using Eq. (5)) and hence, the lack of agreement of the heuristic and exact reduction is not expected. In Fig. 1 we verify that the heuristic and exact reductions agree with each other and with the Finite State Projection (FSP) of the full master equation for large L provided the fast promoter switching limit (large σ_u and σ_b compared to all other parameters) is also met. This is the case for both positive and negative feedback. Note that FSP is a computationally efficient non Monte-Carlo method that solves the master equation to any desirable degree of accuracy [38].

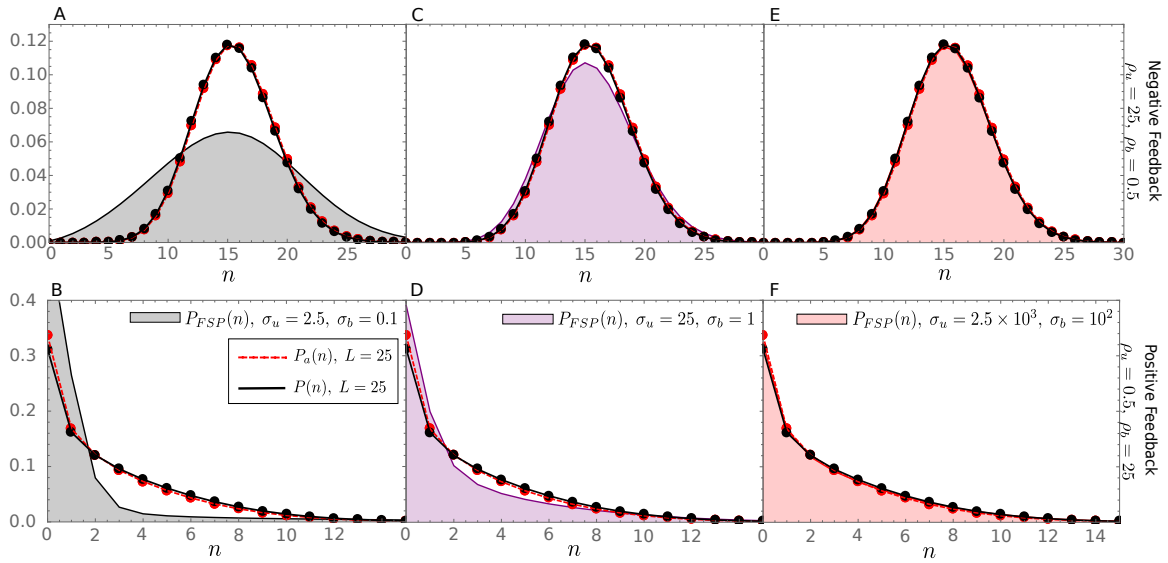

Figure 1: Plots comparing the probability distributions of proteins as predicted by the heuristic reduced master equation, the exact reduced master equation and the full master equation in the limit of large L. Shaded regions indicate the solution of the full master equation Eq. (17) using FSP, dashed red lines indicate the heuristic probability distribution from Eq. (10), and black solid lines indicate the exact solution in the fast promoter switching limit from Eq. (18). Going from left-to-right one can observe how Eqs. (10) and (18) correctly describe the large L limit (here $L = 25$) when both σ_b and σ_u are themselves large, i.e. the fast promoter switching limit. The top row of plots show this for the case of negative feedback ($\rho_u = 25$ and $\rho_b = 0.5$) and the bottom row of plots show this for positive feedback ($\rho_u = 0.5$ and $\rho_b = 25$).

To further understand the differences between the two protein distributions in the limit of small L we now look at the mean protein numbers, the Fano Factor (FF) and the Coefficient of Variation (CV) of protein number fluctuations:

$$\langle n \rangle = \rho_b, \qquad \langle n \rangle_a = \rho_b + \frac{(N-1)\rho_b}{1 + N(\exp(\rho_b) - 1)}, \tag{26}$$

$$\mathrm{FF} = 1, \qquad \mathrm{FF}_a = 1 + \frac{\rho_b(1-N)}{1 + N(\exp(\rho_b) - 1)}, \tag{27}$$

$$\mathrm{CV}^2 = \frac{1}{\rho_b}, \qquad \mathrm{CV}_a^2 = \frac{1 - (1+\rho_b)(1-N^{-1})\exp(-\rho_b)}{\rho_b}. \tag{28}$$

Note that the subscript a denotes calculation using Eq. (16) while no subscript implies calculation using Eq. (24).

From Eq. (26) we deduce that $\langle n \rangle_a < \langle n \rangle$ for $N < 1$ and $\langle n \rangle_a > \langle n \rangle$ for $N > 1$. This means that the solution of the approximate heuristic master equation *underestimates* the mean for positive feedback

($N < 1$) and *overestimates* the mean for negative feedback ($N > 1$). Since the deterministic rate equations also predict a steady-state protein mean of ρ_b for the case $L \to 0$ (see Eq. (7)) it then follows that the approximate heuristic master equation also leads one to believe in noise-induced shifts of the mean which actually do not exist. From Eq. (27) we deduce that the approximate heuristic master equation artificially predicts *sub-Poissonian* ($\mathrm{FF_a} < 1$) fluctuations in molecule numbers for negative feedback ($N > 1$) and *super-Poissonian* ($\mathrm{FF_a} > 1$) fluctuations in molecule numbers for positive feedback ($N < 1$). These deviations from Poissonian behavior are most pronounced for intermediate ρ_b since for small and large ρ_b, $\mathrm{FF}_a \simeq 1$. From Eq. (28) we deduce that $\mathrm{CV}_a^2 > \mathrm{CV}^2$ for $N < 1$ and $\mathrm{CV}_a^2 < \mathrm{CV}^2$ for $N > 1$, i.e. the approximate heuristic master equation overestimates the size of the protein number fluctuations for positive feedback and underestimates them for negative feedback. These observations are confirmed for positive feedback loops in Fig. 2. In particular note that in Fig. 2(A) the heuristic reduced master equation predicts switch-like behavior as ρ_b is increased (from zero for ρ_b below approximately 5 to larger than zero for $\rho_b > 5$) while the exact reduced master equation predicts no such transition – the lack of reliability in predicting the switching characteristics of positive feedback loops is notable because previous studies [26] have used the heuristic reduced master equation to study switching phenomena.

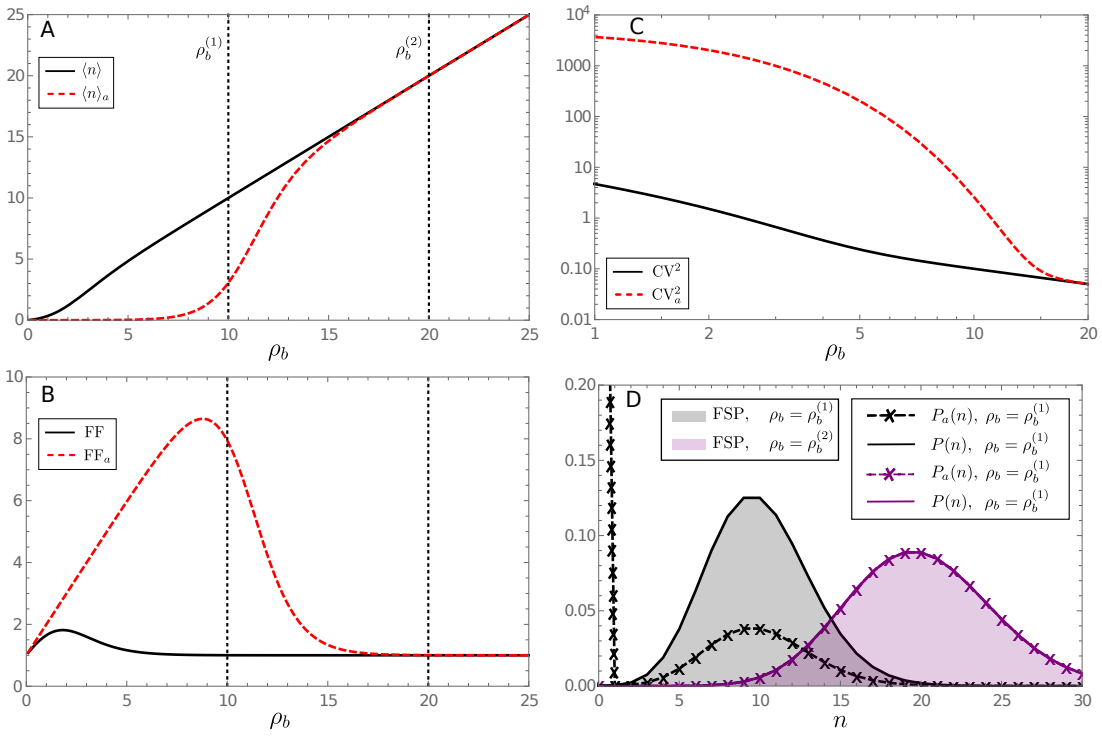

Figure 2: Plots showing the breakdown of the heuristic reduced master equation for fast promoter switching in the limit of small L and positive feedback in steady-state conditions. The plots show the mean protein number (A), the Fano Factor of protein number fluctuations (B) and the Coefficient of Variation squared (C) as a function of ρ_b. In (D) we show the probability distribution of protein numbers corresponding to two different values of ρ_b. The rest of the parameters are fixed to $\sigma_u = 10^2$, $\sigma_b = 10^5$ and $\rho_u = 0.0002$; this implies $L = 10^{-3}$. Note that $\min(\sigma_u, \sigma_b) \gg \max(1, \rho_u, \rho_b)$ and hence fast promoter switching is ensured. Note that $\langle n \rangle_a$, FF_a and CV_a^2 in (A)-(C) are calculated using the solution of the heurestic master equation Eq. (10) while their non-subscript versions are calculated using the solution of the exact reduced master equation Eq. (18). These are in good agreement with the moments calculated in the limit of small L and given by Eq. (26)-(28). The distributions $P_a(n)$ and $P(n)$ in (D) are calculated using Eq. (16) and Eq. (24), respectively. The plots verify the large differences between the heuristic and exact reduced master equation for positive feedback loops (see text for discussion), as well as showing agreement between the theoretical distribution for the exact reduced master equation and that obtained using FSP of the full master equation Eq. (17).

