Limits of noise for autoregulated gene expression

Peter Czuppon[1]   Peter Pfaffelhuber[1]

January 10, 2017

Abstract

Gene expression is influenced by extrinsic noise (involving a fluctuating environment of cellular processes) and intrinsic noise (referring to fluctuations within a cell under constant environment). We study the standard model of gene expression including an (in-)active gene, mRNA and protein. Gene expression is regulated in the sense that the protein feeds back and either represses (negative feedback) or enhances (positive feedback) its production at the stage of transcription. While it is well-known that negative (positive) feedback reduces (increases) intrinsic noise, we give a precise result on the resulting fluctuations in protein numbers. The technique we use is an extension of the Langevin approximation and is an application of a central limit theorem under stochastic averaging for Markov jump processes (Kang, Kurtz and Popovic, 2014). We find that (under our scaling and in equilibrium), negative feedback leads to a reduction in the Fano factor of at most 2, while the noise under positive feedback is potentially unbounded. The fit with simulations is very good and improves on known approximations.

Introduction

It is now widely accepted that gene expression is a stochastic process. The reason is that a single cell is a system with only one or two copies of each gene and of the order tens for mRNA molecules [1, 2, 3]. Experimentally, this stochasticity can even be observed directly by single-cell measurements such as flow cytometry and fluorescence microscopy, which show the inherent fluctuations of protein numbers arising from cell to cell [4].

Usually, noise in gene expression is divided into an intrinsic and an extrinsic part [1, 5]. While the intrinsic part leads to variation of protein numbers from cell to cell in the same environment, the extrinsic part is attributed to the different environmental conditions of the cell. In practice, ensemble averages eliminate intrinsic noise, while single-cell measurements can be thought of having a constant environment, thus eliminating extrinsic noise [6, 7].

Stochasticity in gene expression is not only interesting per se. Today, its role in evolution, development and cell fate decisions is under discussion [8, 9, 10, 11, 12]. In general, noise should be detrimental to cells, since they have to function constantly. Therefore, mechanisms reducing the level of noise are beneficial for real systems.

Under the central dogma of molecular biology, modeling stochasticity of gene expression is straight-forward (see [13] for a review). A gene, which is either turned *on* or *off*, is transcribed into mRNA, which is translated into protein. Both, mRNA and protein are degraded at constant rates. Since the resulting chemical reaction network is linear, the master equation can be solved and all moments can be derived analytically. Most interestingly, the variance can be decomposed into the effects of switching the gene *on* and *off*, noise due to the finite life-time of mRNA, and random fluctuations in the production of protein [13]. It is often stated that gene expression tends

---

[1] Abteilung fur Mathematische Stochastik, Eckerstr. 1, D - 79104 Freiburg, Germany
Keywords and phrases. Intrinsic noise, Langevin approximation, Quasi-steady-state assumption

41 to occur in bursts, which occur due to the short life-time of the *on*-state of the gene and due to
42 the short life-time of mRNA [14].

43   We are interested in the effect of self-regulation on gene expression noise. It is known that a
44 negative feedback loop, i.e. a protein suppressing its own transcription (or translation) leads to a
45 reduced noise, while positive feedback is attributed to increase in noise ([15, 16]). Although these
46 findings are wide-spread, a complete mathematical analysis is lacking. At least, for negative feed-
47 back, [17] and in more generality [18] quantify the effect of negative feedback using a linearization
48 argument. The latter paper further analyzes different feedback models differing between transla-
49 tional and transcriptional autoregulation. Moreover, [19] derive the equilibrium distribution using
50 a multi-scale approach under negative feedback.
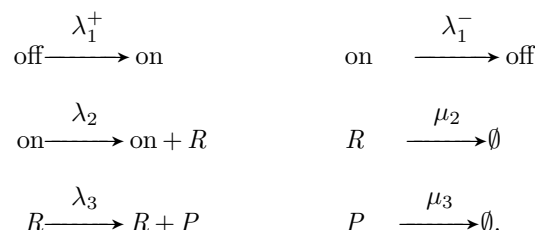
51   Most analyses of noise in unregulated gene expression rely on the master equation (e.g. [13]).
52 By the linearity of this equation, a solution can be given explicitly. Using the approximation that
53 the gene is constantly transcribed to mRNA, this linearity can as well be used under negative
54 feedback [17, 18]. Our approach differs in two ways. First, we are using martingale methods
55 from stochastic analysis in order to describe the chemical system [20]. Second, we can relax the
56 assumption that the gene is transcribed constantly, and therefore derive a more general result.
57 Consequently, we are able to analyze noise in a truly non-linear system under a quasi-steady-state
58 assumption.

59

60   While the full model of regulated gene expression (or any other chemical reaction network) is
61 usually hard to study, considering an ODE approach instead, which approximate the full model,
62 leads to new insights. Formally, a law of large numbers – usually referred to as a fluid limit –
63 can be obtained connecting the stochastic and deterministic model [21, 22]. While such a law of
64 large numbers gives a deterministic limit, fluctuations are studied using central limit results; see
65 [23]. The special situation for gene expression is that the gene and mRNA only have a few copies,
66 while the protein is often in large abundance. Such multi-scale models are often studied under
67 a quasi-steady-state assumption [24]. Here, the species in low abundance are assumed to evolve
68 fast, such that the slow, abundant, species only sense their time-average. For such a stochastic
69 averaging, not only a law of large numbers is given e.g. by [25], but also a central limit result has
70 recently be obtained by [26].

71   While a multi-scale approach to stochastic gene expression is not new (see [27, 19]), the analysis
72 of fluctuations for such systems is not finished yet. In a diffusion setting, fluctuations for a multi-
73 scale system were computed using a Poisson-equation [28, 29]. The results by [26] are similar
74 but are based on Markov jump processes. As our results show, fluctuations take into account all
75 sources of noise and we give explicit formulas for the reduction of noise under negative feedback
76 and the increase in noise under positive feedback.
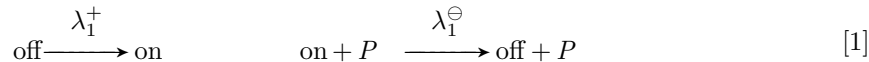
77   1   The model

78 We are dealing with the standard model of gene expression without and with feedback; see e.g.
79 [19]. (All details are found in the SI.) Using the terminology from [13], we write for the model
80 without feedback (or the *neutral* model)

$$\text{off} \xrightarrow{\lambda_1^+} \text{on} \qquad\qquad \text{on} \xrightarrow{\lambda_1^-} \text{off}$$

$$\text{on} \xrightarrow{\lambda_2} \text{on} + R \qquad\qquad R \xrightarrow{\mu_2} \emptyset$$

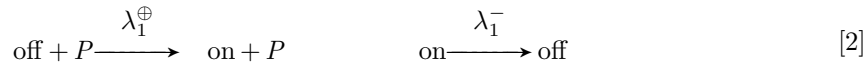$$R \xrightarrow{\lambda_3} R + P \qquad\qquad P \xrightarrow{\mu_3} \emptyset.$$

81   Here, off and on refer to the inactive and the active gene, respectively. mRNA is denoted by
82 R, and the protein by P. Exchanging the first line by

$$\text{off} \xrightarrow{\lambda_1^+} \text{on} \qquad\qquad \text{on} + P \xrightarrow{\lambda_1^{\ominus}} \text{off} + P \qquad\qquad [1]$$

then models a negative feedback while

$$\text{off} + P \xrightarrow{\lambda_1^{\oplus}} \text{on} + P \qquad\qquad \text{on} \xrightarrow{\lambda_1^-} \text{off} \qquad\qquad [2]$$

models a positive feedback. (Note that our results will not change if the protein binds to the gene in the sense that $\text{on} + P \to \text{off}$ or $\text{off} + P \to \text{on}$.) We will refer to these three full models by [*neu*], [*neg*] and [*pos*], respectively. We write $R(t)$ and $P(t)$ for the number of mRNA and protein in the system at time t.

In the sequel, we are interested in the situation where the gene and RNA has $O(1)$ copies, while the protein is much more abundant (i.e. $O(N)$ for some large N). For the parameters in [*neu*], this means that $\mu_3 = O(1), \lambda_1^+, \lambda_1^-, \lambda_2, \mu_2, \lambda_3 \in O(N)$. For [*neg*], we have in contrast that $\lambda_1^{\ominus} = O(1)$ since then $\lambda_1^{\ominus} P = O(N)$, while for [*pos*] we have $\lambda_1^{\oplus} = O(1)$ since then $\lambda_1^{\oplus} P = O(N)$. For this scaling, on, off and R are *fast* molecular species, and quickly equilibrate. Using a quasi-steady-state assumption, we see that, approximately, (see SI, Theorem 1),

$$\frac{d}{dt}\frac{P}{N} = F\Big(\frac{P}{N}\Big)$$

with

$$NF(p/N) = \begin{cases} \dfrac{M\lambda_1^+\lambda_2\lambda_3}{\mu_2(\lambda_1^- + \lambda_1^+)} - \mu_3 p & \text{for } [neu], \\[3mm] \dfrac{M\lambda_1^+\lambda_2\lambda_3}{\mu_2(\lambda_1^{\ominus} p + \lambda_1^+)} - \mu_3 p & \text{for } [neg], \\[3mm] \dfrac{M\lambda_1^{\oplus}\lambda_2\lambda_3 p}{\mu_2(\lambda_1^- + \lambda_1^{\oplus} p)} - \mu_3 p & \text{for } [pos]. \end{cases} \qquad [3]$$

Our goal is derive the variance in protein numbers under [*neu*], [*neg*] and [*pos*]. While [*neu*] is solved explicitly elsewhere, e.g. in [13], some approximations have to be made for [*neg*] and [*pos*]. A first idea is to use a Langevin approximation and write

$$P(t) \approx \int_0^t \lambda_3 R(s) - \mu_3 P(s)ds + \int_0^t \sqrt{\lambda_3 R(s) + \mu_3 P(s)}dW_s$$

and v_P for the exact solution of $\dot{v}_P = F(v_P)$. Then, for $U = \sqrt{N}(P/N - v_P)$ and $b(v_P) = \lambda_3 \mathbb{E}_\pi[R] + \mu_3 v_P$, the average diffusion for fixed v_P, i.e.

$$Nb(v_P) = \begin{cases} \dfrac{M\lambda_1^+\lambda_2\lambda_3}{\mu_2(\lambda_1^- + \lambda_1^+)} + \mu_3 N v_P & \text{for } [neu], \\[3mm] \dfrac{M\lambda_1^+\lambda_2\lambda_3}{\mu_2(\lambda_1^{\ominus} N v_P + \lambda_1^+)} + \mu_3 N v_P & \text{for } [neg], \\[3mm] \dfrac{M\lambda_1^{\oplus}\lambda_2\lambda_3 N v_P}{\mu_2(\lambda_1^- + \lambda_1^{\oplus} N v_P)} + \mu_3 N v_P & \text{for } [pos], \end{cases}$$

we find

$$d\frac{P}{N} = dv_P + \frac{1}{\sqrt{N}}dU \approx \Big(F(v_P) + \frac{1}{\sqrt{N}}F'(v_P)U\Big)dt + \sqrt{\frac{1}{N}b(v_P)}dW. \qquad [4]$$

101 This approach builds on applying a quasi-steady-state approach whenever possible, i.e. when
102 averaging over the on/off-state of genes in order to derive F, and the number of RNA, which is
103 approximated by its mean in order to derive b. Consequently, fluctuations arising from these two
104 mechanisms cannot be accounted for in the resulting variance. In the next section, we give our
105 results which take all reactions into account for deriving noise in gene expression.

106 2 Results

107 Approximate variance and Fano factor for P

108 As an application of [26], we derive in the SI (Theorem 2) the following central limit result:
109 Let v_P be the exact solution of $\dot{v}_P = F(v_P)$. The deviations of P/N from v_P, given through
110 $U = \sqrt{N}(P/N - v_P)$, satisfy for large N the SDE

$$U(t) = U(0) + \int_0^t F'(v_P(s))U(s)ds + \int_0^t \sqrt{c(v_P(s))}dW(s),$$

111 with W a one-dimensional standard Brownian motion and

$$Nc(v_P) = \begin{cases} \frac{M\lambda_1^+\lambda_2\lambda_3}{\mu_2(\lambda_1^-+\lambda_1^+)}\left(\frac{2\lambda_1^-\lambda_2\lambda_3}{\mu_2(\lambda_1^-+\lambda_1^+)^2} + \frac{2\lambda_3}{\mu_2} + 1\right) + \mu_3 N v_P \\ \qquad\qquad\qquad\qquad\qquad \text{for } [neu], \\[2em] \frac{M\lambda_1^+\lambda_2\lambda_3}{\mu_2(\lambda_1^\ominus N v_P+\lambda_1^+)}\left(\frac{2\lambda_1^\ominus N p\lambda_2\lambda_3}{\mu_2(\lambda_1^\ominus N p+\lambda_1^+)^2} + \frac{2\lambda_3}{\mu_2} + 1\right) + \mu_3 N v_P \\ \qquad\qquad\qquad\qquad\qquad \text{for } [neg], \\[2em] \frac{M\lambda_1^\oplus\lambda_2\lambda_3 N v_P}{\mu_2(\lambda_1^-+\lambda_1^\oplus N v_P)}\left(\frac{2\lambda_1^-\lambda_2\lambda_3}{\mu_2(\lambda_1^-+\lambda_1^\oplus N v_P)^2} + \frac{2\lambda_3}{\mu_2} + 1\right) + \mu_3 N v_P \\ \qquad\qquad\qquad\qquad\qquad \text{for } [pos]. \end{cases}$$

