



## Limits of noise for autoregulated gene expression

Peter Czuppon<sup>1,2</sup>  · Peter Pfaffelhuber<sup>1</sup>

Received: 5 September 2017 / Revised: 6 May 2018 / Published online: 24 May 2018  
© The Author(s) 2018

**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 et al. in *Ann Appl Probab* 24:721–759, 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.

**Keywords** Intrinsic noise · Langevin approximation · Quasi-steady-state assumption · Chemical reaction network · Auto-regulated gene expression

**Mathematics Subject Classification** 92C40 · 60J60 · 60F05

---

✉ Peter Czuppon  
czuppon@evolbio.mpg.de

<sup>1</sup> Abteilung für Mathematische Stochastik, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany

<sup>2</sup> Present Address: Department of Evolutionary Theory, Max-Planck Institute for Evolutionary Biology, Plön, Germany

## 1 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 (Swain et al. 2002; Elowitz et al. 2002; Raj and van Oudenaarden 2008). 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 (Li and Xie 2011).

Usually, noise in gene expression is divided into an intrinsic and an extrinsic part (Swain et al. 2002; Raser and O’Shea 2005). 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 over time can be thought of having a constant environment, thus eliminating extrinsic noise (Singh and Soltani 2013; Singh 2014).

Stochasticity in gene expression is not only interesting per se. Today, its role in evolution, development and cell fate decisions is under discussion (Kaern et al. 2005; Maamar et al. 2007; Fraser and Kærn 2009; Eldar and Elowitz 2010; Balázsi et al. 2011; Silva-Rocha and Lorenzo 2010; Wang and Zhang 2011). For instance, noisy gene expression can be detrimental for the survival of cells under harsh conditions (Mitosch et al. 2017; Fraser and Kærn 2009). Still, many cells have to function constantly. Therefore, mechanisms reducing and controlling the level of noise are beneficial for most of real systems.

Under the central dogma of molecular biology, modeling stochasticity of gene expression is straight-forward (see Paulsson 2005 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 (Paulsson 2005). It is often stated that gene expression tends to occur in bursts, which occur due to the short life-time of the *on*-state of the gene and due to the short life-time of mRNA (Kumar et al. 2015).

We are interested in the effect of self-regulation on gene expression noise. It is known that a negative feedback loop, i.e. a protein suppressing its own transcription (or translation) leads to a reduced noise, while positive feedback is attributed to increase noise (Lestas et al. 2010; Hornung and Barkai 2008). Although these findings are widespread, a complete mathematical analysis is lacking. At least, for negative feedback, Thattai and Oudenaarden (2001) and in more generality Swain (2004) quantify the effect of negative feedback using a linearization argument. The latter paper further analyzes different feedback models differing between translational and transcriptional autoregulation. Moreover, Dessalles et al. (2017) derive the equilibrium distribution using a multi-scale approach under negative feedback.

Most analyses of noise in unregulated gene expression rely on the master equation (e.g. Paulsson 2005). By the linearity of this equation, a solution can be given explicitly.

Using the approximation that gene switching is so fast that it is effectively constantly transcribed to mRNA, this linearity can as well be used under negative feedback (Thattai and Oudenaarden 2001; Swain 2004). Our approach and also the one performed in Dessalles et al. (2017) differs in two ways. First, we are using martingale methods from stochastic analysis in order to describe the chemical system (Ethier and Kurtz 1986). Second, we can relax the assumption that the gene is transcribed effectively constantly, and therefore derive a more general result. Consequently, we are able to analyze noise in a truly non-linear system under a quasi-steady-state assumption.

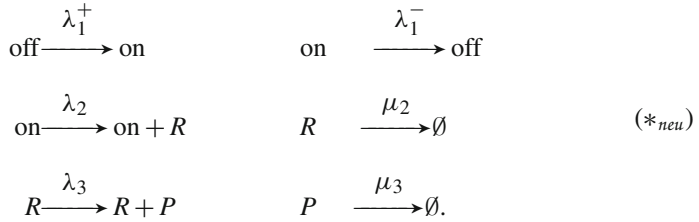
The explicit expression we derive for the noise in the number of proteins is also the main difference between our findings and the results obtained in Dessalles et al. (2017). Since the authors of that paper are interested in the case of not very abundant proteins they compute a stationary distribution for the protein using martingale techniques in the context of birth–death processes as opposed to our stochastic diffusion setting.

While the full model of regulated gene expression (or any other chemical reaction network) is usually hard to study, considering an ODE approach instead, which approximates the full model, leads to new insights. Formally, a law of large numbers—usually referred to as a fluid limit—can be obtained connecting the stochastic and deterministic model (Kurtz 1970b; Darling 2002). While such a law of large numbers gives a deterministic limit, fluctuations are studied using central limit results; see Kurtz (1970a). The special situation for gene expression is that the gene and mRNA only have a few copies, while the protein is often in large abundance. Such multi-scale models are often studied under a quasi-steady-state assumption (Seegel and Slemrod 1989). Here, the species in low abundance are assumed to evolve fast, such that the slow, abundant, species only sense their time-average. For such a stochastic averaging, not only a law of large numbers is given e.g. by Ball et al. (2006), but also a central limit result has recently been obtained by Kang et al. (2014).

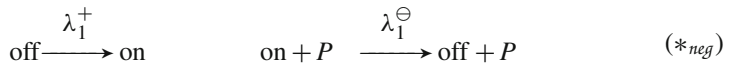
While a multi-scale approach to stochastic gene expression is not new (see Bokes et al. 2012; Dessalles et al. 2017), the analysis of fluctuations for such systems is not finished yet. In the case of multi-scale diffusion systems, Pardoux and Veretennikov (2001, 2003) derive a limit result for the slow components using a Poisson-equation. The results by Kang et al. (2014) are similar but are based on Markov jump processes instead of a diffusion limit framework. We apply the techniques of Kang et al. (2014) on the chemical reaction network of (un-)regulated gene expression. As our results show, fluctuations take into account all sources of noise and we give explicit formulas for the reduction of noise under negative feedback and the increase in noise under positive feedback.