We can also show that generally positive feedback leads to larger deviations of the heuristic from

the exact stochastic model reduction than is the case for negative feedback. Consider strong positive feedback $\rho_u \ll \rho_b$ ($N \ll 1$) with the additional constraint $\rho_u \ll \rho_b \exp(-\rho_b)$. From Eq. (26) it can then be shown that $\langle n \rangle = \rho_b, \langle n \rangle_a \simeq 0$. If we now reverse the values of ρ_u and ρ_b such that we have strong negative feedback then ρ_b is very small and $N \gg 1$ then $\langle n \rangle = \rho_b, \langle n \rangle_a \simeq 1$. Clearly $|\langle n \rangle - \langle n \rangle_a|$ is much larger for positive feedback than negative feedback (with ρ_u, ρ_b interchanged) and this difference is evident in the distributions as well. Finally we compute the conditions for the existence of a mode of the probability distribution at $n = 0$ using Eq. (24) and Eq. (16) respectively:

$$\begin{aligned} \frac{P(1)}{P(0)} &< 1 \Rightarrow \rho_b < 1, \\ \frac{P_a(1)}{P_a(0)} &< 1 \Rightarrow N\rho_b = \rho_u < 1 \end{aligned} \tag{29}$$

This implies that if $\rho_u < 1, \rho_b > 1$ (a special case of positive feedback), the approximate heuristic master equation predicts an artificial mode at $n = 0$ whereas if $\rho_u > 1, \rho_b < 1$ (a special case of negative feedback), the approximate heuristic master equation misses to predict an actual mode at $n = 0$. These predictions, contrasting the differences between the heuristic and exact model reduction for positive and negative feedback loops, are illustrated in Fig. 3. Note that multiple peaks in the protein distribution are often thought to describe switching between different phenotypes and hence of importance to understanding cellular decision-making [39, 40] – the lack of accuracy in the heuristic model predictions for the bimodality of the protein distribution shows that use of this model can lead to incorrect biological predictions.

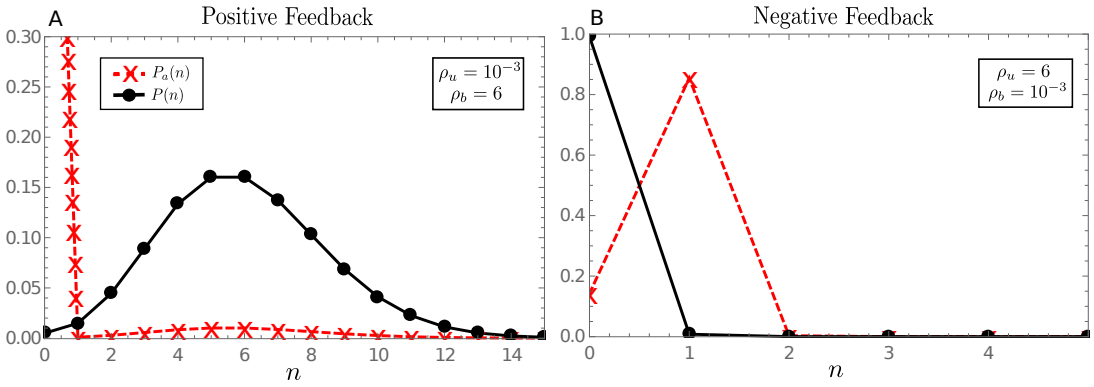

Figure 3: Plots comparing the steady-state protein distributions predicted by the heuristic and exact reduced master equations for strong positive feedback (A) and strong negative feedback (B) for the case of small L. $P_a(n)$ is calculated using Eq. (10) while its subscript version is calculated using the solution of the exact reduced master equation Eq. (18). Note that the heuristic distribution ($P_a(n)$) predicts an artificial mode at zero for positive feedback and misses the prediction of a mode at zero for negative feedback, in line with the conditions Eq. (29). Note also that as predicted by theory, the differences between $P(n)$ and $P_a(n)$ are most significant for positive feedback loops since for negative feedback loops the differences amount to the order of a single molecule. The parameters $\sigma_u = 10^2$ and $\sigma_b = 10^5$ are fixed across both plots (implying $L = 10^{-3}$) while the values of ρ_u and ρ_b are stated on the figure.

The results stated thus far are for steady-state conditions. It would also be interesting to understand the difference between the heuristic reduced master equation Eq. (9) and the exact master equation Eq. (17) for finite time. Since this is analytically intractable we use stochastic simulations to explore this question. Figure 4 summarizes the results of such simulations for two different parameter sets: (i) $\rho_b = \rho_b^{(s)} = 10$ in which case the heuristic predicts a very different steady-state mean number of proteins than the exact reduced master equation; (ii) $\rho_b = \rho_b^{(s)} = 15$ where the heuristic and exact reduced master equations are indistinguishable at steady-state (see Fig. 4(A)). In Fig. 4B we show three independent trajectories of the SSA corresponding to the the exact and heuristic reduced master equations for the two parameter choices; the vast difference between the trajectories of the heuristic and the exact for

$\rho_b = \rho_b^{(s)}$ are particularly striking. In Fig. 4C we show the mean number of proteins as a function of time for the exact and heuristic reduced master equations (solid lines) and compare with the same predicted from the deterministic equations (dashed lines). Two observations can be made: (i) for both parameter sets, the deterministic reaches steady-state at a much earlier time than the stochastic models; (ii) for $\rho_b = \rho_b^{(s)}$ the heuristic predicts that the differences from the deterministic do not decrease with time while the exact predicts that the differences from the deterministic decrease with time (compare top two sub-figures in Fig. 4C). In contrast, for $\rho_b = \rho_b^{(l)}$ both master equations predict that differences from the deterministic decrease with time. Taken together, the results indicate that the full time-dependent solution of the heuristic reduced master equation is an accurate reflection of the exact reduced master equation provided ρ_b, the protein production rate in state G^*, is large enough so that we are far away from the switching point of the positive feedback loop.

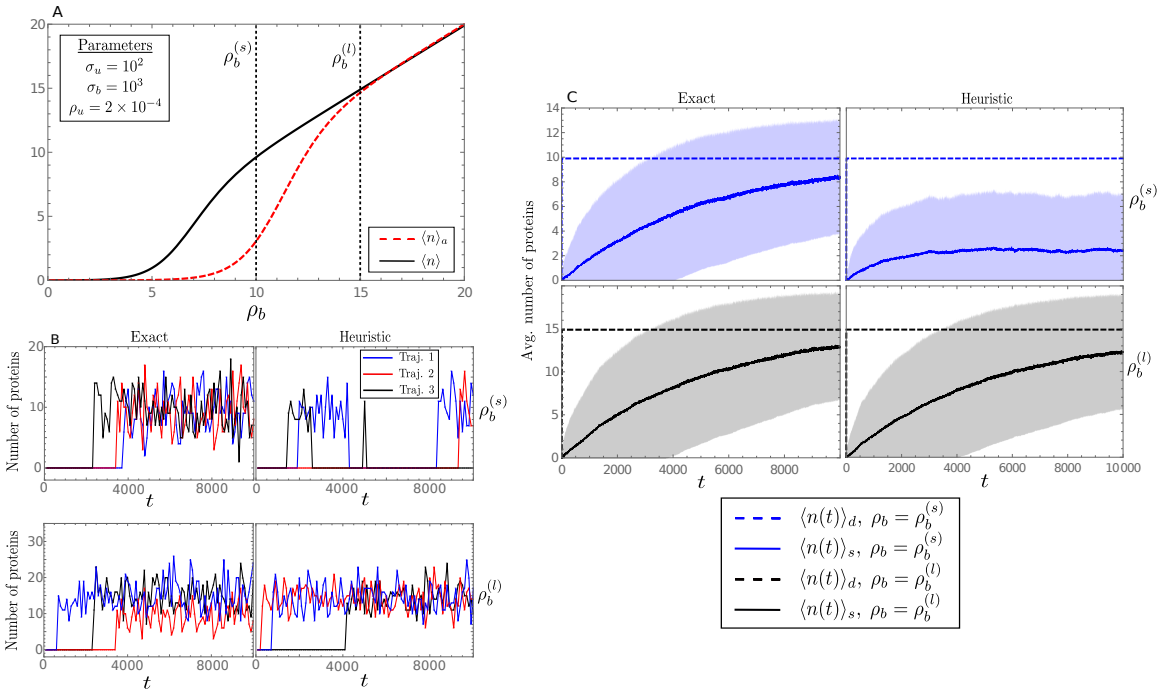

Figure 4: Plots comparing the time-evolution of the heuristic reduced master equation Eq. (9) and the exact master equation Eq. (17) for fast promoter switching conditions, positive feedback and small L (0.1). In (A) we show the switching characteristics of the positive feedback loop as a function of ρ_b for the two master equations in steady-state conditions. In (B) we plot three independent trajectories from the SSA corresponding to the two master equations. Each trajectory shown is down-sampled 1:100 for visual clarity. In (C) we plot the mean number of proteins as a function of time as predicted by the exact and heuristic reduced master equations (shown as solid lines and denoted by the subscript s in the legend) and by the deterministic rate equations (shown as dashed lines and denoted by the subscript d in the legend). The shaded regions show one standard deviation about the mean. The moments were calculated over 2×10^3 SSA trajectories. All sub-figures compare two different parameter sets, one for small ρ_b (where at steady-state the heuristic and exact differ considerably) and one for large ρ_b (where at steady-state the differences are negligible), as indicated in Fig. 4(A). Note that $\min(\sigma_u, \sigma_b) \gg \max(1, \rho_u, \rho_b)$ and hence fast promoter switching is ensured for both parameter sets. See text for discussion.

2.5 Why does conventional model reduction fail?

Having proved that the conventional heuristic model reduction fails for at least the case of small L, the next question is why. To obtain some insight into this problem, we derive the correct form of the effective mean rate of protein production and then contrast with the form assumed by the heuristic

model reduction approach.