112 Solving for P, we obtain, approximately (compare with [4])

$$d\frac{P}{N} = dv_P + \frac{1}{\sqrt{N}}dU = \left(F(v_P) + \frac{1}{\sqrt{N}}F'(v_P)U\right)dt + \sqrt{\frac{1}{N}c(v_P)}dW$$

113 While this limit provides a dynamical result along paths of P, we can also use this approximation
114 and study the process in equilibrium by setting $P(0) = v_P^* N$, where (see SI, Section B)

$$v_P^* = \begin{cases} \frac{M\lambda_1^+\lambda_2\lambda_3}{\mu_2\mu_3(\lambda_1^-+\lambda_1^+)} & \text{for } [neu], \\[1.5em] \frac{\lambda_1^+}{2\lambda_1^\ominus}\left(\sqrt{1 + \frac{4M\lambda_1^\ominus\lambda_2\lambda_3}{\lambda_1^+\mu_2\mu_3}} - 1\right) & \text{for } [neg], \\[1.5em] 0 \vee \frac{M\lambda_1^\oplus\lambda_2\lambda_3 - \lambda_1^-\mu_2\mu_3}{\lambda_1^\oplus\mu_2\mu_3} & \text{for } [pos]. \end{cases} \qquad [5]$$

115 is the unique solution of $F(v_P) = 0$ (and hence $\dot{v}_P = 0$ for $v_P(0) = v_P^*$). We obtain in this case

$$d\frac{P}{N} = F'(v_P^*)\left(\frac{P}{N} - v_P^*\right)dt + \sqrt{\frac{1}{N}c(v_P^*)}dW.$$

116 This is an Ornstein-Uhlenbeck process and it is well-known that its equilibrium has the Fano factor

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} = \frac{N\mathbb{V}[P/N]}{\mathbb{E}[P/N]} = N\frac{\frac{1}{N}c(v_P^*)}{-2F'(v_P^*)v_P^*} = -\frac{c(v_P^*)}{2F'(v_P^*)v_P^*}. \tag{6}$$

117 Since no factor N appears on the right hand side, some authors call the Fano factor dimensionless.
118 Empirically, it was found e.g. by [30], that for all classes of genes and under all conditions, the
119 expression variance was approximately proportional to the mean, which is again reminiscent of
120 the lacking N in the Fano factor above. Plugging in the quantities from above, we find with

$$\alpha_{neg} = \frac{\lambda_1^{\ominus} N v_P^*}{\lambda_1^{\ominus} N v_P^* + \lambda_1^+}, \quad \alpha_{pos} = \frac{\lambda_1^-}{\lambda_1^- + \lambda_1^{\oplus} N v_P^*}$$

121 that (see SI, Remark B.1)

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} = \begin{cases} \frac{\lambda_1^- \lambda_2 \lambda_3}{\mu_2(\lambda_1^- + \lambda_1^+)^2} + \frac{\lambda_3}{\mu_2} + 1 & \text{for } [neu], \\[2ex] \left(\frac{\lambda_1^{\ominus} N v_P^* \lambda_2 \lambda_3}{\mu_2(\lambda_1^{\ominus} N v_P^* + \lambda_1^+)^2} + \frac{\lambda_3}{\mu_2} + 1\right) \Big/ (1 + \alpha_{neg}) & \text{for } [neg], \\[2ex] \left(\frac{\lambda_1^- \lambda_2 \lambda_3}{\mu_2(\lambda_1^- + \lambda_1^{\oplus} N v_P^*)^2} + \frac{\lambda_3}{\mu_2} + 1\right) \Big/ (1 - \alpha_{pos}) & \text{for } [pos]. \end{cases} \tag{7}$$

Refining the Fano factor

123 Using a slightly different model where $\mu_2 = O(1)$ (hence $R = O(N)$) and $\lambda_3 = O(1)$ (hence still
124 $P = O(N)$), it is possible to derive a refined formula for the Fano factor. In SI, Section C, we
125 derive (see (SI.13), (SI.15) and (SI.16)) that

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} = \begin{cases} \frac{\lambda_1^- \lambda_2 \lambda_3}{(\lambda_1^- + \lambda_1^+)^2(\mu_2 + \mu_3)} + \frac{\lambda_3}{\mu_2 + \mu_3} + 1 & \text{for } [neu], \\[2ex] \left(\alpha_{neg}\frac{\mu_3}{\mu_2 + \mu_3} + \frac{\alpha_{neg}^2 \lambda_2 \lambda_3}{\lambda_1^{\ominus} N v_P^*(\mu_2 + \mu_3)} + \frac{\lambda_3}{\mu_2 + \mu_3} + 1\right) \Big/ (1 + \alpha_{neg}) & \text{for } [neg], \\[2ex] \left(-\alpha_{pos}\frac{\mu_3}{\mu_2 + \mu_3} + \frac{\alpha_{pos}^2 \lambda_2 \lambda_3}{\lambda_1^-(\mu_2 + \mu_3)} + \frac{\lambda_3}{\mu_2 + \mu_3} + 1\right) \Big/ (1 - \alpha_{pos}) & \text{for } [pos]. \end{cases} \tag{8}$$

126 Note that this equation approximately gives [7] for $\mu_2, \lambda_3 \gg 1$. In practice (and in our simulations
127 below), the life-time of proteins is much larger than the life-time of mRNA, such that [8] doesn't
128 produce a better fit than [7]. Therefore, we will use [7] in the sequel.

Comparing the noise in $[neu]$, $[neg]$ and $[pos]$

130 It is frequently reported that a negative feedback in gene expression results in a reduced variance
131 (noise) of protein levels, whereas a positive feedback enhances noise. These observations can be
132 made precise by our results from above. Here, we report some consequences on the equilibrium
133 variance and the Fano factor, $\mathbb{V}[P]/\mathbb{E}[P]$. For a fair comparison, we use the models $[neu]$, $[neg]$ and
134 $[pos]$ for equal values of v_P^*. More specifically, we use parameters which satisfy $\lambda_1^- = \lambda_1^{\ominus} v_P^*$ for
135 $[neg]$ and $\lambda_1^+ = \lambda_1^{\oplus} v_P^*$ for $[pos]$. Then, from [7], we find that (see SI, Section B.2)

$$\begin{aligned} \frac{\mathbb{V}_{neg}[P]/\mathbb{E}_{neg}[P]}{\mathbb{V}_{neu}[P]/\mathbb{E}_{neu}[P]} &= \frac{\mathbb{V}_{neg}[P]}{\mathbb{V}_{neu}[P]} = \left(1 + \frac{\lambda_1^{\ominus} N v_P^*}{\lambda_1^{\ominus} N v_P^* + \lambda_1^+}\right)^{-1}, \\[2ex] \frac{\mathbb{V}_{pos}[P]/\mathbb{E}_{pos}[P]}{\mathbb{V}_{neu}[P]/\mathbb{E}_{neu}[P]} &= \frac{\mathbb{V}_{pos}[P]}{\mathbb{V}_{neu}[P]} = \left(1 - \frac{\lambda_1^-}{\lambda_1^- + \lambda_1^+ N v_P^*}\right)^{-1}. \end{aligned} \tag{9}$$
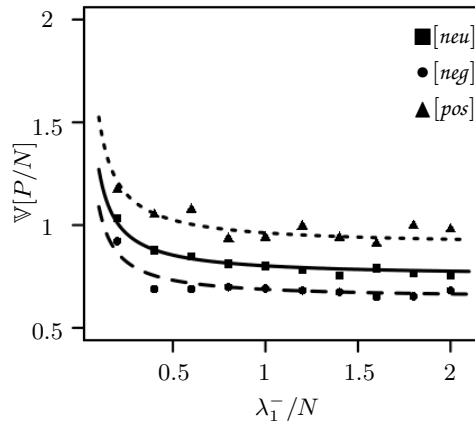
Figure 1: Simulations and theoretical results with a fixed mean of 1250 proteins. The gene association and dissociation rates are varied, i.e. $\lambda_1^- = (20, 40, \ldots, 200)$ and λ_1^+ is chosen such that the protein mean equals 1250 in each case. Furthermore, these rates are adjusted in the cases of negative and positive feedback according to the description right before [9] for the different cases. The dissociation rate λ_1^- is plotted on the x-axis whereas the y-axis represents the variance in protein numbers. The solid, dotted, dashed line are the neutral, positive and negative case, respectively. The other parameter are given by $M = 1, N = 100, \lambda_2 = 300, \mu_2 = 100, \lambda_3 = 500$ and $\mu_3 = 1$. Each data point is derived from 1000 Monte-Carlos simulations (cf. [31]) of the full system given by $[neu], [neg]$ and $[pos]$.

In particular, we see that the variance is reduced in $[neg]$ and increased in $[pos]$, as expected. Furthermore, we performed simulations and compared them to our predictions. As can be seen in Figure 1 the simulated results for relatively small N, i.e. small numbers of proteins, fit the predictions quite well. Of course this fit improves by increasing values of N.

Comparison to previous results

Here, we compare our results in the neutral case with [13], and with [17], [19] and [18] in the case of negative feedback.

The neutral case: Using the approximation $\mu_3 \ll \mu_1, \mu_2$, i.e. the protein life-time is much longer than the duration of constant gene activity and RNA life, we see in the SI, Section D.1, that (4) from [13] and [7] produce the same results.

Negative Feedback, [17]: For $[neg]$, a linearization was studied in [17] for the case of fast switching on and off of the gene. (This will mean that both rates for switching on and off the gene are large; see SI, Section D.2.) Using the further approximation that the gene is switched on most of the time,

$$\lambda_1^\ominus N v_P^* \ll \lambda_1^+, \qquad [10]$$

they find the approximation

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} = 1 + \frac{\lambda_3}{\mu_2 + \mu_3} \frac{1 - \frac{M\lambda_1^\ominus \lambda_2}{\lambda_1^+ \mu_3}}{1 + \frac{M\lambda_1^\ominus \lambda_2 \lambda_3}{\lambda_1^+ \mu_2 \mu_3}}. \qquad [11]$$

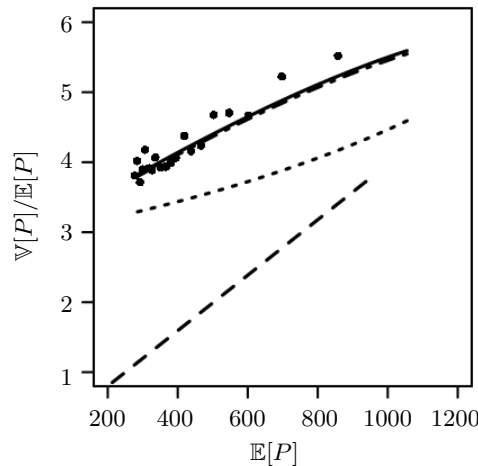Using that $\mu_3 \ll \mu_2$ and [10], we can rearrange this to give

Figure 2: Simulations and theoretical results of gene expression with negative feedback. The mean given on the x-axis is varied and plotted against the Fano factor on the y-axis. The solid line represents [7], the dash-dotted line the result in [8], the dashed line the result from [17] given in [11] and the dotted line the Fano factor calculated in [18] given in [12]. The parameter for the simulations are chosen as follows: $M = 1, N = 100, \lambda_1^+ = 250, \lambda_1^- = (0.2, 0.4, \ldots, 4), \lambda_2 = 300, \mu_2 = 100, \lambda_3 = 500, \mu_3 = 1$. The bullets represent the estimated Fano factors of the full system $[neg]$ obtained from 1000 Monte-Carlo simulations (cf. [31]) for each value of λ_1^-.

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} \approx \left(1 + \frac{\lambda_3}{\mu_2}\right)\left(1 - \frac{\lambda_1^\ominus N v_P^*}{\lambda_1^+}\right).$$

Since fluctuations of switching the gene on and off are not accounted for in this calculation, we therefore find that, in this approximation,

$$\frac{\mathbb{V}_{neg}[P]}{\mathbb{V}_{neu}[P]} \approx \left(1 - \frac{\lambda_1^\ominus N v_P^*}{\lambda_1^+}\right),$$

which approximately equals [9]. Therefore, the results of [17] are similar to our results, but less accurate if gene (in-)activation is slower or if [10] is not satisfied; see also Figure 2.

Negative Feedback, [18]: As explained around [4], the usual Langevin approximation cannot account for all fluctuations when a quasi-steady-state assumption is made. (Precisely, it cannot account for fluctuations in the averaged variables.) In [18], a Langevin approach is carried out in order to analyze fluctuations in autoregulatory gene expression in the cases of transcriptional and translational feedback. The author considers the mRNA and the protein to evolve on the same time-scale whereas the gene (or DNA) is considered to be on a faster time-scale. For transcriptional feedback (which we study here), he obtains in his (5) – see SI, Section D.3 for the transformation of their results into our parameters

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} = \left(\alpha_{neg}\frac{\mu_3}{\mu_2 + \mu_3} + \frac{\lambda_3}{\mu_2 + \mu_3} + 1\right)\Big/\left(1 + \alpha_{neg}\right). \qquad [12]$$

This corresponds exactly to [8] with a missing term in the numerator. This term arises from fluctuations in gene activation, which was averaged out in calculations done [18]. At least, [12] arises from [8] if we assume that $\lambda_1^\ominus N v_P^*, \lambda_1^+ \gg 1$.

