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

J. Holehouse, R. Grima

August 24, 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 distribution of protein numbers is different than that given by standard stochastic models with Hill-type propensities. We show 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.

Statement of Significance

A large number of models of gene regulatory networks in the literature assume that since promoter switching is fast then transcriptional regulation can be effectively modeled using Hill functions. While this approach can be rigorously justified for deterministic models, it is presently unclear if it is also the case for stochastic models. In this article we prove that this is not the case, i.e. stochastic models of gene regulatory systems, namely those with feedback loops, describing transcriptional regulation using Hill functions are only valid in a subset of parameter conditions consistent with fast promoter switching. We identify parameter regimes where these models are correct and where their predictions cannot be trusted.

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 steady-state solutions of the CME are known (see for example [7–10]). 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 [11–13]. In the absence of exact solutions, we become either reliant: (i) on the stochastic simulation algorithm (SSA) [14] or (ii) on approximations of the original network so that analytic results become tractable [15–17]. 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 [18] 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. [19] 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 [20] and Goutsias [21] 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 [22]. There are also several formal results that have been mathematically proven for reaction systems in various scenarios [23–26].

Despite the wide breadth of rigorous approaches (e.g. perturbation theory [27]), by far the most popular approach in the literature of computational and systems biology 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 [28] 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 [29–36]. 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 [37–39] 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 [40]). In contrast, Bundschuh et al. [41] 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 reduced master equations cannot be considered equivalent to the stochastic version of the quasi-steady state approximation (see also [42]), 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 ge-

netic 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 swathes 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. Of course, given that there is only one copy of the gene, one expects fluctuations to be important and a stochastic model to be the most appropriate mathematical description. However for the moment we shall ignore the inherent stochasticity and analyze the system using a deterministic approach. The deterministic rate equations are:

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

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

where $\langle n(t)\rangle_d$ denotes the mean number of molecules of protein P at time t, $\langle g(t)\rangle_d$ denotes the mean number of molecules of gene G at time t and $\langle g^*(t)\rangle_d$ denotes the mean number of molecules of gene G^* at time t. Since the gene can only be in either the bound or unbound state at any one time, one may also interpret $\langle g(t)\rangle_d$ and $\langle g^*(t)\rangle_d$ as the mean fraction of time spent in either gene state respectively. These mean molecule numbers are calculated within the deterministic approximation (hence the subscript d) and will generally be different than the mean molecule numbers of the system obtained from a stochastic description of the system [43]. Note that we have used the relation $\langle g(t)\rangle_d + \langle g^*(t)\rangle_d = 1$, 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. It can also be shown (see Appendix A) that the deterministic equations Eqs. (2)-(3) agree with those from the chemical master equation under the assumption of independence of fluctuations in the protein and gene numbers.

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

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

This is an effective time-evolution equation for the protein numbers, within the deterministic approximation. 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 $\langle n\rangle_d$; this is the case of negative feedback. (ii) If $\rho_u < \rho_b$ then the rate of protein production increases with in-

creasing $\langle n \rangle_d$; this is the case of positive feedback. (iii) If $\rho_u = \rho_b$ then the rate of protein production is independent of $\langle n \rangle_d$ 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:

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

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

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

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

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

Note that $\lambda_{1,2}$ are negative and hence the steady-state of the system is stable to small perturbations. Furthermore Eq. (5) shows 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

As mentioned in the Introduction, one of the most popular stochastic model reduction approaches consists of directly writing the chemical master equation for the reduced reaction scheme deduced from the deterministic analysis in Section 2.1. 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. 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 2.4. The equations for the mean number of protein $\langle n(t)\rangle_a = \sum_n nP_a(n,t)$ can be derived from the master equation:

$$\begin{aligned}\frac{d\langle n\rangle_a}{dt} &= \langle T^+(n)\rangle_a - \langle T^-(n)\rangle_a, \\ &= \left\langle \frac{L\rho_u + \rho_b n}{L+n} \right\rangle_a - \langle n\rangle_a, \\ &\approx \frac{L\rho_u + \rho_b\langle n\rangle_a}{L+\langle n\rangle_a} - \langle n\rangle_a. \end{aligned} \tag{10}$$

Note that in the last line we have made use of the fact that in the limit of small protein number fluctuations n can be replaced by its average. *Hence while the selection of the propensities stems from a heuristic rule with no fundamental microscopic basis, nevertheless 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 (since Eq. (4) and Eq. (10) are the same upon interchanging $\langle n\rangle_d$ by $\langle n\rangle_a$).* Note that however for the general case of non-vanishing protein fluctuations, the mean of the heuristic stochastic model is different than that predicted by the deterministic rate equations.

The exact solution of the one variable master equation Eq. (9) in steady-state conditions can be obtained using standard methods [44] and is given by:

$$P_a(n) = P_a(0)\prod_{y=1}^{n}\frac{T^+(y-1)}{T^-(y)} = \frac{\rho_b^n[LN]_n}{n![L]_n M(LN,L,\rho_b)}, \tag{11}$$

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. The definition of N is used frequently throughout the text. 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:

$$\begin{aligned} T^+(n) &\approx \rho_u, \\ T^-(n) &= n. \end{aligned} \tag{12}$$

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:

$$P_a(n) \approx \frac{\exp(-\rho_u)\rho_u^n}{n!}. \tag{13}$$

Note that this derivation is intuitive but not formally precise because we have implicitly assumed the exchange of two limits: $\lim_{L\to\infty}\lim_{t\to\infty} P_a(n,t) = \lim_{t\to\infty}\lim_{L\to\infty} P_a(n,t)$. A formal proof of this result starting from the exact solution Eq. (11) can be found in Appendix C.

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:

$$\begin{aligned} T^+(n) &\approx (\rho_u - \rho_b)\delta(0,n) + \rho_b, \\ T^-(n) &= n, \end{aligned} \tag{14}$$

where $\delta(0,n)$ 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:

$$\frac{dG(z,t)}{dt} \approx ((\rho_u - \rho_b)G(0,t) + \rho_b G(z,t))(z-1) + (1-z)\frac{dG(z,t)}{dz}, \tag{15}$$

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

$$G(z) = \frac{\rho_b + \rho_u(\exp(\rho_b z) - 1)}{\rho_b + \rho_u(\exp(\rho_b) - 1)}. \tag{16}$$

Hence the steady-state probability distribution is given by:

$$P_a(n) \approx \begin{cases} \frac{1}{1+N(\exp(\rho_b)-1)}, & \text{if } n = 0, \\ \frac{\exp(-\rho_b)\rho_b^n}{n!}\left(1 + \frac{N-1}{1+N(\exp(\rho_b)-1)}\right), & \text{if } n \geq 1. \end{cases} \tag{17}$$

While intuitive, this proof suffers from the same looseness with exchange of limits as in Section 2.2.1. An alternative rigorous proof of this result starting from the exact solution Eq. (11) can be found in Appendix C. Eq. (17) 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 Conditions for the validity of heuristic stochastic model reduction

To obtain further insight into the conditions under which the heuristic model reduction is correct, we consider the stochastic description of the non-bursty feedback loop (1) but ignoring fluctuations in the protein numbers stemming from the reversible binding of protein to gene. The neglection of protein binding fluctuations corresponds to the following reaction scheme:

$$G \xrightarrow{\rho_u} G + P,\ P \xrightarrow{1} \emptyset,\ G \underset{\sigma_u}{\overset{\sigma_b n}{\rightleftharpoons}} G^*,\ G^* \xrightarrow{\rho_b} G^* + P, \tag{18}$$

where n in the reaction rate denotes the number of free proteins. This is a common approximation in the literature [36, 45, 46], the rationale being that since protein numbers are typically much larger than one hence the gain or loss of one molecule via the gene binding reactions can be safely ignored. Given this assumption, the chemical master equation of the non-bursty feedback loop (1) can be conveniently written as a set of two coupled equations:

$$\begin{aligned} \frac{dP_0(n,t)}{dt} &= \rho_u(P_0(n-1,t) - P_0(n,t)) + ((n+1)P_0(n+1,t) - nP_0(n,t)) + \sigma_u P_1(n,t) - \sigma_b n P_0(n,t), \\ \frac{dP_1(n,t)}{dt} &= \rho_b(P_1(n-1,t) - P_1(n,t)) + ((n+1)P_1(n+1,t) - nP_1(n,t)) - \sigma_u P_1(n,t) + \sigma_b n P_0(n,t), \end{aligned} \tag{19}$$

