

# Gene expression noise is affected differentially by feedback in burst frequency and burst size

Pavol Bokes<sup>1</sup> and Abhyudai Singh<sup>2</sup>

<sup>1</sup>Department of Applied Mathematics and Statistics, Comenius University, Bratislava, Slovakia

<sup>2</sup>Department of Electrical and Computer Engineering, University of Delaware, Newark, Delaware, USA

## Abstract

Inside individual cells, expression of genes is stochastic across organisms ranging from bacterial to human cells. A ubiquitous feature of stochastic expression is burst-like synthesis of gene products, which drives considerable intercellular variability in protein levels across an isogenic cell population. One common mechanism by which cells control such stochasticity is negative feedback regulation, where a protein inhibits its own synthesis. For a single gene that is expressed in bursts, negative feedback can affect the burst frequency or the burst size. In order to compare these feedback types, we study a piecewise deterministic model for gene expression of a self-regulating gene. Mathematically tractable steady-state protein distributions are derived and used to compare the noise suppression abilities of the two feedbacks. Results show that in the low noise regime, both feedbacks are similar in term of their noise buffering abilities. Intriguingly, feedback in burst size outperforms the feedback in burst frequency in the high noise regime. Finally, we discuss various regulatory strategies by which cells implement feedback to control burst sizes of expressed proteins at the level of single cells.

## 1 Introduction

Stochastic expression of genes drives significant random fluctuations (noise) in protein copy numbers over time in single cells [1–8]. These fluctuations manifest as cell-to-cell variability in level of a protein, even in genetically-identical populations under the same external conditions. Stochastic gene expression poses a challenge for the precise control of cellular function, placing cells under evolutionary pressure to minimize the noise in vital proteins [9–11]. Not surprisingly, cell use diverse regulatory mechanisms to buffer noise in gene expression [12–16]. Negative feedback, by which the synthesis of gene products is switched off in their excess, and switched on in their absence, is a commonly used mechanism for noise control [17–23].

A major contributor to the overall noise in gene expression is the synthesis of proteins in random bursts, and these bursts can occur both at the transcriptional and translation levels. At the transcriptional level, a promoter can slowly

become active, producing a burst of mRNAs before becoming inactive [24–29]. At the translational level, a short-lived unstable mRNA degrades after synthesizing a burst of protein molecules [30–32]. In the context of such burst-like gene expression, negative feedback can act either by reducing the frequency with which bursts occur, or by reducing their size.

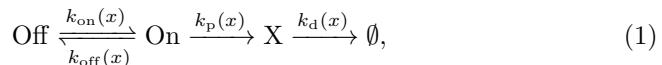
Transcriptional control can reduce the frequency or the size of transcriptional bursts, the former by hindering promoter activation and the latter by enhancing promoter inactivation. By controlling transcription, the frequency of translational bursts can also be regulated; however, their size needs be regulated post-transcriptionally. For example, many RNA binding proteins reduce the size of translational bursts by shortening the half-life of their own mRNA [33–38]. As a specific example, splicing factors typically bind to their own pre-mRNA to create an alternatively spliced transcript that is degraded via non-sense mediated degradation [36; 38].

In this paper, we present a theoretical comparison of the feedback in burst frequency and burst size with regards to their performance in protein noise reduction. We use a piecewise deterministic mathematical framework according to which any decrease (due to decay) in protein concentration is deterministic and continuous, and any increase (due to synthesis) occurs in randomly timed discontinuous jumps of random size [39–48]. This framework yields explicit formulae for protein probability density functions. We utilize these formulae by (i) calculating key noise characteristics by numerical integration and (ii) perform qualitative analysis of noise reduction performance by asymptotic evaluation of integrals.

The outline of the paper is as follows. First, we introduce the chosen modelling framework in Section 2. This is used to study feedback in burst frequency in Section 3 and burst size in Section 4. Then follows a more technical Section 5, in which strong-feedback asymptotics of protein mean and noise are developed. The results of Sections 2–5, and their implications, are summarised in a non-technical language in Section 6. Finally, Section 7 is devoted to discussing our results, especially in the context of previous theoretical comparisons between different types of negative feedback [18; 23; 49–52].

## 2 Modelling framework

We study a random telegraph model for stochastic gene expression with feedback in general form,



according to which the gene transitions between an inactive Off state and an On state, from which the protein  $X$  is synthesised, and eventually degraded.

The reaction rates  $k_{\text{on}}(x)$  of activation,  $k_{\text{off}}(x)$  of inactivation,  $k_p(x)$  of protein production, and  $k_d(x)$  of degradation depend on the current amount  $x$  of protein  $X$  in the system. We shall treat  $x$  as a continuous quantity, i.e. a concentration, which evolves according to the ODE  $dx/dt = k_p(x)$  if the gene is On and according to  $dx/dt = -k_d(x)$  if the gene is Off.

We shall assume that the inactivation rate  $k_{\text{off}}(x)$  and the protein synthesis rate  $k_p(x)$  are much faster than the activation rate  $k_{\text{on}}(x)$  and decay rate  $k_d(x)$ .

In that case, the gene is mostly Off, while the protein level slowly decays, switching momentarily into the On state, upon which a short spell of rapid production of protein, i.e. a burst, ensues, during which the effect of degradation is negligible. Bursts can be either transcriptional, in which case On and Off represent the active and inactive promoter states [cf. 24], or translational, in which case On and Off are meant to indicate the presence or absence of an unstable mRNA transcript [cf. 45].

In order to characterise the dynamics of a single burst, we denote by  $y$  the protein concentration on entering the On state, and let  $G(x, y)$ , where  $x > y$ , be the probability that the protein concentration exceeds  $x$  before the burst is terminated.

For any concentration level  $z$  such that  $x > z > y$ , the ratio  $dz/k_p(z)$  gives the time of gene activity required to produce  $dz$  of protein, while  $k_{\text{off}}(z)$  gives the hazard rate for aborting the burst. The probability that it is not aborted before  $x$  is reached is then determined by exponentiating the cumulative hazard rate [cf. 53; 54],

$$G(x, y) = e^{-\int_y^x \frac{k_{\text{off}}(z)}{k_p(z)} dz}. \quad (2)$$

If  $k_p(x)$  and  $k_{\text{off}}(x)$  are constants, then (2) implies exponential distribution of burst sizes [cf. 39]. We assume that  $\int_y^\infty \frac{k_{\text{off}}(z)}{k_p(z)} dz = \infty$  so that bursts are finite with probability one.

The probability density  $p(x, t)$  of having  $x$  protein at a time  $t$  satisfies a continuity equation

$$\frac{\partial p}{\partial t} + \frac{\partial J}{\partial x} = 0, \quad (3)$$

where

$$J = -k_d(x)p(x, t) + \int_0^x G(x, y)k_{\text{on}}(y)p(y, t)dy. \quad (4)$$

The term  $J$  is the probability flux, which specifies how much probability mass passes through a given point  $x$  (in the positive direction) per unit time. Equation (3) mathematically expresses the fact that all changes in probability density function are due to this flux, i.e. that the total mass remains conserved. By (4), the flux consists of a local term  $-k_d(x)p(x, t)$ , which gives the transfer of probability mass due to protein decay; since decay leads to movement of probability mass in the negative direction, this term takes a negative sign. The other term in (4) is nonlocal, and gives the transfer of probability mass due to bursts that start at a protein concentration  $y$ , which is lower than  $x$ , and end at a concentration which exceeds  $x$ . The probability of a burst being of a sufficient size is equal to  $G(x, y)$ , which needs to be multiplied by the probability  $k_{\text{on}}(y)p(y, t)$  of the burst actually having been initiated, and then integrated over all possible starting concentrations  $y$ ; this indeed yields the second term in (4). Equations (3)–(4) are equivalent to the Chapman–Kolmogorov differential equation [55] for a drift-jump (i.e. diffusion-less) Markov process, where the drift is due to protein degradation and jumps are due to bursts.

The steady-state probability distribution  $p(x)$  is found by setting  $J = 0$  and

$dJ/dx = 0$ , i.e.

$$k_d(x)p(x) = \int_0^x e^{-\int_y^x \frac{k_{\text{off}}(z)}{k_p(z)} dz} k_{\text{on}}(y)p(y)dy, \quad (5)$$

$$\frac{d}{dx}(k_d(x)p(x)) = k_{\text{on}}(x)p(x) - \frac{k_{\text{off}}(x)}{k_p(x)} \int_0^x e^{-\int_y^x \frac{k_{\text{off}}(z)}{k_p(z)} dz} k_{\text{on}}(y)p(y)dy. \quad (6)$$

Eliminating the integral term from (5)–(6), one obtains a linear first-order ordinary differential equation

$$\frac{d}{dx}(k_d(x)p(x)) = \left( \frac{k_{\text{on}}(x)}{k_d(x)} - \frac{k_{\text{off}}(x)}{k_p(x)} \right) k_d(x)p(x),$$

which implies

$$p(x) = \frac{C}{k_d(x)} \exp \left( \int \left( \frac{k_{\text{on}}(x)}{k_d(x)} - \frac{k_{\text{off}}(x)}{k_p(x)} \right) dx \right), \quad (7)$$

where  $C$  is the normalisation constant. In order to guarantee that solutions to the master equation (3)–(4) converge, as time increases to infinity, to (7), one needs to impose, in addition to the integrability condition for (7), a number of additional constraints on the reaction rates to exclude certain degenerate types of behaviour such as extinction due to sublinear decay or infinite waiting for the next burst [see 41]. An alternative derivation of (7), which extends to non-bursting regimes also, can be found in [56]. Additional methodology, such as finding mean first passage times, for problems of this kind can be found in [45].