166   *Negative Feedback, [19]:* Our result for [*neg*] has Theorem 2 of [19] as a limit result. They
167   study a similar model (with $M = 1$), but with a scaling such that $\lambda_3, \mu_3 = O(1)$ and $\lambda_1^\ominus = O(N)$,
168   leading to low (i.e. $O(1)$) abundance of protein in the system. The resulting birth-death process
169   for P has a stationary distribution which they compute explicitly. Moreover, setting

$$\rho = \frac{\lambda_1^+ \lambda_2 \lambda_3}{\lambda_1^\ominus \mu_2 \mu_3},$$

they obtain in their Corollary 4.1 that, in equilibrium,

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} \overset{\rho \to \infty}{\approx} \frac{1}{2} \tag{13}$$

170   For large ρ, it is $v_P^* \approx \sqrt{\rho}$, and [7] gives the same limit.

Simpler model of gene expression

172   In the literature simpler models of gene expression are studied as well. Here, only two molecular
173   species are involved. Either, the gene is constitutively expressed, therefore ignoring the state of
174   the gene, or translation is neglected and the gene is assumed to be transcribed leading to protein
175   in one step [32, 33, 34]. In either case, we have the model

$$\text{off} \xrightarrow{\lambda_1^+} \text{on} \qquad\qquad \text{on} \xrightarrow{\lambda_1^-} \text{off}$$

$$\text{on} \xrightarrow{\lambda_3} \text{on} + P \qquad\qquad P \xrightarrow{\mu_3} \emptyset.$$

176   for [*neu*], whereas for [*neg*] and [*pos*], we take [1] and [2] instead of the first line, respectively. We
177   note that this model arises from the full model described above (see SI, Section D.4), when letting
178   $\mu_2 = \lambda_2 \to \infty$. Hence, we obtain for the simpler model the approximation

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} = \begin{cases} \dfrac{\lambda_1^- \lambda_3}{(\lambda_1^- + \lambda_1^+)^2} + 1 & \text{for } [neu], \\[2ex] \left(1 + \dfrac{\lambda_1^\ominus N v_P^*}{\lambda_1^\ominus N v_P^* + \lambda_1^+}\right)^{-1} \left(\dfrac{\lambda_1^\ominus N v_P^* \lambda_3}{(\lambda_1^\ominus N v_P^* + \lambda_1^+)^2} + 1\right) & \text{for } [neg], \\[2ex] \left(1 - \dfrac{\lambda_1^-}{\lambda_1^- + \lambda_1^\oplus N v_P^*}\right)^{-1} \left(\dfrac{\lambda_1^- \lambda_3}{(\lambda_1^- + \lambda_1^\oplus N v_P^*)^2} + 1\right) & \text{for } [pos]. \end{cases} \tag{14}$$

3   Discussion

180   Quantifying noise in gene expression is essential for understanding regulatory networks in cells [17].
181   Our results give the most complete theory on the intrinsic noise available today. While negative
182   feedback is known to reduce noise under negative feedback, we improve on the quantification of this
183   effect. Moreover, we can provide the same quantification also for positive feedback, where noise is
184   increased. In particular, [9] shows that the average time the gene is *off* determines the reduction
185   of noise in all cases relative to unregulated genes; see also [35]. Both, for negative and positive
186   feedback, the change in noise is maximal if the gene is *off* most of the time, while still having the
187   same amount of protein as in the unregulated (neutral) case. This finding is reminiscent of the fact
188   that gene expression comes in bursts. The burstiness is most extreme if the gene is *on* only for a
189   short time, producing a large amount of mRNA, and afterwards *off* for a long period. Interestingly,
190   previous approaches only gave approximations for noise for negative feedback if switching the gene
191   on and off is very fast [17, 18] and if the gene is *on* most of the time [17]. Hence, all previous
192   papers could not have seen the effects of gene (in-)activation on protein noise. As in previous

results [19], we find that in the limit where the gene is *off* most of the time, the negative feedback reduces noise at most by a factor of two. In addition, noise can increase unboundedly for positive feedback.

Today, quasi-steady-state assumptions are frequently used when analyzing chemical reaction networks. While the intuition suggests the correct approach when approximating the system by a deterministic path, studying fluctuations is apparently much less obvious. In [36], some special cases are studied when a straight-forward approximation of the fluctuations work. In our analysis, we use a new approach by [26] and can also interpret all terms arising in [7]. E.g., for negative feedback, we find – as in the neutral case – contributions from randomness in translation and transcription by

$$\frac{\mathbb{V}_{neg}[P]}{\mathbb{E}_{neg}[P]} = \Big(\underbrace{\frac{\lambda_1^{\ominus} N v_P^* \lambda_2 \lambda_3}{\mu_2 (\lambda_1^{\ominus} N v_P^* + \lambda_1^+)^2}}_{\substack{\text{effect of RNA noise,}\\ \text{originating from gene}\\ \text{(in-)activation}}} + \underbrace{\frac{\lambda_3}{\mu_2}}_{\substack{\text{effect of individ-}\\ \text{ual RNA noise}}} + \underbrace{1}_{\substack{\text{protein}\\ \text{births}\\ \text{and}\\ \text{deaths}}} \Big) \Big/ \underbrace{\Big(1 + \frac{\lambda_1^{\ominus} N v_P^*}{\lambda_1^{\ominus} N v_P^* + \lambda_1^+} \Big)}_{\substack{\text{fast back pushing due to}\\ \text{negative feedback}}}. \qquad [15]$$

Moreover, the negative feedback pushes the amount of protein faster back to its equilibrium value for a burst of gene expression. This results in the denominator in [15], which has the biggest< effect of the noise-reducing effect of negative feedback. In addition, another source of noise comes from switching the gene on and off (term $\lambda_1^{\ominus} N v_P^* \lambda_2 \lambda_3 / (\mu_2 (\lambda_1^{\ominus} N v_P^* + \lambda_1^+)^2)$). It is due to the latter term that the fit of simulations and theory (see e.g. Figure 2) is excellent. Previous studies have averaged out this source of noise, and only the recent approach of [26] reveals the impact of switching gene on and off on the noise in protein numbers.

In their paper, [26] gave as an example an approximation of noise for Michaelis-Menten kinetics and a model for virus infection. Their method relies mostly on solving a Poisson equation $L_2 h = F_n - F$, where L_2 is the generator of the fast subsystem (gene and RNA in our example), F_N and F describe the evolution of the slow system (protein) including all fluctuations and in the limit using the quasi-steady-state assumption, respectively. We stress that this approach is not only useful for equilibrium situations, but also for understanding noise if the slow system has not reached equilibrium yet, e.g. after a cell split.

It was argued that complexity of gene regulatory networks leads to a reduction in the level of noise, while certain network motifs always lead to increased levels of noise [37, 38]. Experimentally, gene expression noise can be used to understand the dynamics of gene regulation [39]. Our analysis should provide an approach for distinguishing between different models of gene regulation based on measurements of noise levels.

Competing Interests

We have no competing interests.

Authors' contributions

PC and PP conducted the research and wrote the manuscript.

Acknowledgements

We thank Jens Timmer, Freiburg, for discussion.

References

[1] Swain PS, Elowitz MB, Siggia ED. Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc Natl Acad Sci USA. 2002;99(20):12795–12800.

[2] Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. Science. 2002;297(5584):1183–1186.

[3] Raj A, A van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell. 2008;135(2):216–226.

[4] Li GW, Xie XS. Central dogma at the single-molecule level in living cells. Nature. 2011;475:308–315.

[5] Raser JM, O'Shea EK. Noise in gene expression: origins, consequences, and control. Science. 2005;309(5743):2010–2013.

[6] Singh A, Soltani M. Quantifying intrinsic and extrinsic variability in stochastic gene expression models. PloS one. 2013;8(12):e84301.

[7] Singh A. Transient changes in intercellular protein variability identify sources of noise in gene expression. Biophysical journal. 2014;107(9):2214–2220.

[8] Kaern M, Elston TC, Blake WJ, Collins JJ. Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet. 2005;6(6):451–464.

[9] Maamar H, Raj A, Dubnau D. Noise in gene expression determines cell fate in Bacillus subtilis. Science. 2007;317(5837):526–529.

[10] Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. Nature. 2010;467(7312):167–173.

[11] Silva-Rocha R, de Lorenzo V. Noise and robustness in prokaryotic regulatory networks. Annual review of microbiology. 2010;64:257–275.

[12] Wang Z, Zhang J. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. Proc Natl Acad Sci. 2011;108(16):E67–E76.

[13] Paulsson J. Models of stochastic gene expression. Physics of Life Review. 2005;2:157–175.

[14] Kumar N, Singh A, Kulkarni RV. Transcriptional Bursting in Gene Expression: Analytical Results for General Stochastic Models. PLoS computational biology. 2015;11(10):e1004292.

[15] Lestas I, Vinnicombe G, Paulsson J. Fundamental limits on the suppression of molecular fluctuations. Nature. 2010;467(7312):174–178.

[16] Hornung G, Barkai N. Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. PLoS computational biology. 2008;4(1):e8.

[17] Thattai M, van Oudenaarden A. Intrinsic noise in gene regulatory networks. Proc Nat Acad Sci. 2001;98(15):8614–8619.

[18] Swain PS. Efficient Attenuation of Stochasticity in Gene Expression Through Post-transcriptional Control. Journal of Molecular Biology. 2004;344(4):965–976.

[19] Dessalles R, Fromion V, Robert P. A Stochastic Analysis of Autoregulation of Gene Expression. Preprint. 2016;p. 1–27.

[20] Ethier SN, Kurtz TG. Markov processes: characterization and convergence. Wiley series in probability and mathematical statistics. New York, Chichester: J. Wiley & Sons; 1986.

[21] Kurtz T. Solutions of ordinary differential equations as limits of pure jump Markov processes. Journal of Applied Probability. 1970;7:49–58.

[22] Darling RWR. Fluid Limits of Pure Jump Markov Processes: a Practical Guide. arxiv preprint. 2002;https://arxiv.org/abs/math/0210109:1–16.

[23] Kurtz T. Limit Theorems for sequences of jump Markov processes approximating ordinary differential processes. Journal of Applied Probability. 1970;8:344–356.

[24] Seegel LA, Slemrod M. The Quasi-Steady-State Assumption: a Case Study in Perturbation. SIAM Review. 1989;31(3):446–477.

[25] Ball K, Kurtz TG, Popovic L, Rempala G. Asymptotic analysis of multiscale approximations to reaction networks. Ann Appl Probab. 2006;16(4):1925–1961.

[26] Kang HW, Kurtz T, Popovic L. Central Limit Theorems and Diffusion Approximations for Multiscale Markov Chain Models. Ann Appl Probab. 2014;24:721–759.

[27] Bokes P, King JR, Wood A, Loose M. Multiscale stochastic modelling of gene expression. J Math Biol. 2012;65(3):493–520.

[28] Pardoux E, Veretennikov AY. On Poisson equation and diffusion approximation 1. Ann Probab. 2001;29:1061–1085.

[29] Pardoux E, Veretennikov AY. On Poisson equation and diffusion approximation 2. Ann Prob. 2003;31:1166–1192.

[30] Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y, et al. Noise in protein expression scales with natural protein abundance. Nature genetics. 2006;38(6):636–643.

[31] Gillespie D. Exact stochastic simulation of coupled chemical reactions. J Phys Chem. 1977;81:2340–2361.

[32] Hornos JEM, Schultz D, Innocentini G, Wang J, Walczak A, Onuchic J, et al. Self-regulating gene: an exact solution. Physical review E, Statistical, nonlinear, and soft matter physics. 2005;72(5 Pt 1):051907.

[33] Shahrezaei V, Swain PS. Analytical distributions for stochastic gene expression. Proc Natl Acad Sci USA. 2008;105(45):17256–17261.

[34] Ramos A, Innocentini G, Hornos J. Exact time-dependent solutions for a self-regulating gene. Physical review E, Statistical, nonlinear, and soft matter physics. 2011;83(6 Pt 1):062902.

[35] Grönlund A, Lötstedt P, Elf J. Transcription factor binding kinetics constrain noise suppression via negative feedback. Nature communications. 2013;4:1864.

[36] Kim JK, Josić K, Bennett MR. The relationship between stochastic and deterministic quasi-steady state approximations. BMC systems biology. 2015;9:87.

[37] Becskei A, Serrano L. Engineering stability in gene networks by autoregulation. Nature. 2000;405(6786):590–593.

[38] Cardelli L, Csikász-Nagy A, Dalchau N, Tribastone M, Tschaikowski M. Noise Reduction in Complex Biological Switches. Scientific reports. 2016;6:20214.

