High cooperativity in negative feedback can amplify noisy gene expression

Pavol Bokes[1], Yen Ting Lin[2], and Abhyudai Singh[3]

[1]Department of Applied Mathematics and Statistics, Comenius University, Bratislava 84248, Slovakia
[2]Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, New Mexico 87545, USA
[3]Department of Electrical and Computer Engineering,, University of Delaware, Newark, Delaware 19716, USA

April 9, 2017

Abstract

Burst-like synthesis of protein is a significant source of cell-to-cell variability in protein levels. Negative feedback is a common example of a regulatory mechanism by which such stochasticity can be controlled. Here we consider a specific kind of negative feedback, which makes bursts smaller in the excess of protein. Increasing the strength of the feedback may lead to dramatically different outcomes depending on a key parameter, the noise load, which is defined as the squared coefficient of variation the protein exhibits in the absence of feedback. Combining stochastic simulation with asymptotic analysis, we identify a critical value of noise load: for noise loads smaller than critical, the coefficient of variation remains bounded with increasing feedback strength; contrastingly, if the noise load is larger than critical, the coefficient of variation diverges to infinity in the limit of ever greater feedback strengths. Interestingly, high-cooperativity feedbacks have lower critical noise loads, implying that low-cooperativity feedbacks in burst size can be preferable for noisy proteins. Finally, we discuss our findings in the context of previous results on the impact of negative feedback in burst size and burst frequency on gene-expression noise.

1 Introduction

Due to the low molecular copy numbers involved, gene expression at a single cell level exhibits a considerable degree of noisiness (Elowitz et al, 2002; Blake et al, 2003). Noise in gene expression is transmitted downstream to the final product

1

of gene expression, the protein, whose concentration can fluctuate widely in time (Golding et al, 2005; Suter et al, 2011). The fluctuations are particularly pronounced if the protein is produced in bursts of multiple molecule copies at a single time (Ozbudak et al, 2002; Dar et al, 2012). The temporal fluctuation of protein concentration manifests itself in a population of isogenic cells as a cell-to-cell variability (Taniguchi et al, 2010). Mathematical modelling can be used to analyse the impact of mechanisms involved in gene expression on the population distribution of protein concentration (Shahrezaei and Swain, 2008).

Many proteins control their own expression with negative feedback (Rosenfeld et al, 2002). By reducing the protein synthesis rate when the protein is in excess, negative feedback reduces the mean protein level and can also decrease the protein noise (Thattai and van Oudenaarden, 2001). Nevertheless, in the case of a strong negative feedback, any decrease in noise due to improved regression to the mean can be eclipsed by an increase due to low reaction frequencies and/or low copy numbers (Singh and Hespanha, 2009). Additionally, the noise-reducing effect of time-averaging of the downstream elements, such as the promoter state or the mRNA copy number, can be cancelled under strong feedback, thus enhancing bursting dynamics and increasing the protein noise even further (Stekel and Jenkins, 2008; Grönlund et al, 2013). Bursting in itself imposes limitations on the noise-reducing ability of a strong negative feedback: individual bursts can grow too large to remain under control of a strong feedback, which is sensitive to small changes in protein copy numbers (Scott et al, 2007; Wang et al, 2014; Bokes and Singh, 2016). It is this final mechanism — the loss of control over large bursts — that we shall investigate in this paper using a minimalistic model which purposely disregards other mechanisms.

Feedback in gene expression can operate on the transcriptional or post-transcriptional level. In transcriptional feedback, the protein regulates the rate with which its mRNA transcript is synthesised (Becskei and Serrano, 2000). Provided that the mRNA life span is short and the translation of mRNA is fast, the transcription of an mRNA molecule is followed by a quick burst of protein translation (Cai et al, 2006). Transcriptional feedback thus regulates the rate, or frequency, with which such translational bursts occur (Friedman et al, 2006). Post-transcriptional feedback, on the other hand, regulates the size of translational bursts (Schikora-Tamarit et al, 2016). As a specific example, RNA binding proteins reduce the size of translational bursts by enhancing the degradation of their mRNA transcript (Yates and Nomura, 1981; Swain, 2004; Schikora-Tamarit et al, 2016). Previous studies indicate that post-transcriptional feedback outperforms transcriptional feedback in terms of its noise-reducing capability (Swain, 2004; Singh, 2011; Bokes and Singh, 2016).

Production delay can lead to dramatic changes in qualitative behaviour of feedback systems (Mackey et al, 1977). Deterministic systems can develop sustained oscillations via a supercritical Hopf bifurcation from a destabilised steady state once a threshold value of the delay is exceeded (Murray, 2003). Stochastic systems with negative feedback become more variable upon the inclusion of a small delay (Lafuerza and Toral, 2011) and exhibit increasingly regular oscillations prior to the onset of the Hopf bifurcation (Barrio et al, 2006). Gene-

expression feedback is necessarily delayed, since both the transcription of a gene into an mRNA and the translation of an mRNA into a protein involve a series of multistep processes before they are completed (Alberts et al, 2002). Large delays in gene expression of a negatively regulating transcription factor can lead to physiologically consequential sustained oscillations (Monk, 2003). In this paper we shall see that the ability of strong negative feedback in burst size to contain gene-expression noise is compromised by an infinitesimally small delay that is implicitly included in our model.

While some gene-expression models are exactly solvable (Biancalani and Assaf, 2015; Smith and Shahrezaei, 2015; Kumar et al, 2014; Grima et al, 2012; Bokes et al, 2012; Ramos et al, 2011; Jia and Kulkarni, 2010; Peccoud and Ycart, 1995), one has often to resort to asymptotic methods to obtain tractable results (Dessalles et al, 2017; Popovic et al, 2016; Be'er and Assaf, 2016; Newby, 2015; Bruna et al, 2014; Leier et al, 2014; Platini et al, 2011). Small-noise approximations, which linearise the model around a deterministic steady state, play an important role among the available asymptotic approaches (Elf and Ehrenberg, 2003; Paulsson, 2004; Komorowski et al, 2013; Wang et al, 2014; Thomas and Grima, 2015; Cardelli et al, 2016). However, expressions obtained by a small-noise approximation need not be uniformly valid across the whole parameter space of the model. Bokes and Singh (2016) used an alternative approximation in the regime of strong negative feedback to gain a thorough understanding of the mean and noise behaviour for an exactly solvable model across its parameter space. Here we shall apply small-noise and strong-feedback approximations in a model for which an exact distribution is, to our best knowledge, unavailable.

We adapt a popular framework which models the protein concentration by a continuous-state drift-jump Markov process in which the deterministic drift represents protein decay and the stochastic jumps model the synthesis of protein in bursts (Friedman et al, 2006; Bokes et al, 2013; Lin and Doering, 2016). In the original formulation, the decay of protein is exponential, the jumps occur randomly in time with a given intensity, and their sizes are drawn at random from an exponential distribution (Friedman et al, 2006; Bokes et al, 2013; Lin and Doering, 2016). Extensions to the original formulation also consider non-exponential decay (Mackey et al, 2013; Bokes and Singh, 2015; Soltani et al, 2015), non-exponential bursting distributions (Jedrak and Ochab-Marcinek, 2016b), multiple gene copies (Jedrak and Ochab-Marcinek, 2016a), and multiple-component systems (Yvinec et al, 2014; Mackey and Tyran-Kaminska, 2015; Lin and Galla, 2016; Pájaro et al, 2017). Most previous studies include feedback in burst frequency: the jump intensity depends on the current level of protein. Bokes and Singh (2016) introduced a particular form of feedback in burst size, in which the hazard rate for burst termination depends on the current protein concentration. Here we shall consider a different formulation of feedback in burst size, in which the protein concentration immediately before a jump determines the mean jump size. We shall argue that the two formulations of burst-size feedback differ in the inclusion or omission of an infinitesimally small delay in protein synthesis. As we shall see, the two formulations have strikingly different noise-reduction
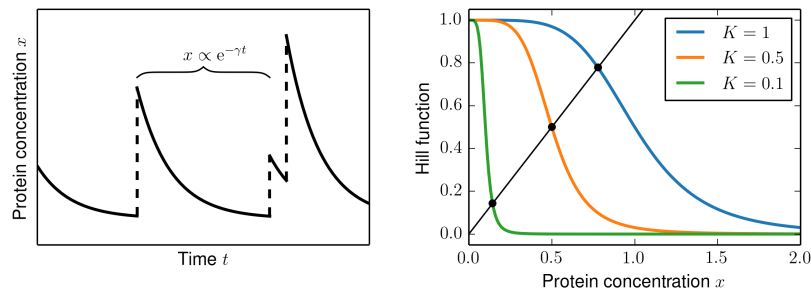
Figure 1: *Left:* A sample trajectory of the drift-jump process modelling the dynamics of protein concentration. *Right:* Hill functions $(1 + (x/K)^5)^{-1}$ and their fixed points for selected values of K.

capabilities.

The modelling framework is first introduced in Section 2 for the simple case of a constitutively expressed protein and then extended by feedback in burst-size in Section 3. The fundamental features of the model behaviour are illustrated in Section 4 using stochastic simulations. Sections 5 and 6 contain a systematic asymptotic analysis of the model behaviour in the small-noise and strong-feedback regimes. Section 7 compares the current model to its alternatives in (Bokes and Singh, 2016) and Section 8 concludes with a summary and discussion of the main findings.

2   Constitutive case

We model the dynamics of protein concentration $x(t)$ by a Markovian drift-jump process (Schuss, 2009), whose trajectories look like the one in Figure 1, left panel. Except for a countable number of discontinuities, the process $x(t)$ decays deterministically with a rate constant γ. The jumps (bursts of protein expression) occur randomly in time with a constant burst frequency a; the waiting time between two consecutive bursts is exponentially distributed with mean $1/a$. The size of the discontinuity (burst size) is always positive and randomly chosen from an exponential distribution with mean b (Friedman et al, 2006).