### 3 Feedback in burst frequency

In this section we assume that the burst frequency  $k_{\text{on}}(x)$  decreases with increasing protein concentration  $x$ , the decay rate  $k_d(x)$  is proportional to the concentration of protein, and the mean burst size is a constant; specifically, we set

$$k_{\text{on}}(x) = \frac{\varepsilon^{-1}}{1 + (x/K)^H}, \quad k_d(x) = x, \quad \frac{k_p(x)}{k_{\text{off}}(x)} = \varepsilon, \quad (8)$$

where the dissociation constant  $K$  and cooperativity coefficient  $H$  parametrise the decreasing Hill-type dependence of the burst frequency on the protein level. The Hill function (8) can be viewed as a quasi-steady-state approximation of a finer-grained regulation mechanism, in which individual protein molecules cooperatively bind to multiple binding sites at the promoter of an inactive gene, whereby they prevent its transition to the active state. Stochastic models for negative autoregulation in the presence of bursting, which explicitly include binding of protein to promoter, have been studied in Grönlund et al. [57] and Kumar et al. [58].

Both time and concentration scales are already nondimensionalised in (8). Time is measured in the units of mean protein lifetime: the decay rate is equal to the concentration of the protein. Concentration is measured in the units of its mean in the absence of self-repression ( $K \rightarrow \infty$ ): the unrepresed (maximal) burst frequency  $\varepsilon^{-1}$  is the reciprocal of the mean burst size  $\varepsilon$ . The parameter  $\varepsilon$  characterises the noisiness in the autoregulatory system. Small  $\varepsilon$  implies frequent and small bursts, large  $\varepsilon$  implies infrequent and large bursts.

Inserting (8) into (7), we find that the steady-state distribution assumes a Wentzel–Kramers–Brillouin (WKB) form [59]

$$p(x) = \frac{C e^{-\frac{\Phi(x)}{\varepsilon}}}{x}, \quad (9)$$

where

$$\Phi(x) = - \int \frac{dx}{x(1 + (x/K)^H)} + x = \frac{\ln(1 + (x/K)^H)}{H} - \ln x + x. \quad (10)$$

The integration constant  $C$ , mean concentration  $\langle x \rangle$  and the variance  $\sigma^2$  can be computed by numerical integration of

$$C = \left( \int_0^\infty \frac{e^{-\frac{\Phi(x)}{\varepsilon}}}{x} dx \right)^{-1}, \quad \langle x \rangle = \int_0^\infty x p(x) dx, \quad \sigma^2 = \int_0^\infty (x - \langle x \rangle)^2 p(x) dx. \quad (11)$$

Some care has to be taken when evaluating in the  $\varepsilon \ll 1$  regime the first integral of (11), which, due to the exponentially small term, can easily become smaller than any absolute error tolerance. Such problems can be circumvented e.g. by multiplying the integrand by a sufficiently large constant, such as  $e^{\Phi(x_s)/\varepsilon}$ , where  $x_s$  is defined as detailed below.

A scale-free characteristic of protein noise is the coefficient of variation defined by

$$CV^2 = \frac{\sigma^2}{\langle x \rangle^2}. \quad (12)$$

We shall compare the coefficient of variation of the regulated protein to that of a constitutively expressed protein with the same mean and burst size; this requires the burst frequency set to  $\langle x \rangle / \varepsilon$ . In the constitutive case, the protein concentration has a gamma distribution with the shape parameter being equal to the burst frequency [39]. The squared coefficient of variation of the gamma distribution is the reciprocal of its shape parameter and hence of the burst frequency. Thus, we define

$$CV_{\text{rel}}^2 = \frac{\langle x \rangle}{\varepsilon} CV^2 \quad (13)$$

as the relative coefficient of variation.

In the small-noise regime ( $\varepsilon \ll 1$ ), explicit asymptotic expression for the noise characteristics can be derived using the linear noise approximation (LNA). Here we can obtain the LNA results easily by expanding the integrals in (11) using Laplace's method [60].

The function  $\Phi(x)$  in (9) is a Lyapunov function [61] corresponding to the deterministic model

$$\frac{dx}{dt} = \frac{1}{1 + (x/K)^H} - x, \quad (14)$$

i.e.  $\Phi(x)$  is minimal when  $x = x_s$ , where  $x_s$  is the single stable steady state of (14), decreasing for  $x < x_s$  and increasing for  $x > x_s$  (cf. Fig 1).

For  $\varepsilon \ll 1$ , the dominant contribution of probability mass in the probability density function (9) comes from the neighbourhood of  $x_s$ , around which we have

$$\Phi(x) = \Phi(x_s) + \frac{1}{2} \Phi''(x_s)(x - x_s)^2 + \dots \quad (15)$$

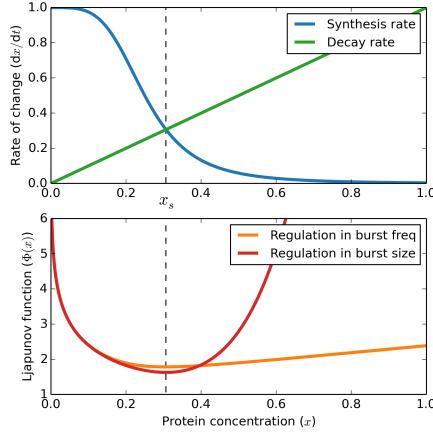


Figure 1: Deterministic model and its two Lyapunov functions. The top panel shows the synthesis rate  $(1 + (x/K)^H)^{-1}$  and the decay rate  $x$  as functions of  $x$ . The point  $x_s$  at which they are equal is the steady state of the deterministic model (horizontal dashed lines). At the same point the Lyapunov functions are minimal (bottom panel). Note the flatness and asymmetry of the Lyapunov function used in the model for regulation via burst frequency in contrast with that used for regulation via burst size. The parameters are  $H = 4$ ,  $K = 1/4$ .

Substituting the parabolic approximation (15) into (9) and neglecting higher order terms in the usual manner [60], we find that the protein concentration is approximately normally distributed with the moments given by

$$\langle x \rangle \sim x_s, \quad \sigma^2 \sim \frac{\varepsilon}{\Phi''(x_s)} \quad \text{if } \varepsilon \ll 1. \quad (16)$$

Notably, the steady-state distribution reduces for  $\varepsilon \rightarrow 0$  to a point mass situated at the steady state  $x_s$  of the ODE model; in Appendix A we provide a more general argument that the (time-dependent) master equation itself reduces, as  $\varepsilon$  tends to zero, to the ODE model (14).

In order to express the variance (16) in terms of the model parameters, we differentiate the Lyapunov function (10) twice, finding

$$\Phi''(x_s) = \frac{H(1 - x_s) + 1}{x_s}. \quad (17)$$

Using (16) and (17) in (12) and (13) we find that asymptotic approximations

$$\text{CV}^2 \sim \frac{\varepsilon}{x_s(H(1 - x_s) + 1)}, \quad \text{CV}_{\text{rel}}^2 \sim \frac{1}{H(1 - x_s) + 1} \quad (18)$$

hold for the coefficients of variation in the small-noise regime ( $\varepsilon \ll 1$ ).

## 4 Feedback in burst size

In this section we focus on the case of

$$k_{\text{on}}(x) = \varepsilon^{-1}, \quad k_{\text{d}}(x) = x, \quad \frac{k_{\text{p}}(x)}{k_{\text{off}}(x)} = \frac{\varepsilon}{1 + (x/K)^H}. \quad (19)$$

In contrast with (8), the burst frequency in (19) is constant, but the burst size is regulated: the burst growth rate  $k_{\text{p}}/k_{\text{off}}$  decreases with increasing protein concentration. The functional form of the decrease is again that of a Hill function parametrised by  $H$  and  $K$ . Small  $\varepsilon$  corresponds to small-noise regime.

If we interpret the Off/On states as indicators of translational, rather than promoter, activity, meaning that Off refers to the absence of transcripts and On indicates the presence of a short-lived mRNA copy [cf. 45], then  $k_{\text{off}}(x)$  acquires the meaning of the mRNA degradation rate. The Hill-type dependency in (19) can be achieved by an RNA-binding protein, which cooperatively catalyses the removal of its mRNA [33–35].

Inserting (19) into (7), we find that the WKB form (9) is still valid for the steady-state distribution, but with a different Lyapunov function

$$\Phi(x) = \frac{x^{H+1}}{(H+1)K^H} - \ln x + x. \quad (20)$$

The difference in the two Lyapunov functions reflects the difference in the two stochastic models. However, both Lyapunov functions correspond to the same ordinary differential equation (14), being minimal at the ODE's steady state  $x_s$  (cf. Fig 1). Thus, regardless of whether the feedback acts on burst size or burst frequency, the steady-state protein concentration is narrowly distributed around the deterministic steady state  $x_s$  in the small-noise regime. In Appendix A, we provide a more general result which holds also outside of the steady-state regime: we show that the master equation (3)–(4) reduces for  $\varepsilon \rightarrow 0$  to the ODE model (14), regardless of whether the feedback acts via burst frequency (8) or burst size (19).

Formulae (9), (11) and (12) can be reused with the new definition (20) of  $\Phi$  to compute numerically the mean, variance, and the squared coefficient of variation of the protein distribution. However, a modification is due in the definition of the relative coefficient of variation. Since the burst frequency is constant but the burst size is regulated, we compare the  $\text{CV}^2$  of a regulated protein to that of a constitutively expressed protein with the same mean  $\langle x \rangle$  and burst frequency  $\varepsilon^{-1}$ , adjusting the burst size to  $\varepsilon/\langle x \rangle$  as required.