[39] Munsky B, Neuert G, van Oudenaarden A. Using gene expression noise to understand gene regulation. Science. 2012;336(6078):183–187.
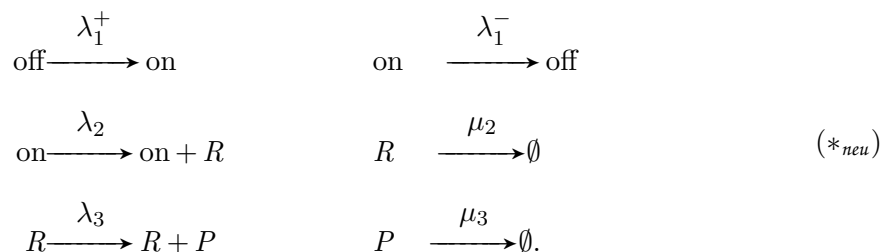
Supporting Information:

Limits of noise for autoregulated gene expression

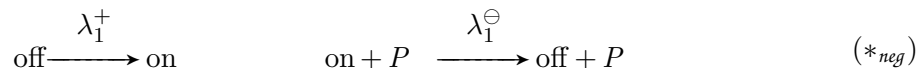by Peter Czuppon and Peter Pfaffelhuber

A   The model

We are dealing with gene expression without and with transcriptional feedback; see e.g. Swain (2004); Dessalles et al. (2016). For general formulations of chemical reaction networks (and their mathematical representations), consult Anderson and Kurtz (2015). Using the terminology from Paulsson (2005), we write for the model without feedback (or the *neutral* model)

$$\text{off} \xrightarrow{\lambda_1^+} \text{on} \qquad\qquad \text{on} \xrightarrow{\lambda_1^-} \text{off}$$

$$\text{on} \xrightarrow{\lambda_2} \text{on} + R \qquad\qquad R \xrightarrow{\mu_2} \emptyset \qquad\qquad (*_{neu})$$

$$R \xrightarrow{\lambda_3} R + P \qquad\qquad P \xrightarrow{\mu_3} \emptyset.$$

Here, *off* and *on* refer to an inactive and an active gene, respectively. The mRNA is given by R, and the protein by P. While the first line of chemical reactions models gene switching from *off* to *on* and back, the second line encodes transcription and degradation of mRNA, while the third line gives translation and degradation of proteins. Exchanging the first line by

$$\text{off} \xrightarrow{\lambda_1^+} \text{on} \qquad\qquad \text{on} + P \xrightarrow{\lambda_1^\ominus} \text{off} + P \qquad\qquad (*_{neg})$$

then models a negative feedback while

$$\text{off} + P \xrightarrow{\lambda_1^\oplus} \text{on} + P \qquad\qquad \text{on} \xrightarrow{\lambda_1^-} \text{off} \qquad\qquad (*_{pos})$$
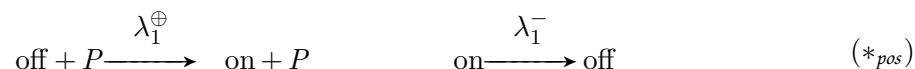
models a positive feedback. In all cases, we number the equations from left to right and from top to bottom by 1–6, so $\mathcal{K} = \{1, ..., 6\}$ is the set of chemical reactions. The species counts are given by $X_{\text{off}}, X_{\text{on}}, X_R$ and X_P for inactive and active gene, mRNA and protein, respectively. We will make use of the following scaling for the abundances of chemical species

$$X_{\text{off}} = O(1), \qquad X_{\text{on}} = O(1), \qquad X_R = O(1), \qquad X_P = O(N),$$

or

$$\alpha_{\text{off}} = \alpha_{\text{on}} = \alpha_R = 0, \qquad \alpha_P = 1,$$

such that $X_i = O(N^{\alpha_i})$ for all i. Reactions are scaled for all models by

$$\lambda_2 = N\kappa_2, \quad \mu_2 = N\nu_2, \quad \lambda_3 = N\kappa_3, \quad \mu_3 = \nu_3.$$

For the neutral model, we also set

$$\lambda_1^+ = N\kappa_1^+, \quad \lambda_1^- = N\kappa_1^-, \qquad (\Box_{neu})$$

whereas for negative feedback

$$\lambda_1^+ = N\kappa_1^+, \quad \lambda_1^\ominus = \kappa_1^\ominus, \qquad (\Box_{neg})$$

and for positive feedback

$$\lambda_1^\oplus = \kappa_1^\oplus, \quad \lambda_1^- = N\kappa_1^-. \qquad (\Box_{pos})$$

(Note that this scaling obeys $\lambda_1^-, \lambda_1^\ominus X_P, \lambda_1^+, \lambda_1^\oplus X_P = O(N)$.) Setting $V_i^N = N^{-\alpha_i} X_i$, and for M copies of the gene, we have in the neutral case

$$
\begin{aligned}
V_{\text{on}}^N(t) &= V_{\text{on}}^N(0) + Y_1\Big(N\int_0^t \kappa_1^+ V_{\text{off}}^N(s)ds\Big) - Y_2\Big(N\int_0^t \kappa_1^- V_{\text{on}}^N(s)ds\Big), \\
V_{\text{off}}^N(t) &= M - V_{\text{on}}^N(t), \\
V_R^N(t) &= V_R^N(0) + Y_3\Big(N\int_0^t \kappa_2 V_{\text{on}}^N(s)ds\Big) - Y_4\Big(N\int_0^t \nu_2 V_R^N(s)ds\Big), \qquad (\bullet_{neu}) \\
V_P^N(t) &= V_P^N(0) + N^{-1}Y_5\Big(N\int_0^t \kappa_3 V_R^N(s)ds\Big) - N^{-1}Y_6\Big(N\int_0^t \nu_3 V_P^N(s)ds\Big),
\end{aligned}
$$

for independent, rate 1 Poisson processes $Y_1, ..., Y_6$. The first equation changes in the case of negative feedback to

$$V_{\text{on}}^N(t) = V_{\text{on}}^N(0) + Y_1\Big(N\int_0^t \kappa_1^+ V_{\text{off}}^N(s)ds\Big) - Y_2\Big(N\int_0^t \kappa_1^\ominus V_{\text{on}}^N(s)V_P^N(s)ds\Big) \qquad (\bullet_{neg})$$

and in the case of positive feedback to

$$V_{\text{on}}^N(t) = V_{\text{on}}^N(0) + Y_1\Big(N\int_0^t \kappa_1^\oplus V_{\text{off}}^N(s)V_P(s)ds\Big) - Y_2\Big(N\int_0^t \kappa_1^- V_{\text{on}}^N(s)ds\Big). \qquad (\bullet_{pos})$$

In the sequel, we will refer to the model without, negative and positive feedback simply as [*neu*], [*neg*] and [*pos*], respectively. We understand all equations $(*_{neu})$, $(\Box_{neu})$, $(\bullet_{neu})$ as the bases for [*neu*], equations $(*_{neg})$, $(\Box_{neg})$, $(\bullet_{neg})$ as the bases for [*neg*] and all equations $(*_{pos})$, $(\Box_{pos})$, $(\bullet_{pos})$ as the bases for [*pos*].

B   Results

B.1   Law of Large Numbers and Central Limit Theorem

The following result can be obtained using a quasi-steady-state assumption. It relies on the method of stochastic averaging; see e.g. Ball et al. (2006).

Theorem 1 (Law of Large Numbers). *We consider the models* [*neu*], [*neg*] *and* [*pos*] *and assume that* $V_P^N(0) \xoverset{N\to\infty}{\Longrightarrow} V_P(0)$. *Then,* $V_P^N \xoverset{N\to\infty}{\Longrightarrow} V_P$, *where* V_P *solves*

$$\dot{V}_P = F(V_P),$$

with (see also [3] in the main text)

$$
F(v_P) = \begin{cases}
\dfrac{M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)} - \nu_3 v_P & \text{for [neu]}, \\[2ex]
\dfrac{M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2(\kappa_1^\ominus v_P + \kappa_1^+)} - \nu_3 v_P & \text{for [neg]}, \\[2ex]
\dfrac{M\kappa_1^\oplus \kappa_2 \kappa_3 v_P}{\nu_2(\kappa_1^- + \kappa_1^\oplus v_P)} - \nu_3 v_P & \text{for [pos]}.
\end{cases}
\tag{SI.1}
$$

In particular, the equilibrium is given by (see also [5] in the main text)

$$
v_P^* = \begin{cases}
\dfrac{M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2 \nu_3(\kappa_1^- + \kappa_1^+)} & \text{for [neu]}, \\[2ex]
\dfrac{\kappa_1^+}{2\kappa_1^\ominus}\left(\sqrt{1 + \dfrac{4M\kappa_1^\ominus \kappa_2 \kappa_3}{\kappa_1^+ \nu_2 \nu_3}} - 1 \right) & \text{for [neg]}, \\[2ex]
0 \vee \dfrac{M\kappa_1^\oplus \kappa_2 \kappa_3 - \kappa_1^- \nu_2 \nu_3}{\kappa_1^\oplus \nu_2 \nu_3} & \text{for [pos]}.
\end{cases}
\tag{SI.2}
$$

Proof. The equilibrium π on the fast scale satisfies, for $V_P = v_P$ fixed,

$$
\mathbb{E}_\pi[V_{\text{on}}] = M - \mathbb{E}_\pi[V_{\text{off}}] = \begin{cases}
\dfrac{M\kappa_1^+}{\kappa_1^- + \kappa_1^+} & \text{for [neu]}, \\[2ex]
\dfrac{M\kappa_1^+}{\kappa_1^\ominus v_P + \kappa_1^+} & \text{for [neg]}, \\[2ex]
\dfrac{M\kappa_1^\oplus v_P}{\kappa_1^- + \kappa_1^\oplus v_P} & \text{for [pos]}.
\end{cases}
$$

and

$$
\mathbb{E}_\pi[V_R] = \frac{\kappa_2}{\nu_2}\mathbb{E}_\pi[V_{\text{on}}] = \begin{cases}
\dfrac{M\kappa_1^+ \kappa_2}{\nu_2(\kappa_1^- + \kappa_1^+)} & \text{for [neu]}, \\[2ex]
\dfrac{M\kappa_1^+ \kappa_2}{\nu_2(\kappa_1^\ominus v_P + \kappa_1^+)} & \text{for [neg]}, \\[2ex]
\dfrac{M\kappa_1^\oplus \kappa_2 v_P}{\nu_2(\kappa_1^- + \kappa_1^\oplus v_P)} & \text{for [pos]}.
\end{cases}
$$

Plugging this equilibrium into the equations for V_P, we obtain that

$$
\dot{V}_P = \kappa_3 \mathbb{E}[V_R] - \nu_3 V_P = F(V_P)
$$

with F as in (SI.1). Computation of the equilibria is standard by solving $F(v_P) = 0$. In particular, we have to solve

$$
-M\kappa_1^+ \kappa_2 \kappa_3 + \kappa_1^+ \nu_2 \nu_3 v_P + \kappa_1^\ominus \nu_2 \nu_3 v_P^2 = 0 \text{ or } v_P^2 + \frac{\kappa_1^+}{\kappa_1^\ominus} v_P - \frac{M\kappa_1^+ \kappa_2 \kappa_3}{\kappa_1^\ominus \nu_2 \nu_3} = 0
$$

for the equilibrium of [neg].   □

Our next goal is to show that $\sqrt{N}(V_P^N - V_P)$ converges and to determine the limiting process. In the proof, we will make use of the method developed by Kang et al. (2014), which we recall in Section E.

Theorem 2 (Central Limit Theorem)**.** *Let* V_P^N, V_P *and* F *be as in Theorem 1 and* $\sqrt{N}(V_P^N(0) - V_P(0)) \xRightarrow{N \to \infty} U(0)$*. Then, for the models* [neu], [neg] *and* [pos]*,* $\sqrt{N}(V_P^N - V_P) \xRightarrow{N \to \infty} U$*, where* U *solves*

$$U(t) = U(0) + \int_0^t \sqrt{c(V_P(s))} dW(s) + \int_0^t F'(V_P(s)) U(s) ds, \tag{SI.3}$$

with W *the one-dimensional standard Brownian motion and*

$$c(v_P) = \begin{cases} \dfrac{M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)} \left(\dfrac{2\kappa_1^- \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)^2} + \dfrac{2\kappa_3}{\nu_2} + 1 \right) + \nu_3 v_P. & \text{for } [neu], \\[4ex] \dfrac{M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2(\kappa_1^\ominus v_P + \kappa_1^+)} \left(\dfrac{2\kappa_1^\ominus v_P \kappa_2 \kappa_3}{\nu_2(\kappa_1^\ominus v_P + \kappa_1^+)^2} + \dfrac{2\kappa_3}{\nu_2} + 1 \right) + \nu_3 p & \text{for } [neg], \\[4ex] \dfrac{M\kappa_1^\oplus \kappa_2 \kappa_3 v_P}{\nu_2(\kappa_1^- + \kappa_1^\oplus v_P)} \left(\dfrac{2\kappa_1^- \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^\oplus v_P)^2} + \dfrac{2\kappa_3}{\nu_2} + 1 \right) + \nu_3 v_P & \text{for } [pos]. \end{cases} \tag{SI.4}$$

Remark B.1 (Deriving the Fano factor in equilibrium)**.** In order to compute the approximate variance of V_P^N, when started in the equilibrium v_P^*, we make use of the fact that the SDE (SI.3) is solved by an Ornstein Uhlenbeck process. In particular, we obtain at late times (see also [6] in the main text)

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} \approx \frac{N\mathbb{V}[V_P^N]}{\mathbb{E}[V_P^N]} \approx \frac{\mathbb{V}[U]}{v_P^*} \approx -\frac{c(v_P^*)}{2F'(v_P^*)v_P^*}. \tag{SI.5}$$

In order to compute the right hand side, note that

$$F'(v_P) = \begin{cases} -\nu_3 & \text{for } [neu], \\[3ex] -\dfrac{M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2} \dfrac{\kappa_1^\ominus}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} - \nu_3 & \text{for } [neg], \\[3ex] \dfrac{M\kappa_1^\oplus \kappa_2 \kappa_3}{\nu_2} \dfrac{\kappa_1^-}{(\kappa_1^- + \kappa_1^\oplus v_P)^2} - \nu_3 & \text{for } [pos]. \end{cases} \tag{SI.6}$$

Plugging in the equilibrium v_P^* from (SI.2) or [5] for [neg] and [pos], we obtain in particular that

$$v_P^* = \begin{cases} \dfrac{M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2 \nu_3(\kappa_1^\ominus v_P^* + \kappa_1^+)} & \text{for } [neg], \\[3ex] \dfrac{M\kappa_1^\oplus v_P^* \kappa_2 \kappa_3}{\nu_2 \nu_3(\kappa_1^- + \kappa_1^\oplus v_P^*)} & \text{for } [pos] \end{cases}$$

and therefore

$$F'(v_P^*) = \begin{cases} -\nu_3 \left(1 + \dfrac{\kappa_1^\ominus v_P^*}{\kappa_1^\ominus v_P^* + \kappa_1^+} \right) & \text{for } [neg], \\[3ex] -\nu_3 \left(1 - \dfrac{\kappa_1^-}{\kappa_1^- + \kappa_1^\oplus v_P^*} \right) & \text{for } [pos]. \end{cases} \tag{SI.7}$$

