Bull Math Biol
https://doi.org/10.1007/s11538-018-0438-y

Society for Mathematical Biology

ORIGINAL ARTICLE

High Cooperativity in Negative Feedback can Amplify Noisy Gene Expression

Pavol Bokes[1] · **Yen Ting Lin**[2,3] · **Abhyudai Singh**[4]

Received: 16 October 2017 / Accepted: 19 April 2018

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

PB is supported by the Slovak Research and Development Agency Grant APVV-14-0378 and also by the VEGA Grant 1/0319/15. YTL thanks for the support from the Center for Nonlinear Studies, Los Alamos National Laboratory and the Engineering and Physical Sciences Research Council EPSRC (UK) (Grant No. EP/K037145/1). AS is supported by the National Science Foundation Grant DMS-1312926.

✉ Pavol Bokes
pavol.bokes@fmph.uniba.sk

Yen Ting Lin
yentingl@lanl.gov

Abhyudai Singh
absingh@udel.edu

1 Department of Applied Mathematics and Statistics, Comenius University, Bratislava 84248, Slovakia

2 Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

3 Statistical Physics and Complex System Group, School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK

4 Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA

Published online: 25 April 2018

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, feedbacks with lower cooperativities have higher critical noise loads, suggesting that they can be preferable for controlling noisy proteins.

Keywords Stochastic gene expression · Protein bursting · Negative feedback · Delayed production · Asymptotic expansions

Mathematics Subject Classification 92C40 · 60K40 · 41A60

1 Introduction

Due to the low molecular copy numbers involved, gene expression at a single cell level exhibits a considerable degree of noisiness (Blake et al. 2003; Elowitz et al. 2002). Noise in gene expression is transmitted downstream to the final product of gene expression, the protein, whose concentration can fluctuate widely in time (Suter et al. 2011; Golding et al. 2005). The fluctuations are particularly pronounced if the protein is produced in bursts of multiple molecule copies at a single time (Dar et al. 2012; Ozbudak et al. 2002). 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 (Munsky et al. 2012; 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 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 (Grönlund et al. 2013; Stekel and Jenkins 2008). 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 (Bokes and Singh 2017; Wang et al. 2014; Scott et al. 2007).

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 (Schikora-Tamarit et al. 2016; Swain 2004; Yates and Nomura 1981). Previous studies indicate that post-transcriptional feedback outperforms transcriptional feedback in terms of its noise-reducing capability (Bokes and Singh 2017; Singh 2011; Swain 2004). Here we shall investigate a stochastic model for burst-like production of a protein which controls its own burst size via a negative (post-transcriptional) feedback loop.

Synthesis of functional protein molecules requires the completion of many consecutive stages of gene expression (Alberts et al. 2002). Consequently, self-regulatory proteins operate their feedback with a delay (Monk 2003). Deterministic models of gene expression with high-cooperativity negative feedback exhibit sustained oscillations if supplied with a sufficient production delay (Griffith 1968). Stochastic models with negative feedback become more variable upon the inclusion of a small delay (Lafuerza and Toral 2011); they exhibit increasingly regular oscillations prior to the onset of a Hopf bifurcation (Barrio et al. 2006). An infinitesimally small delay, it will be argued below, is an integral part of our model for the feedback in burst size. The delay will be shown to be responsible for a failure of the feedback to keep gene-expression noise under control.

While some gene-expression models are exactly solvable (Dattani and Barahona 2017; Yang et al. 2017; Innocentini et al. 2016; Biancalani and Assaf 2015; Smith and Shahrezaei 2015; Kumar et al. 2014), for tractable results one often has to turn to numerical (Johnson and Munsky 2017; Cao et al. 2016; Lester et al. 2016) and asymptotic analysis (Dessalles et al. 2017; Popovic et al. 2016; Bruna et al. 2014; Leier et al. 2014; Platini et al. 2011). A prominent role is played by small-noise approximations, which linearise the model around a deterministic steady state, and WKB approximations, which allow to capture slow transitions between these states or into extinction (Be'er and Assaf 2016; Roberts et al. 2015; Newby 2015; Wang et al. 2014; Komorowski et al. 2013; Elf and Ehrenberg 2003). An alternative approximation has been used by Bokes and Singh (2017) in the regime of strong negative feedback. Here we shall apply small-noise and strong-feedback approximations to a model for feedback in burst size, 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). In its basic 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 (Bokes et al. 2013; Lin and Doering 2016). Extensions to the basic formulation include non-exponential decay (Soltani et al. 2015; Bokes and Singh 2015), non-exponential bursting distributions (Jedrak and Ochab-Marcinek 2016b), multiple gene copies (Jedrak and Ochab-Marcinek 2016a), transcriptional leakage (Ochab-Marcinek and Tabaka 2015), non-cooperative binding (Ochab-Marcinek and Tabaka 2010), and multiple-component systems (Pájaro et al. 2017; Lin and Galla 2016). Most previous studies include feedback in burst frequency: the jump intensity depends on the current level of protein. Bokes and Singh (2017) introduced a particular form of feedback in burst size, in which the hazard rate for the

termination of a growing burst depends on the current protein concentration. Here we shall consider a different formulation, in which the protein concentration immediately before a jump determines the mean jump size. We will 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 capabilities in the strong-feedback regime.

The modelling framework is first introduced in Sect. 2 for the simple case of a constitutively expressed protein and then extended by feedback in burst size in Sect. 3. The fundamental features of the model behaviour are illustrated in Sect. 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 elaborates on the implicit presence of an infinitesimally small delay in the model. Section 8 compares the current model to its alternatives in (Bokes and Singh 2017), and Sect. 9 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 Fig. 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: decay rate constant γ (units of time^{-1}), burst frequency a (units of time^{-1}), and mean burst size b (units of concentration). The typical time scale of protein decay is $1/\gamma$, and the typical (or mean) concentration scale is given by ab/γ. Nondimensionalising the model by measuring time and protein concentration in units of these typical scales is equivalent to setting

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

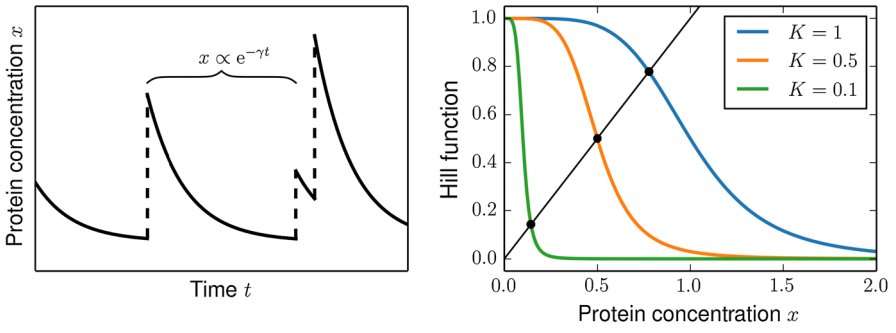

Fig. 1 (Color figure online) Left: a sample trajectory of the drift-jump process modelling the dynamics of protein concentration. Right: Hill functions $(1 + (x/K)^5)^{-1}$ and their fixed points for selected values of K