There are three dimensional parameters in the model so far: decay rate constant γ (units of time^{-1}), burst frequency a (units of time^{-1}), and mean burst size b (units of concentration). We choose the units of time and concentration so that these parameters take the numerical values of

$$\gamma = 1, \quad a = \varepsilon^{-1}, \quad b = \varepsilon,$$

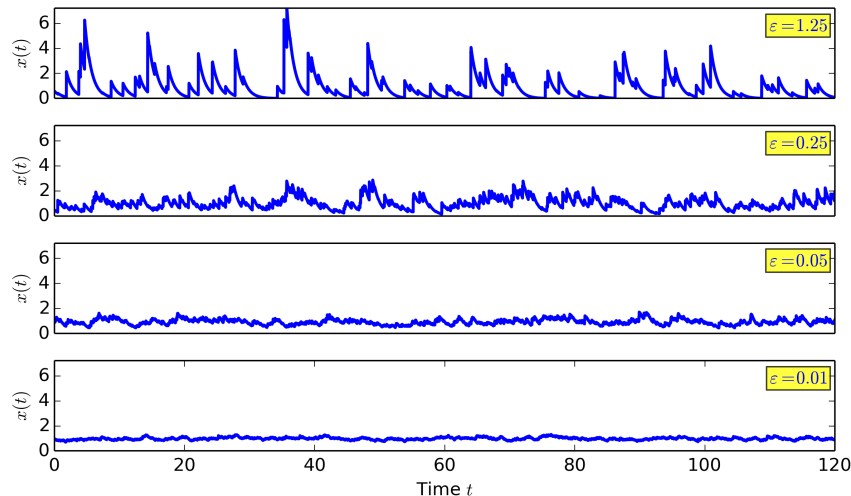where ε is a positive dimensionless real number (Fig. 2).

Figure 2: Sample trajectories for a constitutively expressed protein for selected values of ε.

The probability density function (pdf) $p(x,t)$ of the Markov process $x(t)$ satisfies a master equation (cf. Schuss, 2009)

$$\frac{\partial p}{\partial t} - \frac{\partial(xp)}{\partial x} = \frac{1}{\varepsilon}\left(\int_0^x B(x-x')p(x',t)\mathrm{d}x' - p(x,t)\right), \tag{1}$$

where

$$B(s) = \frac{\mathrm{e}^{-\frac{s}{\varepsilon}}}{\varepsilon} \tag{2}$$

is the exponential burst-size probability density function. The normalised stationary solution to (1)–(2) is given by the gamma probability distribution (cf. Friedman et al, 2006)

$$p(x) \propto \mathrm{e}^{-\frac{x}{\varepsilon}} x^{\frac{1}{\varepsilon}-1}. \tag{3}$$

Here and elsewhere we shall use the same symbol for the independent variable of a pdf and a random variable (or a stationary process) associated with the pdf. For the gamma protein distribution in (3), the mean and variance are given by

$$\langle x \rangle = 1, \quad \mathrm{Var}(x) = \varepsilon. \tag{4}$$

We henceforth call ε the *noise load* and say that the model operates in the *small-noise regime* if $\varepsilon \ll 1$. We can normalise the protein concentration by writing

$$x = 1 + \sqrt{\varepsilon}\xi. \tag{5}$$

5

The stationary pdf satisfies

$$p \propto x^{-1}\mathrm{e}^{-\frac{x-\ln(x)}{\varepsilon}} \propto (1 + \sqrt{\varepsilon}\xi)^{-1}\mathrm{e}^{-\frac{\sqrt{\varepsilon}\xi - \ln(1+\sqrt{\varepsilon}\xi)}{\varepsilon}} = \mathrm{e}^{-\frac{\xi^2}{2}} + O(\sqrt{\varepsilon}),$$

implying that in the small-noise regime the protein concentration is approximately normally distributed.

3   Feedback in burst size

We incorporate negative autoregulation of burst size into the model of Section 2 by making the mean burst size b depend on the concentration x of the protein at the time *immediately before* the burst occurs. Specifically, we shall consider a Hill-type dependence

$$b(x) = \frac{\varepsilon}{1 + (x/K)^H} \tag{6}$$

of the mean burst size on the protein concentration (Fig. 1, right panel). The critical concentration K, which gives the amount of protein that is required to halve the expected burst size, is a reciprocal measure of feedback strength; we say that the model operates in the *strong-feedback regime* if $K \ll 1$. The cooperativity coefficient H determines how steeply the feedback responds to changes in protein concentration.

The master equation for the pdf $p(x, t)$ of the process $x(t)$ with negative feedback in burst size is given by

$$\frac{\partial p}{\partial t} - \frac{\partial(xp)}{\partial x} = \frac{1}{\varepsilon}\left(\int_0^x B(x - x'; x')p(x', t)\mathrm{d}x' - p(x, t)\right), \tag{7}$$

where

$$B(s; x') = \varepsilon^{-1}(1 + (x'/K)^H)\mathrm{e}^{-\varepsilon^{-1}(1+(x'/K)^H)s} \tag{8}$$

gives the burst-size density on condition of x' protein being present immediately before the burst. Notably, the convolution on the right-hand side of (1) of the pdf with an exponential density (2) is replaced in (7) by a more challenging integral kernel (8). Being unaware of the availability or otherwise of explicit solutions to (7)–(8), we shall resort to asymptotic methods to analyse its normalised stationary solution in the small-noise ($\varepsilon \ll 1$) and strong-feedback ($K \ll 1$) regimes.

While solving the master equation becomes problematic after the inclusion of feedback in burst size, simulating stochastic trajectories remains extraordinarily simple. Algorithm 1 loops over a basic procedure which consists of drawing the waiting time τ until the next burst (Line 4), setting the trajectory to an exponential in between the bursts (Line 5), and drawing the next burst's size (Line 6). The waiting time and burst size are both exponentially distributed and can be drawn using the inversion sampling method by multiplying their respective mean values by $-\ln u$, where u is drawn independently of any other drawings from the uniform distribution in the unit interval.

6

Require: Parameter values (ε, H, K).
Return: A sample trajectory $x(t)$, $t \geq 0$.
 1: Initialise the current time and state: $t \leftarrow 0$; $x(0) \leftarrow 0$.
 2: **loop**
 3:     Draw u and $\tilde{u}$ from the standardised uniform distribution.
 4:     Set $\tau \leftarrow -\varepsilon \ln u$.
 5:     Set $x(t+s) \leftarrow x(t)\mathrm{e}^{-s}$ for $0 < s < \tau$.
 6:     Set $x(t+\tau) \leftarrow x(t)\mathrm{e}^{-\tau} - \frac{\varepsilon \ln \tilde{u}}{1+(x(t)\mathrm{e}^{-\tau}/K)^H}$.
 7:     Update the current time $t \leftarrow t + \tau$.
 8: **end loop**

Algorithm 1: Simulation algorithm

4 Critical noise load

Stochastic simulations by Algorithm 1 allow us to investigate the response of protein noise to the strengthening of feedback in burst size. We choose to quantify protein noise by the (squared) coefficient of variation (CV^2), which is defined as the ratio of variance over the mean squared, i.e.

$$\mathrm{CV}^2 = \frac{\langle x^2 \rangle}{\langle x \rangle^2} - 1.$$

By (4), the squared coefficient of variation reduces to the value of the noise load ε in the absence of feedback ($K \to \infty$). The ratio of the CV^2 of a regulated protein to ε, which we shall refer to as the relative CV^2, gives the fold increase in protein noise following the introduction of feedback.

Depending on whether the noise load ε is less or greater than a critical value $\varepsilon_{\mathrm{c}} = 1/2H$, we observe two different responses of protein noise to strengthening feedback. If the noise load is subcritical ($\varepsilon < \varepsilon_{\mathrm{c}}$), the noise in protein concentration remains bounded however strong the feedback may be (Fig. 3); for supercritical noise loads ($\varepsilon > \varepsilon_{\mathrm{c}}$), the noise diverges as K tends to zero (Fig. 4). The vertical (concentration) axes in the panels of Figures 3 and 4 extend from zero to a constant multiple (set to 2 in Fig. 3 and to 10 in Fig. 4) of the protein mean; the apparent variability in the sample trajectory then directly corresponds to the numerical value of the coefficient of variation, which we report in the inset of each panel. Each coefficient of variation is calculated from a single run of simulation which is much longer than the short stretch shown in the individual panels (see Appendix A for details).

As we shall demonstrate in Section 6 systematically using asymptotic analysis, the protein mean $\langle x \rangle$, the mean square $\langle x^2 \rangle$, and the variance all tend to zero, in subcritical and supercritical cases alike, as feedback strengthens. The increase of the coefficient of variation occurs because in supercritical cases the squared mean $\langle x \rangle^2$ converges faster to zero than the mean square $\langle x^2 \rangle$.

Steeper (more cooperative) feedbacks have lower critical noise loads. If we take the cooperativity coefficient H to infinity, the critical noise load becomes
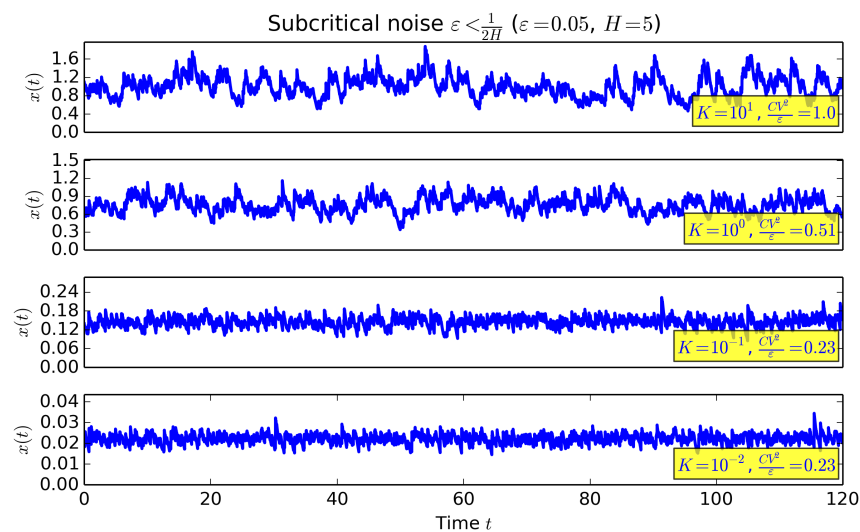
Figure 3: For subcritical noise loads, protein noise remains bounded even under stringent feedback conditions.
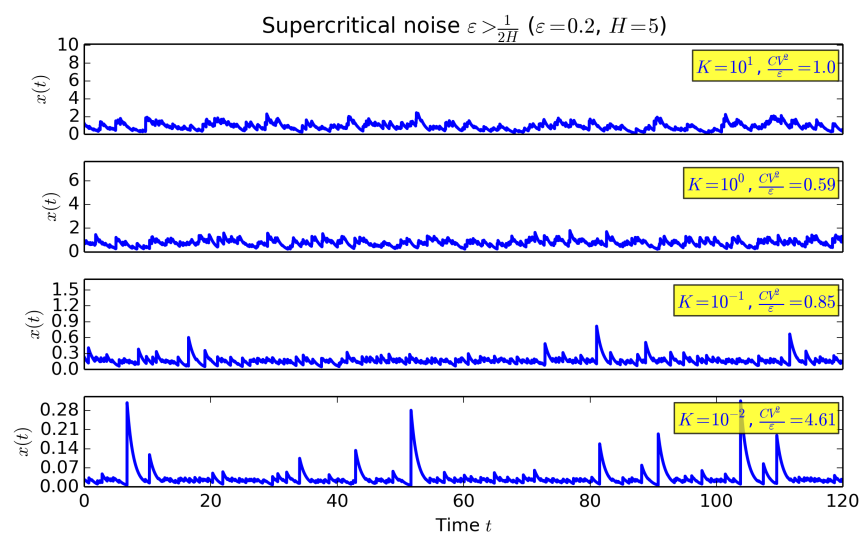