## 2 The model

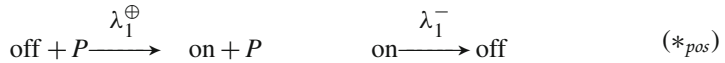
We are dealing with the standard model of gene expression without and with feedback; see e.g. Dessalles et al. (2017). Using the terminology from Paulsson (2005), we write for the model without feedback (or the *neutral* model)



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, and the third line gives translation and degradation of proteins. Exchanging the first line by



then models a negative feedback and



models a positive feedback. In all cases, we number the equations from left to right and from top to bottom by 1–6, so  $\mathcal{K} = \{1, \dots, 6\}$  is the set of chemical reactions. The species counts are given by  $X_i$  for  $i \in \mathcal{S} := \{\text{off}, \text{on}, R, P\}$  for inactive and active gene, mRNA and protein, respectively. In the following we will scale the rates such that the gene switching and the mRNA production happens on a fast time-scale whereas the protein which is also present in higher abundances than the mRNA is evolving on a slower time-scale. This time-scale separation is frequently used in quantitative analyses of gene expression (see for instance Thattai and Oudenaarden 2001; Swain 2004; Ball et al. 2006; Bokes et al. 2012; Dessalles et al. 2017) since it allows to employ a quasi-steady-state assumption for the species evolving on the fast time-scale, cf. Kuehn (2015). It basically means that one first solves for the stationary points in the fast sub-system which are then used to describe the dynamics in the slow sub-system.

Thus, using some large constant  $N$ , we will make use of the following scaling for the abundances of chemical species

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

or  $X_i = O(N^{\alpha_i})$  for  $i \in \mathcal{S}$  with

$$\alpha_{\text{off}} = \alpha_{\text{on}} = \alpha_R = 0, \quad \alpha_P = 1. \tag{1}$$

Reactions are scaled for all models such that we indeed get a time-scale separation, i.e. reaction constants are such that genes and mRNAs evolve much faster than protein numbers. Note, however, that due to  $X_P = O(N)$ , we need that the protein production rate needs to scale with  $N$ . We use the scaled rates  $\kappa_2, \nu_2, \kappa_3, \nu_3 = O(1)$ , which are given through

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

For the neutral model, we also set (with  $\kappa_1^+, \kappa_1^- = O(1)$ )

$$\lambda_1^+ = N\kappa_1^+, \quad \lambda_1^- = N\kappa_1^-, \tag{\square_{neu}}$$

whereas for negative feedback (with  $\kappa_1^+, \kappa_1^\ominus = O(1)$ )

$$\lambda_1^+ = N\kappa_1^+, \quad \lambda_1^\ominus = \kappa_1^\ominus, \tag{\square_{neg}}$$

and for positive feedback (with  $\kappa_1^\oplus, \kappa_1^- = O(1)$ )

$$\lambda_1^\oplus = \kappa_1^\oplus, \quad \lambda_1^- = N\kappa_1^-. \tag{\square_{pos}}$$

(Note that these scalings obey  $\lambda_1^-, \lambda_1^\ominus X_P, \lambda_1^+, \lambda_1^\oplus X_P = O(N)$  which is necessary for the time-scale separation.) Setting (with  $\alpha_i$  from (1))

$$V_i^N = N^{-\alpha_i} X_i, \quad i \in \mathcal{S} \tag{2}$$

as the scaled number of genes, mRNA molecules and proteins, respectively,  $V_i^N(0)$  the corresponding initial value and for  $M$  denoting the total copy number of genes, we have in the neutral case

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

for independent, rate 1 Poisson processes  $Y_1, \dots, Y_6$ . (See (Anderson and Kurtz 2015) for a general introduction on theoretical chemical reaction networks.) The first equation changes in the case of negative feedback to

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

**Table 1** We use  $N$  as a scaling parameter throughout. Reaction rates either come in unscaled ( $\lambda$ 's and  $\mu$ 's) or in scaled ( $\kappa$ 's and  $\nu$ 's) versions

Parameter	Model	Meaning	Scaling
$N$	$[neu], [neg], [pos]$	Scaling parameter	
$\lambda_1^+, \kappa_1^+$	$[neu], [neg]$	Rate of switching genes on	$\lambda_1^+ = \kappa_1^+ N$
$\lambda_1^\oplus, \kappa_1^\oplus$	$[pos]$		$\lambda_1^\oplus = \kappa_1^\oplus$
$\lambda_1^-, \kappa_1^-$	$[neu], [pos]$	Rate of switching genes off	$\lambda_1^- = \kappa_1^- N$
$\lambda_1^\ominus, \kappa_1^\ominus$	$[neg]$		$\lambda_1^\ominus = \kappa_1^\ominus$
$\lambda_2, \kappa_2$	$[neu], [neg], [pos]$	Rate of mRNA production	$\lambda_2 = \kappa_2 N$
$\mu_2, \nu_2$	$[neu], [neg], [pos]$	Rate of mRNA degradation	$\mu_2 = \nu_2 N$
$\lambda_3, \kappa_3$	$[neu], [neg], [pos]$	Rate of protein production	$\lambda_3 = \kappa_3 N$
$\mu_3, \nu_3$	$[neu], [neg], [pos]$	Rate of protein degradation	$\mu_3 = \nu_3$
$M$	$[neu], [neg], [pos]$	Total number of genes	$M = O(1)$

and in the case of positive feedback to

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

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

For a summary of the parameters and their scalings, see Table 1. We will refer to  $\lambda$ 's and  $\mu$ 's as the unscaled parameters and to  $\kappa$ 's and  $\nu$ 's as the scaled parameters.

### 3 Results

The following results are all stated in terms of the scaled parameters ( $\kappa$ 's and  $\nu$ 's and  $V_P^N$ ). For the corresponding formulas using unscaled parameter notation, see Appendix E.

#### 3.1 A limiting process for the amount of protein

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). Basically, it says that, using the law of large numbers for Poisson processes, i.e.  $Y(t) \approx t$  for large  $t$ , and the scaled parameter set we find

$$V_P^N(t) \approx V_P^N(0) + \int_0^t \left( \kappa_3 \mathbb{E}_\pi \left[ V_R^N(s) \right] - v_3 V_P^N(s) \right) ds, \tag{3}$$

where  $\mathbb{E}_\pi[\cdot]$  denotes expectation with respect to the equilibrium dynamics of the fast species  $V_{\text{off}}^N$ ,  $V_{\text{on}}^N$  and  $V_R^N$  for a fixed amount of protein,  $V_P^N$ . Note that we use  $\Rightarrow$  for weak convergence of stochastic processes (Ethier and Kurtz 1986).