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 $P(n,t) = P_0(n,t) + P_1(n,t)$. Defining the generating functions $G_0(z,t) = \sum_n z^n P_0(n,t)$ and $G_1(z,t) = \sum_n z^n P_1(n,t)$, the generating function differential equations corresponding to Eqs. (19) are given by:

$$\frac{\partial G_0(z,t)}{\partial t} = \rho_u(z-1)G_0(z,t) - (z-1)\frac{\partial G_0(z,t)}{\partial z} + \sigma_u G_1(z,t) - \sigma_b z\frac{\partial G_0(z,t)}{\partial z}, \tag{20}$$

$$\frac{\partial G_1(z,t)}{\partial t} = \rho_b(z-1)G_1(z,t) - (z-1)\frac{\partial G_1(z,t)}{\partial z} - \sigma_u G_1(z,t) + \sigma_b z\frac{\partial G_0(z,t)}{\partial z}. \tag{21}$$

We can solve for the total generating function $G(z) = G_0(z) + G_1(z)$ as follows. At steady state $\partial G_i(z,t)/\partial t = 0$, and utilising the relation $G_1(z) = G(z) - G_0(z)$ we can use the sum of Eq. (20) and (21) to find $G_0(z) = G_0(G(z), G'(z), z)$ and $G_0'(z) = G_0'(G'(z), G''(z), z)$ below

$$G_0 = \frac{1}{\rho_b - \rho_u}(\rho_b G - G'), \quad G_0' = \frac{1}{\rho_b - \rho_u}(\rho_b G' - G''), \tag{22}$$

where we suppress the z dependence for brevity. We then substitute Eq. (22) into Eq. (20), again using $G_1 = G - G_0$, to give a second order linear differential equation in terms of G,

$$((1+\sigma_b)z - 1)G'' + ((\sigma_u + \rho_b + \rho_u) - (\rho_u + (1+\sigma_b)\rho_b)z)G' + \rho_u(\rho_b z - \sigma_u - \rho_b)G = 0. \tag{23}$$

This differential equation has two singularities, a regular singularity at $z = 1/(1+\sigma_b)$ and an irregular singularity at $z = \infty$ and hence its solution is amenable to the ${}_1F_1(\alpha;\beta;z)$ hypergeometric function (otherwise known as the Kummer function $M(\alpha;\beta;z)$). Using a change of variable and an exponential transformation we confirm this, and the solution is given as

$$G(z) = \exp\left(\frac{\rho_u(z-1)}{1+\sigma_b}\right)\frac{M(\alpha,\beta,\frac{\gamma(z(1+\sigma_b)-1)}{\sigma_b})}{M(\alpha,\beta,\gamma)}, \tag{24}$$

where

$$\alpha = \frac{\rho_u\sigma_b(\rho_b(1+\sigma_b) - \rho_u + \sigma_u(1+\sigma_b))}{(1+\sigma_b)^2(\rho_b - \rho_u + \rho_b\sigma_b)}, \ \beta = \frac{\sigma_u + \sigma_b(\rho_u + \sigma_u)}{(1+\sigma_b)^2}, \ \gamma = \frac{\sigma_b(\rho_b - \rho_u + \rho_b\sigma_b)}{(1+\sigma_b)^2}. \tag{25}$$

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 leading-order term in the series expansion of Eq. (24) in powers of ϵ is given by:

$$G(z) = \frac{M[LN; L; \rho_b z]}{M[LN; L; \rho_b]}, \tag{26}$$

where we remind the reader of definitions $L = \sigma_u/\sigma_b$ and $N = \rho_u/\rho_u$. It is easy to show that $P(n) = (1/n!)d^nG(z)/dz^n|_{n=0}$ precisely equals Eq. (11).

Hence we have shown that if protein number fluctuations due to reversible binding can be ignored then the stochastic description agrees with that of the heuristic master equation in the limit of fast promoter switching. This indeed gives some credibility to the use of the heuristic master equation and explains the widespread belief, based on stochastic simulations, that the heuristic master equation is correct in the limit of fast promoter switching. This result is however surprising when one considers that the heuristic is often justified from deterministic arguments, and that the deterministic treatment ignores the sizable fluctuations associated with gene switching.

2.4 Exact stochastic model reduction

The master equation we solved in the previous section is not the exact master equation since we have ignored protein binding fluctuations. In what follows we properly take these into account. For the non-bursty feedback loop (1), the stochastic description is given by the chemical master equation which can be conveniently formulated as a set of two coupled equations:

$$\begin{aligned}
\frac{dP_0(n,t)}{dt} &= \rho_u(P_0(n-1,t) - P_0(n,t)) + ((n+1)P_0(n+1,t) - nP_0(n,t)) + \sigma_u P_1(n-1,t) - \sigma_b n P_0(n,t), \\
\frac{dP_1(n,t)}{dt} &= \rho_b(P_1(n-1,t) - P_1(n,t)) + ((n+1)P_1(n+1,t) - nP_1(n,t)) - \sigma_u P_1(n,t) \\
&\quad + \sigma_b(n+1)P_0(n+1,t).
\end{aligned} \tag{27}$$

Note that these equations are the same as Eq. (19) except for the terms describing protein-gene binding, i.e. those proportional to σ_b and σ_u. 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. (27)

(to leading-order in ϵ) is given by:

$$P(n) = \frac{(1+L)N\rho_b^n(n\rho_b + L(L+n+N\rho_b))[1+LN]_n}{AM(1+LN,1+L,\rho_b) + BM(2+LN,2+L,\rho_b)}, \tag{28}$$

where

$$A = (LN+n)n!(1+L)(L+(N-1)\rho_b)[1+L]_n, \tag{29}$$
$$B = (LN+n)n!(1+LN)\rho_b[1+L]_n. \tag{30}$$

See Appendix B for the details of the derivation. Comparing Eq. (28) with the solution of the heuristic reduced master equation Eq. (11), 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. (11) and Eq. (28), we consider two limiting cases of small and large L. Whilst results in these limits can be obtained directly from consideration of Eq. (28) (see Appendix C) 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.

Since $\langle g \rangle$ is the fraction of time spent in state G for the full stochastic model, 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, $\langle g \rangle \approx 1$ (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:

$$P(n) \approx \frac{\exp(-\rho_u)\rho_u^n}{n!}, \text{ if } L \to \infty, \tag{31}$$
$$P(n) \approx \frac{\exp(-\rho_b)\rho_b^n}{n!}, \text{ if } L \to 0. \tag{32}$$

A formal derivation of these results starting from the solution Eq. (28) can be found in Appendix C.

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

Comparing Eqs.(32)-(31) with Eq. (13) and Eq. (17), 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 [47].

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)

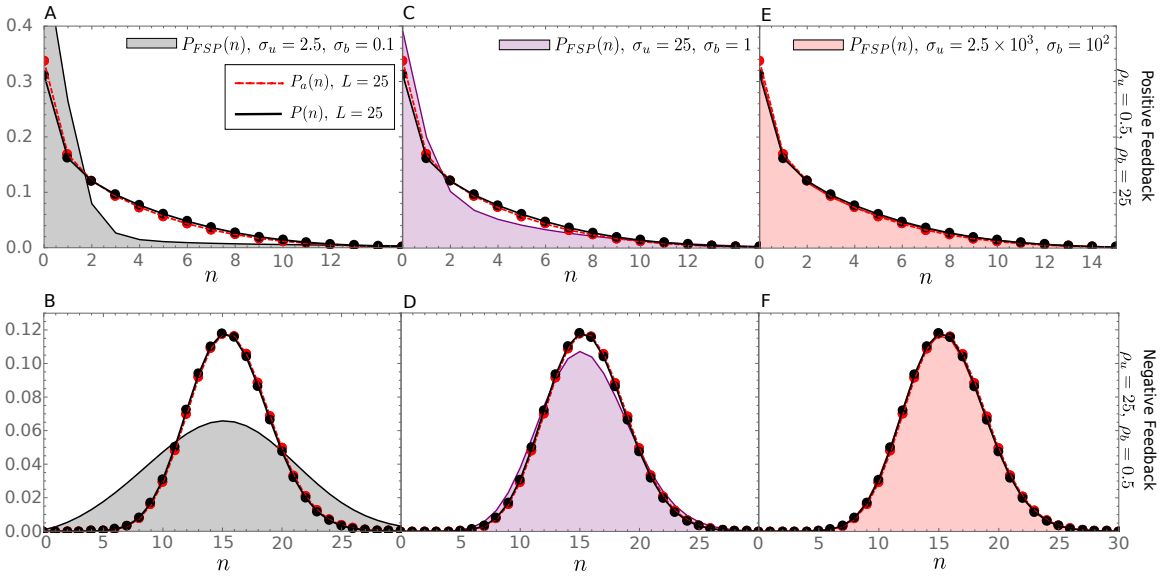

Figure 1: Plots comparing the probability distributions of proteins as predicted by the heuristic reduced master equation ($P_a(n)$), the exact reduced master equation ($P(n)$) and the full master equation in the limit of large L ($P_{FSP}(n)$). L is kept constant throughout the figure. Shaded regions indicate the solution of the full master equation Eq. (27) using FSP, dashed red lines indicate the heuristic probability distribution from Eq. (11), and black solid lines indicate the exact solution in the fast promoter switching limit from Eq. (28). Throughout the paper FSP is used as the benchmark for our analytic results, with a state space truncation chosen such that the probability distributions are indistinguishable from SSA. Going from left-to-right one can observe how Eqs. (11) and (28) 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 positive feedback ($\rho_u = 0.5$ and $\rho_b = 25$) and the bottom row of plots show this for negative feedback ($\rho_u = 25$ and $\rho_b = 0.5$).