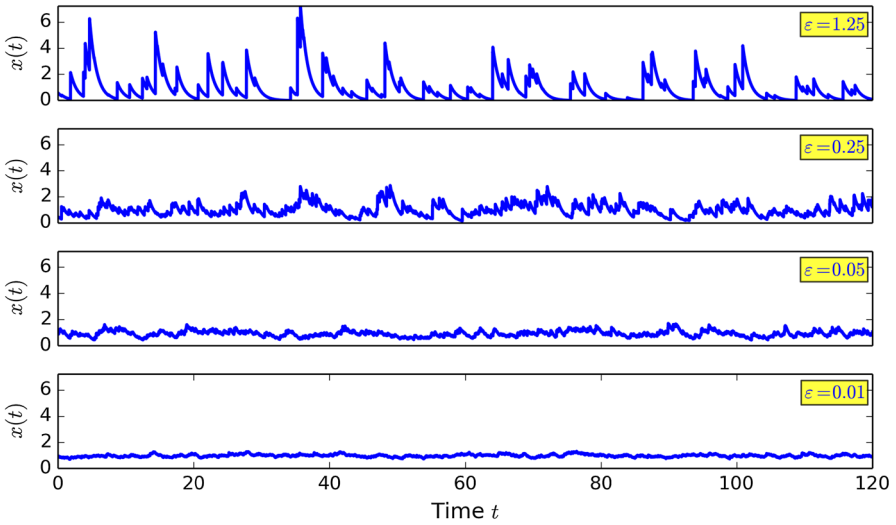

Fig. 2 Sample trajectories for a constitutively expressed protein for selected values of ε

where ε is a positive dimensionless real number, which fully controls the qualitative behaviour of the model (Fig. 2). Small values of ε lead to a limit scenario of small but frequent bursts. We emphasise that small ε does not necessarily imply physically small bursts; rather, it implies that bursts are small in comparison with the mean protein level.

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

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

The stationary pdf, up to a proportionality factor, reduces as $\varepsilon \to 0$ to

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

in terms of the normalised variable ξ. Thus, the protein concentration is approximately normally distributed in the small-noise regime.

3 Feedback in Burst Size

We incorporate negative autoregulation of burst size into the model of Sect. 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 (cf. Fig. 1, right panel)

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

of the mean burst size on the protein concentration.[1] 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

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

[1] Biologically, the mean size of a translational burst is equal to the ratio of the translation rate k_p to the mRNA decay rate γ_m (Friedman et al. 2006). We assume that the degradation of mRNA is catalysed by the protein, such as in

$$\text{mRNA} \xrightarrow{\gamma_0} \emptyset, \quad \text{mRNA} + H \times \text{protein} \xrightarrow{\gamma_1} H \times \text{protein},$$

where γ_0 is the natural decay rate and γ_1 measures the efficiency of the catalysis. The total rate of mRNA decay is then given by $\gamma_\mathrm{m} = \gamma_0 + \gamma_1 x^H$, and the mean burst size is equal to $k_\mathrm{p}/(\gamma_0 + \gamma_1 x^H)$, which can be turned into (6) by a suitable choice of unit concentration. The above reaction channels form the backbone of a full discrete chemical kinetics model, which is examined and compared to our present continuous formulation in "Appendix A".

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

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 turn in Sects. 5 and 6 to asymptotic methods to analyse its 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.

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.

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

In the absence of feedback ($K \to \infty$), the squared coefficient of variation reduces, by (4), to the value ε of the noise load. 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 $\varepsilon_\mathrm{c} = 1/2H$ (which is identified in Sect. 6 using asymptotic methods), 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

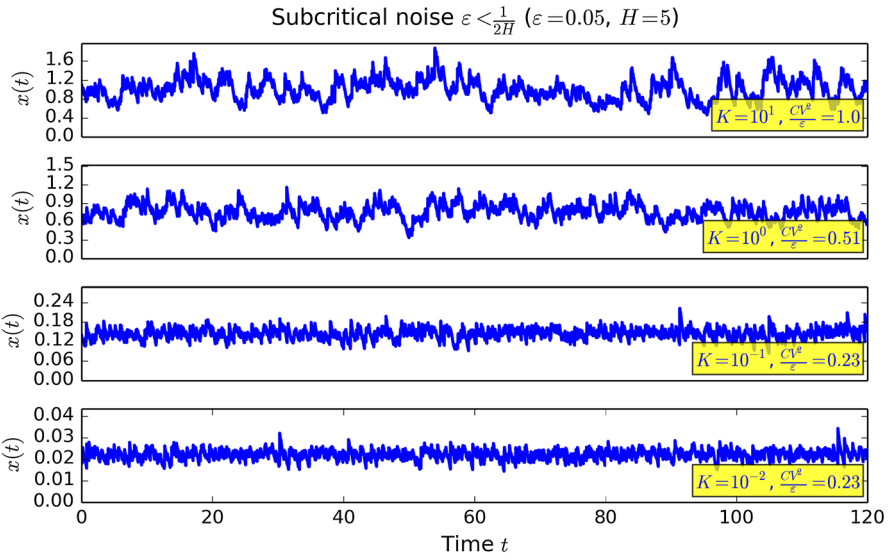

Fig. 3 For subcritical noise loads, protein noise remains bounded even under stringent feedback conditions

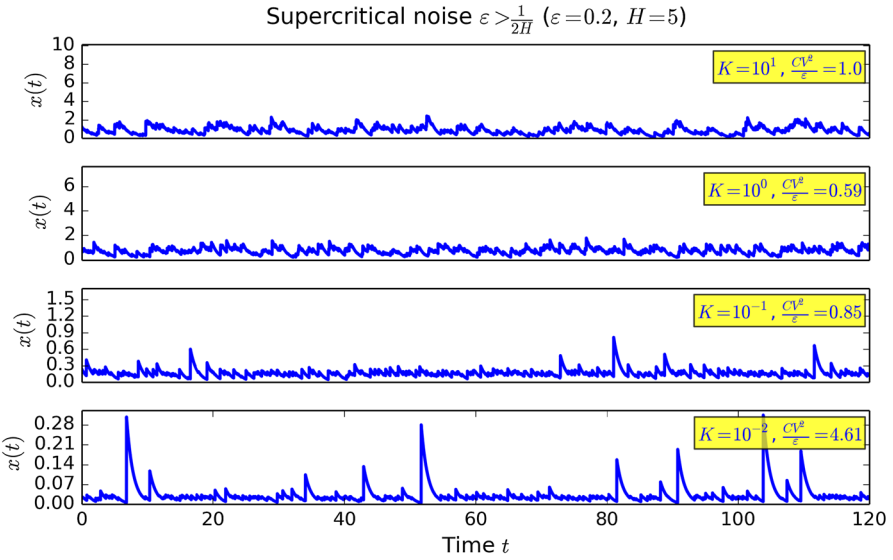

Fig. 4 For supercritical noise loads, protein noise diverges for increasingly strong feedback

strong the feedback may be (Fig. 3); for supercritical noise loads ($\varepsilon > \varepsilon_c$), the noise diverges as K tends to zero (Fig. 4). The vertical (concentration) axes in the panels of Figs. 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 B" for details).