In addition,

$$
c(v_P^*) = \begin{cases}
\dfrac{2M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)} \left(\dfrac{\kappa_1^- \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)^2} + \dfrac{\kappa_3}{\nu_2} + 1 \right) & \text{for } [neu], \\[3ex]
\dfrac{2M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2(\kappa_1^\ominus v_P^* + \kappa_1^+)} \left(\dfrac{\kappa_1^\ominus v_P^* \kappa_2 \kappa_3}{\nu_2(\kappa_1^\ominus v_P^* + \kappa_1^+)^2} + \dfrac{\kappa_3}{\nu_2} + 1 \right) & \text{for } [neg], \\[3ex]
\dfrac{2M\kappa_1^\oplus \kappa_2 \kappa_3 v_P^*}{\nu_2(\kappa_1^- + \kappa_1^\oplus v_P^*)} \left(\dfrac{\kappa_1^- \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^\oplus v_P^*)^2} + \dfrac{\kappa_3}{\nu_2} + 1 \right) & \text{for } [pos].
\end{cases}
$$

Hence, plugging these quantities into (SI.5) gives

$$
\frac{\mathbb{V}[P]}{\mathbb{E}[P]} \approx \begin{cases}
\dfrac{\kappa_1^- \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)^2} + \dfrac{\kappa_3}{\nu_2} + 1 & \text{for } [neu], \\[3ex]
\left(1 + \dfrac{\kappa_1^\ominus v_P^*}{\kappa_1^\ominus v_P^* + \kappa_1^+} \right)^{-1} \left(\dfrac{\kappa_1^\ominus v_P^* \kappa_2 \kappa_3}{\nu_2(\kappa_1^- v_P^* + \kappa_1^+)^2} + \dfrac{\kappa_3}{\nu_2} + 1 \right) & \text{for } [neg], \\[3ex]
\left(1 - \dfrac{\kappa_1^- \nu_2 \nu_3}{M\kappa_1^\oplus \kappa_2 \kappa_3} \right)^{-1} \left(\dfrac{\kappa_1^- \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^\oplus v_P^*)^2} + \dfrac{\kappa_3}{\nu_2} + 1 \right) & \text{for } [pos].
\end{cases}
$$

Proof of Theorem 2. We have to show (❶)–(❹) from Section E in all cases. Note that the function F from Theorem 1 already satisfies (❶). In all cases, the system $(V_{\text{on}}, V_R, V_P)$ is a Markov process with a generator of the form (SI.21) with

$$
L_1^N f(u, r, v_P) = \kappa_3 r N \left(f\left(u, r, v_P + \frac{1}{N}\right) - f(u, r, v_P) \right) + v_P \nu_3 N \left(f\left(u, r, v_P - \frac{1}{N}\right) - f(u, r, v_P) \right)
$$
$$
= (\kappa_3 r - p\nu_3) \frac{\partial f}{\partial v_P}(u, r, v_P) + o(1)
$$

(and different operators L_2^N). This already implies that for all cases

$$
F^N(r, v_P) = \kappa_3 r - \nu_3 v_P.
$$

For $[neu]$,

$$
L_2^N f(u, r, v_P) = (M - u)\kappa_1^+ (f(u+1, r, v_P) - f(u, r, v_P))
$$
$$
+ u\kappa_1^- (f(u-1, r, v_P) - f(u, r, v_P))
$$
$$
+ \kappa_2 u(f(u, r+1, v_P) - f(u, r, v_P)) + \nu_2 r(f(u, r-1, v_P) - f(u, r, v_P)).
$$

From (❷) and (SI.1), we see that we need to solve

$$
L_2^N h^N = \kappa_3 r - \frac{M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)}.
$$

Choosing the Ansatz

$$
h^N(u, r, v_P) = u g_1(v_P) + r g_2(v_P),
$$

we obtain that

$$
(M\kappa_1^+ - u(\kappa_1^+ + \kappa_1^-)) g_1(v_P) + (\kappa_2 u - \nu_2 r) g_2(v_P) \overset{!}{=} \kappa_3 r - \frac{M\kappa_1^+ \kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)}.
$$

Solving for g_2 and then for g_1, we obtain

$$g_2(v_P) = -\frac{\kappa_3}{\nu_2}, \qquad g_1(v_P) = -\frac{\kappa_2 \kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)}.$$

Then, $\varepsilon_1^N \xrightarrow{N \to \infty} 0$ since h^N is bounded in N and $\varepsilon_2^N = 0$ by construction. Hence, we have shown (❸). For (❹), if π is the equilibrium of the fast species U, R for given value v_P of the slow species as in Theorem 1, we have that (using notation introduced in Section E)

$$\frac{1}{N}\sum_{k=1}^{6}\mathbb{E}_\pi\left[\left(\zeta_{kP} + h^N(U, R, v_P) - h^N\left(U + \zeta_{k\mathrm{on}}, R + \zeta_{kr}, v_P + \frac{1}{N}\zeta_{kP}\right)\right)^2 \Lambda_k(v_P)\right]$$

$$= 2g_1(v_P))^2 M\frac{\kappa_1^-\kappa_1^+}{\kappa_1^- + \kappa_1^+} + (g_2(v_P))^2\left(\frac{M\kappa_1^+\kappa_2}{\kappa_1^- + \kappa_1^+} + \frac{M\kappa_1^+\kappa_2}{\nu_2(\kappa_1^- + \kappa_1^+)}\nu_2\right)$$

$$+ \left(\frac{M\kappa_1^+\kappa_2\kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)} + \nu_3 v_P\right)$$

$$= \frac{M\kappa_1^+}{\kappa_1^- + \kappa_1^+}\left(\frac{2\kappa_1^-\kappa_2^2\kappa_3^2}{\nu_2^2(\kappa_1^- + \kappa_2^+)^2} + \frac{2\kappa_2\kappa_3^2}{\nu_2^2} + \frac{\kappa_2\kappa_3}{\nu_2}\right) + \nu_3 v_P.$$

For [neg], all calculations above are the same, but with κ_1^- replaced by $\kappa_1^\ominus v_P$, and for [pos], all calculations are the same with κ_1^+ replaced by $\kappa_1^\oplus v_P$. □

B.2   Comparing [neu], [neg] and [pos] if v_P^* is equal

Consider a model [neu] with parameters $\kappa_1^+, \kappa_1^-, \kappa_2, \kappa_3, \nu_2, \nu_3$ and let v_P^* be the equilibrium from [5] or (SI.2). In addition, consider a model [neg] with $\kappa_1^\ominus := \kappa_1^-/v_P^*$ and all other parameters as above and a model [pos] with $\kappa_1^\oplus := \kappa_1^+/p^*$ and all other parameters as above. Then, from (SI.1), we see that all models have v_P^* as their unique deterministic limit with the same

$$c(v_P^*) = \frac{2M\kappa_1^+\kappa_2\kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)}\left(\frac{\kappa_1^-\kappa_2\kappa_3}{\nu_2(\kappa_1^- + \kappa_1^+)^2} + \frac{\kappa_3}{\nu_2} + 1\right)$$

from (SI.4). Plugging all quantities in (SI.5) (or [6]) then gives

$$\frac{\mathbb{V}_{neg}[V_P^N]}{\mathbb{V}_{neu}[V_P^N]} \approx \left(1 + \frac{\kappa_1^\ominus v_P^*}{\kappa_1^\ominus v_P^* + \kappa_1^+}\right)^{-1},$$

$$\frac{\mathbb{V}_{pos}[V_P^N]}{\mathbb{V}_{neu}[V_P^N]} \approx \left(1 - \frac{\kappa_1^-}{\kappa_1^- + \kappa_1^\oplus v_P^*}\right)^{-1}.$$

C   Slow RNA

C.1   The model

Here, we study the case

$$\lambda_3 = \widetilde{\kappa}_3, \qquad \mu_2 = \widetilde{\nu}_2$$

which leads to $X_R = O(N)$, such that we have the scaling $\alpha_R = 1$ or $V_R = X_R/N$. Hence, V_{on} is fast and (V_R, V_P) are slow. Note that in this case, we have that $(V_R^N, V_P^N) \Rightarrow (V_R, V_P)$ with

$$\frac{d}{dt}(V_R, V_P) = F(V_R, V_P) := \begin{cases} \begin{pmatrix} \dfrac{M\kappa_1^+\kappa_2}{\kappa_1^- + \kappa_1^+} - \widetilde{\nu}_2 V_R \\ \widetilde{\kappa}_3 V_R - \nu_3 V_P \end{pmatrix} & \text{for}[neu], \\[3em] \begin{pmatrix} \dfrac{M\kappa_1^+\kappa_2}{\kappa_1^\ominus V_P + \kappa_1^+} - \widetilde{\nu}_2 V_R \\ \widetilde{\kappa}_3 V_R - \nu_3 V_P \end{pmatrix} & \text{for}[neg], \\[3em] \begin{pmatrix} \dfrac{M\kappa_1^\oplus\kappa_2 V_P}{\kappa_1^- + \kappa_1^\oplus V_P} - \widetilde{\nu}_2 V_R \\ \widetilde{\kappa}_3 V_R - \nu_3 V_P \end{pmatrix} & \text{for}[pos]. \end{cases}$$

Theorem 3 (Central Limit Theorem). *Let* V_R^N, V_R, V_P^N, V_P *and* F *be as above and* $\sqrt{N}((V_R^N(0), V_P^N(0)) - (V_R(0), V_P(0))) \xRightarrow{N\to\infty} (U_R(0), U_P(0))$. *Then, for the models* [neu], [neg] *and* [pos], $\sqrt{N}((V_R^N, V_P^N) - (V_R, V_P)) \xRightarrow{N\to\infty} (U_R, U_P)$, *where* (U_R, U_P) *solves*

$$\begin{aligned} U_R(t) &= U_R(0) + \int_0^t \sqrt{c_R(V_P(s))}dW(s) - \int_0^t \widetilde{\nu}_2 U_R(s) + d_R(V_P(s))U_P(s)ds, \\ U_P(t) &= U_P(0) + \int_0^t \sqrt{\widetilde{\kappa}_3 V_R(s) + \nu_3 V_P(s)}dW'(s) + \int_0^t \widetilde{\kappa}_3 U_R(s) - \nu_3 V_P(s)U_P(s)ds \end{aligned} \tag{SI.8}$$

with W, W' *independent Brownian motions,*

$$c_R(v_R, v_P) = \begin{cases} \dfrac{M\kappa_1^+\kappa_2}{\kappa_1^- + \kappa_1^+}\left(\dfrac{2\kappa_1^-\kappa_2}{(\kappa_1^- + \kappa_1^+)^2} + 1\right) + \widetilde{\nu}_2 v_R & \text{for } [neu], \\[2em] \dfrac{M\kappa_1^+\kappa_2}{\kappa_1^\ominus v_P + \kappa_1^+}\left(\dfrac{2\kappa_1^\ominus v_P\kappa_2}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} + 1\right) + \widetilde{\nu}_2 v_R & \text{for } [neg], \\[2em] \dfrac{M\kappa_1^\oplus v_P\kappa_2}{\kappa_1^- + \kappa_1^\oplus v_P}\left(\dfrac{2\kappa_1^-\kappa_2}{(\kappa_1^- + \kappa_1^\oplus v_P)^2} + 1\right) + \widetilde{\nu}_2 v_R & \text{for } [pos] \end{cases} \tag{SI.9}$$

and

$$d_R(v_P) = \begin{cases} 0 & \text{for } [neu], \\[1.5em] \dfrac{M\kappa_1^\ominus\kappa_1^+\kappa_2}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} & \text{for } [neg], \\[1.5em] -\dfrac{M\kappa_1^-\kappa_1^\oplus\kappa_2}{(\kappa_1^- + \kappa_1^\oplus v_P)^2}. & \text{for } [pos] \end{cases} \tag{SI.10}$$

Remark C.1. In equilibrium, we have

$$
c_R(v_R^*, v_P^*) = \begin{cases} \dfrac{2\widetilde{\nu}_2\nu_3 v_P^*}{\widetilde{\kappa}_3}\left(\dfrac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2} + 1\right) & \text{for } [neu], \\[3mm] \dfrac{2\widetilde{\nu}_2\nu_3 v_P^*}{\widetilde{\kappa}_3}\left(\dfrac{\kappa_1^\ominus v_P \kappa_2}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} + 1\right) & \text{for } [neg], \\[3mm] \dfrac{2\widetilde{\nu}_2\nu_3 v_P^*}{\widetilde{\kappa}_3}\left(\dfrac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^\oplus v_P)^2} + 1\right) & \text{for } [pos]. \end{cases} \tag{SI.11}
$$

Proof. First, note that

$$
DF(v_R, v_P) = \begin{cases} \begin{pmatrix} -\widetilde{\nu}_2 & 0 \\ \widetilde{\kappa}_3 & -\nu_3 \end{pmatrix} & \text{for}[neu], \\[5mm] \begin{pmatrix} -\widetilde{\nu}_2 & -\frac{M\kappa_1^+ \kappa_1^\ominus \kappa_2}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} \\ \widetilde{\kappa}_3 & -\nu_3 \end{pmatrix} & \text{for}[neg], \\[5mm] \begin{pmatrix} -\widetilde{\nu}_2 & \frac{M\kappa_1^- \kappa_1^\oplus \kappa_2}{(\kappa_1^- + \kappa_1^\oplus v_P)^2} \\ \widetilde{\kappa}_3 & -\nu_3 \end{pmatrix} & \text{for}[pos]. \end{cases}
$$