of protein number fluctuations:

$$\langle n \rangle = \rho_b, \qquad \langle n \rangle_a = \rho_b + \frac{(N-1)\rho_b}{1 + N(\exp(\rho_b) - 1)}, \tag{33}$$

$$\mathrm{FF} = 1, \qquad \mathrm{FF}_a = 1 + \frac{\rho_b(1-N)}{1 + N(\exp(\rho_b) - 1)}, \tag{34}$$

$$\mathrm{CV}^2 = \frac{1}{\rho_b}, \qquad \mathrm{CV}_a^2 = \frac{1 - (1+\rho_b)(1-N^{-1})\exp(-\rho_b)}{\rho_b}. \tag{35}$$

Note that the subscript a denotes calculation using Eq. (17) while no subscript implies calculation using Eq. (32).

From Eq. (33) 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. (34) 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, $\mathrm{FF}_a \approx 1$. From Eq. (35) 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 for this set of parameters – the lack of reliability in predicting the switching characteristics of positive feedback loops is notable because previous studies [33] have used the heuristic reduced master equation to study switching phenomena.

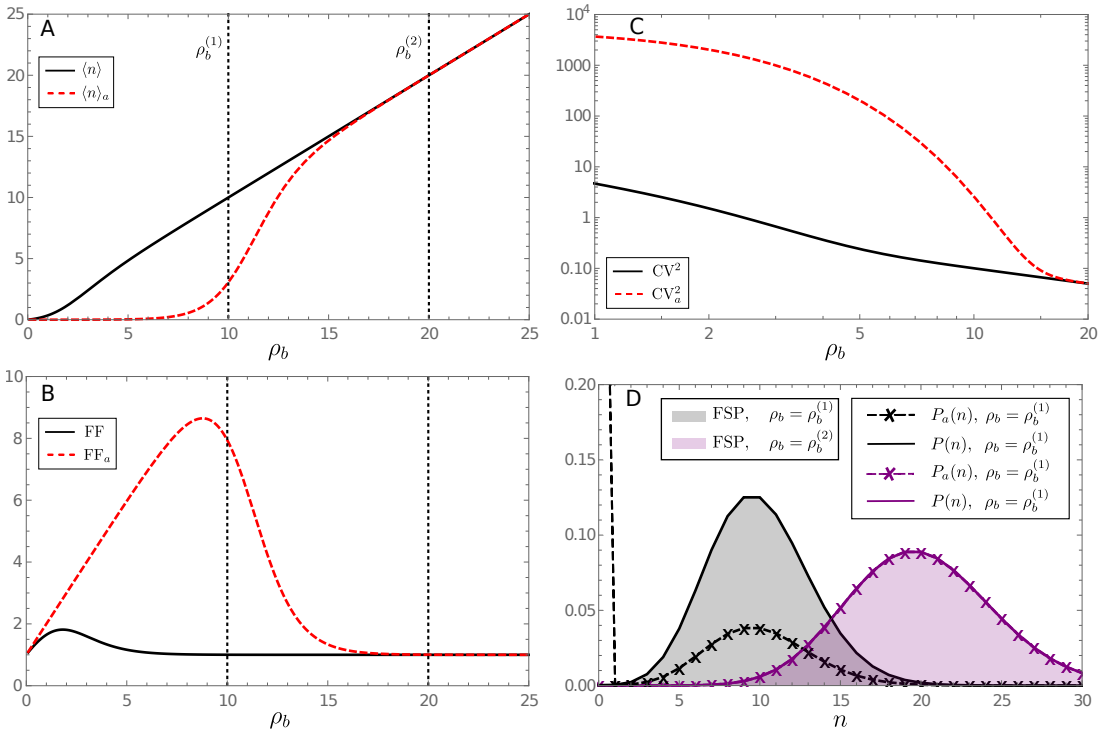

Figure 2: Plots showing the breakdown of the heuristic reduced master equation for fast promoter switching in the limit of small L and positive feedback in steady-state conditions. The plots show the mean protein number (A), the Fano Factor of protein number fluctuations (B) and the Coefficient of Variation 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. (11) while their non-subscript versions are calculated using the solution of the exact reduced master equation Eq. (28). These are in good agreement with the moments calculated in the limit of small L and given by Eq. (33)-(35). The distributions $P_a(n)$ and $P(n)$ in (D) are calculated using Eq. (17) and Eq. (32), 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. (67).

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. (33) it can then be shown that $\langle n \rangle = \rho_b, \langle n \rangle_a \approx 0$. 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 $\langle n \rangle = \rho_b, \langle n \rangle_a \approx 1$. 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. (32) and Eq. (17) respectively:

$$\frac{P(1)}{P(0)} < 1 \Rightarrow \rho_b < 1,$$
$$\frac{P_a(1)}{P_a(0)} < 1 \Rightarrow N\rho_b = \rho_u < 1. \tag{36}$$

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 [48, 49] – 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.

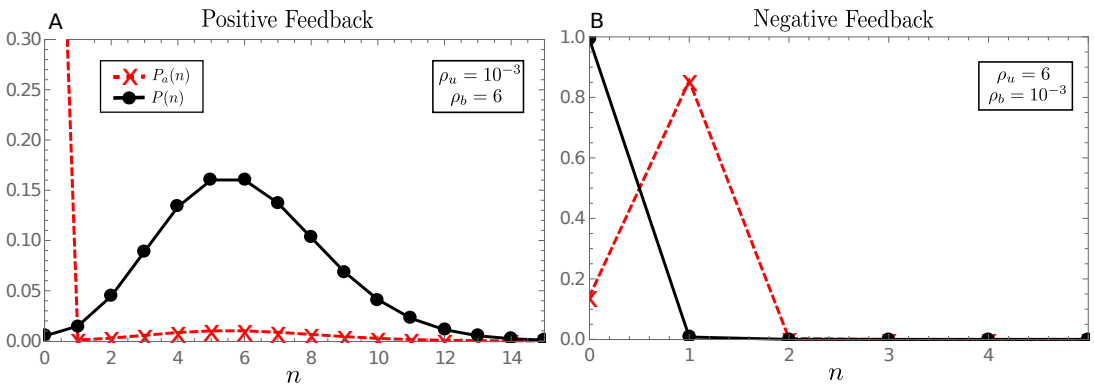

Figure 3: Plots comparing the steady-state protein distributions predicted by the heuristic and exact reduced master equations for strong positive feedback (A) and strong negative feedback (B) for the case of small L. $P_a(n)$ is calculated using Eq. (11) while its subscript version is calculated using the solution of the exact reduced master equation Eq. (28). 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. (36). 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. FSP distributions are indistinguishable from the theoretical distributions shown in the figure.

The differences between the heuristic and exact model reduction can be explained using the results of Section 2.3. There we showed that the heuristic master equation has the same solution, in the limit of fast promoter switching, as the master equation which ignores protein fluctuations due to the reversible protein-DNA binding reaction.

First consider the case of positive feedback (Fig. 3A). When proteins are present in the system there will be rapid switching between the G and G^* states. However, in the rare case of an extinction of proteins in the G^* state and where protein binding fluctuations are neglected, a transition from the bound state G^* to unbound state G does not release a protein (reaction scheme shown in Eq. (18)). The system then must wait for a protein to be produced via the low ρ_u firing rate if it is to leave state G. And hence, the waiting time for a protein to be produced at the low ρ_u firing rate dominates the steady-state dynamics, leading to the mode at zero (Fig. 3A red curve). However, where protein binding fluctuations are included (reaction scheme shown in Eq. (1)), a transition from the G^* to G does release a protein which can immediately bind to G (due to the high σ_b firing rate, meaning $L \ll 1$) and hence the system does not so readily encounter an extinction of proteins (black curve Fig. 3A). We note that even for the black curve there exists a (non-visible) mode at zero, which accounts for the

long waiting times in the extremely rare event (again note $\sigma_b \gg 1$) that the protein released from the G^* state decays before binding to G; clearly however this is not the dominating feature where protein binding fluctuations are taken into consideration.