As we shall demonstrate in Sect. 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 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, nonzero 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 nonzero 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 Sect. 6. Prior to doing that, we perform in the following Sect. 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

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

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

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

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

$$x = x_0 + \sqrt{\varepsilon}\xi, \quad x' = x - \varepsilon u, \quad p(x_0 + \sqrt{\varepsilon}\xi) = \tilde{p}(\xi). \tag{13}$$

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

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

Having extended the upper integration boundary from x/ε to infinity in (14), we have introduced an error that is inconsequential (exponentially small) in the $\varepsilon \ll 1$ regime (Hinch 1991). Inserting the expansions

$$\begin{aligned}
\tilde{p}(\xi - \sqrt{\varepsilon}u) &= \tilde{p}(\xi) - \sqrt{\varepsilon}u\frac{\mathrm{d}\tilde{p}}{\mathrm{d}\xi}(\xi) + O(\varepsilon),\\
\mathrm{e}^{-\frac{u}{\theta(x_0+\sqrt{\varepsilon}\xi-\varepsilon u)}} &= \mathrm{e}^{-\frac{u}{x_0}+\sqrt{\varepsilon}\frac{\theta'(x_0)\xi u}{x_0^2}} + O(\varepsilon)\\
&= \mathrm{e}^{-\frac{u}{x_0}}\left(1 + \sqrt{\varepsilon}\frac{\theta'(x_0)\xi u}{x_0^2}\right) + O(\varepsilon)
\end{aligned}$$

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

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

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.

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

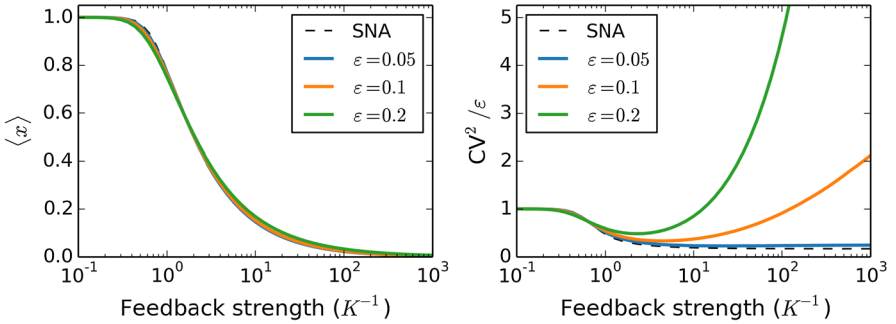

Fig. 5 (Color figure online) Protein concentration mean and the relative CV^2 in response to increasing feedback strength for $H = 5$ and a selection of values of ε, compared with the small-noise approximation (SNA)

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

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

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 B" 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

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

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 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 probability density function $p(x) = p(x; K)$ satisfies the integral equation

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

Inserting an ansatz

$$p(x; K) = \mathrm{e}^{-\frac{x}{\varepsilon}} P\left(K^{-\frac{H}{H+1}}x\right) \tag{21}$$

into (20) and substituting $x = K^{\frac{H}{H+1}}y$ and $x' = K^{\frac{H}{H+1}}y'$, we find that $P(y)$ has to satisfy a simpler equation

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

in order for (21) to satisfy (20). Notably, Eq. (22) is independent of the parameter K, and so is its solution $P(y)$.

While we are unable to find an explicit formula for a solution to (22), we will show that it satisfies

$$P(y) \sim \begin{cases} y^{\frac{1}{\varepsilon}-1} & \text{if } y \ll 1, \quad (23a) \\ \eta y^{-\frac{1}{\varepsilon H}-1} & \text{if } y \gg 1, \quad (23b) \end{cases}$$

where η is a positive constant dependent on ε and H. Relations (23) imply that

$$p(x; K) \sim \begin{cases} K^{-\frac{H(\varepsilon^{-1}-1)}{H+1}} x^{\frac{1}{\varepsilon}-1} & \text{if } x \ll K^{\frac{H}{H+1}}, \quad (24a) \\ \eta K^{\frac{\varepsilon^{-1}+H}{H+1}} \mathrm{e}^{-\frac{x}{\varepsilon}} x^{-\frac{1}{\varepsilon H}-1} & \text{if } x \gg K^{\frac{H}{H+1}}, \quad (24b) \end{cases}$$

hold for the solution (21) to the original integral equation (20).

In order to prove (23a), we note that for $y \ll 1$ the exponential in the integrand in (22) is close to one, implying that $P(y) \sim \phi(y)$, where $\phi(y)$ satisfies $y\phi(y) = \frac{1}{\varepsilon}\int_0^y \phi(y')\mathrm{d}y'$, i.e. $\phi(y) = Cy^{1/\varepsilon-1}$; for uniqueness and algebraic simplicity, we fix the arbitrary constant C to one, which yields (23a).

In order to prove (23b), we substitute $y(y')^H/\varepsilon = z$ in the integral in (22), finding that

$$yP(y) = \frac{\varepsilon^{\frac{1}{H}-1} y^{-\frac{1}{H}}}{H} \int_0^{\frac{y^{H+1}}{\varepsilon}} P\left(\left(\frac{\varepsilon z}{y}\right)^{\frac{1}{H}}\right) \exp\left(-z + \frac{\varepsilon^{\frac{1}{H}} z^{1+\frac{1}{H}}}{y^{1+\frac{1}{H}}}\right) z^{\frac{1}{H}-1} \mathrm{d}z. \quad (25)$$

Since $(\varepsilon z/y)^{\frac{1}{H}}$ becomes increasingly small as $y \to \infty$, we can approximate the unknown part of the integrand in (25) using (23a), i.e.

$$P\left(\left(\frac{\varepsilon z}{y}\right)^{\frac{1}{H}}\right) \sim \left(\frac{\varepsilon z}{y}\right)^{\frac{1}{H}\left(\frac{1}{\varepsilon}-1\right)}, \quad y \gg 1. \quad (26)$$

Inserting (26) into (25), at the same time as extending the upper integration limit in (25) to ∞ and neglecting the $O(y^{-1-1/H})$ small term in the exponential in (25), we find that

$$yP(y) \sim \frac{\varepsilon^{\frac{1}{\varepsilon H}-1} y^{-\frac{1}{\varepsilon H}}}{H} \int_0^{\infty} \mathrm{e}^{-z} z^{\frac{1}{\varepsilon H}-1} \mathrm{d}z,$$

from which (23b) follows, whereby

$$\eta = \frac{\varepsilon^{\frac{1}{\varepsilon H}-1}}{H} \Gamma\left(\frac{1}{\varepsilon H}\right) \quad (27)$$

defines the dependence of the prefactor η in (23b) on the model parameters ε and H.