The reciprocal  $\varepsilon$  of the burst frequency gives the squared coefficient of variation for the referential constitutively expressed protein. Thus, we define

$$\text{CV}_{\text{rel}}^2 = \frac{\text{CV}^2}{\varepsilon} \quad (21)$$

as the relative coefficient of variation. This differs from (13), in which the  $\text{CV}^2$  of a protein with a regulated burst frequency was compared to the  $\text{CV}^2$  of a constitutively expressed protein with the same mean and burst size, adjusting the burst frequency as required.

In the small-noise regime ( $\varepsilon \ll 1$ ), the mean and variance satisfy (16), in which the second derivative of the Lyapunov function is given not by (17) but

$$\bar{\Phi}''(x_s) = \frac{H(1-x_s) + 1}{x_s^2}, \quad (22)$$

as is easily checked by differentiating (20) twice. Using (16) and (22) in the definitions of the  $\text{CV}^2$  (12) and the relative  $\text{CV}^2$  (21), we find that

$$\text{CV}^2 \sim \frac{\varepsilon}{H(1-x_s) + 1}, \quad \text{CV}_{\text{rel}}^2 \sim \frac{1}{H(1-x_s) + 1} \quad (23)$$

hold in the small-noise regime in the case of regulation via burst size.

## 5 Strong feedback asymptotics

Here we present an additional asymptotic analysis that yields explicit predictions for mean and  $\text{CV}^2$  that hold even in the large-noise regime ( $\varepsilon = O(1)$ ), provided that the feedback is very strong ( $K \ll \varepsilon$ ). We focus solely on the case of feedback in burst frequency, for which the strong-feedback asymptotics are more interesting than for feedback in burst size. The latter is nevertheless treated in Appendix B.

By (9)–(10), we have

$$p(x) = C e^{-x/\varepsilon} x^{\frac{1}{\varepsilon}-1} (1 + (x/K)^H)^{-\frac{1}{\varepsilon H}} \quad (24)$$

for the protein pdf.

The protein moments are given by

$$\langle x^n \rangle = \frac{B_n}{B_0}, \quad (25)$$

where

$$B_n = \int_0^\infty e^{-x/\varepsilon} x^{\frac{1}{\varepsilon}-1+n} (1 + (x/K)^H)^{-\frac{1}{\varepsilon H}} dx. \quad (26)$$

Note that  $B_0^{-1} = C$  is the normalisation constant. Substituting  $x = Ky$  in (26) yields

$$B_n = K^{\frac{1}{\varepsilon}+n} A_n, \quad (27)$$

where

$$A_n = \int_0^\infty e^{-Ky/\varepsilon} y^{\frac{1}{\varepsilon}-1+n} (1 + y^H)^{-\frac{1}{\varepsilon H}} dy. \quad (28)$$

The protein mean and the squared coefficient of variation can be expressed in terms of  $A_n$ ,  $n = 0, 1, 2$ , as

$$\langle x \rangle = \frac{KA_1}{A_0}, \quad \text{CV}^2 = \frac{\langle x^2 \rangle}{\langle x \rangle^2} - 1 = \frac{B_0 B_2}{B_1^2} - 1 = \frac{A_0 A_2}{A_1^2} - 1. \quad (29)$$

Thus, we need to establish the limiting behaviour of Laplace transforms

$$A_n = \int_0^\infty e^{-\lambda y} f_n(y) dy, \quad \text{where} \quad f_n(y) = y^{\frac{1}{\varepsilon}-1+n} (1 + y^H)^{-\frac{1}{\varepsilon H}} \quad (30)$$

for small values of the Laplace variable  $\lambda = K/\varepsilon$ .

If  $n \geq 1$ , then  $\lambda \ll 1$  implies  $y \gg 1$ , so that  $f_n(y) \sim y^{-1+n}$ , and

$$A_n \sim \int_0^\infty e^{-\lambda y} y^{-1+n} dy = (n-1)! \lambda^{-n}. \quad (31)$$

The case of  $n = 0$  is an exception because of the divergence of the exponential integral.

For  $n = 0$  we split the integration range [see 62]

$$A_0 = \int_0^\infty e^{-\lambda y} f_0(y) dy = \int_0^\delta e^{-\lambda y} f_0(y) dy + \int_\delta^\infty e^{-\lambda y} f_0(y) dy, \quad (32)$$

where  $\delta$  is chosen so that

$$1 \ll \delta \ll \frac{1}{\lambda} \quad (33)$$

is asymptotically satisfied.

In the second integral of (32),  $y > \delta \gg 1$  implies  $f_0(y) = y^{\varepsilon^{-1}-1} (1 + y^H)^{-1/\varepsilon H} \sim y^{-1}$ . i.e.

$$\int_\delta^\infty e^{-\lambda y} f_0(y) dy \sim \int_\delta^\infty \frac{e^{-\lambda y}}{y} dy = E_1(\lambda \delta) \sim -\ln \delta - \ln \lambda - \gamma, \quad (34)$$

where  $E_1(z)$  is the exponential integral and the right-hand side of (34) is made of the first two terms of its asymptotic expansion:  $\gamma = 0.577\dots$  is the Euler–Mascheroni constant [63].

In the first integral on the right-hand side of (32), we have  $\lambda y < \lambda \delta \ll 1$ , so that

$$\begin{aligned} \int_0^\delta e^{-\lambda y} f_0(y) dy &\sim \int_0^\delta f_0(y) dy \\ &= \int_0^\delta y^{\frac{1}{\varepsilon}-1} (1 + y^H)^{-\frac{1}{\varepsilon H}} dy. \end{aligned} \quad (35)$$

Substitution  $v = y^H/(1 + y^H)$  in (35) yields

$$\int_0^\delta y^{\frac{1}{\varepsilon}-1} (1 + y^H)^{-\frac{1}{\varepsilon H}} dy = \frac{1}{H} \int_0^{\delta^H/(1+\delta^H)} v^{\frac{1}{\varepsilon H}-1} (1 - v)^{-1} dv. \quad (36)$$

Next, we extricate the divergent logarithmic part from the right-hand side of (36) and neglect small terms in the convergent remainder (bearing in mind that  $\delta \gg 1$ ),

$$\begin{aligned} \frac{1}{H} \int_0^{\delta^H/(1+\delta^H)} v^{\frac{1}{\varepsilon H}-1} (1 - v)^{-1} dv \\ &= \frac{1}{H} \left( \ln(1 + \delta^H) - \int_0^{\delta^H/(1+\delta^H)} \frac{1 - v^{\frac{1}{\varepsilon H}-1}}{1 - v} dv \right) \\ &\sim \ln \delta - \frac{1}{H} \int_0^1 \frac{1 - v^{\frac{1}{\varepsilon H}-1}}{1 - v} dv. \end{aligned} \quad (37)$$

The integral on the right-hand side of (37) is Euler's integral representation of the  $(\frac{1}{\varepsilon H} - 1)$ -th harmonic number [64], for which we have

$$\int_0^1 \frac{1 - v^{\frac{1}{\varepsilon H} - 1}}{1 - v} dv = \gamma + \psi\left(\frac{1}{\varepsilon H}\right), \quad (38)$$

where  $\gamma$  is the Euler–Mascheroni constant and  $\psi(s)$  is the digamma function (the logarithmic derivative of the gamma function) [63]. Thus, equations (35)–(38) imply that

$$\int_0^\delta e^{-\lambda y} f_0(y) dy \sim \ln \delta - \frac{1}{H} \left( \gamma + \psi\left(\frac{1}{\varepsilon H}\right) \right) \quad (39)$$

holds for the first integral on the right-hand side of (32).

Inserting the approximations (34) and (39) into (32), we obtain

$$A_0 \sim -\ln \lambda - q, \quad (40)$$

where the constant  $q$  is given by

$$q = \gamma \left( 1 + \frac{1}{H} \right) + \frac{1}{H} \psi\left(\frac{1}{\varepsilon H}\right). \quad (41)$$

The constant  $q$  in (40) asymptotically dominated by the divergent logarithmic term  $-\ln \lambda$  as  $\lambda$  tends to zero; nevertheless, from a practical viewpoint, the constant is not negligible since the slowly convergent logarithmic term is in most situations comparable in magnitude.

Using the asymptotic expressions (31) for  $A_1$  and  $A_2$  and (40) for  $A_0$ , together with the definition  $\lambda = K/\varepsilon$ , in the formulae for the mean and  $\text{CV}^2$  (29), we find

$$\langle x \rangle \sim \frac{\varepsilon}{\ln \frac{\varepsilon}{K} - q}, \quad \text{CV}^2 \sim \ln \frac{\varepsilon}{K} - q - 1, \quad (42)$$

where the constant  $q$  is given by (41); these expressions are valid in the strong feedback regime ( $K \ll \varepsilon$ ). Additionally, we have

$$\text{CV}_{\text{rel}}^2 = \frac{\langle x \rangle}{\varepsilon} \text{CV}^2 \sim 1 - \frac{1}{\ln \frac{\varepsilon}{K} - q} \quad (43)$$

for the ratio  $\text{CV}_{\text{rel}}^2$  of the regulated protein's  $\text{CV}^2$  and that of a constitutively expressed protein with an equal mean expression and mean burst size.

It is interesting to compare the ultimate asymptotics (42) in the strong feedback regime to the intermediate asymptotics obtained by taking  $K$  small in the LNA results. In the latter case, the protein mean is approximated by the deterministic steady state  $x_s$ , which is equal to the fixed point of the function  $(1 + (x/K)^H)^{-1}$ . One sees easily that

$$x_s \sim K^{\frac{H}{1+H}}, \quad K \ll 1, \quad (44)$$

which suggests a faster, power-law, decrease in the protein mean and, if inserted in (16)–(18), a power-law increase in the coefficient of variation. However, the power-law mode is applicable only in the low noise scenario for an intermediate range of  $K$ ; as  $K$  further decreases, the logarithmic law (42) applies.