Given the reaction scheme (1) and its master equation description Eq. (17) it is straightforward to derive the time-evolution equations for the first moments:

$$\frac{d\langle g\rangle}{dt} = -\sigma_b\langle ng\rangle + \sigma_u(1-\langle g\rangle), \tag{30}$$

$$\frac{d\langle n\rangle}{dt} = -\sigma_b\langle ng\rangle + \sigma_u(1-\langle g\rangle) + \rho_u\langle g\rangle + \rho_b(1-\langle g\rangle) - \langle n\rangle, \tag{31}$$

where $\langle n\rangle$ and $\langle g\rangle$ are the mean numbers of proteins and gene respectively, and $\langle ng\rangle$ is the mean of the product of the protein and gene numbers. Note that we have here suppressed the time dependence for convenience. Assuming fast promoter switching, $d\langle g\rangle/dt \simeq 0$, implies that the average number of gene, at all times, is given by Eq. (23). Substituting in Eq. (31) we obtain an effective equation for the time-evolution of the protein numbers under the condition of fast promoter switching:

$$\frac{d\langle n\rangle}{dt} = \frac{L\rho_u + \rho_b\langle n|G\rangle}{L + \langle n|G\rangle} - \langle n\rangle. \tag{32}$$

Contrasting this equation with the effective equation obtained through the deterministic approach, Eq. (4) we see that the two differ only by the effective rate of protein production:

$$r = \frac{L\rho_u + \rho_b\langle n|G\rangle}{L + \langle n|G\rangle}, \tag{33}$$

$$r_d = \frac{L\rho_u + \rho_b\langle n\rangle}{L + \langle n\rangle}. \tag{34}$$

It immediately follows by comparison that the stochastic heuristic reduction based on r_d will be accurate whenever $\langle n|G\rangle \simeq \langle n\rangle$. Using the definition of the mean in terms of the conditional means, $\langle n\rangle = \langle n|G\rangle\langle g\rangle + \langle n|G^*\rangle(1-\langle g\rangle)$ and the definition of $\langle n|G\rangle$, it is easy to show that all the following three statements are equivalent ways to state the condition guaranteeing equivalence between heuristic and exact stochastic model reduction:

$$\langle n|G\rangle = \langle n\rangle, \tag{35}$$

$$\langle ng\rangle = \langle n\rangle\langle g\rangle \tag{36}$$

$$\langle n|G\rangle = \langle n|G^*\rangle, \tag{37}$$

The first states that the mean number of proteins conditional on the gene being in state G is equal to the mean number of proteins, the second states that there is no correlation between protein and gene number fluctuations (zero covariance) while the third states that the mean number of proteins conditional on the gene being in state G is the same as that conditional on the gene being in state G^*.

We can use the first statement to explain our previous predictions in Sections 2.2, 2.3 and 2.4. When L is large, $\sigma_u \gg \sigma_b$ which means that gene spends most of its time in state G hence Eq. (35) is approximately satisfied. Conversely this is not true when L is small ($\sigma_u \ll \sigma_b$) since then the gene spends most of its time in state G^*. If the rate of protein production is the same in both states G and G^* (the case $N = 1$) then the states are indistinguishable and Eq. (37) immediately follows.

2.6 Numerical computation of the distance measure between steady-state distributions

To further understand the regions of parameter space where the heuristic and exact reduced master equations differ, we numerically compute the Hellinger distance (HD) between the exact steady-state solution of the heuristic master equation Eq. (10) and the exact steady-state solution of the master equation Eq. (68) for a large region of parameter space for the positive feedback loop: N varying between 10^{-5} and 1, and L varying between 10^{-4} and 1. The results are shown as a heatmap in Fig. 5(A). Note that the Hellinger distance is a distance measure between two probability distributions; it is convenient for interpretation since the distance is a fraction. In Fig. 5(B) we calculate the Hellinger distance between the steady-state solution of the exact reduced master equation Eq. (18) and the exact steady-state

solution of the master equation Eq. (68). Note that there is clear timescale separation across the whole region of parameter space used for the heatmaps as demonstrated in Fig. 5(E) and hence based on conventional wisdom, one would expect the heuristic reduced master equation to be accurate at all points in this space. However Fig. 5(A) shows this is not the case – the HD between the distributions predicted by the exact and heuristic reduced master equations varies widely between 0 and 1. In contrast Fig. 5(B) shows that the HD between the distributions predicted by the exact and exact reduced master equations is very close to zero across all of parameter space thus verifying that the latter is the correct form of the reduced master equation under fast promoter switching conditions. From Fig. 5(A) we see that there is a trend for the HD between the heuristic and exact master equations to decrease with increasing L which agrees with the theoretical prediction in previous sections that the differences are significant for very small L and disappear in the limit of large L. As well, there is a trend for the HD to decrease with increasing N and to be particularly small close to $N = 1$; this agrees with the the theoretical prediction that for $N = 1$ the heuristic and exact precisely agree because in this case there is no effective feedback mechanism.

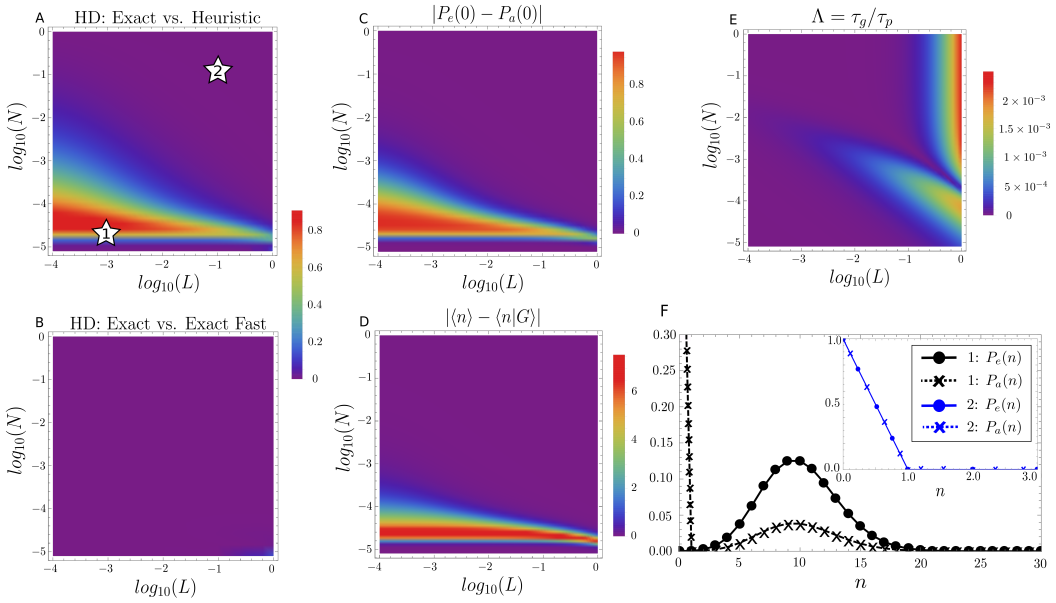

Figure 5: Quantifying the differences between the steady-state protein distributions predicted by the exact master equation Eq. (68), the heuristic reduced master equation Eq. (10) and the exact reduced master equation Eq. (18) for positive feedback loops. In (A) we show the Hellinger distance (HD) between the predictions of the exact master equation and heuristic reduced master equation. In (B) we show the HD between the predictions of the exact master equation and exact reduced master equation (denoted as exact fast). (C) Shows the absolute difference between the probability of zero protein molecules predicted by the exact master equation and the probability of zero protein molecules predicted by the heuristic reduced master equation. (D) Shows the absolute difference $\langle n\rangle$ and $\langle n|G\rangle$ computed using the exact master equation. (E) Shows that the whole region of parameter space chosen has suitable deterministic timescale separation where $\tau_g = 1/\lambda_1, \tau_p = 1/\lambda_2$ and λ_i are the eigenvalues of the Jacobian of the deterministic rate equations Eqs. (2-3) evaluated at steady-state. Note that it is expected from Fig. 4 that the stochastic timescale separation will be much greater than the deterministic. Parameters $\sigma_u = 100$, $\rho_u = 2 \times 10^{-4}$ are fixed throughout the figure, with σ_b varying in the range $100 - 10^6$ (small L) and ρ_b varying in the range $2 \times 10^{-4} - 25$ (positive feedback). Numbered stars in (A) indicate the two points in parameter space whose corresponding probability distributions of protein numbers we show in (F). See text for discussion.

In Fig. 5(F) we plot the protein distributions predicted by the heuristic and exact master equations for the star points labeled (1) and (2) in Fig. 5(A) which are positioned in regions of high and low HD, respectively. Note that for point (1) the heuristic predicts that the probability that the protein numbers

are zero is high whereas the the exact predicts the probability that the protein numbers are zero is very small. Inspired by this observation, as well as the theoretical prediction of modes at zero for small L given by Eq. (29), in Fig. 5(C) we plot a heatmap of the absolute difference between the height of the zero modes of the exact master equation and the heuristic reduced master equation and find that this heatmap is in very good agreement with the heatmap for the HD shown in Fig. 5(A). *This verifies our intuition that the differences between the protein distributions of the two master equations is mostly due to differences in their prediction of the probability of zero proteins at steady-state*. Inspired by the result in the previous section that the heuristic is accurate whenever $\langle n|G\rangle \simeq \langle n\rangle$, we plot in Fig. 5(D) a heatmap of $|\langle n|G\rangle - \langle n\rangle|$ which also shows broad similarity to that in Fig. 5(A).