Now consider the case of strong negative feedback (Fig. 3B). This implies that when a protein is produced in the active G state (now ρ_u is large, ρ_b is very small), the rapid promoter-protein binding reaction will occur forcing the system into the G^* state. Where protein binding fluctuations are neglected no protein is removed upon binding and hence the number of free proteins is still 1; the system will then flip back and forth between the G and G^* states, spending (on average) more time in the G^* state since $\sigma_b \gg \sigma_u$ and hence it is unlikely more than 1 protein will ever be present (due to very small ρ_b and $\sigma_b \gg 1$), hence the mode at $n = 1$ for the red curve in Fig. 3B. In the event of a protein extinction the unbound state G will quickly produce another protein in the G state. For strong negative feedback including the binding fluctuation, the rapid promoter-protein binding reaction will instead remove a protein from the system. Again, since it is unlikely the system will contain more than one protein (bound or otherwise), and since the system spends much more time in the G^* state ($\sigma_b \gg \sigma_u$) the probability distribution for the number of free proteins will have a mode at zero (black curve Fig. 3B).

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. (67) 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) $\rho_b = \rho_b^{(l)} = 15$ 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.

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. (11) and the exact steady-state solution of the master equation Eq. (73) 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, i.e. a HD value of 0 means that two distributions are identical and a HD value approaching 1 means that the distributions are very different from one another. Specifically, a maximum distance 1 is achieved when one of the distributions assigns probability zero to every set to which the other distribution assigns a positive probability. In Fig. 5(B) we calculate the Hellinger distance between the steady-state solution of the exact reduced master equation Eq. (28) and the exact steady-state solution of the master equation Eq. (73). 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

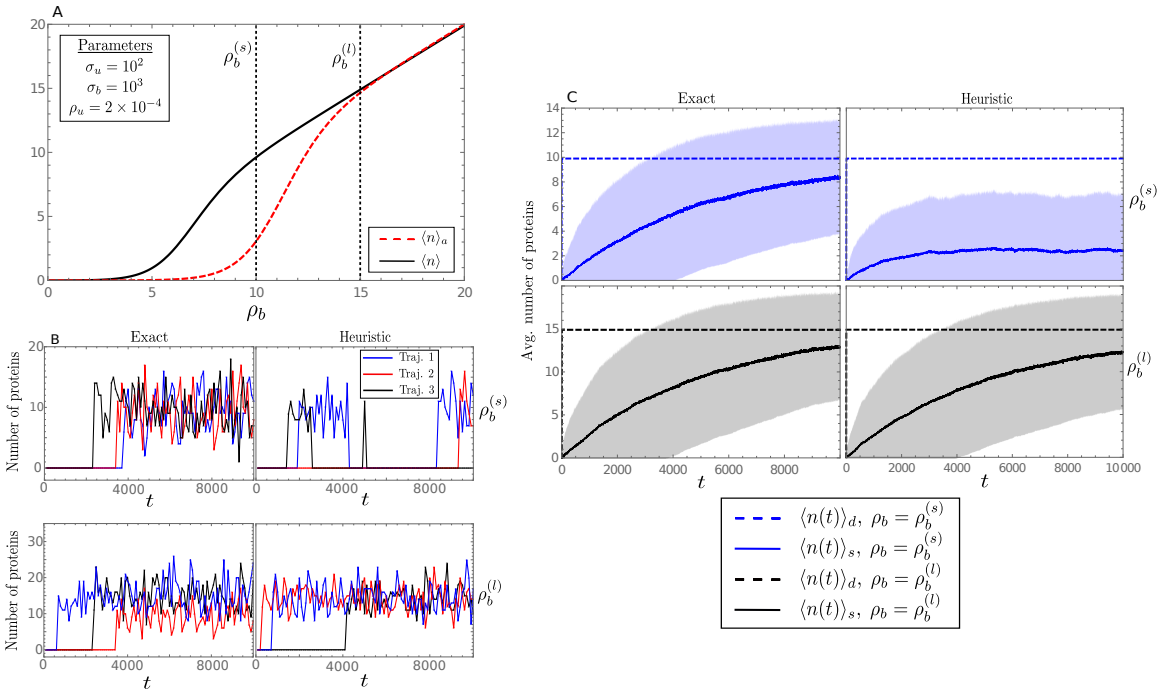

Figure 4: Plots comparing the time-evolution of the heuristic reduced master equation Eq. (9) and the exact master equation Eq. (67) for fast promoter switching conditions, positive feedback and small L (0.1). In (A) we show the switching characteristics of the positive feedback loop as a function of ρ_b for the two master equations in steady-state conditions. Here dotted lines define the values of $\rho_b^{(s)}$ (subscript s for small) and $\rho_b^{(l)}$ (l for large) used throughout the rest of the figure. 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. The top row corresponds to $\rho_b = \rho_b^{(s)}$ and the bottom row corresponds to $\rho_b = \rho_b^{(l)}$. 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.

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.

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

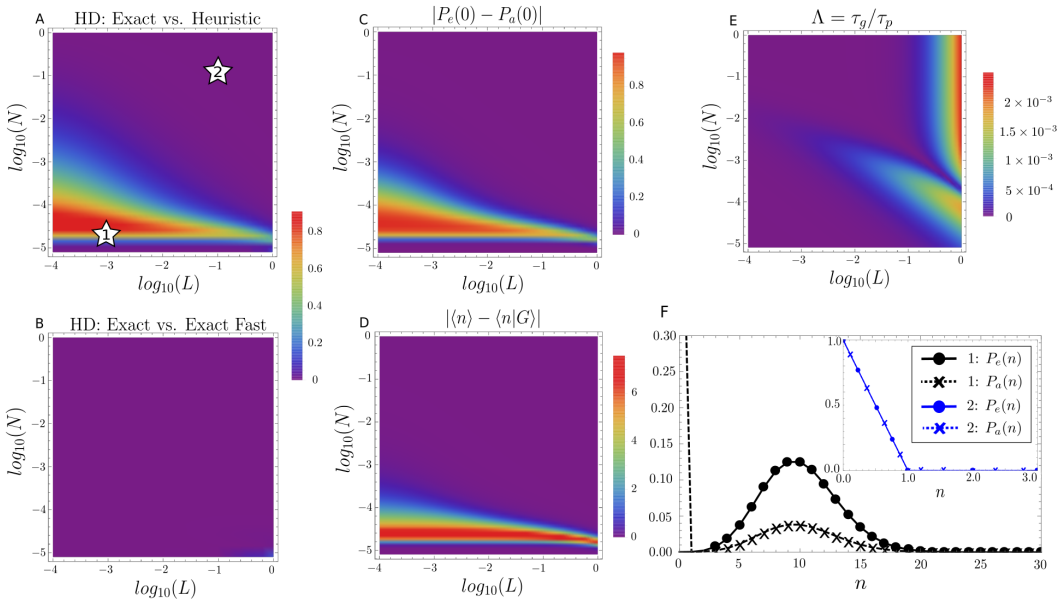

Figure 5: Quantifying the differences between the steady-state protein distributions predicted by the exact master equation Eq. (73), the heuristic reduced master equation Eq. (11) and the exact reduced master equation Eq. (28) for positive feedback loops. $P_e(n)$ denotes the exact steady state solution of the exact master equation (i.e. the exact solution of the reaction scheme from Eq. (1), see derivation in Appendix B) and $P_a(n)$ denotes the distribution from the heuristic reduced master equation. We note that $P_e(n)$ takes the role of P_{FSP} used in previous figures. 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.

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. (36), 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 Appendix A that the deterministic time-evolution equations can be obtained from the full master equation under the assumption $\langle n|G \rangle \approx \langle n \rangle$ (which is equivalent to independence of protein and gene fluctuations), 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).