**Theorem 1** (Law of Large Numbers) *We consider the models [neu], [neg] and [pos] and assume that the initial condition converges in distribution, i.e.  $V_P^N(0) \xrightarrow{N \rightarrow \infty} v_P(0)$ . Then,  $V_P^N \xrightarrow{N \rightarrow \infty} v_P$ , where  $v_P$  solves*

$$\dot{v}_P = F(v_P),$$

with

$$F(v_P) = \begin{cases} \frac{M\kappa_1^+ \kappa_2 \kappa_3}{v_2 (\kappa_1^- + \kappa_1^+)} - v_3 v_P & \text{for [neu],} \\ \frac{M\kappa_1^+ \kappa_2 \kappa_3}{v_2 (\kappa_1^\ominus v_P + \kappa_1^+)} - v_3 v_P & \text{for [neg],} \\ \frac{M\kappa_1^\oplus \kappa_2 \kappa_3 v_P}{v_2 (\kappa_1^- + \kappa_1^\oplus v_P)} - v_3 v_P & \text{for [pos].} \end{cases} \tag{4}$$

In particular, the equilibrium, i.e.  $F(v_P^*) = 0$ , is given by

$$v_P^* = \begin{cases} \frac{M\kappa_1^+ \kappa_2 \kappa_3}{v_2 v_3 (\kappa_1^- + \kappa_1^+)} & \text{for [neu],} \\ \frac{\kappa_1^+}{2\kappa_1^\ominus} \left( \sqrt{1 + \frac{4M\kappa_1^\ominus \kappa_2 \kappa_3}{\kappa_1^+ v_2 v_3}} - 1 \right) & \text{for [neg],} \\ 0 \vee \frac{M\kappa_1^\oplus \kappa_2 \kappa_3 - \kappa_1^- v_2 v_3}{\kappa_1^\oplus v_2 v_3} & \text{for [pos].} \end{cases} \tag{5}$$

*Proof* We apply results from Ball et al. (2006) and only sketch the proof. In order to derive Eq. (4) we need to replace  $V_{\text{on}}$  and  $V_R$  (the fast variables) in the equation for  $V_P^N$  by their equilibria assuming that  $V_P^N$  is constant. Computing these equilibria is done using the corresponding lines in  $(\bullet_{\text{neu}})$ ,  $(\bullet_{\text{neg}})$  and  $(\bullet_{\text{pos}})$ . The resulting distribution  $\pi$  then is the equilibrium on the fast time-scale. For  $v_P$  fixed they read

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

and

$$\mathbb{E}_\pi[V_R] = \frac{\kappa_2}{v_2} \mathbb{E}_\pi[V_{on}] = \begin{cases} \frac{M\kappa_1^+\kappa_2}{v_2(\kappa_1^- + \kappa_1^+)} & \text{for [neu],} \\ \frac{M\kappa_1^+\kappa_2}{v_2(\kappa_1^\ominus v_P + \kappa_1^+)} & \text{for [neg],} \\ \frac{M\kappa_1^\oplus\kappa_2 v_P}{v_2(\kappa_1^- + \kappa_1^\oplus v_P)} & \text{for [pos].} \end{cases}$$

Plugging this equilibrium into (3) which is the limit for large  $N$  of the corresponding equation in (•*neu*), we obtain that

$$\dot{v}_P = \kappa_3 \mathbb{E}_\pi[V_R] - v_3 v_P = F(v_P) \tag{6}$$

with  $F$  as in (4). Computation of the equilibria is standard by solving  $F(v_P) = 0$ . In particular, we have to solve

$$-M\kappa_1^+\kappa_2\kappa_3 + \kappa_1^+ v_2 v_3 v_P + \kappa_1^\ominus v_2 v_3 v_P^2 = 0 \quad \text{or} \quad v_P^2 + \frac{\kappa_1^+}{\kappa_1^\ominus} v_P - \frac{M\kappa_1^+\kappa_2\kappa_3}{\kappa_1^\ominus v_2 v_3} = 0$$

for the equilibrium of [*neg*]. □

### 3.2 Approximate variance and Fano factor for the amount of protein

Our goal is to derive the variance in protein numbers under [*neu*], [*neg*] and [*pos*]. While [*neu*] is solved explicitly elsewhere, e.g. in Paulsson (2005), some approximations have to be made for [*neg*] and [*pos*]. One idea might be to use a Langevin approximation and write

$$\begin{aligned} V_P^N(t) &\approx V_P^N(0) + \int_0^t (\kappa_3 V_R^N(s) - v_3 V_P^N(s)) ds \\ &\quad + \frac{1}{\sqrt{N}} \int_0^t \sqrt{\kappa_3 V_R^N(s) + v_3 V_P^N(s)} dW_s \\ &\approx V_P^N(0) + \int_0^t F(V_P^N(s)) ds + \frac{1}{\sqrt{N}} \int_0^t \sqrt{b(V_P^N(s))} dW_s \end{aligned}$$

with

$$b(V_P) = \kappa_3 \mathbb{E}_\pi[V_R] + v_3 V_P = \begin{cases} \frac{M\kappa_1^+\kappa_2\kappa_3}{v_2(\kappa_1^- + \kappa_1^+)} + v_3 V_P & \text{for [neu],} \\ \frac{M\kappa_1^+\kappa_2\kappa_3}{v_2(\kappa_1^\ominus v_P + \kappa_1^+)} + v_3 V_P & \text{for [neg],} \\ \frac{M\kappa_1^\oplus\kappa_2\kappa_3 V_P}{v_2(\kappa_1^- + \kappa_1^\oplus v_P)} + v_3 V_P & \text{for [pos].} \end{cases}$$



Comparing  $V_P^N$  and  $v_P$ , where  $v_P$  is the exact solution of  $\dot{v}_P = F(v_P) = \kappa_3 \mathbb{E}_\pi[V_R] - v_3 v_P$  with  $F$  from Theorem 1, we assume that  $V_P^N \approx v_P + \frac{1}{\sqrt{N}}U$  for some stochastic process  $U$ . The random process  $U$  will then account for the fluctuations which are not captured by the deterministic approximation above. Hence,

$$\begin{aligned} \frac{1}{\sqrt{N}}dU &= dV_P^N - dv_P \approx \left( F\left(V_P^N\right) - F\left(v_P\right) \right) dt + \frac{1}{\sqrt{N}}\sqrt{b\left(V_P^N\right)}dW \\ &\approx F'\left(v_P\right)\left(V_P^N - v_P\right) dt + \frac{1}{\sqrt{N}}\sqrt{b\left(V_P^N\right)}dW \\ &\approx \frac{1}{\sqrt{N}}\left( F'\left(v_P\right)U dt + \sqrt{b\left(V_P^N\right)}dW \right) \end{aligned}$$

and therefore

$$dV_P^N = dv_P + \frac{1}{\sqrt{N}}dU \approx \left( F\left(v_P\right) + \frac{1}{\sqrt{N}}F'\left(v_P\right)U \right) dt + \frac{1}{\sqrt{N}}\sqrt{b\left(V_P^N\right)}dW. \tag{7}$$

This approach builds on applying a quasi-steady-state assumption whenever possible, i.e. when averaging over the on/off-state of genes in order to derive the deterministic dynamics of  $v_P$  using  $F$ , and the number of mRNA, which is approximated by its mean in order to derive  $b$ . Consequently, fluctuations arising from these two mechanisms cannot be accounted for in the resulting variance. As a result, fluctuations read off from (7) will be too small.