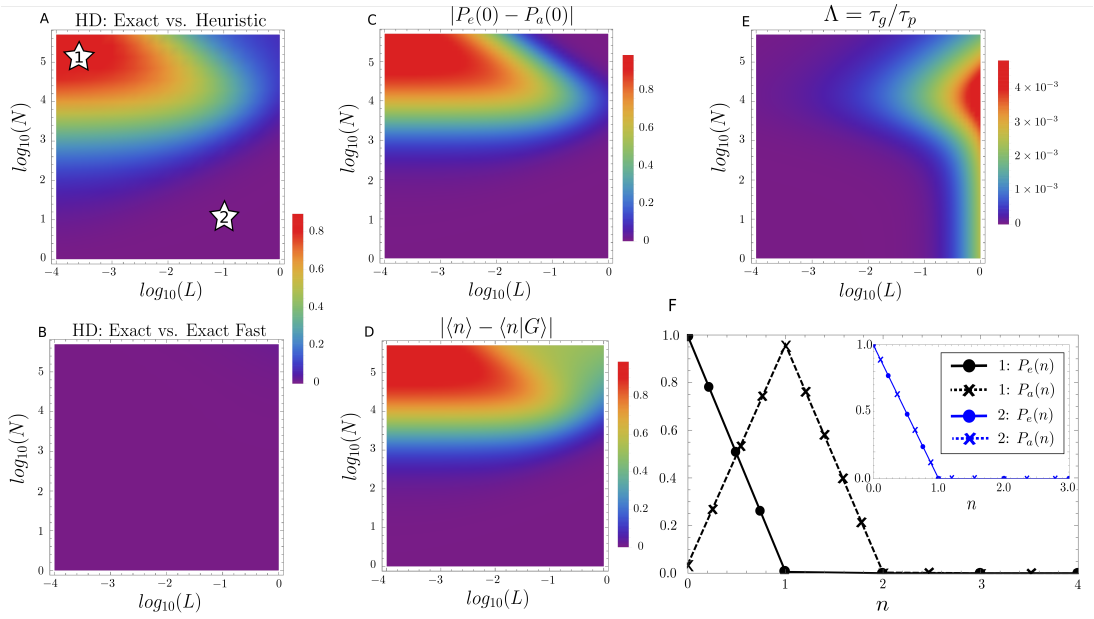

Figure 6: Quantifying the differences between the steady-state protein distributions predicted by the exact master equation Eq. (68), the heuristic reduced master equation Eq. (10) and the exact reduced master equation Eq. (18) for negative feedback loops. In (A) we show the Hellinger distance (HD) between the predictions of the exact master equation and heuristic reduced master equation. In (B) we show the HD between the predictions of the exact master equation and exact reduced master equation. (C) Shows the absolute difference between the probability of zero protein molecules predicted by the exact master equation and the probability of zero protein molecules predicted by the heuristic reduced master equation. (D) Shows the absolute difference $\langle n\rangle$ and $\langle n|G\rangle$ computed using the exact master equation. (E) Shows that the whole region of parameter space chosen has suitable deterministic timescale separation where $\tau_g = 1/\lambda_1, \tau_p = 1/\lambda_2$ and λ_i are the eigenvalues of the Jacobian of the deterministic rate equations Eqs. (2-3) evaluated at steady-state. Note that it is expected from Fig. 4 that the stochastic timescale separation will be much greater than the deterministic. Parameters $\sigma_u = 100$ and $\rho_b = 2 \times 10^{-4}$ are fixed throughout the figure, with σ_b varying between 100 and 10^6 (small L) and ρ_u varying between 2×10^{-4} and 25 (negative feedback). Numbered stars in (A) indicate the two points in parameter space whose corresponding probability distributions of protein numbers we show in (F). See text for discussion.

In Fig. 6 we show the results of the same analysis as in Fig. 5 but now for the case of negative feedback loops. As before, the largest HD between the heuristic reduced master equation and the exact master equation is found for N far away from the trivial case of $N = 1$ and for small L, in line with the theoretical predictions of the previous section. Both the heatmaps of the absolute difference between the heights of the zero modes (Fig. 6(C)) and of the absolute difference between the protein mean and the conditional protein mean (Fig. 6(D)) show high correlation with the HD heatmap in Fig. 6(A). The differences between the heuristic and exact protein distributions for the large HD and small HD in Fig. 6(A) (star point 1, 2 respectively) are shown in Fig. 6(F). Note that the differences between the two cases

amount to the order of a single molecule and are far smaller than the differences found for positive feedback (compare Fig. 5(F)), verifying the theoretical predictions of the previous section, namely that the heuristic fails worst for positive feedback.

2.7 Extending results to the case of multiple protein binding

Here we briefly treat the more general case where multiple protein molecules can bind the promoter, a common case in nature often associated with cooperative behaviour. The reaction scheme is an extension of (1) and reads:

$$\begin{aligned}
&G \xrightarrow{\rho_u} G + P,\ P \xrightarrow{1} \emptyset,\\
&P + G \underset{\sigma_u}{\overset{\sigma_b}{\rightleftharpoons}} G^*,\ P + G^* \underset{\delta_u}{\overset{\delta_b}{\rightleftharpoons}} G^{**},\\
&G^{**} \xrightarrow{\rho_b} G^{**} + P
\end{aligned} \tag{38}$$

Writing the deterministic rate equations for this system and making the assumption of fast promoter switching such that $\partial_t \langle g \rangle \simeq 0, \partial_t \langle g^* \rangle \simeq 0$ and $\partial_t \langle g^{**} \rangle \simeq 0$ it is straightforward to show that the effective deterministic time-evolution equation for the protein numbers has the form:

$$\frac{d\langle n \rangle}{dt} = \frac{LR\rho_u + \rho_b \langle n \rangle^2}{(L + \langle n \rangle)R + \langle n \rangle^2} - \langle n \rangle, \tag{39}$$

where $L = \sigma_u/\sigma_b$ and $R = \delta_u/\delta_b$. It follows by the same reasoning as in Section 2.2 that the corresponding heuristic reduced master equation for protein dynamics is given by Eq. (9) with the effective propensities:

$$\begin{aligned}
T^+(n) &= \frac{LR\rho_u + \rho_b n^2}{(L+n)R + n^2},\\
T^-(n) &= n.
\end{aligned} \tag{40}$$

In the limit of small L and R, Eq. (40) reduces to the effective propensities given by Eq. (13) while in the limit of large L and R, Eq. (40) reduces to the effective propensities given by Eq. (11). Hence the solutions of the heurestic master equation in these two limits are given by Eq. (16) and Eq. (12).

By inspection of the reaction scheme (38) it is obvious that for small L and R, the gene will be mostly in state G^{**} and hence the principal reactions determining the protein dynamics are $G^{**} \xrightarrow{\rho_b} G^{**} + P, P \xrightarrow{1} \emptyset$. By the same reasoning, it follows that for large L and R, the gene will be mostly in state G and the principal reactions are $G \xrightarrow{\rho_u} G + P, P \xrightarrow{1} \emptyset$. Hence the solution of the exact master equation of (38) in the limit of small and large L, R is Poisson and given by Eq. (24) and Eq. (25) respectively. *Hence all the conclusions previously reached regarding the differences between the heuristic reduced master equation and the exact master equation for single protein binding for the case of small and large L also hold for multiple protein binding for the cases of small and large L, R.*

It is interesting to find the general conditions under which the heuristic reduced master equations generally agrees with the exact. For the case of single promoter binding, we showed in Section 2.5 that this condition could be simply stated as $\langle n \rangle = \langle n|G \rangle$. Next we derive a similar condition for the case of multiple protein binding. We start by noting that under fast promoter switching conditions, the reactions $P + G \underset{\sigma_u}{\overset{\sigma_b}{\rightleftharpoons}} G^*, P + G^* \underset{\delta_u}{\overset{\delta_b}{\rightleftharpoons}} G^{**}$ are in equilibrium and hence the deterministic rate equations yield:

$$\begin{aligned}
\langle g \rangle &= \frac{LR}{\langle n \rangle^2 + (L + \langle n \rangle)R},\\
\langle g^{**} \rangle &= \frac{\langle n \rangle^2}{\langle n \rangle^2 + (L + \langle n \rangle)R}.
\end{aligned} \tag{41}$$

Next we derive equations for the same quantities but from the exact master equation. Writing the master

equation for the same two reversible reactions in equilibrium, one can deduce the moment equations:

$$\partial_t \langle g \rangle = 0 = -\sigma_b \langle g \rangle + \sigma_u (1 - \langle g \rangle - \langle g^{**} \rangle), \tag{42}$$

$$\partial_t \langle g^{**} \rangle = 0 = \delta_b \langle n(1 - g - g^{**}) \rangle - \delta_u \langle g^{**} \rangle, \tag{43}$$

from which we can deduce:

$$\begin{aligned} \langle g \rangle &= \frac{L(R + \langle n|G^{**} \rangle - \langle n \rangle)}{\langle n|G \rangle (R + \langle n|G^{**} \rangle) + L(R + \langle n|G^{**} \rangle - \langle n|G \rangle)}, \\ \langle g^{**} \rangle &= \frac{L(\langle n \rangle - \langle n|G \rangle) + \langle n \rangle \langle n|G \rangle}{\langle n|G \rangle (R + \langle n|G^{**} \rangle) + L(R + \langle n|G^{**} \rangle - \langle n|G \rangle)}, \end{aligned} \tag{44}$$

where we used the definitions of conditional means: $\langle n|G \rangle = \langle ng \rangle / \langle g \rangle$ and $\langle n|G^{**} \rangle = \langle ng^{**} \rangle / \langle g^{**} \rangle$. Comparing Eq. (41) and Eq. (44), we see that they can only be equal if the following condition is true:

$$\langle n|G \rangle = \langle n|G^{**} \rangle = \langle n \rangle. \tag{45}$$

By means of the definition of the mean in terms of conditional means $\langle n \rangle = \langle n|G \rangle \langle g \rangle + \langle n|G^* \rangle \langle g^* \rangle + + \langle n|G^{**} \rangle \langle g^{**} \rangle$, one can deduce the final condition required for the solution of the heuristic reduced master equation to equal the exact master equation:

$$\langle n|G \rangle = \langle n|G^* \rangle = \langle n|G^{**} \rangle = \langle n \rangle. \tag{46}$$

In Fig. 7 we verify using simulations that when the above condition is approximately met then the heuristic and exact master equations predict very similar distributions of protein numbers (see Fig. 7(A),(C)) whereas the largest differences between the two master equations (see Fig. 7(B),(D)) correlate with significant differences between the three conditional mean protein numbers $\langle n|G \rangle$, $\langle n|G^* \rangle$, $\langle n|G^{**} \rangle$.