Again, we have to show (❶)–(❹) from Section E in all cases for F as above, which already satisfies (❶). We focus on $[neu]$ first. The system $(V_{\text{on}}, V_R, V_P)$ is a Markov process with a generator of the form (SI.21) with

$$
L_1^N f(u, v_R, v_P) = (\kappa_2 u - v_R\widetilde{\nu}_2)\frac{\partial f}{\partial v_R}(u, v_R, v_P) + (\widetilde{\kappa}_3 v_R - v_P\nu_3)\frac{\partial f}{\partial v_P}(u, v_R, v_P) + o(1),
$$
$$
L_2^N f(u, v_R, v_P) = (M - u)\kappa_1^+(f(u + 1, v_R, v_P) - f(u, v_R, v_P)) \\ + u\kappa_1^-(f(u - 1, v_R, v_P) - f(u, v_R, v_P)).
$$

This implies that

$$
F^N(u, v_R, v_P) = \begin{pmatrix} \kappa_2 u - \widetilde{\nu}_2 v_R \\ \widetilde{\kappa}_3 v_R - \nu_3 v_P \end{pmatrix}.
$$

Hence, for (❷), we have to solve

$$
L_2^N h^N(u, v_R, v_P) = \begin{pmatrix} \kappa_2 u - \dfrac{M\kappa_1^+ \kappa_2}{\kappa_1^- + \kappa_1^+} \\ 0 \end{pmatrix},
$$

which is

$$
h^N(u, v_R, v_P) = \begin{pmatrix} -u\dfrac{\kappa_2}{\kappa_1^- + \kappa_1^+} \\ 0 \end{pmatrix}.
$$

For the quadratic variation in (❹), we find that with $z^{\otimes 2} = zz^\top$

$$N\Big[(V_R, V_P)^\top - \frac{1}{N}h^N(V_{\text{on}}, V_R, V_P)\Big]_t$$

$$= \frac{1}{N}\sum_{k\in\mathcal{K}}\int_0^t \Big((\zeta_{kR}, \zeta_{kP})^\top + h^N\big(V_{\text{on}}(s-), V_R(s-), V_P(s-)\big)$$

$$- h^N\Big(V_{\text{on}}(s-) + \zeta_{k\text{on}}, V_R(s-) + \frac{1}{N}\zeta_{kR}, V_P(s-) + \frac{1}{N}\zeta_{kP}\Big)\Big)^{\otimes 2}$$

$$\Lambda_k^N(V_{\text{on}}(s), V_R(s), V_P(s))ds$$

$$= \int_0^t \begin{pmatrix} \frac{\kappa_2}{\kappa_1^- + \kappa_1^+} \\ 0 \end{pmatrix}^{\otimes 2} \kappa_1^+(M - V_{\text{on}}(s)) + \begin{pmatrix} -\frac{\kappa_2}{\kappa_1^- + \kappa_1^+} \\ 0 \end{pmatrix}^{\otimes 2} \kappa_1^- V_{\text{on}}(s)$$

$$+ \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{\otimes 2} \kappa_2 V_{\text{on}}(s) + \begin{pmatrix} -1 \\ 0 \end{pmatrix}^{\otimes 2} \widetilde{\nu}_2 V_R(s) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}^{\otimes 2} \widetilde{\kappa}_3 V_R(s) + \begin{pmatrix} 0 \\ -1 \end{pmatrix}^{\otimes 2} \nu_3 V_P(s)ds$$

$$= \int_0^t \begin{pmatrix} \frac{\kappa_2^2}{(\kappa_1^- + \kappa_1^+)^2} & 0 \\ 0 & 0 \end{pmatrix}(\kappa_1^+(M - V_{\text{on}}(s)) + \kappa_1^- V_{\text{on}}(s)) + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}(\kappa_2 V_{\text{on}}(s) + \widetilde{\nu}_2 V_R(s))$$

$$+ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}(\widetilde{\kappa}_3 V_R(s) + \nu_3 V_P(s))ds$$

$$\approx \int_0^t \begin{pmatrix} \frac{2M\kappa_1^- \kappa_1^+ \kappa_2^2}{(\kappa_1^- + \kappa_1^+)^3} + \frac{M\kappa_1^+ \kappa_2}{\kappa_1^+ + \kappa_1^-} + \widetilde{\nu}_2 V_R(s) & 0 \\ 0 & \widetilde{\kappa}_3 V_R(s) + \nu_3 V_P(s) \end{pmatrix}ds.$$

This shows the assertion for [*neu*]. The cases [*neg*] and [*pos*] are similar, if we change κ_1^- by $\kappa_1^\ominus v_P$ for [*neu*] and κ_1^+ by $\kappa_1^\oplus v_P$ for [*pos*]. ☐

C.2   Equilibrium Fano factor...

Let us start in equilibrium, i.e. $V_R(0) = v_R^*$, $V_P(0) = v_P^*$. Then, we will plug in c_R from (SI.11), and obtain Ornstein-Uhlenbeck processes in all cases. Since they are two-dimensional, their equilibrium (normal) distribution can be computed (see Section F).

...for [*neu*]

We obtain

$$d\begin{pmatrix} U_R \\ U_P \end{pmatrix} = -\begin{pmatrix} \widetilde{\nu}_2 & 0 \\ -\widetilde{\kappa}_3 & \nu_3 \end{pmatrix}\begin{pmatrix} U_R \\ U_P \end{pmatrix}dt + \sqrt{\frac{2v_P^* \widetilde{\nu}_2 \nu_3}{\widetilde{\kappa}_3}}\begin{pmatrix} \sqrt{\frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2} + 1} & 0 \\ 0 & \sqrt{\frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2}} \end{pmatrix}\begin{pmatrix} dW \\ dW' \end{pmatrix}.$$

Hence, with $U = (U_R, U_P)^\top$, in equilibrium, from Corollary F.1,

$$\mathbf{E}_{neu}[UU^\top] = \frac{v_P^*}{(\widetilde{\nu}_2 + \nu_3)\widetilde{\kappa}_3}\begin{pmatrix} \Big(1 + \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2}\Big)\nu_3(\widetilde{\nu}_2 + \nu_3) & \Big(1 + \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2}\Big)\widetilde{\kappa}_3\nu_3 \\ \Big(1 + \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2}\Big)\widetilde{\kappa}_3\nu_3 & \widetilde{\kappa}_3(\widetilde{\nu}_2 + \nu_3) + \widetilde{\kappa}_3^2\Big(1 + \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2}\Big) \end{pmatrix}.$$

$$\text{(SI.12)}$$

Therefore,

$$\frac{\mathbb{V}_{neu}[V_P]}{v_P^*} = 1 + \frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2 + \nu_3} + \frac{\kappa_1^- \kappa_2 \widetilde{\kappa}_3}{(\kappa_1^- + \kappa_1^+)^2(\widetilde{\nu}_2 + \nu_3)}. \tag{SI.13}$$

...for [*neg*]

Here, we obtain

$$
d\begin{pmatrix} U_R \\ U_P \end{pmatrix} = -\begin{pmatrix} \widetilde{\nu}_2 & \frac{\kappa_1^{\ominus} v_P^* \widetilde{\nu}_2 \nu_3}{(\kappa_1^{\ominus} v_P^* + \kappa_1^+)\widetilde{\kappa}_3} \\ -\widetilde{\kappa}_3 & \nu_3 \end{pmatrix} \begin{pmatrix} U_R \\ U_P \end{pmatrix} dt + \sqrt{\frac{2 v_P^* \widetilde{\nu}_2 \nu_3}{\widetilde{\kappa}_3}} \begin{pmatrix} \sqrt{\frac{\kappa_1^{\ominus} v_P^* \kappa_2}{(\kappa_1^{\ominus} v_P^* + \kappa_1^+)^2} + 1} & 0 \\ 0 & \sqrt{\frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2}} \end{pmatrix} \begin{pmatrix} dW \\ dW' \end{pmatrix}.
$$

Hence, in equilibrium, and if $\cdots$ denote the quantities from (SI.12), and for $b = \frac{\kappa_1^{\ominus} v_P^* \widetilde{\nu}_2 \nu_3}{(\kappa_1^{\ominus} v_P^* + \kappa_1^+)\widetilde{\kappa}_3}$,

$$
\mathbb{E}_{neg}[UU^{\top}] = \frac{v_P^* \widetilde{\nu}_2 \nu_3}{(\widetilde{\nu}_2 \nu_3 + \widetilde{\kappa}_3 b)(\widetilde{\nu}_2 + \nu_3)\widetilde{\kappa}_3} \begin{pmatrix} \cdots + \frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2} b^2 + \left(\frac{\kappa_1^{\ominus} v_P^* \kappa_2}{(\kappa_1^{\ominus} v_P^* + \kappa_1^+)^2} + 1\right)\widetilde{\kappa}_3 b & \cdots - \left(\frac{\kappa_1^{\ominus} v_P^* \kappa_2}{(\kappa_1^{\ominus} v_P^* + \kappa_1^+)^2} + 1\right)\widetilde{\kappa}_3 b \\ \cdots - \widetilde{\kappa}_3 b & \cdots + \frac{\widetilde{\kappa}_3^2}{\widetilde{\nu}_2} b \end{pmatrix}.
$$

$$\text{(SI.14)}$$

So,

$$
\frac{\mathbb{V}_{neg}[V_P]}{v_P^*} = \left(1 + \frac{\kappa_1^{\ominus} v_P^*}{\kappa_1^{\ominus} v_P^* + \kappa_1^+}\right)^{-1}\left(1 + \frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2 + \nu_3} + \frac{\kappa_1^{\ominus} v_P^* \kappa_2 \widetilde{\kappa}_3}{(\kappa_1^{\ominus} v_P^* + \kappa_1^+)^2(\widetilde{\nu}_2 + \nu_3)} \right.
$$
$$
\left. + \frac{\kappa_1^{\ominus} v_P^* \nu_3}{(\kappa_1^{\ominus} v_P^* + \kappa_1^+)(\widetilde{\nu}_2 + \nu_3)}\right).
$$

$$\text{(SI.15)}$$

...for [*pos*]

We obtain

$$
d\begin{pmatrix} U_R \\ U_P \end{pmatrix} = -\begin{pmatrix} \widetilde{\nu}_2 & -\frac{\kappa_1^- \widetilde{\nu}_2 \nu_3}{(\kappa_1^- + \kappa_1^{\oplus} v_P^*)\widetilde{\kappa}_3} \\ -\widetilde{\kappa}_3 & \nu_3 \end{pmatrix} \begin{pmatrix} U_R \\ U_P \end{pmatrix} dt + \sqrt{\frac{2 v_P^* \widetilde{\nu}_2 \nu_3}{\widetilde{\kappa}_3}} \begin{pmatrix} \sqrt{\frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^{\oplus} v_P^*)^2} + 1} & 0 \\ 0 & \sqrt{\frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2}} \end{pmatrix} \begin{pmatrix} dW \\ dW' \end{pmatrix}.
$$

In equilibrium, we now have exactly (SI.14) but with $b = -\frac{\kappa_1^- \widetilde{\nu}_2 \nu_3}{(\kappa_1^- + \kappa_1^{\oplus} v_P^*)\widetilde{\kappa}_3}$. Hence,

$$
\frac{\mathbb{V}_{pos}[V_P]}{v_P^*} = \left(1 - \frac{\kappa_1^-}{\kappa_1^- + \kappa_1^{\oplus} v_P^*}\right)^{-1}\left(1 + \frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2 + \nu_3} + \frac{\kappa_1^- \kappa_2 \widetilde{\kappa}_3}{(\kappa_1^- + \kappa_1^{\oplus} v_P^*)^2(\widetilde{\nu}_2 + \nu_3)} \right.
$$
$$
\left. - \frac{\kappa_1^- \nu_3}{(\kappa_1^- + \kappa_1^{\oplus} v_P^*)(\widetilde{\nu}_2 + \nu_3)}\right).
$$

$$\text{(SI.16)}$$

D   Comparison to previous results

D.1   Comparison for [*neu*]

The neutral case is by now well-studied. From Paulsson (2005), we see that this equilibrium obeys, with

$$
\tau_1 = \frac{1}{\lambda_1^- + \lambda_1^+}, \qquad \tau_2 = \frac{1}{\mu_2}, \qquad \tau_3 = \frac{1}{\mu_3},
$$

the expected life-times of a change in gene activity, mRNA and protein, respectively, that

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} = 1 + \frac{\lambda_3}{\mu_3}\frac{\tau_2}{\tau_2 + \tau_3} + \frac{\mathbb{E}[P]\lambda_1^-}{M\lambda_1^+}\frac{\tau_1}{(\tau_2+\tau_3)(\tau_1+\tau_3)}\frac{\tau_1\tau_2 + \tau_2\tau_3 + \tau_1\tau_3}{\tau_1+\tau_2}$$

$$\overset{\tau_3 \gg \tau_1, \tau_2}{\approx} 1 + \frac{\lambda_3}{\mu_2} + \frac{\lambda_1^- \lambda_2 \lambda_3}{\mu_2\mu_3(\lambda_1^- + \lambda_1^+)}\frac{\tau_1}{\tau_3} = 1 + \frac{\lambda_3}{\mu_2} + \frac{\lambda_1^- \lambda_2 \lambda_3}{\mu_2(\lambda_1^- + \lambda_1^+)^2}.$$

Since our scaling is exactly such that $\tau_3 \gg \tau_1, \tau_2$, i.e. the protein is much more stable than RNA and the state of the gene, this result is in line with [11] in the main text.