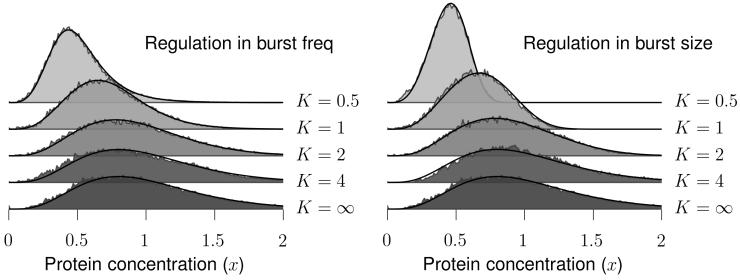


Figure 2: Protein distributions for varied feedback strength.  $\varepsilon = 0.2$ .

## 6 Results

The methods described in the previous sections are used here to study the protein distributions and noise characteristics as a function of strengthening negative feedback, whereby we shall distinguish and juxtapose two cases, the first being the regulation of the burst frequency and the second the regulation of the burst size.

The feedback strength is determined by one key parameter, the dissociation constant  $K$ , which is defined as the concentration of protein required to achieve 50% repression. The lower the dissociation constant, the lower the concentration threshold for effective self-control: the stronger the feedback.

The dissociation constant is measured in the units of protein concentration. In this study, the chosen unit of concentration is equal to the mean protein abundance in the absence of regulation. This natural choice of scale helps minimise the dimension of the parameter space of our models.

In addition to the dimensionless dissociation constant  $K$ , two other key parameters are identified: the cooperativity coefficient  $H$  and the noise parameter  $\varepsilon$ . The cooperativity coefficient determines the steepness of the regulatory response to increasing protein concentrations. All examples in this study use  $H = 4$ , which we consider a satisfactory representative for any  $H > 1$ . The non-cooperative case  $H = 1$  is an exception and is treated in Appendix C. Negative cooperativity ( $0 < H < 1$ ) is not considered.

The noise parameter  $\varepsilon$  determines the size of bursts of protein synthesis in the chosen units of protein concentration. The burst frequency is therefore  $O(\varepsilon^{-1})$  in order that protein concentration be  $O(1)$  as stated. In the small-noise regime ( $\varepsilon \ll 1$ ), analytically tractable expressions for protein noise are obtained using linear noise approximation. These will be contrasted with exact (i.e. not asymptotic) numerical results.

**Right tails of protein distributions are narrower for feedback in burst size.** The response of steady-state protein probability densities to increasing strength of either kind of feedback is investigated in Figure 2. The exact result (9), in which the Lyapunov function  $\Phi$  is given by (10) for feedback in burst frequency and (20) for feedback in burst size, is shown in solid lines, and is compared to histograms obtained by large-scale Gillespie simulations of a finer-grained discrete stochastic model (description of which is found in Appendix

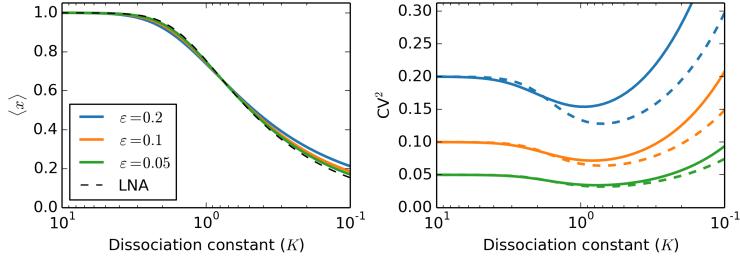


Figure 3: Protein mean and  $CV^2$  in response to strengthening feedback in burst frequency.

D).

Inspecting the distributions in Figure 2, we infer that strengthening negative feedback (decreasing the dissociation constant  $K$ ) of either kind reduces the mode and the width of steady state protein distributions. However, feedback in burst size of medium to high strength ( $K = 0.5$ ) leads to narrower distributions, in particular in their right tail, than feedback in burst frequency. Such differences can intuitively be explained: negative regulation in burst frequency leads to less frequent bursts, which are nevertheless large and contribute towards the right tail; regulation in burst size, on the other hand, leads to smaller burst sizes, thus effectively reducing the tail.

**Noise increases after an initial decrease in response to strengthening feedback in burst frequency.** The impact of increasing the strength of feedback in burst frequency on protein mean and the squared coefficient of variation (12) ( $CV^2$ ) is examined in Figure 3. The horizontal axis in Figure 3 gives the dissociation constant  $K$  on the inverse logarithmic scale, i.e. moving constantly to the right along the axis corresponds to increasing feedback strength exponentially. The values of  $K$  range from  $K = 10$  (low feedback strength) to  $K = 10^{-1}$  (strong feedback).

Solid lines give exact (as opposed to asymptotic) results obtained by numerical integration of the moments of the density (9)–(12) for a selection of values of  $\varepsilon$ . Dashed lines give the linear noise approximation (LNA) results (16)–(18), which are valid asymptotically in the small noise regime of  $\varepsilon \ll 1$ .

Focusing at first on the left panel of Figure 3, we observe that the protein mean  $\langle x \rangle$  monotonically decreases from 1 (for  $K = \infty$ , i.e. without feedback) down to 0 (for  $K = 0$ , i.e. complete repression). In small- to moderate-noise regimes of  $\varepsilon$ , the exact protein mean differs little from the LNA, which is equal to the steady state  $x_s$  of the deterministic model (14). The deterministic steady state is computed numerically as a unique fixed point of the production rate function  $(1 + (x/K)^H)^{-1}$ .

Looking at the right panel of Figure 3, we see that in the absence of regulation ( $K = \infty$ ), we have  $CV^2 = \varepsilon$ . In response to lowering the dissociation constant  $K$ , the  $CV^2$  first decreases and then increases back again. The LNA suggests that  $CV^2$  goes to infinity as  $K$  decreases. For  $\varepsilon \ll 1$ , the minimal  $CV^2$  is achieved for  $x_s = (H + 1)/2H$  and is equal to  $4H\varepsilon/(H + 1)^2$ . Comparing the

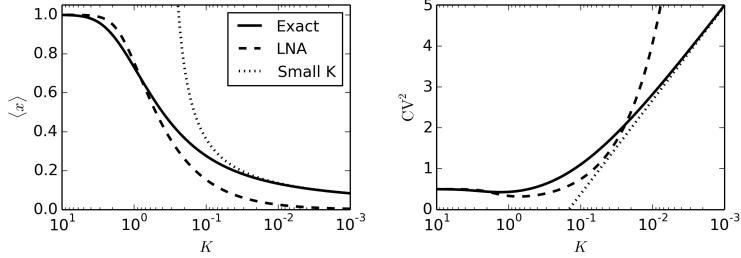


Figure 4: Impact of strengthening feedback in burst frequency on mean and  $CV^2$  of a noisy protein ( $\varepsilon = 0.5$ ). Comparison of exact results (full line) with LNA (dashed line) and small  $K$  (dotted line) asymptotics.

LNA to the exact results, we observe that larger values of  $\varepsilon$  make the initial blip in the  $CV^2$  less pronounced than the LNA predicts.

The initial decrease of noise in response to strengthening feedback strength is intuitively expected: whenever protein is in surplus, additional synthesis of protein is restricted by negative feedback, thus reducing deviations from the mean. However, as well as reducing deviations from the mean, the present type of feedback decreases the overall burst frequency, which implies higher levels of noise. The increase in noise due to lower burst frequencies demonstrably dominates over the decrease due to reduction in deviations from the mean for large feedback strengths. Below, we describe a more refined measure of noise, the relative coefficient of variation, which adjusts for the decrease in overall burst frequency. Additionally, at very low frequencies feedback in burst frequency may lose its ability to control protein fluctuations, whereby each burst overshoots and is followed by a period of complete self-repression, which takes a long time until enough protein is degraded, so that another burst may occur (which, inevitably, overshoots again).

**LNA underestimates the mean and overestimates  $CV^2$  of a noisy protein subject to strong feedback in burst frequency.** As we pointed out above, the LNA predicts that protein  $CV^2$  diverges to infinity as the dissociation constant  $K$  tends to zero. Additionally, it predicts a power-law growth of the  $CV^2$ , which is due to a power-law decay of the mean (44). However, since increasing the feedback strength in burst frequency leads to large levels of noise, the LNA prediction, which assumed little noise, becomes ever less reliable as  $K \rightarrow 0$ . Hence, even for small  $\varepsilon$  the LNA will ultimately fail provided that the feedback is raised to a sufficient strength.

For these reasons, we derived in Section 5 asymptotic expressions for protein mean and  $CV^2$  which are valid in the strong feedback regime even at high noise levels. More precisely, they are valid for  $K \ll \varepsilon$ : noisier proteins require lesser feedback strengths for these results to apply. In contrast with the LNA prediction, a slow logarithmic decrease (42) in the mean and increase in the  $CV^2$  is discovered. The LNA power-law prediction can thus only be taken as an intermediate asymptotic result applicable for small  $\varepsilon$  for intermediate ranges of feedback strengths.