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

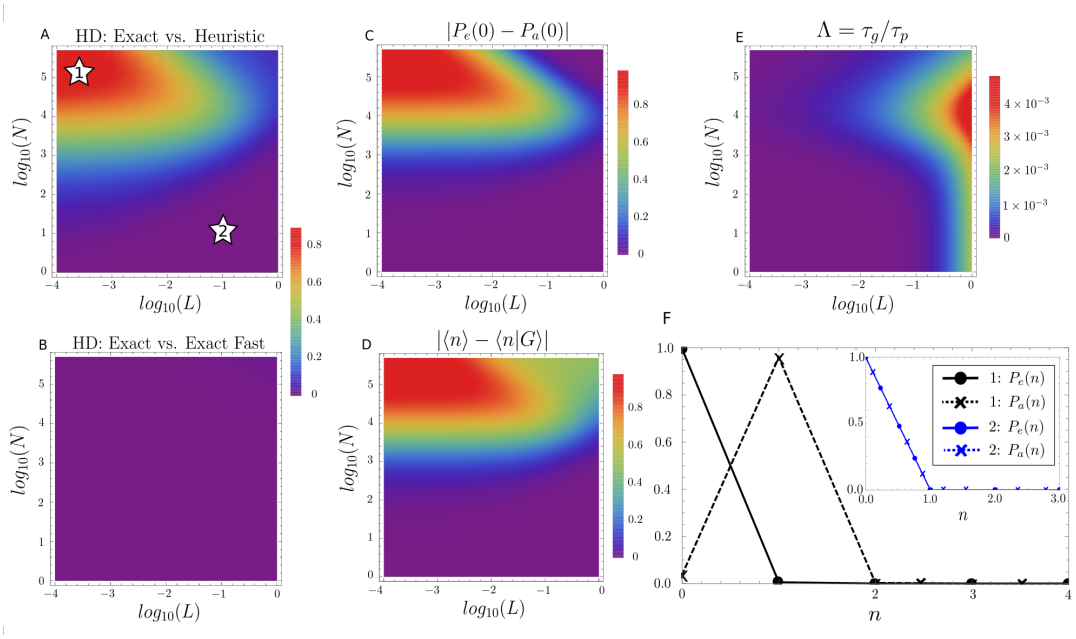

Figure 6: Quantifying the differences between the steady-state protein distributions predicted by the exact master equation Eq. (73), the heuristic reduced master equation Eq. (11) and the exact reduced master equation Eq. (28) for negative feedback loops. In (A) we show the Hellinger distance (HD) between 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.

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 exten-

sion of (1) and reads:

$$
\begin{aligned}
&G \xrightarrow{\rho_u} G + P,\ P \xrightarrow{1} \emptyset, \\
&P + G \underset{\sigma_u}{\overset{\sigma_b}{\rightleftharpoons}} G^*,\ P + G^* \underset{\delta_u}{\overset{\delta_b}{\rightleftharpoons}} G^{**}, \\
&G^{**} \xrightarrow{\rho_b} G^{**} + P.
\end{aligned} \tag{37}
$$

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

$$
\frac{d\langle n \rangle_d}{dt} \approx \frac{LR\rho_u + \rho_b \langle n \rangle_d^2}{(L + \langle n \rangle_d)R + \langle n \rangle_d^2} - \langle n \rangle_d, \tag{38}
$$

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:

$$
\begin{aligned}
T^+(n) &= \frac{LR\rho_u + \rho_b n^2}{(L+n)R + n^2}, \\
T^-(n) &= n.
\end{aligned} \tag{39}
$$

In the limit of small L and R, Eq. (39) reduces to the effective propensities given by Eq. (14) while in the limit of large L and R, Eq. (39) reduces to the effective propensities given by Eq. (12). Hence the solutions of the heurestic master equation in these two limits are given by Eq. (17) and Eq. (13).

By inspection of the reaction scheme (37) 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 (37) in the limit of small and large L, R is Poisson and given by Eq. (32) and Eq. (31) 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.* Note that the derivations here assume the exchangeability of the limits of large/small L, R and large time; hence the proof here presented is not formal but the results are the expected ones and are confirmed by simulations (see later).

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 Appendix A that the deterministic rate equations agreed with the mean of the exact master equation when $\langle n \rangle = \langle n|G \rangle$. 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:

$$
\begin{aligned}
\langle g \rangle_d &= \frac{LR}{\langle n \rangle_d^2 + (L + \langle n \rangle_d)R}, \\
\langle g^{**} \rangle_d &= \frac{\langle n \rangle_d^2}{\langle n \rangle_d^2 + (L + \langle n \rangle_d)R}.
\end{aligned} \tag{40}
$$

Next we derive equations for the same quantities but from the exact master equation (denoted $\langle g \rangle$, $\langle g^* \rangle$ and $\langle g^{**} \rangle$). Writing the master equation for the same two reversible reactions in equilibrium, one can deduce the moment equations:

$$
\partial_t \langle g \rangle = 0 = -\sigma_b \langle g \rangle + \sigma_u (1 - \langle g \rangle - \langle g^{**} \rangle), \tag{41}
$$

$$
\partial_t \langle g^{**} \rangle = 0 = \delta_b \langle n(1 - g - g^{**}) \rangle - \delta_u \langle g^{**} \rangle, \tag{42}
$$

from which we can deduce:

$$\langle g\rangle = \frac{L(R+\langle n|G^{**}\rangle - \langle n\rangle)}{\langle n|G\rangle(R+\langle n|G^{**}\rangle) + L(R+\langle n|G^{**}\rangle - \langle n|G\rangle)},$$
$$\langle g^{**}\rangle = \frac{L(\langle n\rangle - \langle n|G\rangle) + \langle n\rangle\langle n|G\rangle}{\langle n|G\rangle(R+\langle n|G^{**}\rangle) + L(R+\langle n|G^{**}\rangle - \langle n|G\rangle)}, \tag{43}$$

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. (40) and Eq. (43), we see that they can only be equal if the following condition is true:

$$\langle n|G\rangle = \langle n|G^{**}\rangle = \langle n\rangle. \tag{44}$$

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 agreement of the time-evolution equations for the mean protein number according to the deterministic rate equations and the exact master equation:

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

Since the heuristic is based on the deterministic, we expect that this condition is also an indicator of when the heuristic and exact master equations agree. In Fig. 7 we verify this intuition using simulations: 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 [50]. Studies have elucidated the implications of taking into account protein bursting on downstream pathways and shown its importance [51]. Hence we now consider a feedback loop with an implicit mRNA description which has the effective reaction scheme:

$$G \xrightarrow{\rho_u} G + wP,\ P \xrightarrow{1} \emptyset,\ P + G \underset{\sigma_u}{\overset{\sigma_b}{\rightleftharpoons}} G^{*},\ G^{*} \xrightarrow{\rho_b} G^{*} + wP, \tag{46}$$

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 [52] and verified experimentally [53].

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:

$$\frac{dP_a(n,t)}{dt} = \sum_{i=0}^{\infty}(T_i^{+}(n-i)P_a(n-i,t) - T_i^{+}(n)P_a(n,t)) + T^{-}(n+1)P_a(n+1,t) - T^{-}(n)P_a(n,t), \tag{47}$$