In contrast, as an application of Kang et al. (2014) (see also Appendix A), we derive the following central limit result, which takes into account all fluctuations in leading order. Precisely, our next goal is to show that  $\sqrt{N}(V_P^N - v_P)$  converges and to determine the limiting process. This limit will then provide the error due to noise between the deterministic approximation  $v_P$  and the stochastic process  $V_P^N$  of order  $\sqrt{N}$ . In the proof, we will make use of the method developed by Kang et al. (2014).

**Theorem 2** (Central Limit Theorem) *Let  $V_P^N, v_P$  and  $F$  be as in Theorem 1 and assume further weak convergence of the initial conditions:*

$$\sqrt{N}\left(V_P^N(0) - v_P(0)\right) \xrightarrow{N \rightarrow \infty} U(0).$$

*Then, for the models [neu], [neg] and [pos],  $\sqrt{N}(V_P^N - v_P) \xrightarrow{N \rightarrow \infty} U$ , where  $U$  solves*

$$U(t) = U(0) + \int_0^t \sqrt{c(v_P(s))}dW(s) + \int_0^t F'(v_P(s))U(s)ds, \tag{8}$$

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

$$c(v_P) = \begin{cases} \frac{M\kappa_1^+\kappa_2\kappa_3}{v_2(\kappa_1^- + \kappa_1^+)} \left( \frac{2\kappa_1^-\kappa_2\kappa_3}{v_2(\kappa_1^- + \kappa_1^+)^2} + \frac{2\kappa_3}{v_2} + 1 \right) + v_3v_P & \text{for [neu]}, \\ \frac{M\kappa_1^+\kappa_2\kappa_3}{v_2(\kappa_1^\ominus v_P + \kappa_1^+)} \left( \frac{2\kappa_1^\ominus v_P \kappa_2 \kappa_3}{v_2(\kappa_1^\ominus v_P + \kappa_1^+)^2} + \frac{2\kappa_3}{v_2} + 1 \right) + v_3v_P & \text{for [neg]}, \\ \frac{M\kappa_1^\oplus \kappa_2 \kappa_3 v_P}{v_2(\kappa_1^- + \kappa_1^\oplus v_P)} \left( \frac{2\kappa_1^-\kappa_2\kappa_3}{v_2(\kappa_1^- + \kappa_1^\oplus v_P)^2} + \frac{2\kappa_3}{v_2} + 1 \right) + v_3v_P & \text{for [pos]}. \end{cases} \tag{9}$$

Hence, we see that, in contrast to the Langevin approach (7) above, fluctuations arising from gene switching and RNA dynamics are also accounted for in Theorem 2; see also Sect. 3.3 for an interpretation of the individual terms. For the difference between the Langevin approximation and our result and its implications see also Sect. 4.3.

The proof of the Theorem is given in Appendix B. Briefly, we apply the stochastic averaging principle on multiple time scales developed in Kang et al. (2014). The whole approach is revisited in Appendix A. There we also state the conditions which need to be satisfied for the theory to apply. Amongst others these include solving a certain Poisson equation which enables a clean time-scale separation.

*Remark 1 (Deriving the Fano factor in equilibrium)* While (8) provides a dynamical result along paths of  $X_P$ , we can also use this approximation and study the process in equilibrium by setting  $X_P(0) = v_P^*N$ , where  $v_P^*$  is the unique solution of  $F(v_P) = 0$  given by (5).

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 stochastic differential equation (SDE) in (8) is solved by an Ornstein–Uhlenbeck process. In particular, we obtain at late times, i.e. when the process reached its equilibrium (recall that  $X_P^N = NV_P^N$  is the total number of proteins)

$$\frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} \approx \frac{N\mathbb{V}[V_P^N]}{\mathbb{E}[V_P^N]} \approx \frac{\mathbb{V}[U]}{v_P^*} \approx -\frac{c(v_P^*)}{2F'(v_P^*)v_P^*} \tag{10}$$

with  $F(\cdot)$  from (4),  $v_P^*$  from (5) and  $c(\cdot)$  from (9). Since no factor  $N$  appears on the right hand side, some authors call the Fano factor dimensionless. Empirically, it was found e.g. by Bar-Even et al. (2006), that for all classes of genes and under all conditions, the variance in protein numbers was approximately proportional to the mean, which is again reminiscent of the lacking  $N$  in the Fano factor above.

We note here that this approach of computing the Fano factor of  $X_P^N$  in equilibrium was achieved by an unjustified exchange of limits. Namely, for the approximate Fano factor of  $X_P^N$  in equilibrium, we would have to perform  $t \rightarrow \infty$  first, and only then compute  $N \rightarrow \infty$ , but our approach exchanged this limit.

In order to compute the right hand side of (10), note that

$$F'(v_P) = \begin{cases} -v_3 & \text{for [neu],} \\ -\frac{M\kappa_1^+\kappa_2\kappa_3}{v_2} \frac{\kappa_1^\ominus}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} - v_3 & \text{for [neg],} \\ \frac{M\kappa_1^\oplus\kappa_2\kappa_3}{v_2} \frac{\kappa_1^-}{(\kappa_1^- + \kappa_1^\oplus v_P)^2} - v_3 & \text{for [pos].} \end{cases} \quad (11)$$

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

$$v_P^* = \begin{cases} \frac{M\kappa_1^+\kappa_2\kappa_3}{v_2 v_3 (\kappa_1^\ominus v_P^* + \kappa_1^+)} & \text{for [neg],} \\ \frac{M\kappa_1^\oplus v_P^* \kappa_2 \kappa_3}{v_2 v_3 (\kappa_1^- + \kappa_1^\oplus v_P^*)} & \text{for [pos]} \end{cases}$$

and therefore

$$F'(v_P^*) = \begin{cases} -v_3 \left( 1 + \frac{\kappa_1^\ominus v_P^*}{\kappa_1^\ominus v_P^* + \kappa_1^+} \right) & \text{for [neg],} \\ -v_3 \left( 1 - \frac{\kappa_1^-}{\kappa_1^- + \kappa_1^\oplus v_P^*} \right) & \text{for [pos].} \end{cases} \quad (12)$$