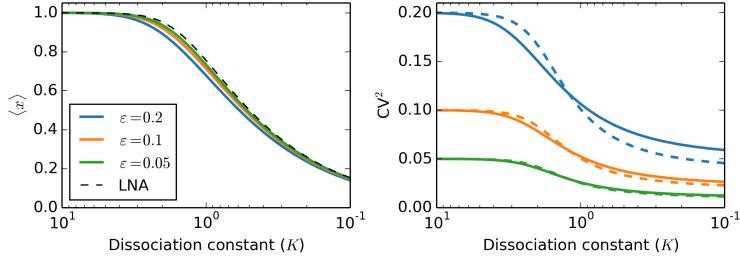


Figure 5: Protein mean and  $CV^2$  in response to strengthening feedback in burst size.

The exact numerics, the linear-noise, and the strong-feedback asymptotics for protein mean and  $CV^2$  are compared in Figure 4 for a relatively noisy protein  $\varepsilon = 0.5$  (this value corresponds to a maximum of average two bursts per protein lifetime). Since  $K$  is measured on a logarithmic scale, the limiting logarithmic dependence (42) of the  $CV^2$  on  $K$  (right-panel, dotted line) looks like a straight line, whose slope is  $\ln 10$  and intercept is  $\ln \varepsilon - q - 1$ , where  $q$  is defined by (41).

**Noise decreases in response to strengthening feedback in burst size.** In case of feedback in burst size, equations (9), (11), and (12) are used with the alternative definition (20) of the Lyapunov function  $\Phi$  to compute the exact mean and  $CV^2$  numerically. The LNA of the mean is again equal to the deterministic steady state  $x_s$ , and the LNA of the  $CV^2$  is given by (23).

In contrast to the previous case, the  $CV^2$  decreases monotonically from the value  $\varepsilon$  in the absence of regulation to a lower value in the limit of complete repression, which is equal to  $\varepsilon/(H+1)$  in the small-noise regime. For moderate values of  $\varepsilon$ , the decrease in the  $CV^2$  is less sigmoidal than predicted by the LNA.

**Comparing noise of a regulated protein to that of a constitutively expressed one with the same mean.** The ability of negative feedback to suppress protein noise can be evaluated by comparing the regulated protein  $CV^2$  to that of a constitutively expressed protein with the same mean level of expression. We refer to the ratio of the regulated  $CV^2$  and constitutive  $CV^2$  as the relative  $CV^2$ , or shortly  $CV_{\text{rel}}$ .

In our modelling framework, a constitutive expression of a protein is modulated by two parameters: the average burst size, which is measured in the chosen units of concentration; and the burst frequency, which is measured in the units of protein decay rate constant. Our condition of equal means implies that the product of these two must be equal to the mean of the regulated protein.

Having made requirement of equal means, one degree of freedom still remains in the parameter space of the constitutively expressed protein, and with this breadth of freedom a continuum of values of  $CV^2$  can be attained. Therefore, an extra condition is required on the constitutive protein to arrive at a well-defined comparison.

This extra condition differs depending whether we investigate feedback in burst frequency or feedback in burst size. If feedback is in burst frequency,

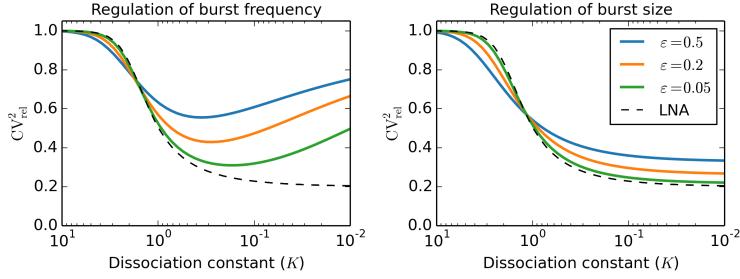


Figure 6: Relative squared coefficient of variation ( $\text{CV}_{\text{rel}}^2$ ), i.e. the ratio of the regulated protein  $\text{CV}^2$  relative to that of a constitutive protein, for feedback in burst frequency and size.

the average burst size is constant, and we require that the constitutive protein have the same average burst size, adjusting its burst frequency to achieve the required mean. On the other hand, if feedback is in burst size, then the burst frequency is constant, and we require that the constitutive protein has the same burst frequency. This difference leads to different constitutive  $\text{CV}^2$ 's and hence different definitions of  $\text{CV}_{\text{rel}}^2$  in the two cases: compare (13) and (21).

**The two feedback types exhibit the same relative noise attenuation in the small noise regime.** In the regime of small but frequent bursts ( $\varepsilon \ll 1$ ), linear noise approximation yields an explicit expression for  $\text{CV}_{\text{rel}}^2$ . Interestingly, the same result is obtained whether feedback is in burst size or frequency, cf. (18) and (23). The asymptotic  $\text{CV}_{\text{rel}}^2$  decreases with increasing feedback strength, converging to  $1/(H + 1)$  as the dissociation constant  $K$  tends to zero (Figure 6, both panels, dashed line).

**Feedback in burst size outperforms feedback in burst frequency in reducing relative noise outside of the small noise regime.** Unlike in the small-noise regime, at moderate noise levels feedback type influences  $\text{CV}_{\text{rel}}^2$ . While feedback of either type brings about a decrease of  $\text{CV}_{\text{rel}}^2$ , feedback in burst frequency is most efficient at intermediate strengths, after which  $\text{CV}_{\text{rel}}^2$  begins to increase again (Figure 6, left panel, solid coloured lines); on the other hand, the response of  $\text{CV}_{\text{rel}}^2$  to strengthening feedback in burst size is monotone, albeit less sigmoidal than in the LNA regime (Figure 6, right panel, solid coloured lines).

**Noise optimisation for feedback in burst frequency.** There is a positive value of the dissociation constant  $K$  which minimises  $\text{CV}_{\text{rel}}^2$  for feedback in burst frequency (Fig 6, left). If feedback is very weak ( $K \gg 1$ ), protein fluctuations are largely free of repression, and protein noise is close to that of an unregulated protein (i.e.  $\text{CV}_{\text{rel}}^2 \approx 1$ ). If feedback is very strong ( $K \ll \varepsilon$ ), then protein time traces consist of separated, regularly spaced, bursts: each burst produces amount of protein that is typically much larger than  $K$ , so that immediately after the burst the propensity for another burst drops dramatically; only after the present amount of protein degrades another burst may occur. In contrast to

an unregulated protein produced with the same average burst frequency, stringently self-repressed protein maintains regular time-spacing between individual bursts. The effect of regular bursting on protein noise becomes less important as feedback strength increases further, since stronger feedback implies a lower average burst frequency and hence a lesser chance of two bursts occurring at similar times by chance in the absence of regulation. Thus, in the strong-feedback limit of  $K \rightarrow 0$ , noise of a strongly self-repressed protein is equal to that of an unregulated protein. In the intermediate range of dissociation constants  $1 \lesssim K \lesssim \varepsilon$ , negative feedback efficiently reduces protein fluctuations, and a regulated protein is less noisy than its unregulated counterpart.

## 7 Discussion

In this paper we aimed to contribute towards the theoretical understanding of the effects of negative feedback on stochastic gene expression. Previous studies used small-noise approximations to obtain tractable expressions for protein noise characteristics as functions of biochemical parameters [18; 49; 51; 52]. Others obtained exact, but perhaps harder to interpret, results, which are valid even at low molecule copy numbers or large-deviation regimes [50; 65–67]. We decided to combine the two approaches, comparing the exact numerical predictions with asymptotic approximations to obtain a complete characterisation for a minimalistic model for the a protein produced in bursts subject to negative feedback.

Our results reinforce previously made observations that downstream feedback (here feedback in burst size) can better perform than upstream feedback (here regulation of burst frequency) in reducing protein variability [18; 23; 49; 51; 52]. For a protein which regulates its burst frequency, increasing feedback strength tends to increase the coefficient of variation, after an initial decrease. On the other hand, strengthening feedback in burst size leads to a monotone decrease in noise.

If instead of focusing on absolute coefficients of variation we measure how does the feedback improve in reducing noise on the performance of an equivalent constitutively expressed protein, we obtain a subtler difference between the two types of feedback, which is indeed indistinguishable in the small-noise regime. However, outside of this regime, even this subtler comparison shows a preference for regulation in burst size, especially in stringent feedback regimes. Hence, our approach suggests a possible role of large deviations in distinguishing between the two regulation mechanisms.

Our paper confirms the useful role asymptotic analysis can play in understanding the minutiae of stochastic gene expression [68–71]. Asymptotics complements numerics, one working well in parameter regimes where the other fails and vice versa. More than one asymptotic regime may be needed to be considered in a given modelling context; our example required two: small noise regime and strong feedback regime. Finding the asymptotics in this two regimes and filling the middle ground with numerical results yielded a satisfactory understanding of the model behaviour across the parameter space.

## Appendix A. Reduction to the deterministic limit

In the main text we showed that, irrespective of whether the feedback acts on the burst frequency or burst size, the steady-state mean  $\langle x \rangle$  of the protein concentration tends in the small-noise limit  $\varepsilon \rightarrow 0$  to the fixed point  $x_s$  of the ordinary differential equation (14). Here we provide a stronger result, showing that for both feedback types the master equation (3)–(4) reduces as  $\varepsilon \rightarrow 0$  to Liouville's partial differential equation associated with the ordinary differential equation (14). Thus, we show that both regulation strategies yield the same deterministic model in the limit of small noise.