Figure 4: For supercritical noise loads, protein noise diverges for increasingly strong feedback.

zero: any positive noise load is supercritical. The Hill function (6) formally reduces for $H = \infty$ to a step function which is equal to ε for $0 < x < K$ and is zero for $x > K$. Under such step-wise regulation, non-zero bursts occur only if x is smaller than K (i.e. very small), and their size is independent of K (i.e. large); after one such large burst no further bursts can occur—or we may say that they still do but their size is necessarily zero—until the protein concentration slowly decays back beneath K. Clearly, however small, but fixed, ε may be, these infrequent but large bursts will drive the variability to infinity if K is lowered sufficiently. However, for finite cooperativities a non-zero burst would typically fire well before the concentration decays down to levels of O(K), and would be of a lesser size as a result: it is exactly this mechanism that enables low-cooperativity feedbacks to cope with ever increasing feedback strengths; the underlying scaling will be identified and analysed by matched asymptotics in Section 6. Prior to doing that, we perform in the following Section 5 a traditional small-noise analysis, obtaining explicit mean and noise results, which are valid as $\varepsilon \to 0$ for bounded ranges of K.

5 Small-noise asymptotics

The master equation (7) can succinctly be written as a continuity equation

$$\frac{\partial p}{\partial t} + \frac{\partial J}{\partial x} = 0, \quad J = \frac{1}{\varepsilon} \int_0^x p(x', t) e^{-\frac{x - x'}{\varepsilon \theta(x')}} \, dx' - xp, \tag{9}$$

where

$$\theta(x) = \frac{1}{1 + (x/K)^H} \tag{10}$$

gives the factor by which the mean burst size is reduced at protein concentration x due to self-regulation. Being the product of the burst frequency ε^{-1} and the mean burst size (6), it can also be interpreted as the mean protein production rate. The function $\theta(x)$ has a (unique) fixed point x_0, i.e. a solution to

$$\theta(x_0) = x_0, \tag{11}$$

which can be interpreted as the protein concentration at which its mean production rate is equal to its rate of decay (cf. Fig. 1, right panel). It is shown below that in the small-noise regime the steady-state protein concentration is constrained with high probability to a small neighbourhood of the fixed point.

The stationary distribution is the time-independent, zero-flux, and normalised solution $p(x)$ to (9), i.e the normalised solution to the integral equation

$$\frac{1}{\varepsilon} \int_0^x p(x') e^{-\frac{x - x'}{\varepsilon \theta(x')}} \, dx' - xp = 0. \tag{12}$$

In order to examine the behaviour of the solution $p(x) = p(x; \varepsilon)$ to (12) in the small-noise regime ($\varepsilon \ll 1$), we use the transformation

$$x = x_0 + \sqrt{\varepsilon}\xi, \quad x' = x - \varepsilon u, \quad p(x_0 + \sqrt{\varepsilon}\xi) = \tilde{p}(\xi). \tag{13}$$
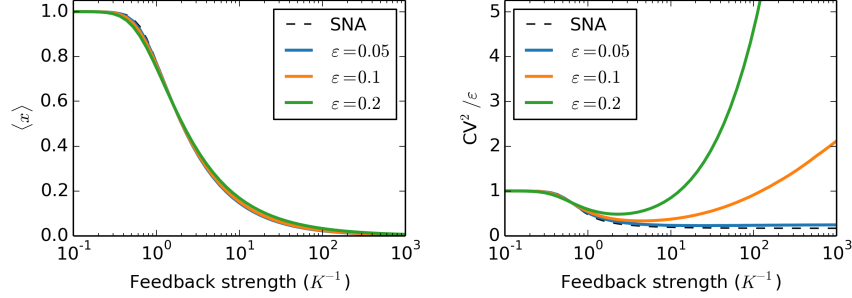
Figure 5: Protein concentration mean and the relative CV^2 in response to increasing feedback strength for $H = 5$ and a selection of values of ε, compared with the small-noise approximation (SNA).

The substitution for x reflects the (presumed) $O(\sqrt{\varepsilon})$ scale of fluctuations of x around the fixed point x_0, cf. Eq. (5); the substitution for x' reflects the $O(\varepsilon)$ scale of individual bursts. Inserting the substitutions (13) into the integral equation (12), we obtain

$$\int_0^{\text{``}\infty\text{''}} \tilde{p}(\xi - \sqrt{\varepsilon}u)\mathrm{e}^{-\frac{u}{\theta(x_0 + \sqrt{\varepsilon}\xi - \varepsilon u)}}\,\mathrm{d}u - x_0\tilde{p} - \sqrt{\varepsilon}\xi\tilde{p} = 0. \tag{14}$$

By extending the upper integration boundary from $\frac{x}{\varepsilon}$ to infinity in (14), we will have introduced an error that is inconsequential (exponentially small) in the $\varepsilon \ll 1$ regime (Hinch, 1991). Inserting the expansions

$$\tilde{p}(\xi - \sqrt{\varepsilon}u) = \tilde{p}(\xi) - \sqrt{\varepsilon}u\frac{\mathrm{d}\tilde{p}}{\mathrm{d}\xi}(\xi) + O(\varepsilon),$$

$$\mathrm{e}^{-\frac{u}{\theta(x_0 + \sqrt{\varepsilon}\xi - \varepsilon u)}} = \mathrm{e}^{-\frac{u}{x_0} + \sqrt{\varepsilon}\frac{\theta'(x_0)\xi u}{x_0^2}} + O(\varepsilon)$$

$$= \mathrm{e}^{-\frac{u}{x_0}}\left(1 + \sqrt{\varepsilon}\frac{\theta'(x_0)\xi u}{x_0^2}\right) + O(\varepsilon)$$

into (14) and carrying out the integration with respect to u we obtain

$$\sqrt{\varepsilon}(-x_0^2\tilde{p}'(\xi) + (\theta'(x_0) - 1)\xi\tilde{p}(\xi)) + O(\varepsilon) = 0, \tag{15}$$

the $O(1)$ terms having cancelled each other out. Dividing (15) by $\sqrt{\varepsilon}$ and neglecting the small $O(\varepsilon)/\sqrt{\varepsilon}$ term, we arrive at a linear first-order differential equation for the leading-order approximation to $\tilde{p}$, solving which yields

$$\tilde{p}(\xi) \sim C\mathrm{e}^{-\frac{(1-\theta'(x_0))\xi^2}{2x_0^2}}, \quad \varepsilon \ll 1, \tag{16}$$

10

where C is the normalisation constant. The result implies that ξ is approximately normally distributed with mean zero and variance $x_0^2/(1-\theta'(x_0))$. Consequently $x = x_0 + \sqrt{\varepsilon}\xi$ has mean x_0 and variance $\varepsilon x_0^2/(1-\theta'(x_0))$, i.e.

$$\langle x \rangle \sim x_0, \quad \frac{\text{CV}^2}{\varepsilon} \sim \frac{1}{1-\theta'(x_0)} = \frac{1}{1+H(1-x_0)}, \quad \varepsilon \ll 1, \qquad (17)$$

hold in the small-noise regime. Also, since the variance is $O(\varepsilon)$, we obtain

$$\langle x^2 \rangle = \langle x \rangle^2 + \text{Var}(x) \sim x_0^2, \quad \varepsilon \ll 1, \qquad (18)$$

for the mean square.

In Fig. 5, we compare for $H = 5$ the small-noise approximations (17) (SNAs; dashed lines in Fig. 5) to the exact (i.e. not asymptotic) mean and relative CV^2 for a range of values of K and three selected values of ε: a subcritical $\varepsilon = 0.05$, the critical $\varepsilon = 0.1$, and a supercritical $\varepsilon = 0.2$. The fixed point x_0 of the function $\theta(x)$ (10) was calculated in the Python programming language with the `fixed_point` routine of the `scipy.optimize` library. For each value of ε, and for the range of values of K, the exact mean and the CV^2 were estimated from the temporal average of a single and long sample path generated by Algorithm 1, which we implemented in the C programming language (see Appendix A for details).

By Fig. 5, left panel, the small-noise approximation of the mean agrees well with the exact mean, decreasing monotonically from one in the absence of regulation ($K = \infty$) to zero under stringent regulation ($K \to 0$). Contrastingly, the small-noise approximation of the CV^2 agrees with the exact CV^2 only if ε is subcritical or K is sufficiently large; in particular, if ε is low and subcritical, the exact relative CV^2 remains close, as $K \to 0$, to the small-noise approximation limit value of $1/(H+1) \approx 0.167$. For supercritical noise loads, however, simulations suggest divergence, as K tends to zero, of the exact CV^2 to infinity, which the small-noise approximation fails to predict.

6 Strong-feedback asymptotics

In the small-noise analysis presented in the previous section, we tacitly assumed that the fixed point x_0 of the Hill function (10) is of order one. However, x_0 becomes increasingly small as K tends to zero; more precisely, the dominant balance in the fixed-point equation $x_0^{-1} = 1 + (x_0/K)^H$ is found between the x_0^{-1} and the $(x_0/K)^H$ terms, the equating of which yields a power-law relationship

$$x_0 \sim K^{\frac{H}{H+1}}, \quad K \ll 1, \qquad (19)$$

between the fixed-point x_0 and the critical concentration K in the strong feedback regime. As x_0 becomes increasingly small, the terms which have been neglected in the small-noise approximation procedure become significant, and

11

the approximation leads to incorrect predictions; in particular, it fails to predict the grow-up of noise under supercritical noise-load conditions (Fig. 5, right panel).