In addition,

$$c(v_P^*) = \begin{cases} \frac{2M\kappa_1^+\kappa_2\kappa_3}{v_2 (\kappa_1^- + \kappa_1^+)} \left( \frac{\kappa_1^- \kappa_2 \kappa_3}{v_2 (\kappa_1^- + \kappa_1^+)^2} + \frac{\kappa_3}{v_2} + 1 \right) & \text{for [neu],} \\ \frac{2M\kappa_1^+\kappa_2\kappa_3}{v_2 (\kappa_1^\ominus v_P^* + \kappa_1^+)} \left( \frac{\kappa_1^\ominus v_P^* \kappa_2 \kappa_3}{v_2 (\kappa_1^\ominus v_P^* + \kappa_1^+)^2} + \frac{\kappa_3}{v_2} + 1 \right) & \text{for [neg],} \\ \frac{2M\kappa_1^\oplus \kappa_2 \kappa_3 v_P^*}{v_2 (\kappa_1^- + \kappa_1^\oplus v_P^*)} \left( \frac{\kappa_1^- \kappa_2 \kappa_3}{v_2 (\kappa_1^- + \kappa_1^\oplus v_P^*)^2} + \frac{\kappa_3}{v_2} + 1 \right) & \text{for [pos].} \end{cases} \quad (13)$$

Hence, plugging these quantities into (10) gives (with  $X_P^* = N v_P^*$ ; see also (42) for the Fano factor in terms of unscaled parameters)

$$\frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} \approx \begin{cases} \frac{\kappa_1^- \kappa_2 \kappa_3}{v_2 (\kappa_1^- + \kappa_1^+)^2} + \frac{\kappa_3}{v_2} + 1 & \text{for [neu],} \\ \left( \frac{\kappa_1^\ominus v_P^* \kappa_2 \kappa_3}{v_2 (\kappa_1^- v_P^* + \kappa_1^+)^2} + \frac{\kappa_3}{v_2} + 1 \right) / \left( 1 + \frac{\kappa_1^\ominus v_P^*}{\kappa_1^\ominus v_P^* + \kappa_1^+} \right) & \text{for [neg],} \\ \left( \frac{\kappa_1^- \kappa_2 \kappa_3}{v_2 (\kappa_1^- + \kappa_1^\oplus v_P^*)^2} + \frac{\kappa_3}{v_2} + 1 \right) / \left( 1 - \frac{\kappa_1^- v_2 v_3}{M \kappa_1^\oplus \kappa_2 \kappa_3} \right) & \text{for [pos].} \end{cases} \tag{14}$$

### 3.3 Interpretation of the Fano factor

The expressions above in Eq. (14) are not only a result from strict calculations but can also be interpreted in biological terms. For instance, for negative feedback, we find—as in the neutral case—contributions from randomness in gene switching, translation and transcription. 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

$$\frac{\mathbb{V}_{neg}[X_P]}{\mathbb{E}_{neg}[X_P]} = \left( \underbrace{\frac{\kappa_1^\ominus v_P^* \kappa_2 \kappa_3}{v_2 (\kappa_1^\ominus v_P^* + \kappa_1^+)^2}}_{\text{effect of RNA noise, originating from switching in gene activity}} + \underbrace{\frac{\kappa_3}{v_2}}_{\text{effect of individual RNA noise}} + \underbrace{1}_{\text{protein births and deaths}} \right) / \underbrace{\left( 1 + \frac{\kappa_1^\ominus v_P^*}{\kappa_1^\ominus v_P^* + \kappa_1^+} \right)}_{\text{fast back pushing due to negative feedback}}, \tag{15}$$

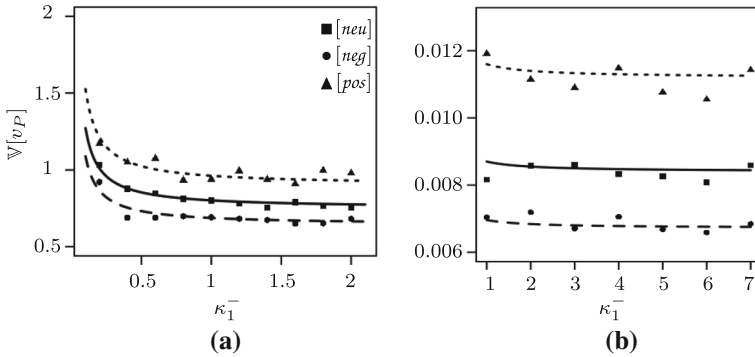
which has the biggest noise-reducing effect of negative feedback which we will study hereafter.

Adjusted explanations hold in the cases of no or positive feedback.

### 3.4 Comparing the noise in [neu], [neg] and [pos]

It is frequently reported that a negative feedback in gene expression results in a reduced variance (noise) of protein levels, whereas a positive feedback enhances noise (Lestas et al. 2010; Hornung and Barkai 2008). These observations can be made precise by our results from above. Here, we report some consequences on the equilibrium variance and the Fano factor,  $\mathbb{V}[X_P]/\mathbb{E}[X_P]$ .

For a fair comparison, we use the models [neu], [neg] and [pos] for equal values of  $v_P^*$ . Consider a model [neu] with parameters  $\kappa_1^+$ ,  $\kappa_1^-$ ,  $\kappa_2$ ,  $\kappa_3$ ,  $v_2$ ,  $v_3$  and let  $v_P^*$  be the equilibrium from (5). 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  $\kappa_1^\oplus := \kappa_1^+/v_P^*$  and all other



**Fig. 1** Simulations and theoretical results with a fixed mean of proteins, **a**  $Nv_P^* = 1250$ , **b**  $Nv_P^* = 60$ . The gene association and dissociation rates are varied, i.e. in **(a)**  $\kappa_1^- = (0.2, 0.4, \dots, 2)$  and in **(b)**  $\kappa_1^- = (1, 2, \dots, 7)$ . The gene association rate  $\kappa_1^+$  is then chosen such that the protein mean equals 1250 or 60 in each case, respectively. Furthermore, these rates are adjusted in the cases of negative and positive feedback according to  $\kappa_1^\ominus := \kappa_1^-/v_P^*$  and  $\kappa_1^\oplus := \kappa_1^+/v_P^*$ , respectively. The other parameters are given by  $M = 1, N = 100, \kappa_2 = 3, v_2 = 1, \kappa_3 = 5$  and  $v_3 = 1$  in **(a)**. For figure **(b)** we chose parameters as in (Anderson and Kurtz 2015, Figure 2.1), i.e.  $M = 1, N = 100, \kappa_2 = 2, v_2 = 0.25, \kappa_3 = 0.1, v_3 = 1$ . The solid, dotted, dashed lines are the theoretical predictions in the no, positive and negative feedback cases, respectively. Each data point is derived from 1000 Monte Carlo simulations (cf. Gillespie 1977) of the full system given by  $[neu], [neg]$  and  $[pos]$

parameters as above. Then, from (4), we see that all models have  $v_P^*$  as their unique deterministic limit with the same  $c(v_P^*)$  from (13). Setting  $\mathbb{V}_{neu}, \mathbb{V}_{neg}$  and  $\mathbb{V}_{pos}$  as the variance of the model without, with negative and with positive feedback, respectively, and plugging in all quantities in (14) then gives (note that the mean cancels out; see (43) for a version with unscaled parameters)