For feedback in burst frequency (8) the probability flux (4) is given by

$$J = -xp(x, t) + \frac{1}{\varepsilon} \int_0^x \frac{e^{-\frac{x-y}{\varepsilon}} p(y, t) dy}{1 + (y/K)^H} \quad (\text{A1})$$

If  $\varepsilon$  is small, a dominant contribution to the integral in (A1) comes from a neighbourhood of the upper integration limit  $y = x$ . Following Watson's lemma [62], we extend the lower integration limit in (A1) to  $-\infty$  and use the approximation

$$\frac{p(y, t)}{1 + (y/K)^H} \sim \frac{p(x, t)}{1 + (x/K)^H} \quad \text{for } y \text{ that is close to } x, \quad (\text{A2})$$

obtaining

$$J \sim -xp + \frac{p}{1 + (x/K)^H} \quad (\text{A3})$$

at the leading order; higher-order terms, which are not required for our present purposes, can be determined by including in (A2) additional terms of the Taylor series expansion in  $y$  around  $x$ . The right-hand side of (A3), being the product of the protein pdf and the right-hand side of the ODE (14), gives the flux of probability induced by the drift of the deterministic model. Inserting (A3) into the probability conservation law (3) yields a Liouville equation [55, p. 213] — a Chapman–Kolmogorov equation without diffusion or jumps — whose solutions are time-dependent pdfs for a variable which evolves deterministically according to (14). Thus, the stochastic model with feedback in burst frequency given by the conservation law (3) and (A1) reduces as  $\varepsilon$  tends to zero to the deterministic model (14).

If feedback acts on burst size (19), the probability flux (4) simplifies to

$$J = -xp(x, t) + \frac{1}{\varepsilon} \int_0^x e^{-\frac{1}{\varepsilon} \int_y^x 1 + (z/K)^H dz} p(y, t) dy. \quad (\text{A4})$$

Again, a neighbourhood of the upper limit  $y = x$  of integration dominates in its contribution to the integral in (A4). Therefore, we extend the lower integration limit to  $-\infty$  without incurring appreciable error; we also use the approximations

$$\int_y^x 1 + (z/K)^H dz \sim (1 + (x/K)^H)(x - y), \quad p(y, t) \sim p(x, t), \quad (\text{A5})$$

which are valid for  $y$  that is close to  $x$ . Inserting (A5) into (A4) and integrating the simple exponential, we obtain the same leading-order approximation (A3) for the probability flux (A4) as we previously did for the flux (A1). Thus, whether the bursting stochastic model operates a feedback in burst frequency (A1) or in burst size (A4), it reduces to the same deterministic model (14) in the small-noise limit of  $\varepsilon \rightarrow 0$ .

## Appendix B. Strong feedback asymptotics (burst size)

Inserting the second Lyapunov function (20) into the WKB form (9), we obtain

$$p(x) = C x^{\frac{1}{\varepsilon}-1} e^{-\frac{1}{\varepsilon} \left( \frac{x^{H+1}}{(H+1)K^H} + x \right)} \quad (\text{B1})$$

for the protein pdf in the case of feedback in burst size.

Similarly as in the Main Text, we express the protein moments as

$$\langle x^n \rangle = \frac{B_n}{B_0}, \quad (\text{B2})$$

where instead of (26) we have

$$B_n = \int_0^\infty x^{\frac{1}{\varepsilon}-1+n} e^{-\frac{1}{\varepsilon} \left( \frac{x^{H+1}}{(H+1)K^H} + x \right)} dx. \quad (\text{B3})$$

Again,  $B_0^{-1} = C$  is the normalisation constant. Substituting  $x = Ky$  in the integral (B3) yields

$$B_n = K^{\frac{1}{\varepsilon}+n} A_n, \quad (\text{B4})$$

where

$$A_n = \int_0^\infty y^{\frac{1}{\varepsilon}-1+n} e^{-\lambda \left( \frac{y^{H+1}}{H+1} + y \right)} dy, \quad (\text{B5})$$

in which  $\lambda = K/\varepsilon$  is an auxiliary parameter.

In the case of strong feedback, we have  $\lambda \ll 1$ , which implies  $y \gg 1$ , and therefore the term  $y^{H+1}/(H+1)$  dominates the term  $y$  in the exponential of (B5), so that

$$A_n \sim \int_0^\infty y^{\frac{1}{\varepsilon}-1+n} e^{-\frac{\lambda y^{H+1}}{H+1}} dy = \frac{1}{\lambda} \left( \frac{H+1}{\lambda} \right)^{\frac{\varepsilon^{-1}+n}{1+H}-1} \Gamma \left( \frac{\varepsilon^{-1}+n}{1+H} \right), \quad (\text{B6})$$

where  $\Gamma(z)$  is the gamma function [63]. Unlike for feedback in burst frequency, here is no need to treat  $A_0$  differently from  $A_1$  or  $A_2$ .

For the protein mean we have

$$\langle x \rangle = \frac{B_1}{B_0} = \frac{KA_1}{A_0} \sim D_\varepsilon K^{\frac{H}{H+1}}, \quad (\text{B7})$$

where the prefactor  $D_\varepsilon$  is given by

$$D_\varepsilon = (\varepsilon(H+1))^{\frac{1}{H+1}} \frac{\Gamma \left( \frac{\varepsilon^{-1}+1}{H+1} \right)}{\Gamma \left( \frac{\varepsilon^{-1}}{H+1} \right)}. \quad (\text{B8})$$

Unlike for feedback in burst frequency, which yielded a slow logarithmic decrease of the mean with decreasing dissociation constant  $K$ , here we obtain a faster power-law decrease, which is consistent with the LNA prediction (44). Additionally, as  $\varepsilon$  tends to zero the prefactor  $D_\varepsilon$  converges to one, which is the

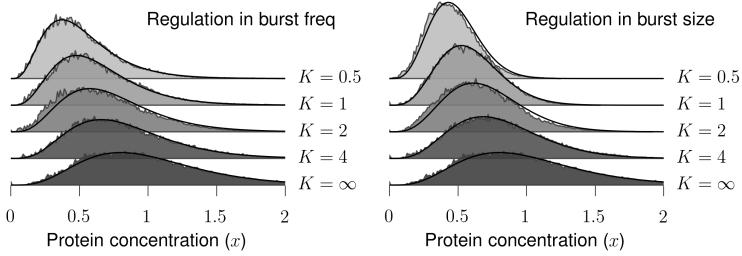


Figure C1: Protein distributions for varied feedback strength.  $\varepsilon = 0.2$ .

prefactor of the LNA-predicted power law. The asymptotics of  $D_\varepsilon$  as  $\varepsilon \rightarrow 0$  follow from

$$\frac{\Gamma(z+a)}{\Gamma(z)} \sim z^a, \quad z \gg 1, \quad (\text{B9})$$

in which we take  $z = \varepsilon^{-1}/(H+1)$  and  $a = 1/(H+1)$ ; see [63] for this and other properties of the gamma function. These results suggest that, unlike for feedback in burst frequency, where the LNA approximation could only be used for intermediate ranges of  $K$ , here the LNA yields a uniform (i.e. valid for all  $K$ ) approximation.

For the protein  $\text{CV}^2$  we have

$$\text{CV}^2 = \frac{B_2 B_0}{B_1^2} - 1 = \frac{A_2 A_0}{A_1^2} - 1 \sim \frac{\Gamma\left(\frac{\varepsilon^{-1}}{H+1}\right) \Gamma\left(\frac{\varepsilon^{-1}+2}{H+1}\right)}{\Gamma^2\left(\frac{\varepsilon^{-1}+1}{H+1}\right)} - 1, \quad (\text{B10})$$

which, for a fixed  $\varepsilon$ , is a constant independent of  $K$ . As  $\varepsilon$  tends to zero, we can again use (B9) to show that the right-hand side of (B10) is equal to  $\varepsilon/(H+1)$  at the leading order in  $\varepsilon$ , which is the same value as that obtained by taking  $K$  very small in the LNA prediction (23). This suggests that the LNA approximation of the coefficient of variation, like that of the mean, can be used uniformly for all  $K$ .

## Appendix C. Noncooperative feedback

Here we present variants of the figures from the Results in the Main Text obtained by taking  $H = 1$  (noncooperative feedback) instead of  $H = 4$ . We shall not repeat the points made in the Main Text, focusing instead on the main differences that occur in the absence of cooperativity.

**$\text{CV}^2$  monotonically increases with strengthening noncooperative feedback in burst frequency.** In contrast with the cooperative case, where a gradual increase in burst-frequency feedback strength led at first to a transient decrease in protein noise (Fig 3), without cooperativity the  $\text{CV}^2$  is strictly increasing (Fig C2).

**Protein mean and  $\text{CV}^2$  are less sensitive to feedback strength.** Wider ranges of dissociation constants are required to achieve similar changes in protein mean and noise as those reported previously. In order to appreciate

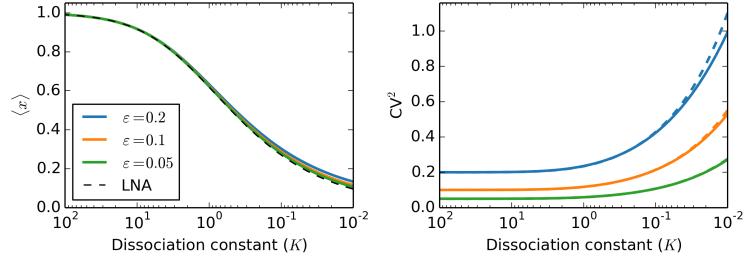


Figure C2: Protein mean and  $CV^2$  in response to strengthening feedback in burst frequency.

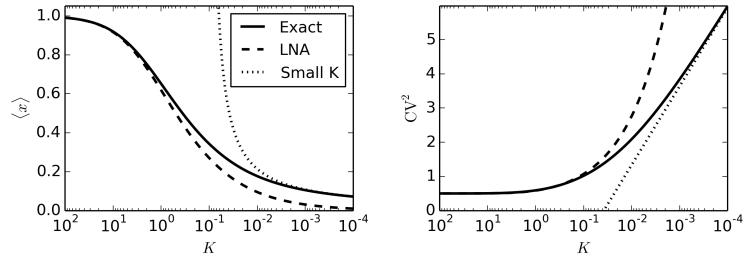


Figure C3: Impact of strengthening feedback in burst frequency on mean and  $CV^2$  of a noisy protein ( $\epsilon = 0.5$ ). Comparison of exact results (full line) with LNA (dashed line) and small  $K$  (dotted line) asymptotics.