3 Model reduction for bursty feedback loops

In this section we consider model reduction for feedback loops in which there is an implicit mRNA description. It has been rigorously shown that when mRNA degrades much quicker than proteins (a common situation for bacteria and yeast cells) then the mRNA does not need to be explicitly described but rather implicitly manifests through protein bursts [41]. Studies have elucidated the implications of taking into account protein bursting on downstream pathways and shown its importance [42]. Hence we now consider a feedback loop with an implicit mRNA description which has the effective reaction scheme:

$$G \xrightarrow{\rho_u} G + wP,\ P \xrightarrow{1} \emptyset,\ P + G \underset{\sigma_u}{\overset{\sigma_b}{\rightleftharpoons}} G^*,\ G^* \xrightarrow{\rho_b} G^* + wP, \tag{47}$$

where w is the protein burst size which is a random positive integer drawn from the geometric distribution with mean $b = k/d_M$, k is the rate at which mRNA is translated into proteins and d_M is the mRNA degradation rate. Note that the geometric form of the protein burst distribution has been shown theoretically [43] and verified experimentally [44].

3.1 Heuristic stochastic model reduction

We are interested in a reduced description of protein fluctuations in the limit of fast promoter switching. Clearly the effective master equation has to have the general form:

$$\frac{dP_a(n,t)}{dt} = \sum_{i=0}^{\infty} (T_i^+(n-i)P_a(n-i,t) - T_i^+(n)P_a(n,t)) + T^-(n+1)P_a(n+1,t) - T^-(n)P_a(n,t), \tag{48}$$

where $T_i^+(n)dt$ is the probability, given n proteins, that a protein burst of size i will occur in the time interval $[t, t+dt)$ and $T^-(n)dt$ is the probability, given n proteins, that a protein degradation event reducing the number of proteins by one will occur in the time interval $[t, t+dt)$. Next we use the deterministic rate equations to guess the equations for $T_i^+(n)$ and $T^-(n)$. The deterministic rate equations

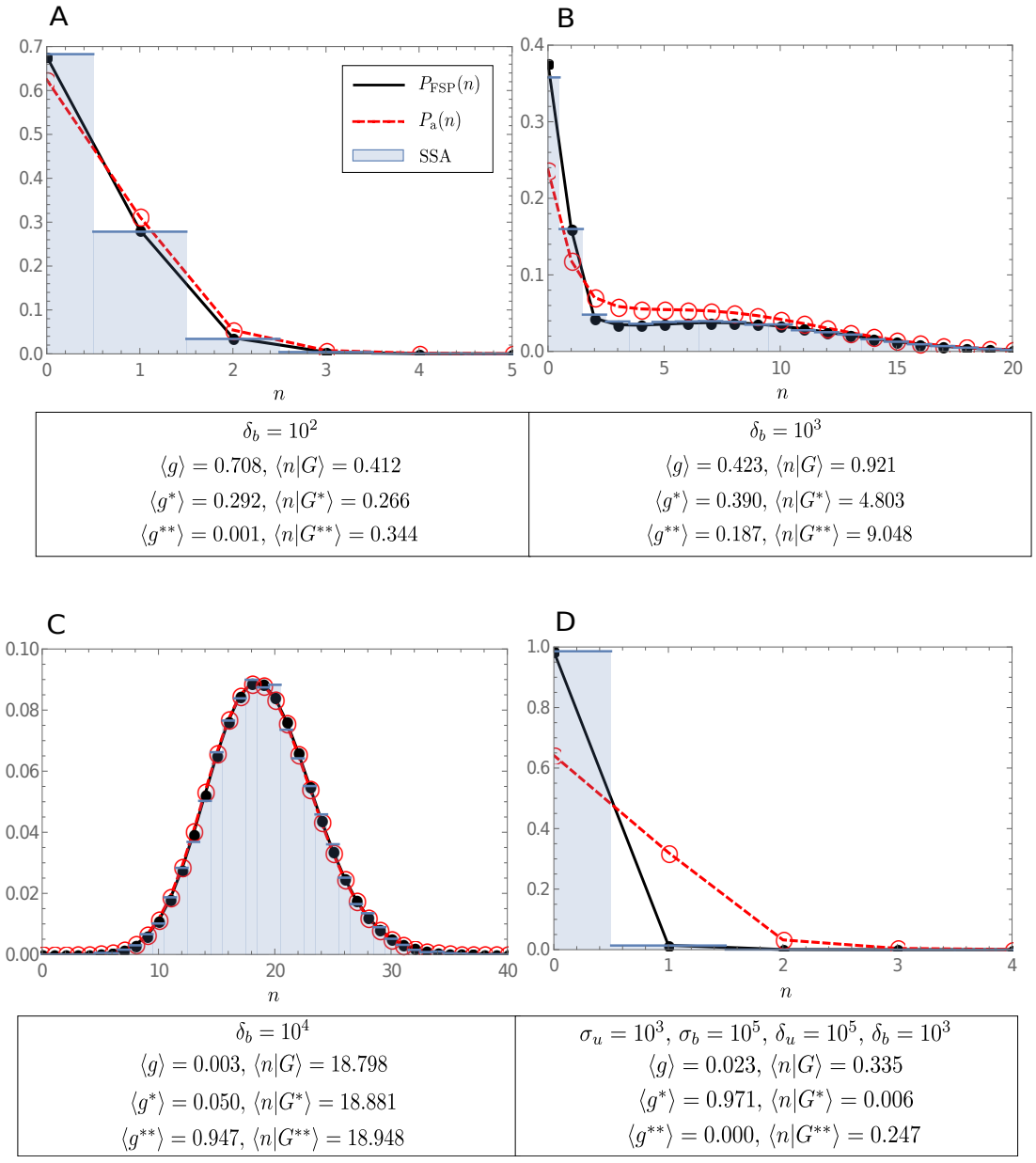

Figure 7: Plots comparing the heuristic master equation (Eq. (9) with Eq. (40)) and exact master equation predictions for the protein number distributions of a positive feedback loop ($\rho_u = 0.5, \rho_b = 20$) with multiple protein binding (38). All calculations done using FSP. Fast promoter switching is enforced by choosing $\min(\sigma_u, \sigma_b, \delta_u, \delta_b) \gg \max(1, \rho_u, \rho_b)$. In (A)-(C) parameter values are fixed to $\sigma_u = 10^4, \sigma_b = 10^4, \delta_u = 10^4$ and δ_b is varied in the range $10^2 - 10^4$. Plot (A) shows a case where the system spends most of its time in state G, in (B) we show a case for which all states G, G^* and G^{**} are frequently accessed by the system and in (C) we show a case where the state G^{**} dominates. Sub-figure (D) shows a case where the systems spends most of its time in state G^*. Note that the differences between the heurestic and exact master equation are reflected in the differences between the values of the mean number of proteins conditional on each state, with the smallest differences occurring for cases (A) and (C), in agreement with the condition given by Eq. (46).

corresponding to (47) are given by:

$$\frac{d\langle g(t)\rangle}{dt} = -\sigma_b\langle g(t)\rangle\langle n(t)\rangle + \sigma_u(1-\langle g(t)\rangle), \tag{49}$$

$$\frac{d\langle n(t)\rangle}{dt} = -\sigma_b\langle g(t)\rangle\langle n(t)\rangle + \sigma_u(1-\langle g(t)\rangle) + \rho_u b\langle g(t)\rangle + \rho_b b(1-\langle g(t)\rangle) - \langle n(t)\rangle. \tag{50}$$

Note that these equations are the same as the deterministic rate equations for the non-bursty case given by Eqs. (2-3) except that ρ_u is replaced by $\rho_u b$ and ρ_b is replaced by $\rho_b b$; this directly follows from the definition of b as the mean protein burst size. Assuming fast promoter switching, $\partial_t\langle g(t)\rangle \simeq 0$, it follows that an effective reduced rate equation for the mean protein numbers is:

$$\frac{d\langle n\rangle}{dt} = \frac{L\rho_u + \rho_b\langle n\rangle}{L+\langle n\rangle}b - \langle n\rangle. \tag{51}$$

The form of this effective equation combined with the fact that we know that burst size is distributed according to a geometric distribution with mean b suggests a one-variable master equation of the form Eq. (48) with effective propensities:

$$\begin{aligned} T_i^+(n) &= \frac{L\rho_u + \rho_b n}{L+n}\psi_i, \\ T^-(n) &= n, \end{aligned} \tag{52}$$

where ψ_i is the probability that a burst has size i which is given by $b^i/(1+b)^{i+1}$. If we denote the angled brackets with subscript a as the statistical averages calculated using the heuristic master equation Eq. (48) with propensities given by Eq. (52) then it follows that:

$$\begin{aligned} \frac{d\langle n\rangle_a}{dt} &= \sum_{i=0}^{\infty} i\langle T_i^+(n)\rangle_a - \langle T^-(n)\rangle_a, \\ &= b\left\langle \frac{L\rho_u + \rho_b n}{L+n}\right\rangle_a - \langle n\rangle_a, \\ &\simeq \frac{L\rho_u + \rho_b\langle n\rangle_a}{L+\langle n\rangle_a}b - \langle n\rangle_a. \end{aligned} \tag{53}$$

Note that this equation is the same as the reduced rate equation Eq. (51) (upon replacing $\langle n\rangle_a$ by $\langle n\rangle$) and hence verifies that the form of the effective propensities given by Eq. (52) guarantees equivalence between the effective equation for the time-evolution of the mean protein numbers of the heuristic master equation and the reduced deterministic rate equation in the limit of small protein number fluctuations when $\langle n\rangle_a \simeq n$.

The heuristic stochastic model given by Eqs. (48,52) is difficult to solve exactly in steady-state because there are no known general solutions for one species reaction systems with multi-step reactions, i.e. reactions leading to the production of more than one molecule at a time [36]. However, as we now show, it can be solved in steady-state in the limits of small and large L.