D.2   Comparison with Thattai

In Thattai and van Oudenaarden (2001), a linearization of [*neg*] was studied in the case of fast switching on and off of the gene. (This will mean that both, $\kappa_1^-, \kappa_1^+ \gg 1$.) Consider as in the proof of Theorem 1 that

$$\mathbb{E}_\pi[V_{\text{on}}] = \frac{M\kappa_1^+}{\kappa_1^\ominus v_P + \kappa_1^+}.$$

Then, study the system (compare with $(\bullet_{neg})$)

$$V_R^N(t) = V_R^N(0) + Y_3\Big(N\int_0^t \kappa_2 \frac{M\kappa_1^+}{\kappa_1^- V_P^N(s) + \kappa_1^+}ds\Big) - Y_4\Big(N\int_0^t \nu_2 V_R^N(s)ds\Big),$$

$$V_P^N(t) = V_P^N(0) + N^{-1}Y_5\Big(N\int_0^t \kappa_3 V_R^N(s)ds\Big) - N^{-1}Y_6\Big(N\int_0^t \nu_3 V_P^N(s)ds\Big), \qquad \text{(SI.17)}$$

which can in the case (8) (i.e. $\lambda_1^\ominus N v_P^* \ll \lambda_1^+$) be approximated by using

$$V_R^N(t) = V_R^N(0) + Y_3\Big(N\int_0^t M\kappa_2\Big(1 - \frac{\kappa_1^\ominus}{\kappa_1^+}V_P^N(s)\Big)ds\Big) - Y_4\Big(N\int_0^t \nu_2 V_R^N(s)ds\Big). \qquad \text{(SI.18)}$$

Now, the system (SI.17) and (SI.18) is exactly as on p. 3 in Thattai and van Oudenaarden (2001) with

$$k_0 = NM\kappa_2, \qquad k_1 = \frac{M\kappa_1^\ominus \kappa_2}{\kappa_1^+}, \qquad k_P = N\kappa_3, \qquad \gamma_R = N\nu_2, \qquad \gamma_P = \nu_3.$$

Plugging these variables into [3] of Thattai and van Oudenaarden (2001) we obtain in equilibrium with

$$\eta = \frac{\gamma_P}{\gamma_R} = \frac{\nu_3}{N\nu_2}, \qquad b = \frac{k_P}{\gamma_R} = \frac{\kappa_3}{\nu_2}, \qquad \phi = \frac{k_1}{\gamma_P} = \frac{M\kappa_1^\ominus \kappa_2}{\kappa_1^+ \nu_3}$$

that (note that η is negligible since N is large and ϕ is small by (8))

$$\frac{1}{N}\mathbb{E}[P] = \frac{1}{N}\Big(\frac{1}{1+b\phi}\Big)\frac{k_0 b}{\gamma_P} \approx \Big(1 - \frac{M\kappa_1^\ominus \kappa_2 \kappa_3}{\kappa_1^+ \nu_2 \nu_3}\Big)\frac{M\kappa_2\kappa_3}{\nu_2\nu_3} = \frac{M\kappa_2\kappa_3}{\nu_2\nu_3} - \frac{M^2\kappa_2^2\kappa_3^2}{\nu_2^2\nu_3^2}\frac{\kappa_1^\ominus}{\kappa_1^+},$$

$$\frac{\mathbb{V}[P]}{\mathbb{E}[P]} = \Big(\frac{1-\phi}{1+b\phi}\cdot\frac{b}{1+\eta} + 1\Big) \approx 1 + b(1-\phi)(1-b\phi) \approx 1 + b(1-(b+1)\phi)$$

$$= (1+b) - (1+b)b\phi = (1+b)(1-b\phi) = \Big(1 + \frac{\kappa_3}{\nu_2}\Big)\Big(1 - M\frac{\kappa_2\kappa_3}{\nu_2\nu_3}\frac{\kappa_1^\ominus}{\kappa_1^+}\Big)$$

$$\approx \Big(1 + \frac{\kappa_3}{\nu_2}\Big)\Big(1 - \frac{\kappa_1^\ominus}{\kappa_1^+}v_P^*\Big) = \Big(1 + \frac{\lambda_3}{\mu_2}\Big)\Big(1 - \frac{\lambda_1^\ominus N v_P^*}{\kappa_1^+}\Big).$$

D.3   Comparison with Swain

In Swain (2004) the author uses $d_0 = \mu_2, d_1 = \mu_3, v_1 = \lambda_3, \langle M \rangle = \frac{\lambda_1^+ \lambda_2}{(\lambda_1^\ominus N v_P^*)\mu_2}$ and

$$\epsilon_c = \frac{2}{1 + \sqrt{1 + 4\frac{\lambda_1^\ominus \lambda_2 \lambda_3}{\lambda_1^+ \mu_2 \mu_3}}}$$

in order to obtain the Fano factor for the auto-regulatory gene expression with negative (transcriptional) feedback. Expressing terms as in Section C, he gets

$$\frac{\mathbb{V}_{neg}[P]}{\mathbb{E}_{neg}[P]} = 1 + \frac{\nu_3 v_P^*}{(\widetilde{\nu}_2 + \nu_3)v_R^*}\left(1 - \frac{\sqrt{1 + 4\frac{\kappa_1^\ominus \kappa_2 \widetilde{\kappa}_3}{\kappa_1^+ \widetilde{\nu}_2 \nu_3}} - 1}{2\sqrt{1 + 4\frac{\kappa_1^\ominus \kappa_2 \widetilde{\kappa}_3}{\kappa_1^+ \widetilde{\nu}_2 \nu_3}}}\left(1 + \frac{\widetilde{\nu}_2}{\widetilde{\kappa}_3}\right)\right)$$

$$= 1 + \frac{\widetilde{\kappa}_3}{(\widetilde{\nu}_2 + \nu_3)}\left(1 - \frac{\kappa_1^- v_P^*}{2\kappa_1^- v_P^* + \kappa_1^+}\left(1 + \frac{\widetilde{\nu}_2}{\widetilde{\kappa}_3}\right)\right)$$

$$= \frac{2\kappa_1^- v_P^* + \kappa_1^+ + \widetilde{\kappa}_3\frac{\kappa_1^- v_P^* + \kappa_1^+}{\widetilde{\nu}_2 + \nu_3} - \frac{\kappa_1^- \widetilde{\nu}_2 v_P^*}{\widetilde{\nu}_2 + \nu_3}}{(2\kappa_1^- v_P^* + \kappa_1^+)}$$

$$= \frac{(\kappa_1^- v_P^* + \kappa_1^+)(1 + \frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2 + \nu_3} + \frac{\kappa_1^- \nu_3 v_P^*}{(\widetilde{\nu}_2 + \nu_3)(\kappa_1^- v_P^* + \kappa_1^+)})}{(2\kappa_1^- v_P^* + \kappa_1^+)}$$

$$= \frac{1 + \frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2 + \nu_3} + \frac{\kappa_1^- \nu_3 v_P^*}{(\widetilde{\nu}_2 + \nu_3)(\kappa_1^- v_P^* + \kappa_1^+)}}{\frac{(\kappa_1^- v_P^* + \kappa_1^+) + \kappa_1^- v_P^*}{\kappa_1^- v_P^* + \kappa_1^+}}$$

$$= \frac{1 + \frac{\widetilde{\kappa}_3}{\widetilde{\nu}_2 + \nu_3} + \frac{\kappa_1^- \nu_3 v_P^*}{(\widetilde{\nu}_2 + \nu_3)(\kappa_1^- v_P^* + \kappa_1^+)}}{1 + \frac{\kappa_1^- v_P^*}{\kappa_1^- v_P^* + \kappa_1^+}}.$$

This is very similar to our equation (SI.15) except for a missing term emerging from gene expression. As explained in the main text, the Langevin approximation Swain uses cannot account for all fluctuations when a quasi-steady-state assumption is made. (Precisely, it cannot account for fluctuations in the averaged variables.)

D.4   Simpler model of gene expression

We want to derive the simpler system without RNA from the full model. Noting that in $(*_{neu})$, for $\kappa_2 = \nu_2 \gg 1$, we have that for every finite N, RNA is approximately in equilibrium, which is Poisson with parameter 1 for every *on*-gene. Feeding this into the equation for P, we obtain that

$$V_P^N(t) = V_P^N(0) + N^{-1}Y_5\left(N\int_0^t \kappa_3 V_{\text{on}}(s)ds\right) - N^{-1}Y_6\left(\int_0^t \nu_3 V_P^N(s)ds\right).$$

This is the same equation which is need for the simple model given in the main text. Hence, all results follow for $\kappa_2 = \nu_2 \to \infty$.

E   Recalling the approach of Kang et al. (2014)

We will consider a general system of chemical reactions with $\mathcal{S}$ as the set of chemical species and $\mathcal{K}$ the set of reactions; see also Anderson and Kurtz (2015) for further reference. The chemical reactions have the form

$$\nu_k \xrightarrow{\lambda_k} \nu_k', \qquad k \in \mathcal{K}, \tag{SI.19}$$

where $\nu_k = (\nu_{ks})_{s \in \mathcal{S}}$ and $\nu_k' = (\nu_{ks}')_{s \in \mathcal{S}}$ are vectors of chemical species, i.e. elements of $\mathbb{N}^{\mathcal{S}}$. For the dynamics, we assume mass action kinetics, i.e. we set

$$\Lambda_k(x) = \lambda_k x^{\nu_k} := \lambda_k \prod_{s \in \mathcal{S}} x_s^{\nu_{ks}}$$

for the reaction rate of reaction $k \in \mathcal{K}$. With

$$\zeta_k := \nu_k' - \nu_k,$$

we can then define the dynamics of the Markov process $X = (X_s)_{s \in \mathcal{S}}$ through the process R_k, which describes the number of occurrences of reaction k up to time t. We have

$$R_k(t) = Y_k\Big(\int_0^t \Lambda_k(X(s))ds \Big)$$

for independent unit rate Poisson processes $Y_k, k \in \mathcal{K}$ and therefore

$$X(t) = X(0) + \sum_{k \in \mathcal{K}} \zeta_k R_k(t) = X(0) + \sum_{k \in \mathcal{K}} \zeta_k Y_k\Big(\int_0^t \Lambda_k(X(s))ds \Big). \tag{SI.20}$$

We will tailor the results of Kang et al. (2014) to the special case we need in our gene expression example. This means that we can make use of several simplifications, e.g. on the form of the generator of the full process.

1. Find $F : \mathbb{R}^{d_s}$ such that
$$F(v_s) = \lim_{N \to \infty} \mathbb{E}_{\rho_N}[L^N v_s],$$

2. Find h^N such that $\sqrt{N}\varepsilon_1^N, \sqrt{N}\varepsilon_2^N \xRightarrow{N \to \infty} 0$ with ε_i^N and ε_2^N from (SI.24).

For some scaling parameter N, assume that (X_s^N, X_f^N) is a Markov jump process with state space $\mathbb{N}^{d_s} \times \mathbb{N}^{d_f}$ such that for $V^N = (V_s^N, V_f^N)$ with $V_f^N = X_f^N$ and $V_s^N = X_s^N/N$, the system V_s^N is a slow (rescaled) sub-system and V_f^N is a fast sub-system. We assume that the generator L^N has the form

$$L^N = L_1^N + N L_2^N, \tag{SI.21}$$

where L_2^N describes the dynamics of V_f^N, i.e. $L_2^N f = 0$ if f only depends on v_s. Our goal is to show convergence

$$V_s^N \xRightarrow{N \to \infty} V_s, \tag{SI.22}$$

$$U^N := \sqrt{N}(V_s^N - V_s) \xRightarrow{N \to \infty} U \tag{SI.23}$$

for some V_s and U. Therefore, we proceed as follows.

1. We have that (with the projection π_{s} on the slow species and $F^N := L^N \pi_{\mathrm{s}} = L_1^N \pi_{\mathrm{s}}$)

$$M_1^N(t) := V_{\mathrm{s}}^N(t) - V_{\mathrm{s}}^N(0) - \int_0^t F^N(V_{\mathrm{s}}^N(s), V_{\mathrm{f}}^N(s))ds$$

is a (local) martingale. For the convergence (SI.22), we assume that

$$\int_0^t F^N(V_{\mathrm{s}}^N(s), V_{\mathrm{f}}^N(s)) \xrightarrow{N \to \infty} \int_0^t F(V_{\mathrm{s}}(s))ds$$

for some unique process V_{s}, which holds for

$$F(v_{\mathrm{s}}) = \lim_{N \to \infty} \mathbb{E}_{\rho_N}[L^N v_{\mathrm{s}}], \tag{❶}$$

where ρ_N is the equilibrium of V_{f}^N for fixed slow species v_s. Thus, the convergence $V_{\mathrm{s}}^N \xrightarrow{N \to \infty} V_{\mathrm{s}}$ holds with

$$V_{\mathrm{s}}(t) = V_{\mathrm{s}}(0) + \int_0^t F(V_{\mathrm{s}}(s))ds$$

and we have shown (SI.22).