Approximating the moments The n-th moment is defined by

$$m_n = \int_0^{\infty} p(x; K) x^n \mathrm{d}x. \quad (28)$$

Using (24b), we obtain

$$m_n \sim \eta K^{\frac{\varepsilon^{-1}+H}{H+1}} \int_0^{\infty} \mathrm{e}^{-\frac{x}{\varepsilon}} x^{n-\frac{1}{\varepsilon H}-1} \mathrm{d}x = \mu_n K^{\frac{\varepsilon^{-1}+H}{H+1}}, \quad \text{if } n > \frac{1}{\varepsilon H}, \quad (29)$$

where

$$\mu_n = \eta \varepsilon^{n-\frac{1}{\varepsilon H}} \Gamma\left(n - \frac{1}{\varepsilon H}\right) = \frac{\varepsilon^{n-1}}{H} \Gamma\left(\frac{1}{\varepsilon H}\right) \Gamma\left(n - \frac{1}{\varepsilon H}\right) \quad (30)$$

is finite if $n > 1/\varepsilon H$ and does not depend on K.

If $n < 1/\varepsilon H$, the $x = O(K^{H/(H+1)})$ boundary layer of the solution (21) becomes important, and (29) cannot be used. Instead, we write

$$m_n = \int_0^\infty \mathrm{e}^{-x/\varepsilon} P\left(K^{-\frac{H}{H+1}} x\right) x^n \mathrm{d}x = K^{\frac{(n+1)H}{H+1}} \int_0^\infty \mathrm{e}^{-K^{\frac{H}{H+1}} y/\varepsilon} P(y) y^n \mathrm{d}y$$
$$\sim \nu_n K^{\frac{(n+1)H}{H+1}}, \quad \text{if } n < \frac{1}{\varepsilon H}, \tag{31}$$

where

$$\nu_n = \int_0^\infty P(y) y^n \mathrm{d}y \tag{32}$$

is finite if $n < 1/\varepsilon H$ owing to (23) and does not depend on K. If $n = 1/\varepsilon H$, the $x = O(1)$ domain and the $x = O(K^{H/(H+1)})$ boundary layer are equally important in contributing to the leading-order asymptotics of the n-th moment (see "Appendix C").

Normalisation The normalised steady-state protein pdf is obtained by dividing the solution $p(x)$ by its norm $m_0 = \int_0^\infty p(x)\mathrm{d}x$. Combining (24) and (31) for $n = 0$, we arrive at asymptotic approximations

$$\frac{p(x)}{m_0} \sim \begin{cases} \nu_0^{-1} K^{-\frac{H}{\varepsilon(H+1)}} x^{\frac{1}{\varepsilon}-1} & \text{if } x \ll K^{\frac{H}{H+1}}, \quad (33a) \\ \eta \nu_0^{-1} K^{\frac{\varepsilon^{-1}}{H+1}} \mathrm{e}^{-\frac{x}{\varepsilon}} x^{-\frac{1}{\varepsilon H}-1} & \text{if } x \gg K^{\frac{H}{H+1}}. \quad (33b) \end{cases}$$

We refer to (33a) and (33b) as the *inner* and *outer* scalings, respectively. Figure 6 shows an excellent agreement between the theoretical scalings (33) and probability distributions estimated by extensive kinetic Monte Carlo simulation with Algorithm 1. Numerically, the probabilities are calculated from 10^6 sample paths (ensemble average). After the random process reaches stationarity (which we set to be 250 units of time), we record the state of sample paths every unit of time for another 1750 units of time (temporal average), and use the cumulated data to compute the stationary distributions. The low values of K used in Fig. 6 and elsewhere have been chosen so as to numerically cross-validate the asymptotic analysis; they are not necessarily meant to represent biologically realisable situations.

Approximating the protein mean and CV^2 Using the asymptotic results (29) and (31) 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

$$\langle x \rangle = \frac{m_1}{m_0}, \quad \langle x^2 \rangle = \frac{m_2}{m_0}. \tag{34}$$

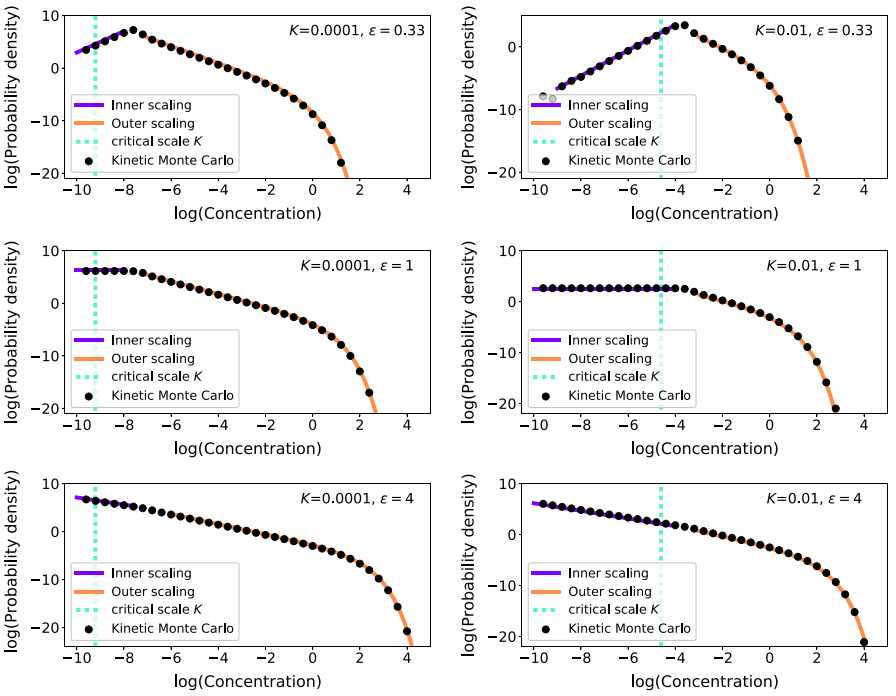

Fig. 6 (Color figure online) Protein distributions for selected parameter sets with $H = 5$. Natural-logarithm scales on both axes are used. Discrete markers: probability distributions estimated by kinetic Monte Carlo simulation with Algorithm 1; dashed line: the critical concentration K; solid lines: theoretical predictions for the inner (33a) (purple colour) and outer (33b) (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

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

$$\langle x \rangle \sim \begin{cases} \dfrac{\nu_1}{\nu_0} K^{\frac{H}{H+1}} & \text{if } \varepsilon < \frac{1}{H}, \quad (35a) \\ \dfrac{\mu_1}{\nu_0} K^{\frac{\varepsilon^{-1}}{H+1}} & \text{if } \varepsilon > \frac{1}{H}. \quad (35b) \end{cases}$$

By (35a), under low-noise conditions ($\varepsilon < 1/H$), 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 $1/H$, the exponent of the power law (35b) 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

$$\langle x^2 \rangle \sim \begin{cases} \dfrac{\nu_2}{\nu_0} K^{\frac{2H}{H+1}} & \text{if } \varepsilon < \frac{1}{2H}, \quad (36a) \\ \dfrac{\mu_2}{\nu_0} K^{\frac{\varepsilon^{-1}}{H+1}} & \text{if } \varepsilon > \frac{1}{2H}. \quad (36b) \end{cases}$$

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 $1/2H$, which is half the threshold in (36) for the mean and which we previously in Sect. 4 referred to as the critical noise load.