3.1.1 The limit of large L

In this limit, Eq. (52) reduces to the simpler form:

$$\begin{aligned} T_i^+(n) &= \rho_u\psi_i, \\ T^-(n) &= n. \end{aligned} \tag{54}$$

Substituting these in the heuristic reduced master equation Eq. (48), multiplying both sides by z^n and taking the sum over n on both sides of this equation we get the generating function equation:

$$\frac{dG(z,t)}{dt} = \rho_u G(z,t)\left(\frac{1}{1+b(1-z)} - 1\right) + (1-z)\frac{dG(z,t)}{dz}, \tag{55}$$

where $G(z,t) = \sum_n z^n P_a(n,t)$. This equation can be solved in steady-state yielding $G(z) = (1 - b(z - 1))^{-\rho_u}$ which implies that:

$$P_a(n) = \frac{1}{n!}\frac{d^n G(z)}{dz^n}\bigg|_{z=0} = \left(\frac{b}{1+b}\right)^n \left(1 - \frac{b}{1+b}\right)^{\rho_u} \frac{\Gamma(\rho_u + n)}{\Gamma(n+1)\Gamma(\rho_u)} = NB\left(\rho_u, \frac{b}{(1+b)}\right), \tag{56}$$

where $NB(x,y)$ stands for a negative binomial distribution with parameters x, y and mean $xy/(1-y)$.

3.1.2 The limit of small L

In this limit, Eq. (52) reduces to the simpler form:

$$\begin{aligned} T_i^+(n) &= ((\rho_u - \rho_b)\delta(n) + \rho_b)\psi_i, \\ T^-(n) &= n, \end{aligned} \tag{57}$$

where $\delta(n)$ is the Kronecker delta. Substituting these in the heuristic reduced master equation Eq. (48), multiplying both sides by z^n and taking the sum over n on both sides of this equation we get the corresponding generating function equation:

$$\frac{dG(z,t)}{dt} = ((\rho_u - \rho_b)G(0,t) + \rho_b G(z,t))\left(\frac{1}{1+b(1-z)} - 1\right) + (1-z)\frac{dG(z,t)}{dz}. \tag{58}$$

In steady-state, this equation has the solution:

$$G(z) = \frac{1 + N(-1 + (1+b)^{\rho_b}(1 - b(z-1))^{-\rho_b})}{1 + N(-1 + (1+b)^{\rho_b})}. \tag{59}$$

Hence the steady-state probability distribution is given by:

$$P_a(n) = \frac{1}{n!}\frac{d^n G(z)}{dz^n}\bigg|_{z=0} = \begin{cases} \frac{1}{1+N((1+b)^{\rho_b}-1)}, & \text{if } n = 0, \\ NB\left(\rho_b, \frac{b}{1+b}\right)\frac{N}{N-(N-1)(1+b)^{-\rho_b}}, & \text{if } n \geq 1 \end{cases} \tag{60}$$

3.2 Exact stochastic model reduction

To determine how accurate is the heuristic model reduction we need to compare it with the reduction done on the exact model in the limit of fast promoter switching. Unlike the case of a non-bursty feedback loop, the exact solution of the the chemical master equation for reaction scheme (47) is unknown. However by taking the same approach as we did in Section 2, it is easy to find the the solution of the chemical master equation for the case of fast promoter switching and L being either very small or very large.

The limit of fast promoter switching implies that the reaction $P + G \underset{\sigma_u}{\overset{\sigma_b}{\rightleftharpoons}} G^*$ in the reaction scheme (47) is approximately in equilibrium for all times. From the chemical master equation for this reversible reaction one finds the mean number of gene at all times is then given by Eq. (23). Hence it follows that in the limit of small L, $\langle g \rangle$ is also very small, the gene spends most of its time in state G^* and consequently the only effective reactions determining the protein dynamics are:

$$G^* \xrightarrow{\rho_b} G^* + wP, P \xrightarrow{1} \emptyset, \tag{61}$$

where w is the protein burst size which is a random positive integer drawn from the geometric distribution with mean b. The chemical master equation for these two reactions can be easily solved in steady-state leading to a negative binomial distribution, $P(n) = NB(\rho_b, b/(1+b))$. In the opposite limit of large L, $\langle g \rangle$ is approximately 1, the gene spends most of its time in state G and consequently the only effective reactions determining the protein dynamics are:

$$G \xrightarrow{\rho_u} G + wP, P \xrightarrow{1} \emptyset. \tag{62}$$

Solving the chemical master equation in steady-state (using the generating function method) for these two reactions leads to another negative binomial solution, $P(n) = NB(\rho_u, b/(1+b))$.

Hence summarizing the results of Sections 3.1 and 3.2, we can state that the heuristic and exact stochastic model reduction in the limit of fast promoter switching agree for large L (both predict a negative binomial distribution, $P(n) = P_a(n) = NB(\rho_u, b/(1+b))$) but disagree for small L: the exact reduction predicts a negative binomial distribution, $P(n) = NB(\rho_b, b/(1+b))$ while the heuristic reduction predicts the different distribution given by Eq. (60). These results qualitatively parallel those previously obtained for a non-bursty feedback loop.

To further understand the differences between these two distributions for small L we now look at the mean protein numbers, the Fano Factor and the Coefficient of Variation of protein number fluctuations:

$$\langle n \rangle = b\rho_b, \quad \langle n \rangle_a = b\rho_b + \frac{b(N-1)\rho_b}{1+((1+b)^{\rho_b}-1)N}, \tag{63}$$

$$\mathrm{FF} = 1+b, \quad \mathrm{FF}_a = 1+b+\frac{b(1-N)\rho_b}{1+((1+b)^{\rho_b}-1)N}, \tag{64}$$

$$\mathrm{CV}^2 = \frac{1+b}{b\rho_b}, \quad \mathrm{CV}_a^2 = \frac{1+b}{b\rho_b} - \frac{(1+b)^{-\rho_b}(1-N^{-1})(1+b+b\rho_b)}{b\rho_b}. \tag{65}$$

Hence for $N < 1$ (positive feedback), the heuristic underestimates the mean protein number and overestimates the Fano Factor and the Coefficient of Variation of protein number fluctuations and the opposite occurs when $N > 1$ (negative feedback). The deterministic rate equations predict a mean of $b\rho_b$ (can be deduced from Eq. (51) in limit of small L) which agrees with $\langle n \rangle$ but not with $\langle n \rangle_a$ and hence the heuristic artificially predicts noise-induced deviations from the deterministic mean. These are the same conclusions that we reached in Section 2.4 for the case of a non-bursty feedback loop. The exact reduction predicts super-Poissonian fluctuations (FF > 1) while the heuristic predicts the same for $N < 1$ and either super- or sub-Poissonian fluctuations for $N > 1$ ($FF_a > 1$ and $FF_a < 1$ respectively). We note that the prediction of sub-Poissonian fluctuations is a surprising illogical output of the heuristic model since naturally the production of proteins in bursts has to lead to number distributions which are wider than Poisson. The theoretical predictions for the mean, FF, CV are corroborated using FSP in Fig. 8(A)-(C).

The relative errors (made by the heuristic reduction method) for the mean, FF and CV^2 can be computed using $e_m = |\langle n \rangle - \langle n \rangle_a| / \langle n \rangle$, $e_{FF} = |FF - FF_a|/FF$ and $e_{CVs} = |CV^2 - CV_a^2|/CV^2$, respectively. The errors for the bursty feedback loop (computed using Eqs. (63-65)) are smaller than the errors for the non-bursty feedback loop (computed using Eqs. (26-28)) provided the mean burst size $b \gg 1$. *Hence a major prediction of our theory is that bursts in protein expression generally reduce the size of the discrepancies between heuristic and exact stochastic model reduction in the limit of small L.* This theoretical prediction is verified using FSP in Fig. 8(D) .

Finally we compute the conditions for the existence of a mode of the probability distribution at $n = 0$ using the distribution obtained from the exact method $P(n) = NB(\rho_b, b/(1+b))$ and the distribution from the heuristic reduction Eq. (60) respectively:

$$\frac{P(1)}{P(0)} < 1 \Rightarrow \rho_b < \frac{1+b}{b}, \tag{66}$$

$$\frac{P_a(1)}{P_a(0)} < 1 \Rightarrow N\rho_b = \rho_u < \frac{1+b}{b} \tag{67}$$

This implies that if $\rho_u < (1+b)/b$, $\rho_b > (1+b)/b$ (a special case of positive feedback), the approximate heuristic master equation predicts an artificial mode at $n = 0$ whereas if $\rho_u > (1+b)/b$, $\rho_b < (1+b)/b$ (a special case of negative feedback), the approximate heuristic master equation misses to predict an actual mode at $n = 0$.

4 Conclusion

In this paper we have conclusively shown that heuristic stochastic models with Hill-type propensities for transcriptional regulation are not generally valid under fast promoter switching conditions, as commonly assumed. Rather we show that they are valid only over a subset of parameter space consistent