In order to obtain a truthful characterisation of the model behaviour in the strong-feedback regime, we consider the steady-state protein probability density function in this section as a function of K which tends to zero, whereby the value of ε is fixed and not necessarily small. By (12) and (10), the steady-state protein pdf satisfies the integral equation

$$xp(x) = \frac{1}{\varepsilon} \int_0^x p(x')e^{-\frac{(x-x')(1+(x'/K)^H)}{\varepsilon}} \, \mathrm{d}x'. \tag{20}$$

Below we show that a solution $p(x) = p(x; K)$ to (20) tends to two separate distinguished limits, one occurring on the inner $x = O(K^{H/(H+1)})$ concentration scale associated with the fixed-point asymptotics (19), the other emerging on the outer $O(1)$ scale associated with large, uncontrolled, bursts. The relative importance of the inner or outer limits will determine, in particular, whether the CV^2 converges or diverges as K tends to zero.

Inner solution. We rescale the variables of (20) according to

$$x = K^{\frac{H}{H+1}}y, \quad x' = K^{\frac{H}{H+1}}y', \quad p(x) = P(y), \tag{21}$$

obtaining

$$yP(y) = \frac{1}{\varepsilon} \int_0^y P(y')e^{-\frac{(y-y')(K^{H/(H+1)}+(y')^H)}{\varepsilon}} \, \mathrm{d}y'. \tag{22}$$

Inserting

$$P(y; K) = P_0(y) + O(K^{H/(H+1)}) \tag{23}$$

into (22) and collecting leading-order terms yield

$$yP_0(y) = \frac{1}{\varepsilon} \int_0^y P_0(y')e^{-\frac{(y-y')(y')^H}{\varepsilon}} \, \mathrm{d}y'. \tag{24}$$

If $y \ll 1$, then the exponential in the integrand is close to one, implying that $P_0(y) \sim \phi_0(y)$, where $\phi_0(y)$ satisfies $y\phi_0(y) = \frac{1}{\varepsilon} \int_0^y \phi_0(y')\mathrm{d}y'$, i.e. $\phi_0(y) = Cy^{1/\varepsilon-1}$; for uniqueness and algebraic simplicity we fix the arbitrary constant C to one, thus obtaining

$$P_0(y) \sim y^{\frac{1}{\varepsilon}-1}, \quad y \ll 1. \tag{25}$$

We refer to the (uniquely determined) solution to (24) which additionally satisfies (25) as the *inner solution*; it depends on ε and H but is independent of the value of K.

The integral equation (20) determines an exact solution $p(x)$ up to a multiplicative constant. The inner solution approximates the exact solution which additionally satisfies

$$p(x) = P\left(K^{-\frac{H}{H+1}}x\right) \sim K^{-\frac{H(\varepsilon^{-1}-1)}{H+1}}x^{\frac{1}{\varepsilon}-1}, \quad x \ll K^{\frac{H}{H+1}}. \tag{26}$$

12

The (uniquely determined) solution $p(x)$ to (20) and (26) is not a probability density function, as it does not necessarily integrate to one. Prior to normalising the solution, we characterise its approximation on the outer scale and evaluate the asymptotic behaviour of its moments.

Outer solution. We substitute $x' = Kz$ into (20), while leaving $x = O(1)$, which gives

$$xp(x) = \frac{1}{\varepsilon} \int_0^{x/K} p(Kz) K e^{-\frac{(x-Kz)(1+z^H)}{\varepsilon}} \, \mathrm{d}z. \tag{27}$$

Recalling that x' has the interpretation of the protein concentration immediately before a burst and x represents (a lower bound for) the concentration after the burst, the current scaling relates to the possibility that the protein may deviate to very low $O(K)$ levels, at which the burst size ceases to be under regulation, hence the following burst is $O(1)$ large.

Since the $O(K)$ concentration scale is smaller than the $O(K^{H/(H+1)})$ inner concentration scale, we can approximate $p(Kz)$ on the right-hand side of (27) by the asymptotic expansion (26), i.e.

$$p(Kz) \sim K^{\frac{\varepsilon^{-1}-1}{H+1}} z^{\frac{1}{\varepsilon}-1}. \tag{28}$$

We insert (28) into (27), extend the upper limit of integration in (27) to infinity and neglect small $O(K)$ terms in the argument of the exponential in (27), obtaining

$$xp(x) \sim \varepsilon^{-1} K^{\frac{\varepsilon^{-1}+H}{H+1}} e^{-\frac{x}{\varepsilon}} \int_0^\infty z^{\frac{1}{\varepsilon}-1} e^{-\frac{xz^H}{\varepsilon}} \, \mathrm{d}z. \tag{29}$$

Substituting $u = xz^H/\varepsilon$ in the integral in (29) and dividing both sides of the relation by x, we obtain

$$p(x) \sim p_0(x) := \eta K^{\frac{\varepsilon^{-1}+H}{H+1}} e^{-\frac{x}{\varepsilon}} x^{-\frac{1}{\varepsilon H}-1}, \tag{30}$$

where

$$\eta = \eta(\varepsilon, H) = \frac{\varepsilon^{\frac{1}{\varepsilon H}-1}}{H} \Gamma\left(\frac{1}{\varepsilon H}\right) \tag{31}$$

is a constant which depends on ε and H only. We refer to the function $p_0(x)$ defined by the right-hand side of (30) as the *outer solution*.

Approximating the moments. The n-th moment of the solution $p(x)$ to (20) satisfying (26) is defined by

$$m_n = \int_0^\infty p(x) x^n \mathrm{d}x. \tag{32}$$

By (30), the outer approximation $p_0(x)x^n$ of the integrand in (32) is of the order of $x^{n-\frac{1}{\varepsilon H}-1}$ as x tends to zero; if $n - \frac{1}{\varepsilon H} > 0$, it is integrable: it increases too slowly as $x \to 0$ for the boundary layer to make a significant contribution to

13

the integral (Hinch, 1991). Thus, the n-th moment, for $n > \frac{1}{\varepsilon H}$, of the exact solution $p(x)$ can be approximated for small values of K by the n-th moment of its outer approximation (30), i.e.

$$m_n \sim \int_0^\infty p_0(x) x^n \mathrm{d}x = \mu_n K^{\frac{\varepsilon^{-1}+H}{H+1}}, \quad \text{for } n > \frac{1}{\varepsilon H}, \tag{33}$$

where

$$\mu_n = \eta \varepsilon^{n-\frac{1}{\varepsilon H}} \Gamma\left(n - \frac{1}{\varepsilon H}\right) = \frac{\varepsilon^{n-1}}{H} \Gamma\left(\frac{1}{\varepsilon H}\right) \Gamma\left(n - \frac{1}{\varepsilon H}\right) \tag{34}$$

is a constant which does not depend on K.

Substituting $x = K^{\frac{H}{H+1}} y$ in (32), we express m_n for $n < \frac{1}{\varepsilon H}$ in terms of the inner variable,

$$m_n = K^{\frac{(n+1)H}{H+1}} \int_0^\infty P(y) y^n \mathrm{d}y. \tag{35}$$

By asymptotic matching (Kevorkian and Cole, 1981), the asymptotic behaviour of the inner solution $P_0(y)$ as $y \to \infty$ coincides with that of the outer solution $p_0(x)$ as $x \to 0$. Consequently, the inner approximation $P_0(y) y^n$ of the integrand in (35) is of the order of $y^{n-\frac{1}{\varepsilon H}-1}$ as $y \to \infty$. If $n < \frac{1}{\varepsilon H}$, the inner solution is integrable, decreasing too rapidly as $y \to \infty$ for the outer region to contribute significantly to the integral (Hinch, 1991).

Thus, we can write

$$m_n \sim \nu_n K^{\frac{(n+1)H}{H+1}}, \quad \text{for } n < \frac{1}{\varepsilon H}, \tag{36}$$

where

$$\nu_n = \int_0^\infty P_0(y) y^n \mathrm{d}y \tag{37}$$

is the n-th moment of the inner solution, which does not depend on K and depends on ε and H only.

Normalisation. The normalised steady-state protein pdf is obtained by dividing the solution $p(x)$ to (20) satisfying (26) by its norm $m_0 = \int_0^\infty p(x) \mathrm{d}x$. Combining (26), (30) and (36) we arrive at asymptotic approximations

$$\frac{p(x)}{m_0} \sim \begin{cases} \nu_0^{-1} K^{-\frac{H}{\varepsilon(H+1)}} x^{\frac{1}{\varepsilon}-1} & \text{if } x \ll K^{\frac{H}{H+1}}, & \text{(38a)} \\ \eta \nu_0^{-1} K^{\frac{\varepsilon-1}{H+1}} \mathrm{e}^{-\frac{x}{\varepsilon}} x^{-\frac{1}{\varepsilon H}-1} & \text{if } x \gg K^{\frac{H}{H+1}}. & \text{(38b)} \end{cases}$$

Figure 6 shows an excellent agreement between the theoretical scalings (38a)–(38b) and probability distributions estimated by extensive kinetic Monte Carlo simulation with Algorithm 1.
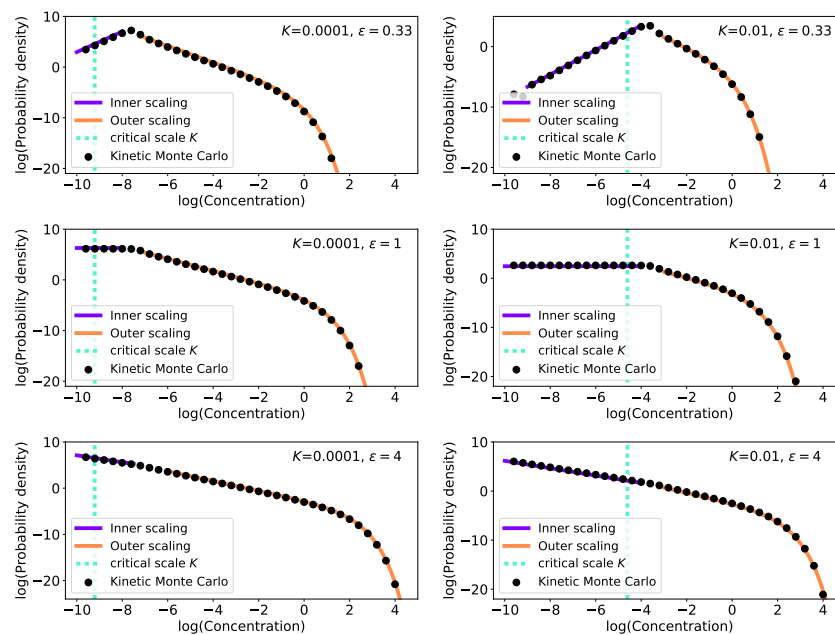
14

Figure 6: Protein distributions for selected parameter sets. Discrete markers: probability distributions estimated by kinetic Monte Carlo simulation with Algorithm 1; dashed line: the critical concentration K; solid lines: theoretical predictions for the inner (38a) (purple colour) and outer (38b) (orange colour) scalings of the probability distributions with manually tuned prefactors $\nu_0^{-1} = O(1)$. The simulation data were estimated from 10^6 sample paths after the stationarities of the process were reached.