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

$$\mathrm{CV}^2 = \frac{\langle x^2 \rangle}{\langle x \rangle^2} - 1 \sim \begin{cases} \dfrac{\nu_2 \nu_0}{\nu_1^2} - 1 & \text{if } \varepsilon < \frac{1}{2H}, \quad (37a) \\ \dfrac{\mu_2 \nu_0}{\nu_1^2} K^{-\frac{2H-\frac{1}{\varepsilon}}{H+1}} & \text{if } \frac{1}{2H} < \varepsilon < \frac{1}{H}, \quad (37b) \\ \dfrac{\mu_2 \nu_0}{\mu_1^2} K^{-\frac{\varepsilon^{-1}}{H+1}} & \text{if } \varepsilon > \frac{1}{H}. \quad (37c) \end{cases}$$

Asymptotic approximations (37) prove the observations made in Sect. 4: CV^2 converges to a constant value as $K \to 0$ for subcritical noise loads ($\varepsilon < 1/2H$). For supercritical noise loads ($\varepsilon > 1/2H$), the CV^2 exhibits a power-law increase as K decreases to zero. Additionally, approximations (37) imply that as ε increases beyond the critical value $1/2H$, the absolute value of the exponent of the power law in (37b) first increases, reaching a maximum of $H/(H+1)$ for $\varepsilon = 1/H$; after ε exceeds $1/H$, the exponent of the power law in (37c) begins to decrease in absolute value. Thus, the fastest grow-up in the coefficient of variation as $K \to 0$ is achieved for $\varepsilon = 1/H$ by a combination of a relatively fast decrease in the mean and a relatively slow decrease in the variance. Under excessive noise conditions ($\varepsilon \gg 1/H$), 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 was reached, after which time the moments were measured every unit of time for another 1750 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 (35)–(37) (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 (35) and $\langle x^2 \rangle$ in (36), some discrepancies are observed around the critical values of ε (which are $\varepsilon = 0.2$ for $\langle x \rangle^2$ and $\varepsilon = 0.1$

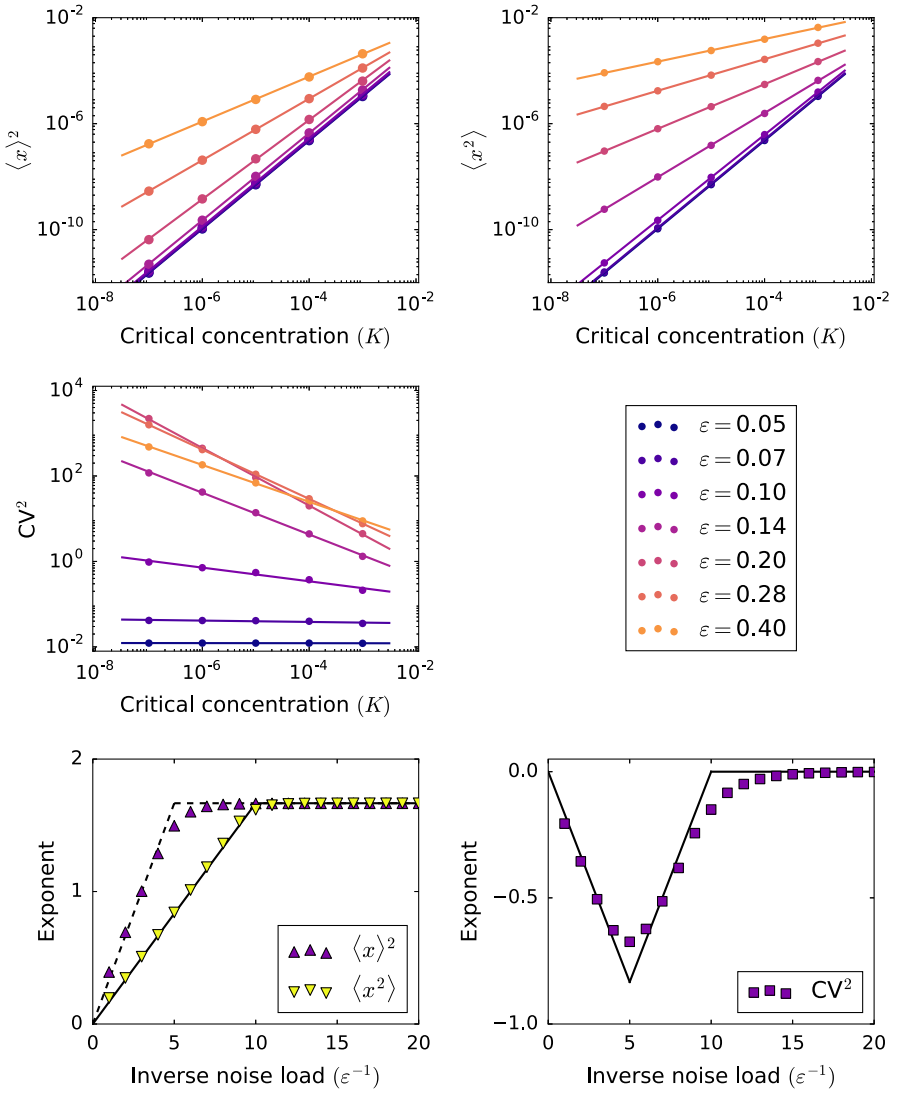

Fig. 7 The squared mean (top left), the mean square (top right) and the CV^2 (centre left), obtained by stochastic simulation (Algorithm 1), as functions of critical concentrations $K \ll 1$ for $H = 5$ and a selection of values of ε (centre right). The discrete markers are the numerically measured results. 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 (35)–(37), as functions of the inverse noise load ε^{-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 C"). 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 deviations of the simulation-based exponents from the asymptotic predictions become more pronounced after taking the difference. An

additional source of discrepancy is in the neglecting of the constant term -1 in the asymptotic approximations (37b) and (37c), which becomes important if the absolute value of the exponent is low. Despite these differences, 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 Infinitesimal Delay

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 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.

Here we would like to discuss, without delving deeply into the intricacies of the underlying multi-scale analysis, how this delay compares with other time-scales explicitly or implicitly present in the model. It is useful to think of the protein bursting model as a singular limit of a piecewise deterministic Markov process (PDMP) with two internal states 0 and 1, indicating, for definiteness, the absence or presence of an mRNA transcript (Lin and Doering 2016). The limit is achieved by making the 1 to 0 transition very fast (i.e. considering short mRNA lifetimes) while assuming that in state 1 protein concentration grows rapidly (i.e. translation is fast). Regulation of burst size, as suggested previously in Sect. 3, can be modelled by making the 1 to 0 transition rate depend on the protein concentration. However, a straightforward implementation of such a scheme implies that newly produced proteins immediately engage in feedback (Bokes and Singh 2017), including within the same burst, which is at odds with our current assumption of memoryless burst sizes.

In order to avoid immediate regulation, we need to distinguish between immature proteins that are unable to partake in feedback and mature proteins and consider an (exponential) delaying step required for protein maturation. We expect that our current bursting model can be recovered as a singular limit of such a two-component PDMP provided that

$$\text{mRNA lifetime} \ll \text{maturation delay} \ll \text{waiting time until the next burst} \tag{38}$$

hold. (We expect that the two components decouple due to time-scale separation and that we end up with our one-component bursting model for mature proteins.) Since the maturation delay is by (38) assumed to be much larger that the typical waiting time of any 1 to 0 transition, the delay will carry over any single episode of translational activity; only those proteins that have been present at its beginning can regulate how much will be produced during it. On the other hand, since the maturation delay is by (38) assumed to be much shorter than the expected waiting time until the next translational burst, all proteins produced in previous bursts will have matured before the next burst occurs, thanks to which the limiting model is one-component

and Markovian. The maturation delay in the limiting model is infinitesimally small, yet it bestows crucial qualitative features, as will become apparent in the next section from a comparison to a non-delayed model. The PDMP model itself can be thought of as an asymptotic reduction of a fine-grained discrete chemical kinetics model, which we examine in "Appendix A".