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

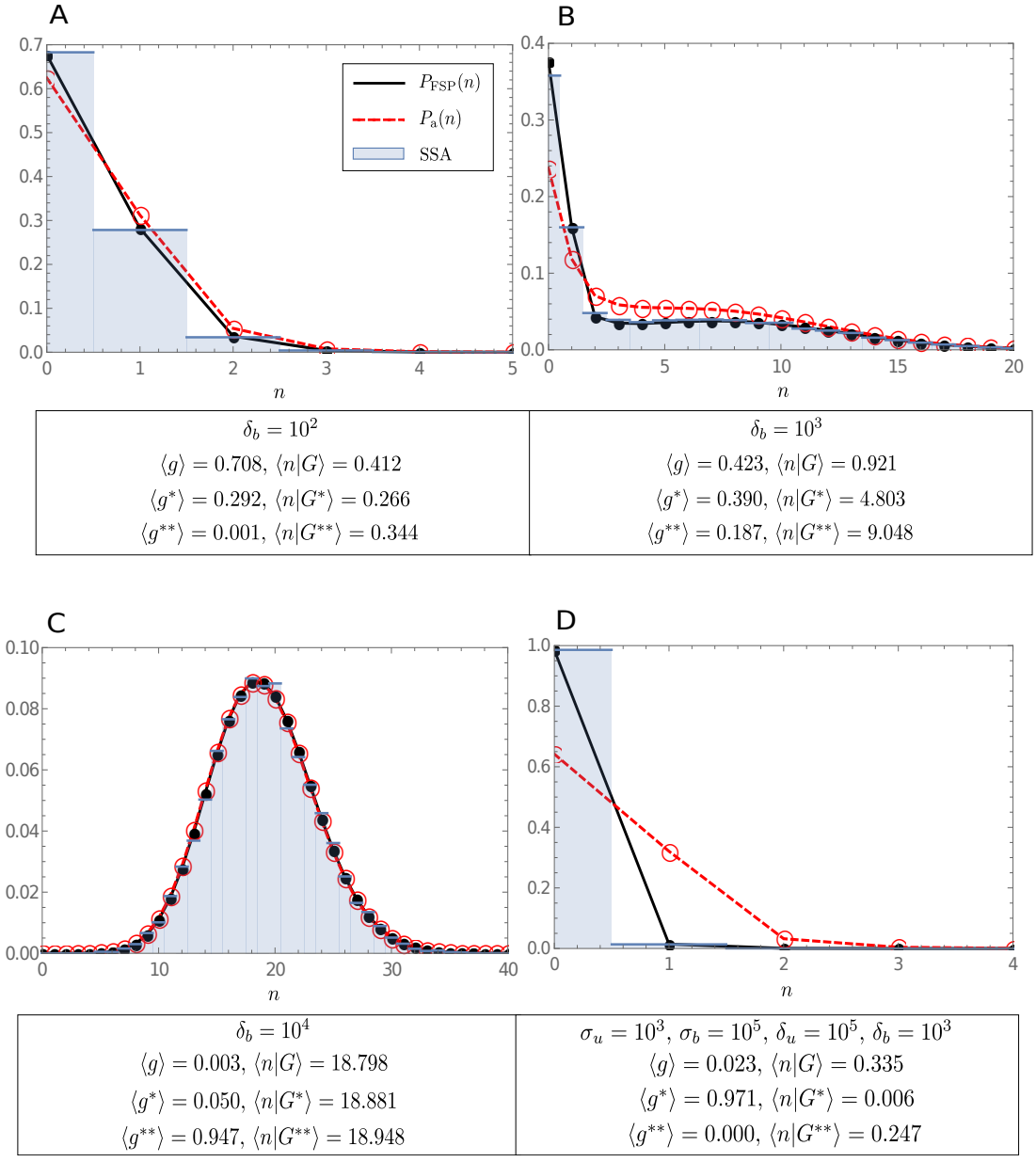

Figure 7: Plots comparing the heuristic master equation (Eq. (9) with Eq. (39)) and exact master equation predictions for the protein number distributions of a positive feedback loop ($\rho_u = 0.5, \rho_b = 20$) with multiple protein binding (37). All calculations done using FSP. Fast promoter switching 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. Subfigure (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. (45).

corresponding to (46) are given by:

$$\frac{d\langle g(t)\rangle_d}{dt} = -\sigma_b\langle g(t)\rangle_d\langle n(t)\rangle_d + \sigma_u(1-\langle g(t)\rangle_d), \tag{48}$$

$$\frac{d\langle n(t)\rangle_d}{dt} = -\sigma_b\langle g(t)\rangle_d\langle n(t)\rangle_d + \sigma_u(1-\langle g(t)\rangle_d) + \rho_u b\langle g(t)\rangle_d + \rho_b b(1-\langle g(t)\rangle_d) - \langle n(t)\rangle_d. \tag{49}$$

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, $\partial_t\langle g(t)\rangle_d \approx 0$, it follows that an effective reduced rate equation for the mean protein numbers is:

$$\frac{d\langle n\rangle_d}{dt} \approx \frac{L\rho_u + \rho_b\langle n\rangle_d}{L+\langle n\rangle_d} b - \langle n\rangle_d. \tag{50}$$

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. (47) with effective propensities:

$$\begin{aligned} T_i^+(n) &= \frac{L\rho_u + \rho_b n}{L+n}\psi_i, \\ T^-(n) &= n, \end{aligned} \tag{51}$$

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. (47) with propensities given by Eq. (51) then it follows that:

$$\begin{aligned} \frac{d\langle n\rangle_a}{dt} &= \sum_{i=0}^{\infty} i\langle T_i^+(n)\rangle_a - \langle T^-(n)\rangle_a, \\ &= b\left\langle \frac{L\rho_u + \rho_b n}{L+n}\right\rangle_a - \langle n\rangle_a, \\ &\approx \frac{L\rho_u + \rho_b\langle n\rangle_a}{L+\langle n\rangle_a} b - \langle n\rangle_a. \end{aligned} \tag{52}$$

Note that this equation is the same as the reduced rate equation Eq. (50) (upon replacing $\langle n\rangle_a$ by $\langle n\rangle$) and hence verifies that the form of the effective propensities given by Eq. (51) 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 $\langle n\rangle_a \approx n$.

The heuristic stochastic model given by Eqs. (47,51) 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 [44]. However, as we now show, provided we can assume exchangeability of the limits of large/small L and large time then closed-form solutions can be obtained for the steady-state distributions of protein numbers.

3.1.1 The limit of large L

In this limit, Eq. (51) reduces to the simpler form:

$$\begin{aligned} T_i^+(n) &\approx \rho_u\psi_i, \\ T^-(n) &= n. \end{aligned} \tag{53}$$

Substituting these in the heuristic reduced master equation Eq. (47), multiplying both sides by z^n and taking the sum over n on both sides of this equation we get the generating function equation:

$$\frac{dG(z,t)}{dt} \approx \rho_u G(z,t)\left(\frac{1}{1+b(1-z)} - 1\right) + (1-z)\frac{dG(z,t)}{dz}, \tag{54}$$

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:

$$P_a(n) = \frac{1}{n!}\frac{d^n G(z)}{dz^n}\bigg|_{z=0} \approx \left(\frac{b}{1+b}\right)^n \left(1 - \frac{b}{1+b}\right)^{\rho_u} \frac{\Gamma(\rho_u + n)}{\Gamma(n+1)\Gamma(\rho_u)} = \mathrm{NB}\left(\rho_u, \frac{b}{(1+b)}\right), \tag{55}$$

where $\mathrm{NB}(x,y)$ 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. (51) reduces to the simpler form:

$$\begin{aligned} T_i^+(n) &\approx ((\rho_u - \rho_b)\delta(0,n) + \rho_b)\psi_i, \\ T^-(n) &= n, \end{aligned} \tag{56}$$

where $\delta(0,n)$ is the Kronecker delta. Substituting these in the heuristic reduced master equation Eq. (47), 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:

$$\frac{dG(z,t)}{dt} \approx ((\rho_u - \rho_b)G(0,t) + \rho_b G(z,t))\left(\frac{1}{1+b(1-z)} - 1\right) + (1-z)\frac{dG(z,t)}{dz}. \tag{57}$$

In steady-state, this equation has the solution:

$$G(z) = \frac{1 + N(-1 + (1+b)^{\rho_b}(1 - b(z-1))^{-\rho_b})}{1 + N(-1 + (1+b)^{\rho_b})}. \tag{58}$$

Hence the steady-state probability distribution is given by:

$$P_a(n) \approx \frac{1}{n!}\frac{d^n G(z)}{dz^n}\bigg|_{z=0} = \begin{cases} \frac{1}{1+N((1+b)^{\rho_b}-1)}, & \text{if } n = 0, \\ \mathrm{NB}\left(\rho_b, \frac{b}{1+b}\right)\frac{N}{N-(N-1)(1+b)^{-\rho_b}}, & \text{if } n \geq 1. \end{cases} \tag{59}$$

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 (46) 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 (46) is approximately in equilibrium for all times. From the chemical master equation for this reversible reaction one finds that the fraction of time that the gene is ON is given by Eq. (71). 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:

$$G^* \xrightarrow{\rho_b} G^* + wP, P \xrightarrow{1} \emptyset, \tag{60}$$

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, $P(n) \approx \mathrm{NB}(\rho_b, b/(1+b))$. 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:

$$G \xrightarrow{\rho_u} G + wP, P \xrightarrow{1} \emptyset. \tag{61}$$

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

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, $P(n) = P_a(n) = \mathrm{NB}(\rho_u, b/(1+b))$) but disagree for small L: the exact reduction predicts a negative binomial distribution, $P(n) = \mathrm{NB}(\rho_b, b/(1+b))$ while the heuristic reduction predicts the different distribution given by Eq. (59). 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:

$$\langle n \rangle = b\rho_b, \quad \langle n \rangle_a = b\rho_b + \frac{b(N-1)\rho_b}{1+((1+b)^{\rho_b}-1)N}, \tag{62}$$

$$\mathrm{FF} = 1+b, \quad \mathrm{FF}_a = 1+b+\frac{b(1-N)\rho_b}{1+((1+b)^{\rho_b}-1)N}, \tag{63}$$

$$\mathrm{CV}^2 = \frac{1+b}{b\rho_b}, \quad \mathrm{CV}_a^2 = \frac{1+b}{b\rho_b} - \frac{(1+b)^{-\rho_b}(1-N^{-1})(1+b+b\rho_b)}{b\rho_b}. \tag{64}$$