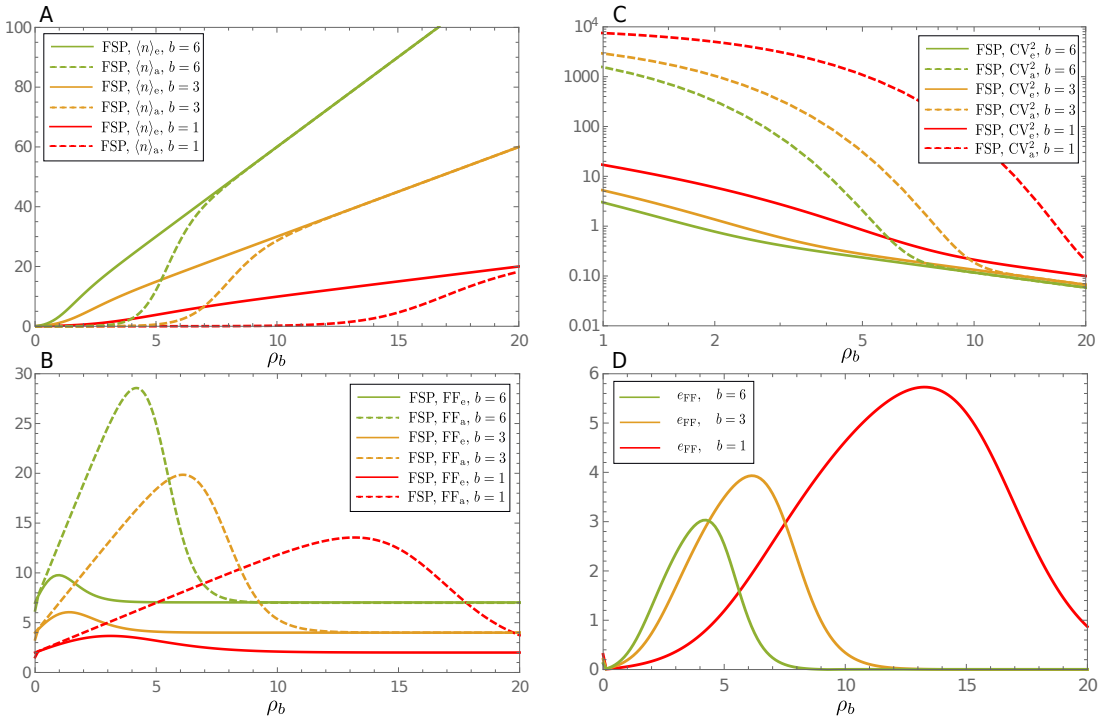

Figure 8: Plots showing the breakdown of the heuristic reduced master equation for fast promoter switching in the limit of small L for a bursty positive feedback loop. The plots show the mean protein number number (A), the Fano Factor of protein number fluctuations (B) and the Coefficient of Variation squared (C) as a function of ρ_b and the mean burst size b. In these plots we compare the FSP solution of the full master equation corresponding to reaction scheme (47) with the FSP solution of the heuristic reduced master equation Eq. (48) with Eq. (52); the latter is distinguished from the former by the subscript e. These are in good agreement with the moments calculated in the limit of small L and given by Eq. (63)-(65) – for example the means of the exact solution in (A) are very well approximated by $\langle n\rangle = b\rho_b$. In (D) we show the relative error in the heuristic reduced master equation's FF computed using the data in (B) where $e_{FF} = |FF_e - FF_a|/FF_e$. Note that the relative error decreases with increasing mean burst size b. In all cases $\rho_u = 0.0002$, $\sigma_u = 100$ and $\sigma_b = 10^5$ were chosen such that L is small, there is positive feedback and fast promoter switching is guaranteed.

with the fast promoter switching assumption, namely when the rate of protein-DNA binding reaction is much less than the unbinding reaction. Our work shows that when this condition is not met, the protein distributions predicted by the heuristic models can be considerably different than the true protein distributions. These differences exist for both negative and positive feedback loops but are particularly pronounced for the latter – in this case we have shown that the heuristic model can predict an artificial mode at zero proteins, an incorrect switching point from low to high protein expression as a parameter is varied, artificial deviations of the mean number of proteins from that predicted by the rate equations and a huge overestimation of the size of protein number fluctuations and of the Fano Factor. Our work further builds on previous work by other authors [33, 34, 45] but has the advantage of using theory to precisely deduce the region of validity of the heuristic approach.

A number of open answered questions remain: (i) Is there a simple way of constructing a different type of reduced stochastic model which avoids the pitfalls of the common heuristic models and which also avoids the use of sophisticated mathematical analysis to derive it? The requirement of simplicity is essential because typically only such methods are widely adopted and indeed this is a main reason why the problematic heuristic reduced stochastic models treated in this paper are so widespread. (ii) What would be the differences between heuristic and correctly reduced stochastic spatial models of genetic feedback loops in the limit of fast gene switching? Would the differences between the two models increase or decrease with the diffusion coefficient of protein molecules? Spatial modeling of such systems

is relatively rare but recent work in this direction [46–50] shows that these models are richer in complex behavior than their non-spatial counterparts and of course they are closer to reality. (iii) Say one used a heuristic reduced stochastic model to construct a likelihood function and then use the latter within a Bayesian approach to infer parameters of auto-transcriptional feedback loops from experimental data: how would these differ from parameters inferred using a likelihood built from a non-reduced model? A recent study [51] shows that inference from moment-based approaches is very sensitive to the type of approximation used to construct the likelihood function and hence suggests large differences between the parameters inferred using heuristic reduced or exact master equations.

In conclusion, our study shows that care must be exerted in the interpretation of the results of heuristic stochastic models. It is now essential that more research is done in this area to elucidate simple rules that can be used to ascertain the biological and physical meaningfulness of the results of commonly used stochastic reduced models of gene regulatory networks.

Author Contributions

J. H. implemented and ran all simulations, produced the figures and checked all the calculations. R. G. conceived the study, directed it, performed the calculations and wrote the paper. All the authors reviewed and edited the final manuscript.

Acknowledgments

This work was supported by a BBSRC EASTBIO PhD studentship to J. H and BBSRC grant BB/M025551/1 to R. G.

A Exact steady-state solution of non-bursty feedback loop in limit of fast gene switching

The exact steady-state solution of Eq. (17) has been previously reported in the literature [7, 52] and is given by:

$$P_e(n) = \frac{1}{n!}\frac{d(G_0(z)+G_1(z))^n}{dz^n}\bigg|_{z=0}, \tag{68}$$

$$G_0(z) = A^{-1}\exp\left(\rho_b(z-1)\right)\left(\frac{1+\sigma_b}{\sigma_b}\frac{\alpha}{\rho_u}M(1+\alpha,\beta,w(z)) - \frac{\alpha}{\rho_u-\rho_b}M(\alpha,\beta,w(z))\right), \tag{69}$$

$$G_1(z) = A^{-1}\exp\left(\rho_b(z-1)\right)M(\alpha,\beta,w(z)), \tag{70}$$

with the definitions:

$$A = \frac{1+\sigma_b}{\sigma_b}\frac{\alpha}{\rho_u}M(1+\alpha,\beta,w(1)) + M(\alpha,\beta,w(1))(1-\frac{\alpha}{\rho_u-\rho_b}), \tag{71}$$

$$\alpha = \frac{\sigma_u(\rho_u-\rho_b)}{\rho_u-\rho_b(1+\sigma_b)}, \tag{72}$$

$$\beta = 1 + \frac{\sigma_u + \sigma_b\frac{\rho_u}{1+\sigma_b}}{1+\sigma_b}, \tag{73}$$

$$w(z) = (\rho_u - \rho_b(1+\sigma_b))\frac{(1+\sigma_b)z-1}{(1+\sigma_b)^2}. \tag{74}$$

Note that $M(x,y,z)$ is the Kummer confluent hypergeometric function and $G_i(z)$ is the generating function $\sum_n z^n P_i(n)$. Replacing σ_u by σ_u/ϵ and σ_b by σ_b/ϵ and taking the limit of $\epsilon \to 0$ (the fast switching limit), we find that $G(z)$ becomes a function of only three non-dimensional parameters $L = \sigma_u/\sigma_b$, $N = \rho_u/\rho_b, \rho_b$ and the corresponding steady-state distribution of protein numbers (to leading-order in

ϵ) has the form:

$$P(n) = \frac{(1+L)N\rho_b^n(n\rho_b + L(L+n+N\rho_b))[1+LN]_n}{AM(1+LN,1+L,\rho_b) + BM(2+LN,2+L,\rho_b))}, \tag{75}$$

where

$$A = (LN+n)n!(1+L)(L+(N-1)\rho_b)[1+L]_n, \tag{76}$$
$$B = (LN+n)n!(1+LN)\rho_b[1+L]_n. \tag{77}$$

B Limits of small and large L from exact steady-state solutions

B.1 The limit of large L

We here consider the limit $L \to \infty$ at constant N, ρ_b. Making use of a standard result [53]:

$$\frac{\Gamma(x+n)}{\Gamma(x)x^n} \simeq 1, \quad x \to \infty, \quad \forall n \in \mathbb{Z}, \tag{78}$$

it follows that $[x]_n = \Gamma[x+n]/\Gamma[x] \simeq x^n$ as $x \to \infty$. Hence we have that for any $a \in \mathbb{R}$, in the limit of large L:

$$\frac{[a+LN]_n}{[a+L]_n} \simeq N^n, \quad M(a+LN, a+L, \rho_b) = \sum_{n=0}^{\infty} \frac{[a+LN]_n}{[a+L]_n} \frac{\rho_b^n}{n!} \simeq \exp(N\rho_b). \tag{79}$$

Using Eq. (79) and $N = \rho_u/\rho_b$ it follows that in the limit of large L, the steady-state distribution of molecule numbers as predicted by the reduced master equation Eq. (10) and by the full master equation in the limit of fast promoter switching Eq. (18) is a Poissonian with mean ρ_u:

$$P(n) = P_a(n) \simeq \frac{\rho_u^n \exp(-\rho_u)}{n!} \tag{80}$$

B.2 The limit of small L

Using the definition of the Pochhammer symbol and the standard result $\Gamma[x] \simeq 1/x$ as $x \to 0$, one can show that in the limit $L \to 0$ at constant N, ρ_b:

$$\frac{[a+LN]_n}{[a+L]_n} = \frac{\Gamma(a+LN+n)\Gamma(a+L)}{\Gamma(a+LN)\Gamma(a+L+n)} \simeq 1, \ \forall n \quad if\ a \neq 0,$$

and

$$\frac{[LN]_n}{[L]_n} = \frac{\Gamma(LN+n)\Gamma(L)}{\Gamma(LN)\Gamma(L+n)} \simeq \begin{cases} 1, & \text{if } n = 0, \\ N, & \text{if } n \geq 1 \end{cases}$$

Using these two results, it then follows that:

$$M(a+LN, a+L, \rho_b) = \sum_{n=0}^{\infty} \frac{[a+LN]_n}{[a+L]_n} \frac{\rho_b^n}{n!} \simeq \exp(\rho_b) \quad if\ a \neq 0$$

and

$$M(LN, L, \rho_b) = \sum_{n=0}^{\infty} \frac{[LN]_n}{[L]_n} \frac{\rho_b^n}{n!} \simeq 1 + N(\exp(\rho_b) - 1) \quad if\ a = 0$$

Using these results it is straightforward to show that in the limit of small L, the steady-state distribution of molecule numbers predicted by the heuristic master equation Eq. (10) reduces to:

$$P_a(n) \simeq \begin{cases} \frac{1}{1+N(\exp(\rho_b)-1)}, & \text{if } n = 0, \\ \frac{\exp(-\rho_b)\rho_b^n}{n!}\left(1 + \frac{N-1}{1+N(\exp(\rho_b)-1)}\right), & \text{if } n \geq 1 \end{cases} \tag{81}$$

while the steady-state distribution of molecule numbers predicted by the master equation in the limit of fast promoter switching Eq. (18) reduces to:

$$P(n) \simeq \frac{\rho_b^n \exp\left(-\rho_b\right)}{n!}. \tag{82}$$

Clearly $P(n) \neq P_a(n)$ since the latter is not-Poissonian and hence shows that in the small L limit, the approximate heuristic master equation gives the incorrect answer.