Biologically, maturation 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. Other examples of processes which can be the source of a sufficiently large production delay include post-translational modifications, folding to the functional secondary structure, and assembly into protein complexes.

8 Comparison with Previous Results

In an alternative version of our model for feedback in burst size, which was analysed in a previous paper (Bokes and Singh 2017), 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 2017) is considerably easier to treat analytically: the master equation of the undelayed model reads

$$\frac{\partial p}{\partial t}+\frac{\partial J}{\partial x}=0,\quad J=\frac{1}{\varepsilon}\int_0^x p(x',t)\mathrm{e}^{-\frac{1}{\varepsilon}\int_{x'}^x\frac{\mathrm{d}\xi}{\theta(\xi)}}\mathrm{d}x'-xp, \tag{39}$$

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

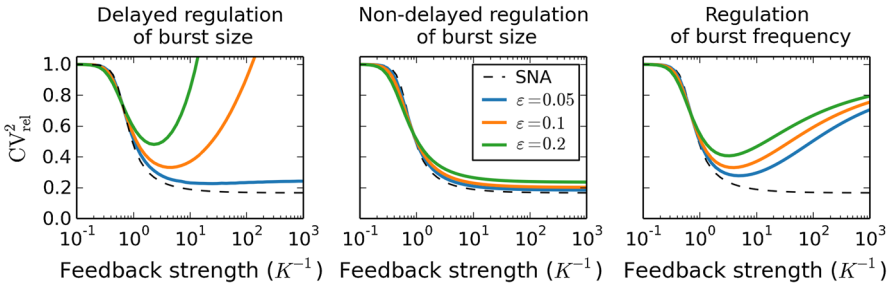

Fig. 8 (Color figure online) The relative CV^2 of a protein produced in bursts subject to different types of feedback in response to increasing feedback strength K^{-1}. We fix $H = 5$ and use a selection of values of ε which are given in the legend within the central panel. Exact numerical/simulational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

in the small-noise and strong-feedback regimes (Bokes and Singh 2017). 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). However, the non-delayed model exhibits a monotonic decrease in the coefficient of variation as feedback strengthens for any combination of ε and H (Fig. 8, central panel, coloured lines). 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 infinitesimal delay which carries over the bursting time scale.

While we have so far focused exclusively on feedback in burst size, more attention has traditionally been paid in the literature to feedback in burst frequency. If one adheres to the fundamentals of the model of Sect. 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 2017). The master equation of the model reads

$$\frac{\partial p}{\partial t} + \frac{\partial J}{\partial x} = 0, \quad J = \frac{1}{\varepsilon}\int_0^x p(x',t)\theta(x')\mathrm{e}^{-\frac{x-x'}{\varepsilon}}\mathrm{d}x' - xp, \tag{40}$$

which admits an explicit steady-state solution, see (Lin and Doering 2016; Bokes and Singh 2017; Friedman et al. 2006), given by

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

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 (2017) 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 2017) to be larger by a factor of $1/x_0$ 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. Sect. 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 $1/\varepsilon$ to the regulated value of $\langle x\rangle/\varepsilon$ 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 (2017) 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 $\mathrm{CV}^2_{\mathrm{rel}} = \langle x\rangle \mathrm{CV}^2/\varepsilon$. Interestingly, the latter definition yields an expression that is proportional to the Fano factor, a noise metric that is popular in particular in discrete stochastic systems. 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 (2017) 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 (2017) 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. In the present work, we focus exclusively on the effects of vanishingly small delays. If feedback in burst frequency is delayed by a positive finite amount of time, then a number of successive bursts may occur before the burst frequency drops. While we expect that this particular effect will vanish in the limit of very small delays, additional bursts following shortly after the first one are expected to lead to an increase in the reported level of noise.

9 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.

We should cautiously remark that the assumptions of continuity of concentration and instantaneity of bursts do not generally hold in real biological systems. In particular, under strong (self-)repression the effects of discrete protein copy numbers can be significant, such as indicated by the analysis of "Appendix A". Also, uncontrolled large bursts can take substantially longer to accumulate than smaller controlled ones. From a standpoint of mathematical modelling, we nevertheless believe that it is important to classify in detail the qualitative behaviour of simple models such as presented herein. More detailed models, after all, are bound to exhibit the same or similar qualitative features in intermediate asymptotic regimes (Barenblatt 1996).

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 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 $\varepsilon_c = 1/2H$. 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.

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 $1/(H+1)$, 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 2017) provided that one compensates for the decrease in the overall (time-averaged) burst frequency. However, 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 but supercritical ($\varepsilon > \varepsilon_c = 1/2H$). 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 so that within-burst feedback does not take place. 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.

Acknowledgements We thank an anonymous referee for useful comments and important insights, in particular those leading to the analysis of "Appendix A".

Appendix A: Discrete Model

Here we compare the results of the Main Text to a fine-grained chemical kinetics model consisting of three discrete species (mRNA, PreProtein, and Protein) that are subject to reaction channels

$$\begin{array}{lll} \emptyset \xrightarrow{\frac{1}{\varepsilon}} \mathrm{mRNA}, & \mathrm{mRNA} \xrightarrow{\frac{1}{\delta}} \emptyset, & \mathrm{mRNA} + \mathrm{Protein} \xrightarrow{\frac{\text{Protein}}{\delta\Omega K}} \mathrm{Protein}, \\ \mathrm{Protein} \xrightarrow{1} \emptyset, & \mathrm{PreProtein} \xrightarrow{\frac{1}{\lambda}} \mathrm{Protein}, & \mathrm{mRNA} \xrightarrow{\frac{\varepsilon\Omega}{\delta}} \mathrm{mRNA} + \mathrm{PreProtein}. \end{array} \tag{41}$$

These six reactions represent the following processes: (i) mRNA synthesis (burst initiation); (ii) natural mRNA decay; (iii) mRNA decay catalysed by the protein; (iv) protein decay (v) protein maturation; (vi) translation. We restrict ourselves, for simplicity, to non-cooperative feedback in the third reaction, i.e. $H = 1$ in the sense of the Main Text. The reaction constants in (41) are parameterised in a manner that relates the discrete model to its continuous counterpart of the Main Text. In addition to the familiar parameters ε (noise load) and K (critical concentration), three new parameters appear: the system-size parameter Ω gives the number of protein molecules in a unit of concentration; δ and λ give the mRNA lifetime and the mean protein maturation delay, respectively, in units of the protein lifetime.