15

Approximating the protein mean and CV^2. Using the asymptotic results (33) and (36) for the moments m_0, m_1, and m_2, we can also derive approximations for the mean and the mean square, which are given by the ratios

$$\langle x \rangle = \frac{m_1}{m_0}, \quad \langle x^2 \rangle = \frac{m_2}{m_0}. \tag{39}$$

The mean satisfies, as $K \to 0$,

$$\langle x \rangle \sim \begin{cases} \dfrac{\nu_1}{\nu_0} K^{\frac{H}{H+1}} & \text{if } \varepsilon < \frac{1}{H}, & \text{(40a)} \\[3mm] \dfrac{\mu_1}{\nu_0} K^{\frac{\varepsilon^{-1}}{H+1}} & \text{if } \varepsilon > \frac{1}{H}. & \text{(40b)} \end{cases}$$

By (40a), under low-noise conditions ($\varepsilon < \frac{1}{H}$), the mean exhibits the same power-law decrease as $K \to 0$ as its small-noise prediction x_0 (19), albeit with a prefactor which is different from one. As ε exceeds the threshold $\frac{1}{H}$, the exponent of the power law (40b) becomes smaller with increasing ε, implying a slower decrease of the mean as $K \to 0$ as that predicted by the small-noise approximation.

For the mean square we obtain in the small-K regime the asymptotics

$$\langle x^2 \rangle \sim \begin{cases} \dfrac{\nu_2}{\nu_0} K^{\frac{2H}{H+1}} & \text{if } \varepsilon < \frac{1}{2H}, & \text{(41a)} \\[3mm] \dfrac{\mu_2}{\nu_0} K^{\frac{\varepsilon^{-1}}{H+1}} & \text{if } \varepsilon > \frac{1}{2H}. & \text{(41b)} \end{cases}$$

Thus, for small values of ε, the mean square $\langle x^2 \rangle$ decreases as $K \to 0$ with the same exponent $2H/(H+1)$ as its small-noise prediction x_0^2 (18). The alternative, slower, exponent $\varepsilon^{-1}/(H+1)$ applies if ε exceeds the threshold $\frac{1}{2H}$, which is half the threshold in (41) for the mean and which we previously in Section 4 referred to as the critical noise load.

The coefficient of variation can be approximated for small values of K by

$$CV^2 = \frac{\langle x^2 \rangle}{\langle x \rangle^2} - 1 \sim \begin{cases} \dfrac{\nu_2 \nu_0}{\nu_1^2} - 1 & \text{if } \varepsilon < \frac{1}{2H}, & \text{(42a)} \\[3mm] \dfrac{\mu_2 \nu_0}{\nu_1^2} K^{-\frac{2H-\frac{1}{\varepsilon}}{H+1}} & \text{if } \frac{1}{2H} < \varepsilon < \frac{1}{H}, & \text{(42b)} \\[3mm] \dfrac{\mu_2 \nu_0}{\mu_1^2} K^{-\frac{\varepsilon^{-1}}{H+1}} & \text{if } \varepsilon > \frac{1}{H}. & \text{(42c)} \end{cases}$$

By (42a), CV^2 converges to a constant value as $K \to 0$ for subcritical noise loads ($\varepsilon < \frac{1}{2H}$). For supercritical noise loads ($\varepsilon > \frac{1}{2H}$), the CV^2 exhibits a power-law increase as K decreases to zero. As ε increases beyond the critical value $\frac{1}{2H}$, the absolute value of the exponent of the power law in (42b) first increases, reaching a maximum of $H/(H+1)$ for $\varepsilon = \frac{1}{H}$; after ε exceeds $\frac{1}{H}$, the exponent of the power law in (42c) begins to decrease in absolute value. Thus,
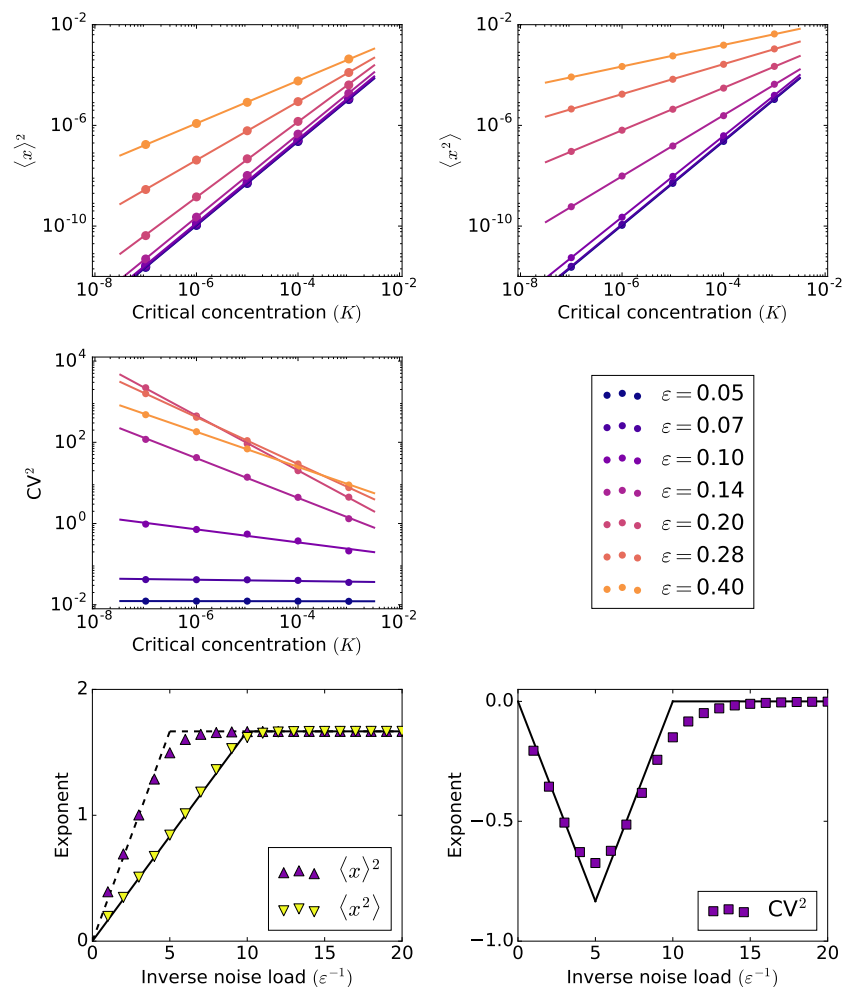
Figure 7: The squared mean (top left), the mean square (top right) and the CV^2 (centre left), obtained by stochastic simulation (Algorithm 1), as functions of critical concentrations $K \ll 1$ for $H = 5$ and a selection of values of ε (centre right). The discrete markers are the numerically measured results. The exponents of the power-law dependence on K of $\langle x \rangle^2$, $\langle x^2 \rangle$ (bottom left), and CV^2 (bottom right), obtained by linear regression (solid lines in top left, top right, and centre left) and asymptotic analysis (40)–(42), as functions of the inverse noise load ε^{-1}.

17

the fastest grow-up in the coefficient of variation as $K \to 0$ is achieved for $\varepsilon = \frac{1}{H}$ by a combination of a relatively fast decrease in the mean and a relatively slow decrease in the variance. Under excessive noise conditions ($\varepsilon \gg \frac{1}{H}$), feedback plays a limited role, which is evidenced by low grow-up rates of CV^2.

For numerical analysis, we fixed $H = 5$, used a selection of 20 values for ε ranging from 0.05 to 1, and a geometric sequence of 5 values for K ranging from 10^{-7} to 10^{-3}. For each of the 20×5 parameter combinations from the (H, ε, K) parameter space of the model, we estimated the steady-state mean concentration value $\langle x \rangle$ and the mean square value $\langle x^2 \rangle$ by generating 10^6 sample paths using Algorithm 1. The *ensemble average* of the first and the second moments were measured after stationarity of the process is reached, after which time the moments were measured every 0.2 units of time for another 5×10^3 units of time to perform an additional *temporal averaging*.

The simulation results are visualised in Fig. 7, where we show the squared mean $\langle x \rangle^2$ (Fig. 7, top left), the mean square $\langle x^2 \rangle$ (Fig. 7, top right), and the CV^2 (Fig. 7, centre left) as functions of the critical concentration K. We use (decimal) logarithmic scale for both axes in all panels; any power-law relationship then appears as a straight line with slope which is equal to the power-law's exponent. Using simple linear regression, we estimate the power-law exponents (Fig. 7, bottom, shown as triangular and square markers) and compare the estimates to the asymptotic predictions (40)–(42) (Fig. 7, bottom, black lines).

While the simulation-based exponent estimates in Fig. 7 are in a general agreement with the asymptotic results for $\langle x \rangle$ in (40) and $\langle x^2 \rangle$ in (41), some discrepancies are observed around the critical values of ε (which are $\varepsilon = 0.2$ for $\langle x \rangle^2$ and $\varepsilon = 0.1$ for $\langle x^2 \rangle$). We attribute these errors to logarithmic correction terms which appear in the leading-order asymptotic approximations for the moments at critical values (see Appendix B).