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

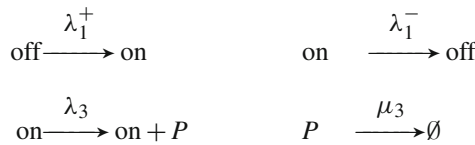
$$\frac{\mathbb{V}_{pos}[X_P^N]}{\mathbb{V}_{neu}[X_P^N]} \approx \frac{\mathbb{V}_{pos}[V_P^N]}{\mathbb{V}_{neu}[V_P^N]} \approx \left( 1 - \frac{\kappa_1^-}{\kappa_1^- + \kappa_1^\oplus v_P^*} \right)^{-1}.$$
(16)

In particular, we see that the variance is reduced in  $[neg]$  and increased in  $[pos]$ , as expected; see also Fig. 1. Moreover, the graphs show that the performed simulations fit our predictions well for both higher (a) and lower (b) values of  $v_P^*$ .

Additionally, we see from Eq. (16) that the change in noise is maximal if the gene is *off* most of the time, while still having the same amount of protein as in the unregulated (neutral) case. This finding is reminiscent of the fact that gene expression comes in bursts. The burstiness is most extreme if the gene is *on* only for a short time, producing a large amount of mRNA, and afterwards *off* for a long period. Especially, we see that for  $\kappa_1^\ominus \rightarrow \infty$  the maximal reduction due to negative feedback is twofold while the increase in noise is unbounded for  $\kappa_1^- \rightarrow \infty$  in case of positive auto-regulation of the protein.

### 3.5 Negative feedback in a simpler model

In the literature simpler models of gene expression are studied as well. Here, only two molecular species are involved. Either, the gene is constitutively expressed, therefore ignoring the state of the gene, or translation is neglected and the gene is assumed to be transcribed leading to protein in one step (Hornos et al. 2005; Ramos et al. 2015, 2011; Shahrezaei and Swain 2008). In either case, we have the model similar to



for  $[neu]$ , whereas for  $[neg]$  and  $[pos]$ , we take  $(*_{neg})$  and  $(*_{pos})$  instead of the first line, respectively. We note that this model arises from the full model described in  $(*_{neu})$ ,  $(*_{neg})$  and  $(*_{pos})$ , when letting  $v_2 = \kappa_2 \rightarrow \infty$ . Hence, we obtain the following approximation for the simpler model

$$\frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} = \begin{cases} \frac{\kappa_1^- \kappa_3}{(\kappa_1^- + \kappa_1^+)^2} + 1 & \text{for } [neu], \\ \left( \frac{\kappa_1^\ominus v_p^* \kappa_3}{(\kappa_1^\ominus v_p^* + \kappa_1^+)^2} + 1 \right) / \left( 1 + \frac{\kappa_1^\ominus v_p^*}{\kappa_1^\ominus v_p^* + \kappa_1^+} \right) & \text{for } [neg], \\ \left( \frac{\kappa_1^- \kappa_3}{(\kappa_1^- + \kappa_1^\oplus v_p^*)^2} + 1 \right) / \left( 1 - \frac{\kappa_1^-}{\kappa_1^- + \kappa_1^\oplus v_p^*} \right) & \text{for } [pos]. \end{cases} \quad (17)$$

When comparing these equations with (15) we see that all the terms containing explicit mRNA noise disappeared. Thus, we still have contributions from gene activation and protein production and degradation as well as the scaling factor due to the auto-regulation.

### 3.6 Refining the Fano factor

Here, we again consider the original model given by equations  $(*_{neu})$ ,  $(*_{neg})$  and  $(*_{pos})$ , but use a different scaling for the mRNA. We assume that not only the protein evolves on the slow time-scale but also the mRNA production and degradation. This is a slightly more complex model since we cannot average the number of mRNA molecules when analyzing the protein fluctuations. This model allows us to compare our results in a more straightforward way with results obtained previously in Thattai and Oudenaarden (2001), Swain (2004); see Sects. 4.2 and 4.3.

In order to account for the new scaling we set  $\mu_2 = O(1)$  and  $\lambda_3 = O(1)$ . As scaled variables, we introduce  $\tilde{v}_2 = \mu_2$  and  $\tilde{\kappa}_3 = \lambda_3$ . In this model, we have  $X_R = O(N)$ , but still  $X_P = O(N)$ . With these assumptions it is possible to derive a refined formula for the Fano factor for  $X_P$  in equilibrium. Precisely, we compute in Appendix D—see Eqs. (37), (39) and (40)

$$\frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} \approx \begin{cases} \frac{\kappa_1^- \kappa_2 \tilde{\kappa}_3}{(\kappa_1^- + \kappa_1^+)^2 (\tilde{v}_2 + v_3)} + \frac{\tilde{\kappa}_3}{\tilde{v}_2 + v_3} + 1 & \text{for } [neu], \\ \left( \alpha_{neg} \frac{v_3}{\tilde{v}_2 + v_3} + \frac{\alpha_{neg}^2 \kappa_2 \tilde{\kappa}_3}{\kappa_1^\ominus v_P^* (\tilde{v}_2 + v_3)} + \frac{\tilde{\kappa}_3}{\tilde{v}_2 + v_3} + 1 \right) / (1 + \alpha_{neg}) & \text{for } [neg], \\ \left( -\alpha_{pos} \frac{v_3}{\tilde{v}_2 + v_3} + \frac{\alpha_{pos}^2 \kappa_2 \tilde{\kappa}_3}{\kappa_1^- (\tilde{v}_2 + v_3)} + \frac{\tilde{\kappa}_3}{\tilde{v}_2 + v_3} + 1 \right) / (1 - \alpha_{pos}) & \text{for } [pos], \end{cases} \tag{18}$$

with

$$\alpha_{neg} = \frac{\kappa_1^\ominus v_P^*}{\kappa_1^\ominus v_P^* + \kappa_1^+}, \quad \alpha_{pos} = \frac{\kappa_1^-}{\kappa_1^- + \kappa_1^\oplus v_P^*}.$$

Note that this equation approximately gives (14) for  $\tilde{v}_2, \tilde{\kappa}_3 \gg 1$ . However, adding interpretations as in Sect. 3.3 is not straight-forward since the additional terms in (18) stem from interactions between RNA and protein dynamics. In practice (and in our simulations below), the life-time of proteins is much larger than the life-time of mRNA, such that (18) does not produce a better fit than (14); see also Fig. 2 and compare the dash-dotted and the solid line. Therefore, when not stated otherwise, we will use (14) in the sequel.