While a systematic asymptotic analysis of (41) is beyond the scope of this paper, we expect that in the limit of $\Omega \to \infty$, the discrete model reduces to a piecewise deterministic Markov process (PDMP), with one discrete species (mRNA) and two continuous species (pre-protein and mature protein). In the limit of $\delta \to 0$, whilst keeping λ finite positive, mRNA is adiabatically eliminated from the model, and pre-protein is produced in instantaneous bursts. The size of a burst is drawn from a geometric distribution, the mean of which decreases with the amount of mature protein in the system. If one additionally takes $\lambda \to 0$ *after* the limit of $\delta \to 0$ has already been taken, then the system features an infinitesimal delay in the sense of Sect. 7. Combining the above limits of $\Omega \to \infty$ and $\delta \to 0$ before $\lambda \to 0$, one expects to

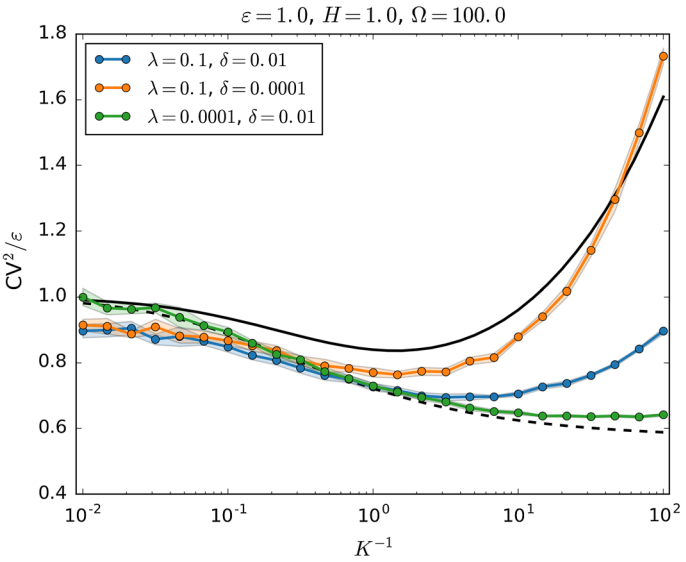

Fig. 9 (Color figure online) Noise ($\mathrm{CV}^2/\varepsilon$) as function of feedback strength (K^{-1}). We use the continuous-state drift-jump model with infinitesimal delay (solid black) or without infinitesimal delay (dashed black), and the discrete model (41) (coloured circles). Shaded bands indicate 95% confidence intervals

arrive at the continuous-state bursting model with infinitesimal delay, on which we focused throughout the Main Text.

The species means and standard deviations were calculated using Stochpy's (Maarleveld et al. 2013) implementation of Gillespie's direct method (Gillespie 1976). We skipped the first 30 units of time to reach stationarity and used the next 10^6 iterations to estimate the moments. The moments were used, in particular, to estimate the protein coefficient of variation. The procedure was repeated ten times to improve precision by additional ensemble averaging and to obtain approximate confidence intervals (assuming asymptotic normality).

In the present simulation, we fixed $\varepsilon = 1$, $H = 1$, and varied K between 10^{-2} and 10^2. The noise load $\varepsilon = 1$ exceeds the critical value $\varepsilon_c = 1/2H = 0.5$. Comparing the results of the full discrete and continuous bursting simulation in Fig. 9, we are able to draw the following conclusions:

- Finiteness of $1/\lambda$ reduces the noise. It takes λ units of time, on average, for a pre-protein to mature. The intrinsic variability of maturation times attenuates the effective burst size and reduces protein noise. Finiteness of $1/\lambda$ is responsible, in particular, for the relative decrease of noise in weak feedback conditions (Fig. 9, blue and orange markers). Under strong-feedback conditions, other effects, which are listed below, dominate.
- Finiteness of λ/δ also reduces the noise. Large values of λ/δ mean that few proteins will have matured over the duration of a single burst. Nevertheless, if feedback is strong, these few proteins may suffice to stop the burst and ameliorate noise

(Fig. 9, blue markers). An additional increase in λ/δ is required to offset the effect (Fig. 9, orange markers).
– Finiteness of Ω increases the noise. A unit of concentration comprises Ω copies of protein. Strong feedback decreases the mean protein concentration, making the low-copy-number Poissonian noise increasingly more important. We attribute to this effect, in particular, the excess noise exhibited by the discrete model under strong-feedback conditions (Fig. 9, orange and green markers).

Aside from exhibiting the aforementioned additional effects, the discrete model reflects the essential features of the continuous bursting descriptions of the Main Text, in particular with respect to the interaction between the bursting and maturation timescales, and their effect on protein noise.

Appendix B: 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

$$\frac{1}{T}\int_0^T x^n(t)\mathrm{d}t = \frac{1}{T}\sum_{i=0}^{N-1}\int_0^{\tau_i} x_i^n \mathrm{e}^{-nt}\mathrm{d}t = \frac{1}{nT}\sum_{i=0}^{N-1} x_i^n(1-\mathrm{e}^{-\tau_i n}), \tag{42}$$

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

$$x_{i+1} = x_i\mathrm{e}^{-\tau_i} - \frac{\varepsilon \ln\tilde{u}_i}{1+(x_i\mathrm{e}^{-\tau_i}/K)^H}, \quad x_0 = 0, \quad \tau_i = -\varepsilon\ln u_i, \quad i = 0, 1, \ldots, \tag{43}$$

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 (42), and shifting the time frame to reduce the effect of the transient behaviour, we arrive at an estimate

$$\widehat{\langle x^n\rangle} = \frac{1}{\varepsilon n N}\sum_{i=M}^{N+M-1} x_i^n(1-\mathrm{e}^{-\tau_i n}), \tag{44}$$

where x_i and τ_i are given by (43). We used (44) with $N = 10^9$ and $M = 10^7$ to estimate the theoretical value of $\langle x^n\rangle$ in Figs. 3, 4, 5 and 8. In the strong-feedback limit (Figs. 6, 7), ensemble averaging across a large number of sample paths was needed for more precise measurements of the power-law exponents.

Appendix C: 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 $n < 1/\varepsilon H$ or from the outer $O(1)$ concentration scale if $n > 1/\varepsilon H$. In the borderline case of $n = 1/\varepsilon H$, 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),

$$m_n = \int_0^\infty p(x)x^n \mathrm{d}x = \int_0^\delta p(x)x^n \mathrm{d}x + \int_\delta^\infty p(x)x^n \mathrm{d}x, \tag{45}$$

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 (45), we substitute $x = K^{H/(H+1)}y$ and, neglecting the slowly decaying exponential in (21), we obtain

$$\int_0^\delta p(x)x^n \mathrm{d}x \sim K^{\frac{(n+1)H}{H+1}} \int_0^{K^{-\frac{H}{H+1}}\delta} P(y)y^n \mathrm{d}y. \tag{46}$$