The exponent for the CV^2 is given by the difference of the exponents for $\langle x^2 \rangle$ and that for $\langle x \rangle^2$. The simulation-based estimate of the exponent for the CV^2 incorporates the errors incurred in estimating the exponents for both $\langle x^2 \rangle$ and that for $\langle x \rangle^2$: the resulting error can be quite large relative to the absolute value of the exponent itself. Additionally, the constant term -1, which we neglected as higher-order in the asymptotic approximations (42b) and (42c), can nevertheless be relatively important if the absolute value of the exponent is low; this adds yet another source of error in the estimate of the exponent. Despite these inaccuracies, the simulation-based results for the CV^2 are in qualitative agreement with the leading-order asymptotic predictions: the CV^2 remains bounded if $\varepsilon < 0.1$ and exhibits the fastest grow-up for $\varepsilon = 0.2$.

7 Comparison with previous results

In our present model for feedback in burst size, the concentration of protein immediately before a burst occurs determines the expected size of the burst. The distribution of burst sizes is exponential, whose defining property is its memorylessness: at any stage of the growth of a burst, the amount of protein

18

yet to be produced is independent of how much has already been produced. Such lack of self-control on a single-burst level implies that our model implicitly includes a delaying step, which newly produced molecules have to undergo before they can take part in the self-regulation. In order that our model be applicable, the delay must be neither too short nor too long: on one hand, it needs to last long enough to carry over the duration of a single burst; on the other hand, it must be short enough so that all protein produced in the current burst will have matured by the time of the next burst. Biologically, such delays can easily be accounted by the time it takes to complete the synthesis of a gene product. As a specific example, translational bursts in prokaryotes occur when a short-lived mRNA is transcribed from a gene and repeatedly translated before it is degraded (McAdams and Arkin, 1997). Each translation begins when the mRNA is bound by a ribosome and continues by the ribosome sliding down the mRNA molecule. As it moves along the mRNA, the ribosome forms an elongating chain of amino acids, which is to become a protein after the ribosome reaches the end of the mRNA code. Since the binding of new ribosome molecules occurs while the previous ones are still elongating, the proteins whose translation initiated earlier cannot inhibit those initiated later within a single burst.

In an alternative version of our model for feedback in burst size, which was analysed in a previous paper (Bokes and Singh, 2016), the growth of a burst depends on the current value of the protein concentration, including the molecules already synthesised within the same burst, rather than on its pre-burst level. Consequently, the distribution of burst sizes is not exponential and harder to draw random variates from, which complicates the implementation of a fast and exact simulation algorithm. On the other hand, the model in (Bokes and Singh, 2016) is considerably easier to treat analytically: the master equation of the undelayed model reads

$$\frac{\partial p}{\partial t} + \frac{\partial J}{\partial x} = 0, \quad J = \frac{1}{\varepsilon}\int_0^x p(x',t)\mathrm{e}^{-\frac{1}{\varepsilon}\int_{x'}^x \frac{\mathrm{d}\xi}{\theta(\xi)}}\,\mathrm{d}x' - xp, \tag{43}$$

where $\theta(x) = (1 + (x/K)^H)^{-1}$ is the Hill function. The integral term in the probability flux J can be recognised as the variation-of-constants solution $f(x)$ to the differential equation $\mathrm{d}f/\mathrm{d}x + f/\varepsilon\theta = p$ subject to $f(0) = 0$. Applying the differential operator $\mathrm{d}/\mathrm{d}x + 1/\varepsilon\theta$ on the steady state equation $J = 0$ transforms this integral equation into an ordinary differential equation (Lin and Doering, 2016), solving which yields an explicit steady-state probability density function

$$p(x) = Cx^{\frac{1}{\varepsilon}-1}\mathrm{e}^{-\frac{1}{\varepsilon}\left(\frac{x^{H+1}}{(H+1)K^H}+x\right)},$$

where C is a normalisation constant. Steady-state protein moments can be evaluated by numerical integration of the density or by asymptotic approximation of the integrals in the small-noise and strong feedback regimes (Bokes and Singh, 2016). In the small-noise limit, the results that follow are identical with (17) obtained for the current model (Fig. 8, left and central panels, dashed lines). The non-delayed model exhibits a monotonic decrease in the coefficient of variation for any combination of ε and H (Fig. 8, central panel, coloured lines).
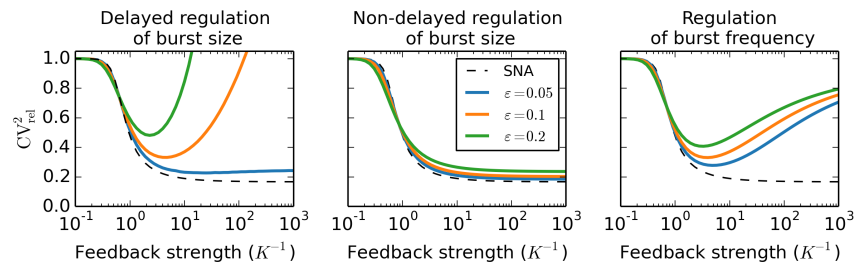
19

Figure 8: The relative CV^2 of a protein produced in bursts subject to different types of feedback in response to increasing feedback strength K^{-1}. We fix $H = 5$ and use a selection of values of ε which are given in the legend within the central panel. Exact numerical/simulational results are compared to the small-noise approximation (SNA).

The loss of control over protein noise, which occurs for supercritical noise loads in the strong-feedback regime of our current model (Fig. 8, left panel, coloured lines), can therefore be attributed to the delay which carries over the bursting timescale.

While so far we have focused exclusively on feedback in burst size, more attention has traditionally been paid in literature to feedback in burst frequency. If one adheres to the fundamentals of the model of Section 2, but uses the Hill function to reduce the burst frequency in the excess of protein, rather than reducing the expected burst size, one arrives at a model for negative feedback in burst frequency which was also analysed in (Bokes and Singh, 2016). The master equation of the model reads

$$\frac{\partial p}{\partial t} + \frac{\partial J}{\partial x} = 0, \quad J = \frac{1}{\varepsilon} \int_0^x p(x',t)\theta(x')e^{-\frac{x-x'}{\varepsilon}} \, dx' - xp, \tag{44}$$

which admits an explicit steady-state solution, see (Lin and Doering, 2016; Bokes and Singh, 2016; Friedman et al, 2006), given by

$$Ce^{-\frac{x}{\varepsilon}} x^{\frac{1}{\varepsilon}-1} (1 + (x/K)^H)^{-\frac{1}{\varepsilon H}},$$

where C is a normalisation constant. In the small-noise limit, the steady-state mean converges to the fixed point x_0 of the Hill function $\theta(x) = (1+(x/K)^H)^{-1}$, consult Bokes and Singh (2016) for details, which is the same value as was obtained in (17) for feedback in burst size. The squared coefficient of variation, on the other hand, was found in (Bokes and Singh, 2016) to be larger by a factor of $\frac{1}{x_0}$ than the value (17) obtained for feedback in burst size. We should note, however, that, even in the absence of any regulation, the squared coefficient of variation is inversely proportional to the burst frequency (cf. Section 2). The increase in the coefficient of variation in the small-noise limit can therefore be

attributed to the a decrease in the overall (time-averaged) burst frequency from the unregulated value of $\frac{1}{\varepsilon}$ to the regulated value of $\frac{\langle x \rangle}{\varepsilon}$ rather than to a loss of control over protein noise by the regulation of burst frequency.

Adjusting for the decrease in the overall burst frequency, Bokes and Singh (2016) defined the relative coefficient of variation as the ratio of the coefficient of variation of the self-regulating protein and that of a constitutively expressed protein with the same overall, time-averaged, frequency of bursts. While for the feedback in burst size this definition trivially reduces to $\mathrm{CV}^2_{\mathrm{rel}} = \mathrm{CV}^2/\varepsilon$, for feedback in burst frequency, the relative coefficient of variation is given by $\mathrm{CV}^2_{\mathrm{rel}} = \langle x \rangle \mathrm{CV}^2/\varepsilon$. In the small-noise limit, both feedback types yield the same approximation for the relative coefficient of variation, which decreases monotonically with increasing feedback strength (Fig. 8, dashed lines). For feedback in burst frequency, the relative coefficient of variation diverges for $K \ll 1$ from the small-noise prediction, but in a manner which is qualitatively different from the one which we identified in this paper to occur for feedback in burst size (Fig. 8, left and right panels, coloured lines).

Consistently with the small-noise prediction, the relative coefficient of variation initially decreases from the value of one as feedback in burst frequency strengthens, until the critical concentration K drops to $O(\varepsilon)$ levels, at which point the relative coefficient of variation begins to increase again, converging back to one as K tends to zero (Fig. 8, right panel; consult Bokes and Singh (2016) for asymptotics). Although at lower noise loads the turnaround occurs at higher feedback strengths, a critical noise load below which the small-noise approximation would be uniform does not exist, and an eventual loss of control over noise is inevitable in the small-K limit (Fig. 8, right panel). On the other hand, a protein which regulates its burst frequency is never noisier than a constitutively expressed with the same average burst frequency, whereas a protein with feedback in burst size will indeed be noisier if subjected to supercritical noise loads. Bokes and Singh (2016) previously compared the noise-reduction performance of feedback in burst frequency and (undelayed) feedback in burst size, reporting that the latter always performs better. In light of the present results, we conclude that depending on parametric conditions such as indicated above, either (delayed) feedback in burst size or in burst frequency can be optimal in terms of minimising the noise.

8   Discussion

The synthesis of protein molecules has been shown to occur in bursts of rapid production which alternate with periods of inactivity. In a minimalistic model for burst-like protein expression, bursts are represented by randomly occurring, randomly sized, discontinuous jumps in protein concentration; these are counterbalanced by deterministic decay of the concentration due dilution by cell growth and/or active degradation. We extended this minimalistic model by a specific kind of negative feedback, making the expected burst size decrease with increasing protein concentration. We investigated the ability of such kind of

negative feedback to control stochastic variability in protein levels.

Three positive dimensionless parameters — the noise load ε, the cooperativity coefficient H, and the critical protein concentration K — completely determine the behaviour of the model at steady state. The noise load ε is equal to the squared coefficient of variation the protein exhibits without any feedback. Its reciprocal ε^{-1} is equal to the average number of bursts per protein lifetime. The cooperativity coefficient H measures the steepness of decrease in the expected burst size in response to increasing protein concentration. Biologically, a positive integer value of H means that H protein molecules interact to form a complex which interferes with the transcription or translation machinery to reduce the burst size. The critical concentration K, measured in the units of the protein concentration mean in the absence of regulation, gives the threshold value which is required to reduce the expected burst size by half. We use the parameter K as an inverse measure of feedback strength: the lower the threshold for efficient self-repression, the stronger the feedback.