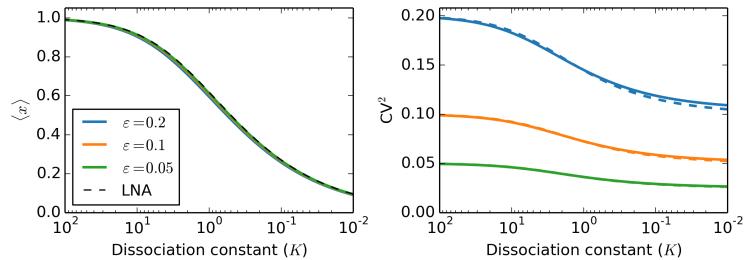


Figure C4: Protein mean and  $CV^2$  in response to strengthening feedback in burst size.

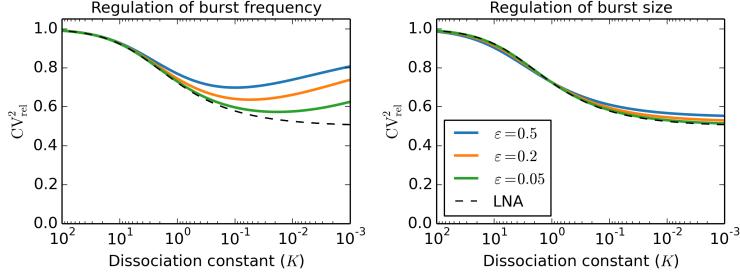


Figure C5: Relative squared coefficient of variation ( $CV_{\text{rel}}^2$ ), i.e. the ratio of the regulated protein  $CV^2$  relative to that of a constitutive protein for feedback in burst frequency and size.

this, one needs to compare the scales on the horizontal axes of Figures C2–C5 with those of their counterparts in the Main Text.

**Noncooperative performs worse than the cooperative in reducing noise.** Noncooperative feedback, even if acting through burst size, leads at best to a 50% reduction in  $CV^2$  (Fig C5), which is inferior to a 80% reduction achievable in the cooperative case with  $H = 4$  (Fig 6).

**The main conclusion of the Main Text holds also in noncooperative case.** The regulation in burst size performs better in reducing noise, especially for noisy proteins subject to strong self-repression (Fig C5).

## Appendix D. Discrete simulations

The discrete model is a chemical system of two species [72],  $A$  and  $P$ , whereby  $A \in \{0, 1\}$  is an indicator variable describing whether the gene is active ( $A = 1$ ) or inactive ( $A = 0$ ) and  $P$  gives the number of protein.

The two species are subject to four reactions, gene activation, gene inactivation, protein production, and protein decay. Each reaction is characterised by the change in copy numbers that a single occurrence of the reaction induces and by the stochastic rate with which the reaction occurs (Table D1).

The dependence of the rates of activation  $\tilde{k}_{\text{on}}(P)$ , inactivation  $\tilde{k}_{\text{off}}(P)$ , protein production  $\tilde{k}_{\text{p}}(P)$  and protein decay  $\tilde{k}_{\text{d}}(P)$  is as yet undefined in Table D1, but is specified below for feedbacks in burst frequency and burst size. We

Reaction name	Copy number change	Stochastic rate
Activation	$A \rightarrow A + 1$	$(1 - A)\tilde{k}_{\text{on}}(P)$
Inactivation	$A \rightarrow A - 1$	$A\tilde{k}_{\text{off}}(P)$
Protein production	$P \rightarrow P + 1$	$A\tilde{k}_{\text{p}}(P)$
Protein decay	$P \rightarrow P - 1$	$\tilde{k}_{\text{d}}(P)$

Table D1: Reactions, their stoichiometries, and rates for the discrete stochastic model.

use tildes to distinguish the microscopic rates (expressed in terms of individual molecules) from the macroscopic ones (expressed in terms of concentrations) which were used throughout the Main Text.

For feedback in burst frequency, we choose

$$\tilde{k}_{\text{on}}(P) = \frac{\varepsilon^{-1}}{1 + (P/K\Omega)^H}, \quad \tilde{k}_{\text{p}}(P) = \frac{\varepsilon\Omega}{\delta}, \quad \tilde{k}_{\text{off}}(P) = \frac{1}{\delta}, \quad \tilde{k}_{\text{d}}(P) = P, \quad (\text{D1})$$

while for feedback in burst size, we use

$$\tilde{k}_{\text{on}}(P) = \varepsilon^{-1}, \quad \tilde{k}_{\text{p}}(P) = \frac{\varepsilon\Omega}{\delta(1 + (P/K\Omega)^H)}, \quad \tilde{k}_{\text{off}}(P) = \frac{1}{\delta}, \quad \tilde{k}_{\text{d}}(P) = P. \quad (\text{D2})$$

In addition to the noise parameter  $\varepsilon$ , dimensionless dissociation constant  $K$  and the cooperativity coefficient  $H$ , which have been introduced in the Main Text, cf. Eq. (8) and (19), we have in (D1) and (D2) two new parameters:  $\delta$  and  $\Omega$ . The parameter  $\delta$  compares the time scale of gene activity to that of protein turnover, and  $\Omega$  is the system size parameter: the number of proteins corresponding to the unit of concentration.

Provided that  $\delta \ll 1$  and  $\Omega \gg 1$ , the protein concentration defined as  $x = P/\Omega$  can be compared to the predictions of the continuous bursting model (7). For mathematical analysis of the bursting asymptotics ( $\delta \ll 1$ ) as well as system-size asymptotics ( $\Omega \gg 1$ ), we refer the reader to [70].

In Figure 2, we used  $\varepsilon = 0.2$ ,  $H = 4$ , a range of values of  $K$  (detailed within the figure panels),  $\delta = 0.01$  and  $\Omega = 100$ . Each distribution was estimated from a single large run ( $10^5$  iterations) of Gillespie's direct method [73] implemented in the StochPy stochastic modelling software package [74].

## Acknowledgements

PB is supported by the Slovak Research and Development Agency grant APVV-14-0378 and also by the VEGA grant 1/0319/15. AS is supported by the National Science Foundation grant DMS-1312926. The authors thank Daniel Ševčovič and Branislav Novotný for discussion on some of the ideas contained herein.

## References

- [1] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain. Stochastic gene expression in a single cell. *Science*, 297:1183–1186, 2002.
- [2] A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O’Shea, Y. Pilpel, and N. Barkai. Noise in protein expression scales with natural protein abundance. *Nat. Genet.*, 38:636–643, 2006.
- [3] J. M. Raser and E. K. O’Shea. Noise in gene expression: Origins, consequences, and control. *Science*, 309:2010–2013, 2005.
- [4] Y. Taniguchi, P. J. Choi, G. W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science*, 329:533–538, 2010.

- [5] W. J. Blake, M. Kaern, C. R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422:633–637, 2003.
- [6] M. Kærn, T.C. Elston, W.J. Blake, and J.J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.*, 6:451–464, 2005.
- [7] A. Eldar and M.B. Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467:167–173, 2010.
- [8] B. Munsky, G. Neuert, and A. van Oudenaarden. Using gene expression noise to understand gene regulation. *Science*, 336:183–187, 2012.
- [9] B. Lehner. Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Mol. Syst. Biol.*, 4:170, 2008.
- [10] A. Singh and J. P. Hespanha. Evolution of gene auto-regulation in the presence of noise. *Systems Biol., IET*, 3:368–378, 2009.
- [11] E. Libby, T. J. Perkins, and P. S. Swain. Noisy information processing through transcriptional regulation. *P. Natl. Acad. Sci. USA*, 104:7151–7156, 2007.
- [12] I. Lestas, G. Vinnicombe, and J. Paulsson. Fundamental limits on the suppression of molecular fluctuations. *Nature*, 467:174–178, 2010.
- [13] M. Soltani, P. Bokes, Z. Fox, and A. Singh. Nonspecific transcription factor binding can reduce noise in the expression of downstream proteins. *Phys. Biol.*, 12:055002, 2015.
- [14] H. El-Samad and M. Khammash. Regulated degradation is a mechanism for suppressing stochastic fluctuations in gene regulatory networks. *Biophys. J.*, 90:3749–3761, 2006.
- [15] J. M. Pedraza and J. Paulsson. Effects of molecular memory and bursting on fluctuations in gene expression. *Science*, 319:339–343, 2008.
- [16] R. Bundschuh, F. Hayot, and C. Jayaprakash. The role of dimerization in noise reduction of simple genetic networks. *J. Theor. Biol.*, 220:261–269, 2003.
- [17] A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405:590–593, 2000.
- [18] P. S. Swain. Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *J. Mol. Biol.*, 344:965–976, 2004.
- [19] T. B. Kepler and T. C. Elston. Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophys. J.*, 81: 3116–3136, 2001.
- [20] A. Singh and J. P. Hespanha. Optimal feedback strength for noise suppression in autoregulatory gene networks. *Biophys. J.*, 96:4013–4023, 2009.