2. Note that

$$U^N(t) - U^N(0)$$
$$= \sqrt{N}(V_{\mathrm{s}}^N(t) - V_{\mathrm{s}}(t)) - \sqrt{N}(V_{\mathrm{s}}^N(0) - V_{\mathrm{s}}(0))$$
$$= \sqrt{N}\left(M^{N,1}(t) + \int_0^t F^N(V_{\mathrm{s}}^N(s), V_{\mathrm{f}}^N(s)) - F(V_{\mathrm{s}}^N(s))ds + \int_0^t F(V_{\mathrm{s}}^N) - F(V_{\mathrm{s}})ds\right).$$

Assume that we

$$\text{find } h^N \text{ such that } L_2^N h^N \approx F^N - F. \tag{❷}$$

(The '$\approx$' is controlled by ε_2^N below. Note that this is a Poisson equation.) With

$$\varepsilon_1^N(t) := \frac{1}{N}\left(h^N(V_{\mathrm{s}}^N(t), V_{\mathrm{f}}^N(t)) - h^N(V_{\mathrm{s}}^N(0), V_{\mathrm{f}}^N(0)) - \int_0^t L_1^N h^N(V_{\mathrm{s}}^N(s), V_{\mathrm{f}}^N(s))ds\right),$$

$$\varepsilon_2^N(t) := -\int_0^t L_2^N h^N(V_{\mathrm{s}}^N(s), V_{\mathrm{f}}^N(s)) - (F^N(V_{\mathrm{s}}^N(s), V_{\mathrm{f}}^N(s)) - F(V_{\mathrm{s}}^N(s)))ds,$$

$$\tag{SI.24}$$

we obtain that

$$M_2^N(t) := \varepsilon_1^N(t) + \varepsilon_2^N(t) - \int_0^t F^N(V_{\mathrm{s}}^N(s), V_{\mathrm{f}}^N(s)) - F(V_{\mathrm{s}}^N(s))ds$$

is a (local) martingale. Hence

$$\sqrt{N}(V_{\mathrm{s}}^N(t) - V_{\mathrm{s}}(t)) - \sqrt{N}(V_{\mathrm{s}}^N(0) - V_{\mathrm{s}}(0))$$
$$= \sqrt{N}\left(M^{N,1}(t) - M^{N,2}(t) + \varepsilon_1^N(t) + \varepsilon_2^N(t) + \int_0^t F(V_{\mathrm{s}}^N) - F(V_{\mathrm{s}})ds\right)$$

We assume for $\varepsilon_1^N, \varepsilon_2^N$ from (SI.24) that

$$\sqrt{N}\varepsilon_1^N, \sqrt{N}\varepsilon_2^N \xrightarrow{N \to \infty} 0. \tag{❸}$$

3. In order to show convergence of $M_1^N - M_2^N$, we use the Martingale Central Limit Theorem; see Theorem 5. Note that since the quadratic variation of all integrals $\int dt$ vanishes, we find that (recall the notation for Chemical Reaction Networks from (SI.19)–(SI.20), which we now equip with a superscript N to account for the scaling constant), with $z^{\otimes 2} = zz^\top$

$$[\sqrt{N}(M_1^N - M_2^N)]_t = N\Big[V_s^N - \frac{1}{N}(h^N(V_s^N, V_f^N))\Big]_t$$

$$= \frac{1}{N}\sum_{k \in \mathcal{K}}\int_0^t \Big(\zeta_{ks} + h^N\big(V_s^N(s-), V_f^N(s-)\big)$$

$$- h^N\Big(V_s^N(s-) + \frac{1}{N}\zeta_{ks}, V_f^N(s-) + \zeta_{kf}\Big)\Big)^{\otimes 2} dR_k^N(s),$$

where ζ_{ks} and ζ_{kf} are the stochiometric changes of the $k-$th reaction in the slow and fast subsystem, respectively. Note that in all applications, we will have that R_k^N either changes slowly, or changes fast and can thus be approximated by a deterministic curve, such that

$$\lim_{N\to\infty}[\sqrt{N}(M_1^N - M_2^N)]_t = \lim_{N\to\infty}\frac{1}{N}\sum_{k \in \mathcal{K}}\int_0^t \Big(\zeta_{ks} + h^N\big(V_s^N(s), V_f^N(s-)\big)$$

$$- h^N\Big(V_s^N(s) + \frac{1}{N}\zeta_{ks}, V_f^N(s-) + \zeta_{kf}\Big)\Big)^{\otimes 2}\Lambda_k^N(s)ds.$$

Now, for the equilibrium ρ_N of the fast species for given concentration of slow species, v_s, if

$$\frac{1}{N}\sum_{k \in \mathcal{K}}\mathbb{E}_{\rho^N}\Big[\Big(\zeta_{ks} + h^N\big(v_s, V_f^N\big) - h^N\Big(v_s + \frac{1}{N}\zeta_{ks}, V_f^N + \zeta_{kf}\Big)\Big)^{\otimes 2}\Lambda_k^N(V_s)\Big] \xRightarrow{N\to\infty} c(v_s), \quad \text{(❹)}$$

we have that

$$\frac{1}{N}\sum_{k \in \mathcal{K}}\int_0^t \Big(\zeta_{ks} + h^N\big(V_s^N(s), V_f^N(s-)\big)$$

$$- h^N\Big(V_s^N(s) + \frac{1}{N}\zeta_{ks}, V_f^N(s-) + \zeta_{kf}\Big)\Big)^{\otimes 2}\overline{\Lambda}_k^N(V_s^N)ds \xRightarrow{N\to\infty} \int_0^t c(V(s))ds,$$

where the right hand side is a deterministic, absolutely continuous, $\mathbb{R}^{s\times s}$-valued function with non-negative time-derivative. Hence we know from Theorem 5 that

$$\sqrt{N}(M^{N,1} - M^{N,2}) \xRightarrow{N\to\infty} M$$

where M satisfies

$$dM = \sqrt{c(V(t))}dW.$$

4. Concluding, if $F \in \mathcal{C}^1(\mathbb{R}^{d_s})$ with

$$F(V_s^N) - F(V_s) = \frac{1}{\sqrt{N}}\big(\nabla F(V_s)U^N + o(1)\big)$$

we find that, if $U^N \xRightarrow{N\to\infty} U$, then

$$U(t) - U(0) = \int_0^t \nabla F(V_s(r))U_r dr + \int_0^t \sqrt{c(V(r))}dW.$$

This gives (SI.23).

F The two-dimensional Ornstein-Uhlenbeck process

We recall results for the two-dimensional Ornstein-Uhlenbeck process; see e.g. Gardiner (2009).

Theorem 4 (Stationary variance of two-dimensional Ornstein-Uhlenbeck process)**.** *Let* $X = (X_1, X_2)$ *solve*

$$dX = -AX dt + BdW,$$

where $A, B \in \mathbb{R}^{2 \times 2}$. *Then, if all eigenvalues of* A *have positive real part, the stationary solution* X_0 *of the SDE has* $\mathbb{E}[X_0] = 0$ *and*

$$\mathbf{E}[X_0 X_0^\top] = \frac{(A - (\mathrm{tr}A)E_2)BB^\top(A^\top - (\mathrm{tr}A)E_2) + (\det A)BB^\top}{2(\det A)(\mathrm{tr}A)}.$$

Proof. Using partial integration, it is easy to see that this SDE is solved by

$$X_t = e^{-At}X_0 + \int_0^t e^{-A(t-s)}BdW.$$

If all eigenvalues of A have positive real part, the stationary solution of the SDE has the distribution

$$X_0 = \int_{-\infty}^0 e^{As}BdW_s.$$

In particular, $\mathbb{E}[X_0] = 0$ and

$$\mathbb{E}[X_0 X_0^\top] = \int_{-\infty}^0 e^{As}BB^\top e^{A^\top s}ds.$$

In order to compute the right hand side, we note that, for any 2×2-matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have that (for the unit matrix I_2)

$$A^2 = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & cb + d^2 \end{pmatrix} = \begin{pmatrix} (a+d)a - (ad - bc) & (a+d)b \\ (a+d)c & (a+d)d - (ad - bc) \end{pmatrix}$$
$$= (\mathrm{tr}A)A - (\det A)I_2.$$

Hence, we can write

$$e^{As} = \alpha_s + \beta_s A, \qquad e^{A^\top s} = \alpha_s + \beta_s A^\top,$$

i.e.

$$\mathbb{E}[X_0 X_0^\top] = \int_{-\infty}^0 (\alpha_s + \beta_s A)BB^\top(\alpha_s + \beta_s A^\top)ds$$
$$= \alpha BB^\top + \beta(ABB^\top + BB^\top A^\top) + \gamma ABB^\top A^\top \qquad \text{(SI.25)}$$
$$= \gamma(A + \frac{\beta}{\gamma})(BB^\top)(A^\top + \frac{\beta}{\gamma}) + (\alpha - \frac{\beta^2}{\gamma})BB^\top$$

for

$$\alpha = \int_{-\infty}^0 \alpha_s^2 ds, \qquad \beta = \int_{-\infty}^0 \alpha_s \beta_s ds, \qquad \gamma = \int_{-\infty}^0 \beta_s^2 ds.$$

We then write

$$
\begin{aligned}
BB^\top &= \int_{-\infty}^0 \frac{d}{ds} e^{As} BB^\top e^{A^\top s} ds = A\mathbb{E}[X_0 X_0^\top] + \mathbb{E}[X_0 X_0^\top] A^\top \\
&= \alpha(ABB^\top + BB^\top A) + 2\beta ABB^\top A^\top + \beta(A^2 BB^\top + BB^\top (A^\top)^2) \\
&\qquad\qquad\qquad\qquad + \gamma(A^2 BB^\top A^\top + ABB^\top (A^\top)^2) \\
&= (\alpha - \gamma(\det A))(ABB^\top + BB^\top A) + 2(\beta + \gamma(\mathrm{tr}A))ABB^\top A^\top \\
&\qquad\qquad + \beta(\mathrm{tr}A)(ABB^\top + BB^\top A^\top) - 2\beta(\det A)BB^\top \\
&= -2\beta(\det A)BB^\top + 2(\beta + \gamma(\mathrm{tr}A))ABB^\top A^\top \\
&\qquad\qquad + (\alpha + \beta(\mathrm{tr}A) - \gamma(\det A))(ABB^\top + BB^\top A),
\end{aligned}
$$

which is only possible if

$$
\begin{aligned}
1 + 2\beta(\det A) &= 0, \\
\beta + \gamma(\mathrm{tr}A) &= 0, \\
\alpha + \beta(\mathrm{tr}A) - \gamma(\det A) &= 0,
\end{aligned}
$$

i.e.

$$
\alpha = \frac{\det A + (\mathrm{tr}A)^2}{2(\det A)(\mathrm{tr}A)}, \qquad \beta = -\frac{1}{2\det A}, \qquad \gamma = \frac{1}{2(\det A)(\mathrm{tr}A)}
$$

Combining this with (SI.25) then gives the result. $\square$

Corollary F.1 (Diagonal matrix B). *Note that if*

$$
A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad B = \begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\rho} \end{pmatrix},
$$

then

$$
\mathbb{E}[X_0 X_0^\top] = \frac{\begin{pmatrix} \rho b^2 + \lambda(d \cdot \mathrm{tr}(A) - bc) & -(\rho ab + \lambda cd) \\ -(\rho ab + \lambda cd) & \rho(a \cdot \mathrm{tr}(A) - bc) + \lambda c^2 \end{pmatrix}}{2(ad - bc)(a + d)}.
$$

Proof. Indeed, $A - \mathrm{tr}(A)E_2 = \begin{pmatrix} -d & b \\ c & -a \end{pmatrix}$, so

$$
\begin{aligned}
(A - (\mathrm{tr}A)E_2)BB^\top (A^\top - (\mathrm{tr}A)E_2) &= \begin{pmatrix} -\lambda d & \rho b \\ \lambda c & -\rho a \end{pmatrix} \begin{pmatrix} -d & c \\ b & -a \end{pmatrix} \\
&= \begin{pmatrix} \rho b^2 + \lambda d^2 & -(\rho ab + \lambda cd) \\ -(\rho ab + \lambda cd) & \rho a^2 + \lambda c^2 \end{pmatrix}.
\end{aligned}
$$

$\square$

G   Martingale Central Limit Theorem

Here, we give a version of the martingale central limit theorem as given in Ethier and Kurtz (1986), Theorem 7.1.4.

Theorem 5. *Let (M_N) be a sequence of $\mathbb{R}^d$-valued martingales. Suppose that*

$$\lim_{N \to \infty} \mathbb{E}\left[\sup_{s \leq t} |M_N(s) - M_N(s-)|\right] = 0$$

and

$$[M_N^i, M_N^j]_t \to c_{ij}(t)$$

for all $t \geq 0$, where $C = (c_{ij})$ is deterministic and continuous. Then $M_N \Rightarrow M$ for a stochastic process M which is Gaussian with independent increments and $\mathbb{E}[M(t)M(t)^T] = C(t)$.

   Furthermore, since C is non-negative and symmetric and if C is absolutely continuous its derivative is non-negative and thus has a non-negative square-root. This implies that M can be written as

$$M(t) = \int_0^t \sqrt{\dot{C}(s)} dW(s),$$

where W is the d-dimensional standard Brownian motion.

References

Anderson, D. and T. G. Kurtz (2015). *Stochastic Analysis of Biochemical Systems.* Springer.

Ball, K., T. G. Kurtz, L. Popovic, and G. Rempala (2006). Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Probab. 16*(4), 1925–1961.

Dessalles, R., V. Fromion, and P. Robert (2016). A stochastic analysis of autoregulation of gene expression. *Preprint*, 1–27.

Ethier, S. N. and T. G. Kurtz (1986). *Markov processes: characterization and convergence.* Wiley series in probability and mathematical statistics. New York, Chichester: J. Wiley & Sons.

Gardiner, C. (2009). *Stochastic Methods. A Handbook for the Natural and Social Sciences.* Springer.

Kang, H.-W., T. Kurtz, and L. Popovic (2014). Central limit theorems and diffusion approximations for multiscale Markov chain models. *Ann. Appl. Probab. 24*, 721–759.

Paulsson, J. (2005). Models of stochastic gene expression. *Physics of Life Review 2*, 157–175.

Swain, P. S. (2004). Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *Journal of Molecular Biology 344*(4), 965–976.

Thattai, M. and A. van Oudenaarden (2001). Intrinsic noise in gene regulatory networks. *Proc. Nat. Acad. Sci. 98*(15), 8614–8619.