The central result of the paper lies in characterising the response of the steady-state protein statistics — the coefficient of variation in particular — to the strengthening of feedback, i.e. to decreasing K, while keeping the noise load ε and the cooperativity coefficient H constant. We found dramatically different responses depending on whether the noise load is less than or greater than a critical value $\varepsilon_{\mathrm{c}} = \frac{1}{2H}$. For subcritical noise loads, the coefficient of variation remains bounded with increasing feedback strength. Contrastingly, for supercritical noise loads, the coefficient of variation diverges to infinity as a power of K with a negative exponent.

The identification of the critical noise load is one important consequence of a bi-scale behaviour of the protein probability density function $p(x) = p(x; K)$ as $K \to 0$ characterised in Section 6. In addition to the outer $x = O(1)$ scale of large uncontrolled bursts, there is also an inner $x = O(K^{H/(1+H)})$ scale of small controlled bursts, on which the density tends to two distinguished limits, the outer and the inner solutions. The small-K behaviour of the moments $\langle x^n \rangle$ is decided in the overlap region $K^{H/(1+H)} \ll x \ll 1$, in which the density is proportional to the power $x^{-\frac{1}{\varepsilon H} - 1}$. The integrability of $x^{n - \frac{1}{\varepsilon H} - 1}$ as $x \to \infty$ (if $n < \frac{1}{\varepsilon H}$) or as $x \to 0$ (if $n > \frac{1}{\varepsilon H}$) decides whether the dominant contribution to the n-th moment comes from the inner or the outer region. If the inner region is dominant, the n-th moment tends to zero as $K^{nH/(1+H)}$; if the outer region dominates, the convergence of the n-th moment is still polynomial but slower. The critical exponent $\varepsilon_{\mathrm{c}} = \frac{1}{2H}$ thus gives the value of noise value at which the second moment ceases to be determined in the inner region of controlled bursts and begins to be driven by the outer region of uncontrolled bursts. The small-K behaviour of the second moment, and to a lesser extend that of the first moment (the mean), determine whether the coefficient of variation converges or diverges, as well as the rate of the divergence, as K tends to zero. We anticipate a potential criticism that our emphasis on the coefficient of variation is arbitrary and obscures the effect of the noise load and cooperativity coefficient on the polynomial decay of the density in the overlap region. Nevertheless, the coefficient of variation has widely been used both in experimental and theo-

retical analyses of stochastic gene expression. By focusing on the well-known measure of noise, we make our asymptotic analysis relatable to previous work and potentially interesting to a wider audience.

Our results have implications for the role of high cooperativity in negative feedback systems. Using a standard small-noise ($\varepsilon \ll 1$) approximation approach, we have shown that feedback in burst size can reduce the squared coefficient of variation by a factor of $H + 1$, thus confirming previous reports that cooperativity leads to improved attenuation of noise (Singh and Hespanha, 2009). This observation is not specific to feedback in burst size but also holds for feedback in burst frequency (Bokes and Singh, 2016) provided that one compensates for the drop in the overall burst frequency, which is larger for cooperative feedbacks. Contrastingly, using an alternative strong-feedback ($K \ll 1$) approximation approach, we have identified, specifically for feedback in burst size, an adverse effect of cooperativity on protein noise: high cooperativity can lead to a significant amplification of protein noise in response to increasing feedback strength even if the underlying noise load is relatively low ($\varepsilon > \varepsilon_{\mathrm{c}} = \frac{1}{2H}$). It is well known that high-cooperativity feedbacks are prone to instability if they operate with a sufficiently large delay (Murray, 2003). Interestingly, the loss of control over noise in our present model can also be attributed to a delay, albeit an infinitesimally small one: it has been introduced into the model by assuming that the mean burst size is determined by the protein concentration immediately before the burst starts. We expect that the interesting interplay between bursting noise, cooperativity, and delay, which we have illustrated here within a minimalistic modelling framework, will have implications in other, more complex, systems also.

Appendix A. Estimating protein moments from simulations

We can estimate the n-th steady-state moment by the time average $\frac{1}{T} \int_0^T x^n(t)\mathrm{d}t$, where $T \gg 1$, of a sample trajectory $x(t)$ generated by Algorithm 1 and raised to the power of n. Since $x(t)$ is piecewise exponential, we have

$$\frac{1}{T} \int_0^T x^n(t)\mathrm{d}t = \frac{1}{T} \sum_{i=0}^{N-1} \int_0^{\tau_i} x_i^n \mathrm{e}^{-nt}\mathrm{d}t = \frac{1}{nT} \sum_{i=0}^{N-1} x_i^n(1 - \mathrm{e}^{-\tau_i n}), \qquad \text{(A1)}$$

where x_i, $i \geq 1$, is the protein concentration immediately after the i-th burst, x_0 is the initial protein concentration, τ_i, $i \geq 1$, is the waiting time from the i-th burst until the $(i+1)$-th burst, τ_0 is the waiting time from the initial time until the first burst, and $T = \sum_{i=0}^{N-1} \tau_i$, where N is a large integer. By Algorithm 1, the values of x_i and τ_i are obtained by

$$x_{i+1} = x_i \mathrm{e}^{-\tau_i} - \frac{\varepsilon \ln \tilde{u}_i}{1 + (x_i \mathrm{e}^{-\tau_i}/K)^H}, \quad x_0 = 0, \quad \tau_i = -\varepsilon \ln u_i, \quad i = 0, 1, \dots, \tag{A2}$$

where u_i and $\tilde{u}_i$, $i = 0, 1, \ldots$, are random variates drawn independently of each other from the uniform distribution in the unit interval.

Inserting $T \approx \varepsilon N$, which holds by the law of large numbers, into (A1), and shifting the time frame to reduce the effect of the transient behaviour, we arrive at an estimate

$$\widehat{\langle x^n \rangle} = \frac{1}{\varepsilon n N} \sum_{i=M}^{N+M-1} x_i^n (1 - \mathrm{e}^{-\tau_i n}), \tag{A3}$$

where x_i and τ_i are given by (A2). We used (A3) with $N = 10^9$ and $M = 10^7$ to estimate the theoretical value of $\langle x^n \rangle$ in Figures 3–5 and Fig. 8. In the strong feedback limit (Figures 6 and 7), ensemble averaging across a large number of sample paths was needed for more precise measurements of the power-law exponents.

Appendix B. Strong-feedback asymptotics in critical cases

The dominant contribution to the n-th moment of a solution $p(x)$ to (20) comes exclusively from the inner $O(K^{H/(H+1)})$ concentration scale if $n < \frac{1}{\varepsilon H}$ or from the outer $O(1)$ concentration scale if $n > \frac{1}{\varepsilon H}$. In the borderline case of $n = \frac{1}{\varepsilon H}$, which we explore in this Appendix, the inner and outer regions both contribute to the leading-order behaviour of the n-th moment. The individual contributions can be identified by splitting the range of integration (Hinch, 1991),

$$m_n = \int_0^\infty p(x) x^n \mathrm{d}x = \int_0^\delta p(x) x^n \mathrm{d}x + \int_\delta^\infty p(x) x^n \mathrm{d}x, \tag{B1}$$

where δ is a value taken from an intermediate scale ($K^{H/(H+1)} \ll \delta \ll 1$).

In the first integral on the right-hand side of (B1), we substitute $x = K^{H/(H+1)} y$ and replace the integrand by the inner approximation,

$$\int_0^\delta p(x) x^n \mathrm{d}x \sim K^{\frac{(n+1)H}{H+1}} \int_0^{K^{-\frac{H}{H+1}} \delta} P_0(y) y^n \mathrm{d}y. \tag{B2}$$

Since $\delta \gg K^{H/(H+1)}$, the upper integration limit of the integral on the right-hand side of (B2) tends to infinity as K becomes increasingly small. The integral itself then diverges to infinity, since by (30)–(31) and van Dyke's matching rule an asymptotic expansion

$$P_0(y) y^n \sim \eta y^{-1}, \quad y \gg 1, \tag{B3}$$

holds for the integrand. Extricating the divergent part from the integral in (B2),

we obtain an asymptotic approximation

$$
\int_0^{K^{-\frac{H}{H+1}}\delta} P_0(y)y^n \mathrm{d}y
$$

$$
= \int_0^{K^{-\frac{H}{H+1}}\delta} \left(P_0(y)y^n - \frac{\eta}{y+1} \right) \mathrm{d}y + \eta \int_0^{K^{-\frac{H}{H+1}}\delta} \frac{\mathrm{d}y}{y+1}
$$

$$
\sim \int_0^\infty \left(P_0(y)y^n - \frac{\eta}{y+1} \right) \mathrm{d}y - \frac{\eta H \ln K}{H+1} + \eta \ln\delta. \tag{B4}
$$

In the second integral on the right-hand side of (B1), we replace the integrand by the outer approximation (30), obtaining

$$
\int_\delta^\infty p(x)x^n \mathrm{d}x \sim \int_\delta^\infty p_0(x)x^n \mathrm{d}x = \eta K^{\frac{(n+1)H}{H+1}} \int_\delta^\infty \frac{\mathrm{e}^{-\frac{x}{\varepsilon}}}{x} \mathrm{d}x
$$

$$
= \eta K^{\frac{(n+1)H}{H+1}} E_1\left(\frac{\delta}{\varepsilon} \right) \sim \eta K^{\frac{(n+1)H}{H+1}} (-\ln\delta + \ln\varepsilon - \gamma), \tag{B5}
$$

where $E_1(t)$ is the exponential integral and $\gamma = 0.577\ldots$ is the Euler–Mascheroni constant.

Collecting (B2), (B4), and (B5), we find that the n-the moment can be expanded into

$$
m_n \sim K^{\frac{(n+1)H}{H+1}} \left(\int_0^\infty \left(P_0(y)y^n - \frac{\eta}{y+1} \right) \mathrm{d}y - \frac{\eta H}{H+1}\ln K + \eta(\ln\varepsilon - \gamma) \right). \tag{B6}
$$

Although the logarithmic term in (B6) asymptotically dominates, as K tends to zero, the neighbouring constant terms, in practice the magnitudes of the logarithm and the constant terms are similar so the latter cannot be neglected.