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. (50) 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 $N > 1$ ($\mathrm{FF}_a > 1$ and $\mathrm{FF}_a < 1$ 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 $e_m = |\langle n \rangle - \langle n \rangle_a|/\langle n \rangle$, $e_{\mathrm{FF}} = |\mathrm{FF} - \mathrm{FF}_a|/\mathrm{FF}$ and $e_{\mathrm{CVs}} = |\mathrm{CV}^2 - \mathrm{CV}_a^2|/\mathrm{CV}^2$, respectively. The errors for the bursty feedback loop (computed using Eqs. (62-64)) are smaller than the errors for the non-bursty feedback loop (computed using Eqs. (33-35)) 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 $P(n) = \mathrm{NB}(\rho_b, b/(1+b))$ and the distribution from the heuristic reduction Eq. (59) respectively:

$$\frac{P(1)}{P(0)} < 1 \Rightarrow \rho_b < \frac{1+b}{b}, \tag{65}$$

$$\frac{P_a(1)}{P_a(0)} < 1 \Rightarrow N\rho_b = \rho_u < \frac{1+b}{b}. \tag{66}$$

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 $n = 0$.

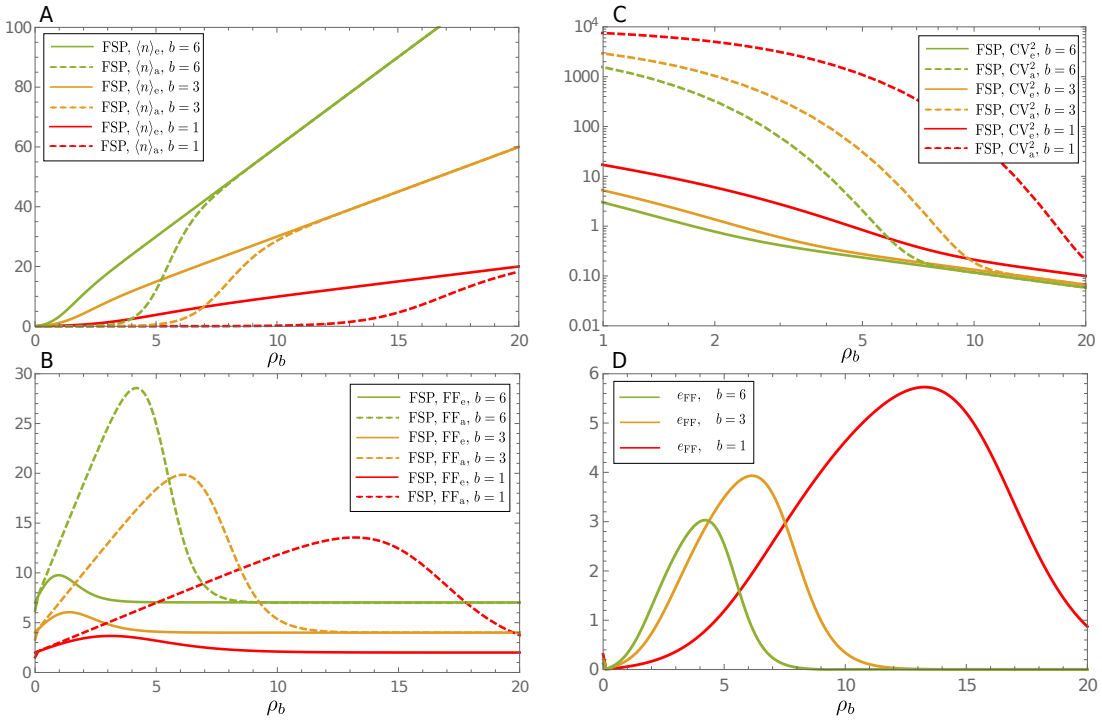

Figure 8: Plots showing the breakdown of the heuristic reduced master equation for fast promoter switching in the limit of small L for a bursty positive feedback loop. The plots show the mean protein number number (A), the Fano Factor of protein number fluctuations (B) and the Coefficient of Variation 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 (46) with the FSP solution of the heuristic reduced master equation Eq. (47) with Eq. (51); 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. (62)-(64) – 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 $e_{\mathrm{FF}} = |\mathrm{FF}_e - \mathrm{FF}_a|/\mathrm{FF}_e$. 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.

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 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. Surprisingly, we found that the heuristic solution exactly corresponds to the fast gene switching limit of the autoregulatory system that ignores protein number fluctuations due to the protein-promoter binding reaction. Our work further builds on previous work by other authors [40,41,54] 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 [55–59] 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 [60] 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.

Author Contributions

J. H. implemented and ran all simulations, produced the figures and performed some of the calculations. R. G. conceived the study, directed it, performed most of 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 Relationship between the deterministic and stochastic models

For the non-bursty feedback loop (1), the stochastic description is given by the chemical master equation which can be formulated as a set of two coupled equations:

$$\begin{aligned}\frac{dP_0(n,t)}{dt} &= \rho_u(P_0(n-1,t) - P_0(n,t)) + ((n+1)P_0(n+1,t) - nP_0(n,t)) + \sigma_u P_1(n-1,t) - \sigma_b n P_0(n,t),\\ \frac{dP_1(n,t)}{dt} &= \rho_b(P_1(n-1,t) - P_1(n,t)) + ((n+1)P_1(n+1,t) - nP_1(n,t)) - \sigma_u P_1(n,t)\\ &\quad + \sigma_b(n+1)P_0(n+1,t),\end{aligned} \tag{67}$$

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 $P(n,t) = P_0(n,t) + P_1(n,t)$. Using these equations it is straightforward to show that the time-evolution equations for the first moments:

$$\frac{d\langle g\rangle}{dt} = -\sigma_b\langle ng\rangle + \sigma_u(1-\langle g\rangle), \tag{68}$$

$$\frac{d\langle n\rangle}{dt} = -\sigma_b\langle ng\rangle + \sigma_u(1-\langle g\rangle) + \rho_u\langle g\rangle + \rho_b(1-\langle g\rangle) - \langle n\rangle, \tag{69}$$

where $\langle n\rangle = \sum_n nP(n)$ is the mean numbers of proteins, $\langle g\rangle = \sum_n P_0(n)$ is the fraction of time the gene is in the ON state (or equivalently the average number of gene in ON state) and $\langle ng\rangle = \sum_n nP_0(n)$. Note that we have here suppressed the time dependence for convenience. A comparison of Eqs. (2)-(3) with Eqs. (68)-(69) shows that the two are the same if $\langle ng\rangle = \langle n\rangle\langle g\rangle$, i.e. the deterministic and exact stochastic models agree in the means if the fluctuations in the gene and protein numbers are independent of each other.

We next show that in the limit of fast promoter switching and when there is a non-zero correlation between the fluctuations of protein and gene, the exact stochastic model gives a time-evolution equation for the protein number mean which is different than that given by the deterministic analysis. The limit of fast promoter switching implies that $d\langle g\rangle/dt \approx 0$ and hence using Eq. (68) it follows that the mean number of proteins conditional on the gene being in state G can be written as:

$$\langle n|G\rangle = \frac{\sum_{n=0}^{\infty} nP_0(n)}{\sum_{n=0}^{\infty} P_0(n)} = \frac{\langle ng\rangle}{\langle g\rangle} \approx L\frac{(1-\langle g\rangle)}{\langle g\rangle}, \tag{70}$$

from which we obtain after rearrangement:

$$\langle g\rangle \approx \frac{L}{L+\langle n|G\rangle}. \tag{71}$$

Substituting in Eq. (69) we obtain an effective equation for the time-evolution of the protein numbers under the condition of fast promoter switching:

$$\frac{d\langle n\rangle}{dt} \approx \frac{L\rho_u + \rho_b\langle n|G\rangle}{L+\langle n|G\rangle} - \langle n\rangle. \tag{72}$$

Contrasting this equation with the effective equation obtained through the deterministic approach, Eq. (4) we see that the two are generally different. They are only the same when $\langle n|G\rangle = \langle n\rangle$, i.e. the mean number of proteins conditional on the gene being in state G is equal to the mean number of proteins which occurs when gene and protein number fluctuations are independent.