### 4 Comparison to previous results

Here, we compare our results in the neutral case with those obtained in Paulsson (2005), and in the case of negative feedback with the formulas for the Fano factor derived in Dessalles et al. (2017), Ramos et al. (2015), Swain (2004) and Thattai and Oudenaarden (2001).

#### 4.1 The neutral case, Paulsson (2005)

In Paulsson (2005) the neutral model  $[neu]$  was studied without assuming any scalings of the parameter  $\lambda_i, \mu_i$  or of the number of mRNA molecules or proteins. Setting

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

as the expected life-times of a change in gene activity, mRNA and protein, respectively, we see that the Fano factor in equilibrium obeys (see equation (4) in Paulsson (2005))

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

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

The approximation in the last line corresponds to our scaling, which is exactly such that  $\tau_3 \gg \tau_1, \tau_2$  (since  $\mu_3 \ll \mu_1, \mu_2$ ), i.e. the protein is much more stable than mRNA and the state of the gene. Thus, our approximation of the Fano factor in Eq. (14) (or rather its version in unscaled parameters in Eq. (42)) is in line with (Paulsson 2005).

### 4.2 Negative feedback, Thattai and Oudenaarden (2001)

In Thattai and Oudenaarden (2001), a linearization of [neg] was studied in the case of fast on- and off-switching of the gene. In particular, this will mean that both,  $\kappa_1^-, \kappa_1^+ \gg 1$ . In the following, we derive their result within our framework. Therefore, 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, averaging out the gene state, we obtain the following system (compare with ( $\bullet_{\text{neg}}$ ))

$$V_R^N(t) = V_R^N(0) + Y_3 \left( N \int_0^t \kappa_2 \frac{M \kappa_1^+}{\kappa_1^\ominus V_P^N(s) + \kappa_1^+} ds \right) - Y_4 \left( N \int_0^t v_2 V_R^N(s) ds \right),$$

$$V_P^N(t) = V_P^N(0) + N^{-1} Y_5 \left( N \int_0^t \kappa_3 V_R^N(s) ds \right) - N^{-1} Y_6 \left( N \int_0^t v_3 V_P^N(s) ds \right).$$

(19)

Assuming, like in Thattai and Oudenaarden (2001), that the equilibrium effective mRNA-production is a linear function, i.e. letting

$$\kappa_1^\ominus v_P^* \ll \kappa_1^+, \tag{20}$$

such that we basically model a constitutively expressing gene, we can further approximate  $V_R^N$  by

$$V_R^N(t) = V_R^N(0) + Y_3 \left( N \int_0^t M \kappa_2 \left( 1 - \frac{\kappa_1^\ominus}{\kappa_1^+} V_P^N(s) \right) ds \right) - Y_4 \left( N \int_0^t v_2 V_R^N(s) ds \right).$$

(21)



Now, the system (19) and (21) is exactly as on p. 3 in Thattai and Oudenaarden (2001) with

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

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

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

that (note that  $\eta$  is negligible since  $N$  is large and  $\phi$  is small by (8))

$$\begin{aligned} \frac{1}{N}\mathbb{E}[X_P] &= \frac{1}{N} \left( \frac{1}{1+b\phi} \right) \frac{k_0 b}{\gamma_P} \approx \left( 1 - \frac{M\kappa_1^\ominus \kappa_2 \kappa_3}{\kappa_1^+ \nu_2 \nu_3} \right) \frac{M\kappa_2 \kappa_3}{\nu_2 \nu_3} \\ \frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} &= \left( \frac{1-\phi}{1+b\phi} \cdot \frac{b}{1+\eta} + 1 \right) \approx 1 + b(1-\phi)(1-b\phi) \\ &\approx (1+b)(1-b\phi) = \left( 1 + \frac{\kappa_3}{\nu_2} \right) \left( 1 - M \frac{\kappa_2 \kappa_3}{\nu_2 \nu_3} \frac{\kappa_1^\ominus}{\kappa_1^+} \right) \\ &\approx \left( 1 + \frac{\kappa_3}{\nu_2} \right) \left( 1 - \frac{\kappa_1^\ominus}{\kappa_1^+} v_P^* \right). \end{aligned} \tag{22}$$

Considering our scaling assumption, and further assuming (20), Eq. (14) can be simplified to

$$\begin{aligned} \frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} &\approx \left( \frac{\kappa_1^\ominus v_P^* \kappa_2 \kappa_3}{(\kappa_1^+ + \kappa_1^\ominus v_P^*)^2 \nu_2} + \frac{\kappa_3}{\nu_2} + 1 \right) \left( 1 + \frac{\kappa_1^\ominus v_P^*}{\kappa_1^\ominus v_P^* + \kappa_1^+} \right)^{-1} \\ &\approx \left( 1 + \frac{\kappa_3}{\nu_2} \right) \left( 1 - \frac{\kappa_1^\ominus v_P^*}{\kappa_1^+} \right) \end{aligned}$$

which equals the expression from (22). The results of this approximation are compared to our result in Fig. 2.

Recalling our exact result for the Fano factor from Eq. (14), we note that due to the linearization of the mRNA expression and thus a basically constant mRNA production, the noise emerging from the random gene switches is not adequately represented in the formula obtained in Thattai and Oudenaarden (2001). To be more precise, in contrast to our formula in (14), the effect of mRNA noise due to gene switching (first term in first bracket) is not taken into account at all. Additionally, the negative feedback (last bracket in (22)) does not affect the noise in the same way as it does in the exact formula (14). As can be seen in Fig. 2, when comparing the solid and the dashed lines,

these effects lead to an underestimation of the actual noise which is produced by an exact simulation of [neg].