As a particular application of the expansion (B6), we evaluate the small-K behaviour of the protein coefficient of variation subject to a critical noise load $\varepsilon = 1/2H$. The leading-order approximations to m_0 and m_1 involve contributions from the inner scale only and are given by (36)–(37). On the other hand, the leading-order approximation to m_2 combines contributions from either scale and is given by (B6). For the coefficient of variation we obtain

$$
\mathrm{CV}^2 = \frac{m_2 m_0}{m_1^2} - 1 \sim \frac{\nu_0}{\nu_1^2} \left(\int_0^\infty \left(P_0(y)y^2 - \frac{1}{2H^2(y+1)} \right) \mathrm{d}y \right.
$$

$$
\left. - \frac{\ln K}{2H(1+H)} - \frac{\ln 2 + \ln H + \gamma}{2H^2} \right) - 1.
$$

In the Main Text, we showed that the coefficient of variation remains bounded as K goes to zero if $\varepsilon < 1/2H$ and increases polynomially if $\varepsilon > 1/2H$. The above result implies that in the critical case $\varepsilon = 1/2H$ exhibits a slow logarithmic increase as K tends to zero.

References

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2002) Molecular biology of the cell. Garland Science New York

Barrio M, Burrage K, Leier A, Tian T (2006) Oscillatory regulation of hes1: discrete stochastic delay modelling and simulation. PLoS Comput Biol 2:e117

Becskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. Nature 405:590–593

Be'er S, Assaf M (2016) Rare events in stochastic populations under bursty reproduction. J Stat Mech Theory E 2016(11):113,501

Biancalani T, Assaf M (2015) Genetic toggle switch in the absence of cooperative binding: Exact results. Phys Rev Lett 115(20):208,101

Blake W, Kaern M, Cantor C, Collins J (2003) Noise in eukaryotic gene expression. Nature 422:633–637

Bokes P, Singh A (2015) Protein copy number distributions for a self-regulating gene in the presence of decoy binding sites. PloS one 10:e0120,555

Bokes P, Singh A (2016) Gene expression noise is affected differentially by feedback in burst frequency and burst size. J Math Biol DOI 10.1007/s00285-016-1059-4

Bokes P, King JR, Wood AT, Loose M (2012) Exact and approximate distributions of protein and mRNA levels in the low-copy regime of gene expression. J Math Biol 64:829–854

Bokes P, King J, Wood A, Loose M (2013) Transcriptional bursting diversifies the behaviour of a toggle switch: hybrid simulation of stochastic gene expression. B Math Biol 75:351–371

Bruna M, Chapman SJ, Smith MJ (2014) Model reduction for slow–fast stochastic systems with metastable behaviour. J Chem Phys 140:174,107

Cai L, Friedman N, Xie X (2006) Stochastic protein expression in individual cells at the single molecule level. Nature 440:358–362

Cardelli L, Kwiatkowska M, Laurenti L (2016) Stochastic analysis of chemical reaction networks using linear noise approximation. Biosystems 149:26–33

Dar RD, Razooky BS, Singh A, Trimeloni TV, McCollum JM, Cox CD, Simpson ML, Weinberger LS (2012) Transcriptional burst frequency and burst size are equally modulated across the human genome. P Natl Acad Sci USA 109:17,454–17,459

Dessalles R, Fromion V, Robert P (2017) A stochastic analysis of autoregulation of gene expression. J Math Biol DOI 10.1007/s00285-017-1116-7

Elf J, Ehrenberg M (2003) Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. Genome Res 13(11):2475–2484

Elowitz M, Levine A, Siggia E, Swain P (2002) Stochastic gene expression in a single cell. Science 297:1183–1186

Friedman N, Cai L, Xie X (2006) Linking stochastic dynamics to population distribution: an analytical framework of gene expression. Phys Rev Lett 97:168,302

Golding I, Paulsson J, Zawilski S, Cox E (2005) Real-time kinetics of gene activity in individual bacteria. Cell 123:1025–1036

Grima R, Schmidt DR, Newman TJ (2012) Steady-state fluctuations of a genetic feedback loop: An exact solution. J Chem Phys 137:035,104

Grönlund A, Lötstedt P, Elf J (2013) Transcription factor binding kinetics constrain noise suppression via negative feedback. Nature Comms 4:1864

Hinch EJ (1991) Perturbation methods. Cambridge university press

Jedrak J, Ochab-Marcinek A (2016a) Influence of gene copy number on self-regulated gene expression. J Theor Biol 408:222–236

Jedrak J, Ochab-Marcinek A (2016b) Time-dependent solutions for a stochastic model of gene expression with molecule production in the form of a compound poisson process. Phys Rev E 94:032,401

Jia T, Kulkarni RV (2010) Post-transcriptional regulation of noise in protein distributions during gene expression. Phys Rev Lett 105(1):018,101

Kevorkian J, Cole J (1981) Perturbation Methods in Applied Mathematics. Springer-Verlag New York

Komorowski M, Miekisz J, Stumpf MP (2013) Decomposing noise in biochemical signaling systems highlights the role of protein degradation. Biophys J 104(8):1783–1793

Kumar N, Platini T, Kulkarni RV (2014) Exact distributions for stochastic gene expression models with bursting and feedback. Phys Rev Lett 113:268,105

Lafuerza L, Toral R (2011) Role of delay in the stochastic creation process. Phys Rev E 84:021,128

Leier A, Barrio M, Marquez-Lago TT (2014) Exact model reduction with delays: closed-form distributions and extensions to fully bi-directional monomolecular reactions. J Roy Soc Interface 11:20140,108

Lin YT, Doering CR (2016) Gene expression dynamics with stochastic bursts: Construction and exact results for a coarse-grained model. Phys Rev E 93:022,409

Lin YT, Galla T (2016) Bursting noise in gene expression dynamics: linking microscopic and mesoscopic models. J Roy Soc Interface 13:20150,772

Mackey MC, Tyran-Kaminska M (2015) The limiting dynamics of a bistable molecular switch with and without noise. J Math Biol 73:367

Mackey MC, Glass L, et al (1977) Oscillation and chaos in physiological control systems. Science 197:287–289

Mackey MC, Tyran-Kaminska M, Yvinec R (2013) Dynamic behavior of stochastic gene expression models in the presence of bursting. SIAM J Appl Math 73:1830–1852

McAdams H, Arkin A (1997) Stochastic mechanisms in gene expression. P Natl Acad Sci USA 94:814–819

Monk N (2003) Oscillatory expression of hes1, p53, and nf-κb driven by transcriptional time delays. Curr Biol 13:1409–1413

Murray J (2003) Mathematical Biology. Springer

Newby J (2015) Bistable switching asymptotics for the self regulating gene. J Phys A-math Gen 48:185,001

Ozbudak EM, Thattai M, Kurtser I, Grossman AD, Van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. Nature genetics 31(1):69–73

Pájaro M, Alonso AA, Otero-Muras I, Vázquez C (2017) Stochastic modeling and numerical simulation of gene regulatory networks with protein bursting. J Theor Biol 421:51–70

Paulsson J (2004) Summing up the noise in gene networks. Nature 427:415–418

Peccoud J, Ycart B (1995) Markovian modeling of gene-product synthesis. Theor Popul Biol 48:222–234

Platini T, Jia T, Kulkarni RV (2011) Regulation by small rnas via coupled degradation: mean-field and variational approaches. Phys Rev E 84(2):021,928

Popovic N, Marr C, Swain PS (2016) A geometric analysis of fast-slow models for stochastic gene expression. J Math Biol 72:87–122

Ramos AF, Innocentini GCP, Hornos JEM (2011) Exact time-dependent solutions for a self-regulating gene. Phys Review E 83:062,902

Rosenfeld N, Elowitz MB, Alon U (2002) Negative autoregulation speeds the response times of transcription networks. J Mol Biol 323:785–793

Schikora-Tamarit MA, Toscano-Ochoa C, Espinos JD, Espinar L, Carey LB (2016) A synthetic gene circuit for measuring autoregulatory feedback control. Integr Biol 8(4):546–555

Schuss Z (2009) Theory and applications of stochastic processes: an analytical approach. Springer Science & Business Media

Scott M, Hwa T, Ingalls B (2007) Deterministic characterization of stochastic genetic circuits. P Natl Acad Sci USA 104(18):7402–7407

Shahrezaei V, Swain P (2008) The stochastic nature of biochemical networks. Curr Opin Biotechnol 19:369–374

Singh A (2011) Negative feedback through mRNA provides the best control of gene-expression noise. IEEE T NanoBiosci 10:194–200

Singh A, Hespanha JP (2009) Optimal feedback strength for noise suppression in autoregulatory gene networks. Biophys J 96:4013–4023

Smith S, Shahrezaei V (2015) General transient solution of the one-step master equation in one dimension. Phys Rev E 91(6):062,119

Soltani M, Bokes P, Fox Z, Singh A (2015) Nonspecific transcription factor binding can reduce noise in the expression of downstream proteins. Phys Biol 12:055,002

Stekel DJ, Jenkins DJ (2008) Strong negative self regulation of prokaryotic transcription factors increases the intrinsic noise of protein expression. Bmc Syst Biol 2:6

Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F (2011) Mammalian genes are transcribed with widely different bursting kinetics. Science 332:472–474

Swain PS (2004) Efficient attenuation of stochasticity in gene expression through post-transcriptional control. J Mol Biol 344:965–976

Taniguchi Y, Choi P, Li G, Chen H, Babu M, Hearn J, Emili A, Xie X (2010) Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. Science 329:533–538

Thattai M, van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. P Natl Acad Sci USA 98:151588,598

Thomas P, Grima R (2015) Approximate probability distributions of the master equation. Phys Rev E 92:012,120

Wang J, Lefranc M, Thommen Q (2014) Stochastic oscillations induced by intrinsic fluctuations in a self-repressing gene. Biophys J 107(10):2403–2416

Yates JL, Nomura M (1981) Feedback regulation of ribosomal protein synthesis in e. coli: localization of the mRNA target sites for repressor action of ribosomal protein l1. Cell 24:243–249

Yvinec R, Zhuge C, Lei J, Mackey MC (2014) Adiabatic reduction of a model of stochastic gene expression with jump markov process. J Math Biol 68:1051–1070