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

The exact steady-state solution of Eq. (67) has been previously reported in the literature [7, 61] and is given by:

$$P_e(n) = \frac{1}{n!}\frac{d(G_0(z)+G_1(z))^n}{dz^n}\bigg|_{z=0}, \tag{73}$$

$$G_0(z) = A^{-1}\exp\left(\rho_b(z-1)\right)\left(\frac{1+\sigma_b}{\sigma_b}\frac{\alpha}{\rho_u}M(1+\alpha,\beta,w(z)) - \frac{\alpha}{\rho_u-\rho_b}M(\alpha,\beta,w(z))\right), \tag{74}$$

$$G_1(z) = A^{-1}\exp\left(\rho_b(z-1)\right)M(\alpha,\beta,w(z)), \tag{75}$$

with the definitions:

$$A = \frac{1+\sigma_b}{\sigma_b}\frac{\alpha}{\rho_u}M(1+\alpha,\beta,w(1)) + M(\alpha,\beta,w(1))(1-\frac{\alpha}{\rho_u-\rho_b}), \tag{76}$$

$$\alpha = \frac{\sigma_u(\rho_u-\rho_b)}{\rho_u-\rho_b(1+\sigma_b)}, \tag{77}$$

$$\beta = 1 + \frac{\sigma_u+\sigma_b\frac{\rho_u}{1+\sigma_b}}{1+\sigma_b}, \tag{78}$$

$$w(z) = (\rho_u-\rho_b(1+\sigma_b))\frac{(1+\sigma_b)z-1}{(1+\sigma_b)^2}. \tag{79}$$

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:

$$P(n) = \frac{(1+L)N\rho_b^n(n\rho_b + L(L+n+N\rho_b))[1+LN]_n}{AM(1+LN,1+L,\rho_b)+BM(2+LN,2+L,\rho_b))}, \tag{80}$$

where

$$A = (LN+n)n!(1+L)(L+(N-1)\rho_b)[1+L]_n, \tag{81}$$

$$B = (LN+n)n!(1+LN)\rho_b[1+L]_n. \tag{82}$$

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

C.1 Interchanging the sum and the limit

Here we prove a result which will be used in Sections C.2 and C.3, for the purpose of interchanging limits of L with the infinite sum that defines the Kummer function (for example that in Eq. (11) and Eq. (28)). For reference, the Kummer function is defined through the sum:

$$M(\alpha,\beta,x) = \sum_{n=0}^{\infty}\frac{[\alpha]_n}{[\beta]_n}\frac{x^n}{n!}. \tag{83}$$

We will prove that:

$$\lim_{L\to A}\sum_{n=0}^{\infty}\frac{[a+LN]_n}{[a+L]_n}\frac{\rho_b^n}{n!} = \sum_{n=0}^{\infty}\lim_{L\to A}\frac{[a+LN]_n}{[a+L]_n}\frac{\rho_b^n}{n!}, \tag{84}$$

where A is some real number.

Let $f_n = [a+LN]_n/[a+L]_n$ which implies $f_{n+1} = f_n(a+n+LN)/(a+n+L)$. Consider first the case $N \geq 1$. Since $(a+n+LN)/(a+n+L) \leq N$ then if $f_n \leq N^n$ this implies that $f_{n+1} \leq N^{n+1}$. Also

it is easy to check that $f_1 \leq N$. Hence by induction it follows that if $N \geq 1$ then $f_n \leq N^n$. Similarly it is straightforward to prove by induction that if $N < 1$ then $f_n < 1$. It then follows that:

$$\sum_{n=0}^{\infty} \frac{[a+LN]_n}{[a+L]_n} \frac{\rho_b^n}{n!} \leq \exp(N\rho_b), \quad N \geq 1, \tag{85}$$

$$\sum_{n=0}^{\infty} \frac{[a+LN]_n}{[a+L]_n} \frac{\rho_b^n}{n!} \leq \exp(\rho_b), \quad N < 1 \tag{86}$$

Since the sums are bounded by a finite value, it follows by the dominated convergence theorem that the limit and sum can be switched, i.e. Eq. (84) holds.

C.2 The limit of large L

Making use of a standard result [62]:

$$\lim_{x\to\infty} \frac{\Gamma(x+n)}{\Gamma(x)x^n} = 1, \quad \forall n \in \mathbb{Z}, \tag{87}$$

and $[x]_n = \Gamma[x+n]/\Gamma[x]$, we have that for any $a \in \mathbb{R}$:

$$\lim_{L\to\infty} \frac{[a+LN]_n}{[a+L]_n} = N^n, \quad \lim_{L\to\infty} M(a+LN, a+L, \rho_b) = \sum_{n=0}^{\infty} \lim_{L\to\infty} \frac{[a+LN]_n}{[a+L]_n} \frac{\rho_b^n}{n!} = \exp(N\rho_b). \tag{88}$$

Note that here we have used the interchange of sum and limit as given by Eq. (84). Using Eq. (88) 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. (11) and by the full master equation in the limit of fast promoter switching Eq. (28) is a Poissonian with mean ρ_u:

$$\lim_{L\to\infty} P(n) = \lim_{L\to\infty} P_a(n) = \frac{\rho_u^n \exp(-\rho_u)}{n!} \tag{89}$$

C.3 The limit of small L

Using the definition of the Pochhammer symbol and the standard result $\Gamma(x) \sim 1/x$ as $x \to 0$, one can show that:

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

and

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

Using these two results, it then follows that:

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

and

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

Note that here we have used the interchange of sum and limit as given by Eq. (84). 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. (11) reduces to:

$$\lim_{L\to 0} P_a(n) = \begin{cases} \frac{1}{1+N(\exp(\rho_b)-1)}, & \text{if } n = 0, \\ \frac{\exp(-\rho_b)\rho_b^n}{n!}\left(1 + \frac{N-1}{1+N(\exp(\rho_b)-1)}\right), & \text{if } n \geq 1 \end{cases} \tag{90}$$

while the steady-state distribution of molecule numbers predicted by the master equation in the limit of fast promoter switching Eq. (28) reduces to:

$$\lim_{L\to 0} P(n) = \frac{\rho_b^n \exp(-\rho_b)}{n!}. \tag{91}$$

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] David F Anderson, Gheorghe Craciun, and Thomas G Kurtz. Product-form stationary distributions for deficiency zero chemical reaction networks. Bulletin of mathematical biology, 72(8):1947–1970, 2010.

[10] Daniele Cappelletti and Carsten Wiuf. Product-form poisson-like distributions and complex balanced reaction systems. SIAM Journal on Applied Mathematics, 76(1):411–432, 2016.

[11] 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.

[12] 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.

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

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

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

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

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

[18] 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.

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

[20] 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.

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

[22] 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.

[23] Karen Ball, Thomas G Kurtz, Lea Popovic, Greg Rempala, et al. Asymptotic analysis of multiscale approximations to reaction networks. The Annals of Applied Probability, 16(4):1925–1961, 2006.

[24] Hye-Won Kang, Thomas G Kurtz, et al. Separation of time-scales and model reduction for stochastic reaction networks. The Annals of Applied Probability, 23(2):529–583, 2013.

[25] Daniele Cappelletti, Carsten Wiuf, et al. Elimination of intermediate species in multiscale stochastic reaction networks. The Annals of Applied Probability, 26(5):2915–2958, 2016.

[26] Jae Kyoung Kim, Grzegorz A Rempala, and Hye-Won Kang. Reduction for stochastic biochemical reaction networks with multiscale conservations. Multiscale Modeling & Simulation, 15(4):1376–1403, 2017.

[27] Jay Newby. Bistable switching asymptotics for the self regulating gene. Journal of Physics A: Mathematical and Theoretical, 48(18):185001, 2015.

[28] 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.

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

[30] 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.

[31] 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.

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

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

[34] 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.

[35] 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.

[36] 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.

[37] 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.

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

[39] 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.

[40] 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.

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

[42] 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.

[43] Ramon Grima. An effective rate equation approach to reaction kinetics in small volumes: Theory and application to biochemical reactions in nonequilibrium steady-state conditions. The Journal of chemical physics, 133(3):07B604, 2010.

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

[45] Yves Vandecan and Ralf Blossey. Self-regulatory gene: an exact solution for the gene gate model. Physical Review E, 87(4):042705, 2013.

[46] Frits Veerman, Carsten Marr, and Nikola Popović. Time-dependent propagators for stochastic models of gene expression: an analytical method. Journal of mathematical biology, 77(2):261–312, 2018.

[47] 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.

[48] 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.

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

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

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

[52] 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.

[53] 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.

[54] 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.

[55] 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.

[56] 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.

[57] 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.

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

[59] 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.

[60] 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.

[61] 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.

[62] *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.