References

[1] Jonathan M Raser and Erin K O'shea. Noise in gene expression: origins, consequences, and control. Science, 309(5743):2010–2013, 2005.

[2] Arren Bar-Even, Johan Paulsson, Narendra Maheshri, Miri Carmi, Erin O'Shea, Yitzhak Pilpel, and Naama Barkai. Noise in protein expression scales with natural protein abundance. Nature genetics, 38(6):636, 2006.

[3] John RS Newman, Sina Ghaemmaghami, Jan Ihmels, David K Breslow, Matthew Noble, Joseph L DeRisi, and Jonathan S Weissman. Single-cell proteomic analysis of s. cerevisiae reveals the architecture of biological noise. Nature, 441(7095):840, 2006.

[4] Rajesh Ramaswamy, Nélido González-Segredo, Ivo F Sbalzarini, and Ramon Grima. Discreteness-induced concentration inversion in mesoscopic chemical systems. Nature communications, 3:779, 2012.

[5] Nicolaas Godfried Van Kampen. Stochastic processes in physics and chemistry, volume 1. Elsevier, 1992.

[6] Tobias Jahnke and Wilhelm Huisinga. Solving the chemical master equation for monomolecular reaction systems analytically. Journal of mathematical biology, 54(1):1–26, 2007.

[7] Ramon Grima, Deena R Schmidt, and Timothy J Newman. Steady-state fluctuations of a genetic feedback loop: An exact solution. The Journal of chemical physics, 137(3):035104, 2012.

[8] Niraj Kumar, Thierry Platini, and Rahul V Kulkarni. Exact distributions for stochastic gene expression models with bursting and feedback. Physical review letters, 113(26):268105, 2014.

[9] Abhyudai Singh and Joao P Hespanha. Optimal feedback strength for noise suppression in autoregulatory gene networks. Biophysical journal, 96(10):4013–4023, 2009.

[10] Shai S Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon. Network motifs in the transcriptional regulation network of escherichia coli. Nature genetics, 31(1):64, 2002.

[11] Attila Becskei, Bertrand Séraphin, and Luis Serrano. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. The EMBO journal, 20(10):2528–2535, 2001.

[12] Daniel T Gillespie. Stochastic simulation of chemical kinetics. Annu. Rev. Phys. Chem., 58:35–55, 2007.

[13] David Schnoerr, Guido Sanguinetti, and Ramon Grima. Approximation and inference methods for stochastic biochemical kinetics—a tutorial review. Journal of Physics A: Mathematical and Theoretical, 50(9):093001, 2017.

[14] Philipp Thomas and Ramon Grima. Approximate probability distributions of the master equation. Physical Review E, 92(1):012120, 2015.

[15] Zhixing Cao and Ramon Grima. Linear mapping approximation of gene regulatory networks with stochastic dynamics. Nature communications, 9(1):3305, 2018.

[16] Eduardo S Zeron and Moisés Santillán. Distributions for negative-feedback-regulated stochastic gene expression: Dimension reduction and numerical solution of the chemical master equation. Journal of theoretical biology, 264(2):377–385, 2010.

[17] Yang Cao, Daniel T Gillespie, and Linda R Petzold. The slow-scale stochastic simulation algorithm. The Journal of chemical physics, 122(1):014116, 2005.

[18] Eric L Haseltine and James B Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. The Journal of chemical physics, 117(15):6959–6969, 2002.

[19] John Goutsias. Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems. The Journal of chemical physics, 122(18):184102, 2005.

[20] Xingye Kan, Chang Hyeong Lee, and Hans G Othmer. A multi-time-scale analysis of chemical reaction networks: Ii. stochastic systems. Journal of mathematical biology, 73(5):1081–1129, 2016.

[21] Leonardo A Sepúlveda, Heng Xu, Jing Zhang, Mengyu Wang, and Ido Golding. Measurement of gene regulation in individual cells reveals rapid switching between promoter states. Science, 351(6278):1218–1222, 2016.

[22] Tomás Aquino, Elsa Abranches, and Ana Nunes. Stochastic single-gene autoregulation. Physical Review E, 85(6):061913, 2012.

[23] Manuel Pájaro, Irene Otero-Muras, Carlos Vázquez, and Antonio A Alonso. Inherent stochasticity precludes hysteresis in gene regulatory networks. arXiv preprint arXiv:1810.10409, 2018.

[24] Michael Assaf, Elijah Roberts, and Zaida Luthey-Schulten. Determining the stability of genetic switches: explicitly accounting for mrna noise. Physical review letters, 106(24):248102, 2011.

[25] JaeJun Lee and Julian Lee. Quantitative analysis of a transient dynamics of a gene regulatory network. Physical Review E, 98(6):062404, 2018.

[26] Rutger Hermsen, David W Erickson, and Terence Hwa. Speed, sensitivity, and bistability in auto-activating signaling circuits. PLoS computational biology, 7(11):e1002265, 2011.

[27] Nir Friedman, Long Cai, and X Sunney Xie. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. Physical review letters, 97(16):168302, 2006.

[28] Djomangan A Ouattara, Wassim Abou-Jaoudé, and Marcelle Kaufman. From structure to dynamics: Frequency tuning in the p53-mdm2 network. ii: Differential and stochastic approaches. Journal of theoretical biology, 264(4):1177–1189, 2010.

[29] Hao Ge, Hong Qian, and X Sunney Xie. Stochastic phenotype transition of a single cell in an intermediate region of gene state switching. Physical review letters, 114(7):078101, 2015.

[30] Philipp Thomas, Arthur V Straube, and Ramon Grima. The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions. BMC systems biology, 6(1):39, 2012.

[31] Philipp Thomas, Arthur V Straube, and Ramon Grima. Communication: limitations of the stochastic quasi-steady-state approximation in open biochemical reaction networks, 2011.

[32] Philipp Thomas, Ramon Grima, and Arthur V Straube. Rigorous elimination of fast stochastic variables from the linear noise approximation using projection operators. Physical Review E, 86(4):041110, 2012.

[33] Jae Kyoung Kim, Krešimir Josić, and Matthew R Bennett. The relationship between stochastic and deterministic quasi-steady state approximations. BMC systems biology, 9(1):87, 2015.

[34] R Bundschuh, F Hayot, and C Jayaprakash. Fluctuations and slow variables in genetic networks. Biophysical journal, 84(3):1606–1615, 2003.

[35] Tomás Alarcón. Stochastic quasi-steady state approximations for asymptotic solutions of the chemical master equation. The Journal of chemical physics, 140(18):05B609_1, 2014.

[36] C Gardiner. Stochastic methods: a handbook for the natural and social sciences 4th ed.(2009).

[37] Ramon Grima. A study of the accuracy of moment-closure approximations for stochastic chemical kinetics. The Journal of chemical physics, 136(15):04B616, 2012.

[38] Brian Munsky and Mustafa Khammash. The finite state projection algorithm for the solution of the chemical master equation. The Journal of chemical physics, 124(4):044104, 2006.

[39] Philipp Thomas, Nikola Popović, and Ramon Grima. Phenotypic switching in gene regulatory networks. Proceedings of the National Academy of Sciences, 111(19):6994–6999, 2014.

[40] Paul C Bressloff. Stochastic switching in biology: from genotype to phenotype. Journal of Physics A: Mathematical and Theoretical, 50(13):133001, 2017.

[41] Vahid Shahrezaei and Peter S Swain. Analytical distributions for stochastic gene expression. Proceedings of the National Academy of Sciences, 105(45):17256–17261, 2008.

[42] DLK Toner and R Grima. Effects of bursty protein production on the noisy oscillatory properties of downstream pathways. Scientific reports, 3:2438, 2013.

[43] Otto G Berg. A model for the statistical fluctuations of protein numbers in a microbial population. Journal of theoretical biology, 71(4):587–603, 1978.

[44] Ji Yu, Jie Xiao, Xiaojia Ren, Kaiqin Lao, and X Sunney Xie. Probing gene expression in live cells, one protein molecule at a time. Science, 311(5767):1600–1603, 2006.

[45] Marco J Morelli, Rosalind J Allen, Sorin Tănase-Nicola, and Pieter Rein ten Wolde. Eliminating fast reactions in stochastic simulations of biochemical networks: a bistable genetic switch. The Journal of chemical physics, 128(4):01B620, 2008.

[46] Stephen Smith, Claudia Cianci, and Ramon Grima. Analytical approximations for spatial stochastic gene expression in single cells and tissues. Journal of The Royal Society Interface, 13(118):20151051, 2016.

[47] Marc Sturrock, Andreas Hellander, Sahar Aldakheel, Linda Petzold, and Mark AJ Chaplain. The role of dimerisation and nuclear transport in the hes1 gene regulatory network. Bulletin of mathematical biology, 76(4):766–798, 2014.

[48] Michael J Lawson, Linda Petzold, and Andreas Hellander. Accuracy of the michaelis–menten approximation when analysing effects of molecular noise. Journal of The Royal Society Interface, 12(106):20150054, 2015.

[49] Stephen Smith and Ramon Grima. Breakdown of the reaction-diffusion master equation with nonelementary rates. Physical Review E, 93(5):052135, 2016.

[50] Koichi Takahashi, Sorin Tănase-Nicola, and Pieter Rein Ten Wolde. Spatio-temporal correlations can drastically change the response of a mapk pathway. Proceedings of the National Academy of Sciences, 107(6):2473–2478, 2010.

[51] Zhixing Cao and Ramon Grima. Accuracy of parameter estimation for auto-regulatory transcriptional feedback loops from noisy data. Journal of The Royal Society Interface, 16:20180967, 2019.

[52] Guilherme CP Innocentini, Alexandre F Ramos, and José Eduardo M Hornos. Comment on "steady-state fluctuations of a genetic feedback loop: An exact solution"[j. chem. phys. 137, 035104 (2012)]. The Journal of chemical physics, 142(2):035104, 2015.

[53] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.22 of 2019-03-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.