- [21] D. Nevozhay, R.M. Adams, K.F. Murphy, K. Josic, and G. Balazsi. Negative autoregulation linearizes the dose response and suppresses the heterogeneity of gene expression. *P. Natl. Acad. Sci. USA*, 106:5123–5128, 2009.
- [22] M. Voliotis and C.G. Bowsher. The magnitude and colour of noise in genetic negative feedback systems. *Nucleic Acids Res.*, page gks385, 2012.
- [23] L. Bandiera, A. Pasini, L. Pasotti, S. Zucca, G. Mazzini, P. Magni, E. Giordano, and S. Furini. Experimental measurements and mathematical modeling of biological noise arising from transcriptional and translational regulation of basic synthetic gene circuits. *J. Theor. Biol.*, 395:153–160, 2016.
- [24] A. Raj, C. S. Peskin, D. Tranchina, D.Y. Vargas, and S. Tyagi. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.*, 4:e309, 2006.
- [25] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef. Mammalian genes are transcribed with widely different bursting kinetics. *Science*, 332:472–474, 2011.
- [26] J.P. Bothma, H.G. Garcia, E. Esposito, G. Schlissel, T. Gregor, and M. Levine. Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living drosophila embryos. *P. Natl. Acad. Sci. USA*, 111:10598–10603, 2014.
- [27] R. D. Dar, B. S. Razooky, A. Singh, T. V. Trimeloni, J. M. McCollum, C. D. Cox, M. L. Simpson, and L. S. Weinberger. Transcriptional burst frequency and burst size are equally modulated across the human genome. *P. Natl. Acad. Sci. USA*, 109:17454–17459, 2012.
- [28] A. Singh. Transient changes in intercellular protein variability identify sources of noise in gene expression. *Biophys. J.*, 107:2214–2220, 2014.
- [29] N. Kumar, A. Singh, and R.V. Kulkarni. Transcriptional bursting in gene expression: Analytical results for general stochastic models. *PLOS Comput. Biol.*, 11:e1004292, 2015.
- [30] L. Cai, N. Friedman, and X. S. Xie. Stochastic protein expression in individual cells at the single molecule level. *Nature*, 440:358–362, 2006.
- [31] J. Yu, J. Xiao, X. Ren, K. Lao, and X. S. Xie. Probing gene expression in live cells, one protein molecule at a time. *Science*, 311:1600–1603, 2006.
- [32] J. Paulsson. Model of stochastic gene expression. *Phys. Life Rev.*, 2:157–175, 2005.
- [33] Supriya Kumar and A Javier Lopez. Negative feedback regulation among sr splicing factors encoded by rbp1 and rbp1-like in drosophila. *EMBO J.*, 24:2646–2655, 2005.
- [34] M. Jangi, P.L. Boutz, P. Paul, and P.A. Sharp. Rbfox2 controls autoregulation in RNA-binding protein networks. *Gene. Dev.*, 28:637–651, 2014.

- [35] O. Kolesnikova, R. Back, M. Graille, and B. Séraphin. Identification of the Rps28 binding motif from yeast Edc3 involved in the autoregulatory feedback loop controlling RPS28b mRNA decay. *Nucleic Acids Res.*, 41: 9514–9523, 2013.
- [36] E. Buratti and F.E. Baralle. TDP-43: new aspects of autoregulation mechanisms in RNA binding proteins and their connection with human disease. *FEBS J.*, 278:3530–3538, 2011.
- [37] D. Matelska, E. Purta, S. Panek, M.J. Boniecki, J.M. Bujnicki, and S. Dunin-Horkawicz. S6:S18 ribosomal protein complex interacts with a structural motif present in its own mRNA. *RNA*, 19:1341–1348, 2013.
- [38] O. Rossbach, L.-H. Hung, S. Schreiner, I. Grishina, M. Heiner, J. Hui, and A. Bindereif. Auto- and cross-regulation of the hnRNP L proteins by alternative splicing. *Mol. Cell. Biol.*, 29:1442–1451, 2009.
- [39] N. Friedman, L. Cai, and X. S. Xie. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Phys. Rev. Lett.*, 97:168302, 2006.
- [40] M. C. Mackey, M. Tyran-Kaminska, and R. Yvinec. Molecular distributions in gene regulatory dynamics. *J. Theor. Biol.*, 274:84–96, 2011.
- [41] M. C. Mackey, M. Tyran-Kaminska, and R. Yvinec. Dynamic behavior of stochastic gene expression models in the presence of bursting. *SIAM J. Appl. Math.*, 73:1830–1852, 2013.
- [42] P. Bokes, J. R. King, A. T. A. Wood, and M. Loose. Transcriptional bursting diversifies the behaviour of a toggle switch: hybrid simulation of stochastic gene expression. *B. Math. Biol.*, 75:351–371, 2013.
- [43] M. C. Mackey and M. Tyran-Kaminska. The limiting dynamics of a bistable molecular switch with and without noise. *J. Math. Biol.*, 73:367, 2015.
- [44] Y. T. Lin and T. Galla. Bursting noise in gene expression dynamics: linking microscopic and mesoscopic models. *J. Roy. Soc. Interface*, 13:20150772, 2016.
- [45] Y. T. Lin and C. R. Doering. Gene expression dynamics with stochastic bursts: Construction and exact results for a coarse-grained model. *Physical Review E*, 93:022409, 2016.
- [46] P. Bokes and A. Singh. Protein copy number distributions for a self-regulating gene in the presence of decoy binding sites. *PloS one*, 10: e0120555, 2015.
- [47] A. Ochab-Marcinek and M. Tabaka. Bimodal gene expression in noncooperative regulatory systems. *P. Natl. Acad. Sci. USA*, 107:22096–22101, 2010.
- [48] A. Ochab-Marcinek and M. Tabaka. Transcriptional leakage versus noise: A simple mechanism of conversion between binary and graded response in autoregulated genes. *Phys. Rev. E*, 91:012704, 2015.

- [49] A. Singh. Negative feedback through mrna provides the best control of gene-expression noise. *IEEE T. NanoBiosci.*, 10:194–200, 2011.
- [50] S. Zeiser, U. Franz, J. Müller, and V. Liebscher. Hybrid modeling of noise reduction by a negatively autoregulated system. *B. Math. Biol.*, 71:1006–1024, 2009.
- [51] A. Singh and J. P. Hespanha. Reducing noise through translational control in an auto-regulatory gene network. In *American Control Conference*, pages 1712–1717, 2009.
- [52] A. Singh. Genetic negative feedback circuits for filtering stochasticity in gene expression. In *50th IEEE Conference on Decision and Control and European Control Conference*, pages 4366–4370, 2011.
- [53] A. Crudu, A. Debussche, and O. Radulescu. Hybrid stochastic simplifications for multiscale gene networks. *BMC Syst. Biol.*, 3:89, 2009.
- [54] A. Crudu, A. Debussche, A. Muller, O. Radulescu, et al. Convergence of stochastic gene networks to hybrid piecewise deterministic processes. *Ann. Appl. Probab.*, 22:1822–1859, 2012.
- [55] Z. Schuss. *Theory and applications of stochastic processes: an analytical approach*. Springer Science & Business Media, 2009.
- [56] P. G. Hufnagel, Y. T. Lin, T. Galla, and A. J McKane. Intrinsic noise in systems with switching environments. *Phys. Rev. E*, 93:052119, 2016.
- [57] A. Grönlund, P. Lötstedt, and J. Elf. Transcription factor binding kinetics constrain noise suppression via negative feedback. *Nature Comms.*, 4:1864, 2013.
- [58] N. Kumar, T. Platini, and R. V. Kulkarni. Exact distributions for stochastic gene expression models with bursting and feedback. *Phys. Rev. Lett.*, 113: 268105, 2014.
- [59] P. C. Bressloff. *Stochastic processes in cell biology*. Springer, 2014.
- [60] Ali H Nayfeh. *Perturbation methods*. John Wiley & Sons, 2008.
- [61] S. H. Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Westview press, 2014.
- [62] E. J. Hinch. *Perturbation methods*. Cambridge university press, 1991.
- [63] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Washington, D.C., 1972.
- [64] C. E. Sandifer. *How Euler did it*. MAA Spectrum, Mathematical Association of America, 2007.
- [65] J. E. M. Hornos, D. Schultz, G. C. P. Innocentini, J. A. M. W. Wang, A. M. Walczak, J. N. Onuchic, and P. G. Wolynes. Self-regulating gene: an exact solution. *Phys. Rev. E*, 72:051907, 2005.

- [66] T. Fournier, J-P. Gabriel, C. Mazza, J. Pasquier, J. L. Galbete, and N. Mermoud. Steady-state expression of self-regulated genes. *Bioinformatics*, 23: 3185–3192, 2007.
- [67] R. Grima, D. R. Schmidt, and T. J. Newman. Steady-state fluctuations of a genetic feedback loop: An exact solution. *J. Chem. Phys.*, 137:035104, 2012.
- [68] C. Kuehn. *Multiple time scale dynamics*, volume 191. Springer, 2015.
- [69] N. Popovic, C. Marr, and P. S. Swain. A geometric analysis of fast-slow models for stochastic gene expression. *J. Math. Biol.*, 72:87–122, 2016.
- [70] P. Bokes, J. R. King, A. T. A. Wood, and M. Loose. Multiscale stochastic modelling of gene expression. *J. Math. Biol.*, 65:493–520, 2012.
- [71] J. M. Newby. Bistable switching asymptotics for the self regulating gene. *J. Phys. A: Math. Theor.*, 48:185001, 2015.
- [72] D. J. Wilkinson. *Stochastic Modelling for Systems Biology*. CRC Press, 2006.
- [73] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81:2340–2361, 1977.
- [74] T. R. Maarleveld, B. G. Olivier, and F. J. Bruggeman. Stochpy: A comprehensive, user-friendly tool for simulating stochastic biological processes. *PloS one*, 8:e79345, 2013.