Since $\delta \gg K^{H/(H+1)}$, the upper integration limit of the integral on the right-hand side of (46) is large. The integral itself diverges to infinity, since $n = 1/\varepsilon H$ together with (23b) imply that $P(y)y^n \sim \eta y^{-1}$ for $y \gg 1$. Extricating the divergent part from the integral in (46), we obtain an asymptotic approximation

$$\begin{aligned}\int_0^{K^{-\frac{H}{H+1}}\delta} P(y)y^n \mathrm{d}y &= \int_0^{K^{-\frac{H}{H+1}}\delta} \left(P(y)y^n - \frac{\eta}{y+1}\right)\mathrm{d}y + \eta \int_0^{K^{-\frac{H}{H+1}}\delta} \frac{\mathrm{d}y}{y+1} \\ &\sim \int_0^\infty \left(P(y)y^n - \frac{\eta}{y+1}\right)\mathrm{d}y - \frac{\eta H}{H+1}\ln K + \eta \ln\delta. \end{aligned} \tag{47}$$

In the second integral on the right-hand side of (45), we approximate the integrand with (24b), finding that

$$\int_\delta^\infty p(x)x^n \mathrm{d}x \sim \eta K^{\frac{(n+1)H}{H+1}} E_1\left(\frac{\delta}{\varepsilon}\right) \sim \eta K^{\frac{(n+1)H}{H+1}} (-\ln\delta + \ln\varepsilon - \gamma), \tag{48}$$

where $E_1(t)$ is the exponential integral and $\gamma = 0.577\ldots$ is the Euler–Mascheroni constant (Abramowitz and Stegun 1972).

Combining (45)–(48), we find that the n-the moment can be expanded into

$$m_n \sim K^{\frac{(n+1)H}{H+1}} \left(\int_0^\infty \left(P(y)y^n - \frac{\eta}{y+1}\right)\mathrm{d}y - \frac{\eta H}{H+1}\ln K + \eta(\ln\varepsilon - \gamma)\right). \tag{49}$$

Although the logarithmic term in (49) 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 (49), 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 $O(K^{H/(H+1)})$ scale only and are given by (31). On the other hand, the leading-order approximation to m_2 combines contributions from both inner and outer scales and is given by (49). For the coefficient of variation we obtain

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

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

Abramowitz M, Stegun I (1972) Handbook of mathematical functions with formulas, graphs, and mathematical tables. National Bureau of Standards, Washington, DC

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

Barenblatt GI (1996) Scaling, self-similarity, and intermediate asymptotics: dimensional analysis and intermediate asymptotics. Cambridge University Press, Cambridge

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 Exp 2016:113501

Biancalani T, Assaf M (2015) Genetic toggle switch in the absence of cooperative binding: exact results. Phys Rev Lett 115:208101

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:e0120555

Bokes P, Singh A (2017) Gene expression noise is affected differentially by feedback in burst frequency and burst size. J Math Biol 74:1483–1509

Bokes P, King J, Wood A, Loose M (2013) Transcriptional bursting diversifies the behaviour of a toggle switch: hybrid simulation of stochastic gene expression. Bull 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 140:174107

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

Cao Y, Terebus A, Liang J (2016) State space truncation with quantified errors for accurate solutions to discrete chemical master equation. Bull Math Biol 78:617–661

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. Proc Natl Acad Sci USA 109:17454–17459

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

Dessalles R, Fromion V, Robert P (2017) A stochastic analysis of autoregulation of gene expression. J Math Biol. https://doi.org/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 13:2475–2484
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 97:168302
Gillespie D (1976) A general method for numerically simulating stochastic time evolution of coupled chemical reactions. J Comput Phys 22:403–434
Golding I, Paulsson J, Zawilski S, Cox E (2005) Real-time kinetics of gene activity in individual bacteria. Cell 123:1025–1036
Griffith J (1968) Mathematics of cellular control processes I. Negative feedback to one gene. J Theor Biol 20:202–208
Grönlund A, Lötstedt P, Elf J (2013) Transcription factor binding kinetics constrain noise suppression via negative feedback. Nat Commun 4:1864
Hinch EJ (1991) Perturbation methods. Cambridge University Press, Cambridge
Innocentini GC, Forger M, Radulescu O, Antoneli F (2016) Protein synthesis driven by dynamical stochastic transcription. Bull Math Biol 78:110–131
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 94:032401
Johnson R, Munsky B (2017) The finite state projection approach to analyze dynamics of heterogeneous populations. Phys Biol 14:035002
Komorowski M, Miekisz J, Stumpf MP (2013) Decomposing noise in biochemical signaling systems highlights the role of protein degradation. Biophys J 104:1783–1793
Kumar N, Platini T, Kulkarni RV (2014) Exact distributions for stochastic gene expression models with bursting and feedback. Phys Rev Lett 113:268105
Lafuerza L, Toral R (2011) Role of delay in the stochastic creation process. Phys Rev E 84:021128
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 R Soc Interface 11:20140108
Lester C, Baker RE, Giles MB, Yates CA (2016) Extending the multi-level method for the simulation of stochastic biological systems. Bull Math Biol 78:1640–1677
Lin YT, Doering CR (2016) Gene expression dynamics with stochastic bursts: construction and exact results for a coarse-grained model. Phys Rev E 93:022409
Lin YT, Galla T (2016) Bursting noise in gene expression dynamics: linking microscopic and mesoscopic models. J R Soc Interface 13:20150772
Maarleveld TR, Olivier BG, Bruggeman FJ (2013) Stochpy: a comprehensive, user-friendly tool for simulating stochastic biological processes. PLoS ONE 8:e79345
McAdams H, Arkin A (1997) Stochastic mechanisms in gene expression. Proc 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
Munsky B, Neuert G, Van Oudenaarden A (2012) Using gene expression noise to understand gene regulation. Science 336:183–187
Murray J (2003) Mathematical biology: I introduction. Springer, New York
Newby J (2015) Bistable switching asymptotics for the self regulating gene. J Phys A Math Theor 48:185001
Ochab-Marcinek A, Tabaka M (2010) Bimodal gene expression in noncooperative regulatory systems. Proc Natl Acad Sci USA 107:22096–22101
Ochab-Marcinek A, Tabaka M (2015) Transcriptional leakage versus noise: a simple mechanism of conversion between binary and graded response in autoregulated genes. Phys Rev E 91:012704
Ozbudak EM, Thattai M, Kurtser I, Grossman AD, Van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. Nat Genet 31:69–73
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

Platini T, Jia T, Kulkarni RV (2011) Regulation by small rnas via coupled degradation: mean-field and variational approaches. Phys Rev E 84:021928
Popovic N, Marr C, Swain PS (2016) A geometric analysis of fast-slow models for stochastic gene expression. J Math Biol 72:87–122
Roberts E, Be'er S, Bohrer C, Sharma R, Assaf M (2015) Dynamics of simple gene-network motifs subject to extrinsic fluctuations. Phys Rev E 92:062717
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 8:546–555
Schuss Z (2009) Theory and applications of stochastic processes: an analytical approach. Springer Science & Business Media, Berlin
Scott M, Hwa T, Ingalls B (2007) Deterministic characterization of stochastic genetic circuits. Proc 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 Trans 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):062119
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. Proc Natl Acad Sci USA 98:151588598
Wang J, Lefranc M, Thommen Q (2014) Stochastic oscillations induced by intrinsic fluctuations in a self-repressing gene. Biophys J 107(10):2403–2416
Yang X, Wu Y, Yuan Z (2017) Characteristics of mrna dynamics in a multi-on model of stochastic transcription with regulation. Chin. J Phys 55:508–518
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