### 4.3 Negative feedback, Swain (2004)

As explained around (7), 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 Swain (2004), 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, see also Sect. 3.6. For transcriptional feedback (which we study here), Swain obtains in his equation (5)—see below for the transformation of his results into our parameters (unscaled, and scaled)

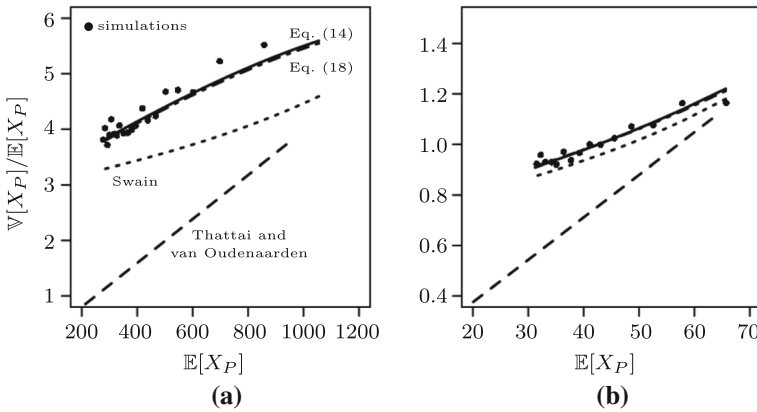
$$\begin{aligned} \frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} &= \left( \alpha_{neg} \frac{\mu_3}{\mu_2 + \mu_3} + \frac{\lambda_3}{\mu_2 + \mu_3} + 1 \right) / (1 + \alpha_{neg}) \\ &= \left( \alpha_{neg} \frac{v_3}{\tilde{v}_2 + v_3} + \frac{\tilde{\kappa}_3}{\tilde{v}_2 + v_3} + 1 \right) / (1 + \alpha_{neg}) \end{aligned} \quad (23)$$

with  $\alpha_{neg} = \frac{\kappa_1^\ominus v_p^*}{\kappa_1^\ominus v_p^* + \kappa_1^+}$ . This corresponds exactly to (18)—see also Eq. (44) in Appendix E for the unscaled version—with a missing term in the numerator, namely  $\frac{\alpha_{neg}^2 \kappa_2 \tilde{\kappa}_3}{\kappa_1^\ominus v_p^* (\tilde{v}_2 + v_3)}$ . This term arises from fluctuations in gene activation, which was averaged out in calculations done in Swain (2004). At least, (23) arises from (18) if we assume that  $\kappa_1^\ominus v_p^*, \kappa_1^+ \gg 1$ , i.e. fast gene switching.

For a comparison of the two results we refer to the solid and dotted lines in Fig. 2 where we see that due to the missing term in the approximation obtained by Swain, his result slightly underestimates the Fano factor resulting from simulations of [neg]. However, the difference becomes smaller for a lower number of expected proteins, see Fig. 2b. This can be explained by the difference between Eqs. (23) and (18). The missing term in Swain's derivation can be related to noise emerging from the gene switching. These processes however are given by the overall parameter configuration. Hence the larger difference between the solid and dotted lines in (a) when compared to (b) simply emerges from the corresponding term in (a) being larger than in (b). Thus, it therefore has a stronger effect on the overall fluctuations in protein numbers.

In order to obtain Eq. (23), reference (Swain 2004) uses  $d_0 = \tilde{v}_2$ ,  $d_1 = v_3$ ,  $v_1 = \tilde{\kappa}_3$ ,  $\langle M \rangle = \frac{\kappa_1^+ \kappa_2}{(\kappa_1^\ominus v_p^*) \tilde{v}_2}$  and

$$\epsilon_c = \frac{2}{1 + \sqrt{1 + 4 \frac{\kappa_1^\ominus \kappa_2 \tilde{\kappa}_3}{\kappa_1^+ \tilde{v}_2 v_3}}}$$



**Fig. 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 (14), the dash-dotted line the result in (18), the dashed line the result from Thattai and Oudenaarden (2001) given in (22) and the dotted line the Fano factor calculated in Swain (2004) given in (23). The parameters for the simulations are chosen as follows: **a**  $M = 1, N = 100, \lambda_1^+ = 250, \lambda_1^- = (0.2, 0.4, \dots, 4), \lambda_2 = 300, \mu_2 = 100, \lambda_3 = 500, \mu_3 = 1$ ; **b**  $M = 1, N = 100, \lambda_1^+ = 300, \lambda_1^- = (1, 2, \dots, 15), \lambda_2 = 200, \mu_2 = 25, \lambda_3 = 10, \mu_3 = 1$ . The bullets represent the estimated Fano factors of the full system  $[neg]$  obtained from 1000 Monte Carlo simulations (cf. Gillespie 1977) in (a) and 5000 in (b) for each value of  $\lambda_1^-$

for the Fano factor for the auto-regulatory gene expression with negative (transcriptional) feedback. Expressing terms as in Sect. 3.6, we rearrange his equation (5) (recall (5) and note that  $v_p^*/E_\pi[V_R] = \tilde{\kappa}_3/v_3$ ; see (6)) with  $M = 1$

$$\begin{aligned} \frac{V_{neg}[X_P]}{E_{neg}[X_P]} &= 1 + \frac{v_3 u_p^*}{(\tilde{v}_2 + v_3) E_\pi[V_R]} \left( 1 - \frac{\sqrt{1 + 4 \frac{\kappa_1^- \kappa_2 \tilde{\kappa}_3}{\kappa_1^+ \tilde{v}_2 v_3}} - 1}{2 \sqrt{1 + 4 \frac{\kappa_1^- \kappa_2 \tilde{\kappa}_3}{\kappa_1^+ \tilde{v}_2 v_3}}} \left( 1 + \frac{\tilde{v}_2}{\tilde{\kappa}_3} \right) \right) \\ &= 1 + \frac{\tilde{\kappa}_3}{(\tilde{v}_2 + v_3)} \left( 1 - \frac{\kappa_1^- v_p^*}{2 \kappa_1^- v_p^* + \kappa_1^+} \left( 1 + \frac{\tilde{v}_2}{\tilde{\kappa}_3} \right) \right) \\ &= \left( \kappa_1^- v_p^* + \kappa_1^+ \right) \left( 1 + \frac{\tilde{\kappa}_3}{\tilde{v}_2 + v_3} + \frac{\kappa_1^- v_3 v_p^*}{(\tilde{v}_2 + v_3) (\kappa_1^- v_p^* + \kappa_1^+)} \right) / (2 \kappa_1^- v_p^* + \kappa_1^+) \\ &= \left( 1 + \frac{\tilde{\kappa}_3}{\tilde{v}_2 + v_3} + \frac{\kappa_1^- v_3 v_p^*}{(\tilde{v}_2 + v_3) (\kappa_1^- v_p^* + \kappa_1^+)} \right) / \left( 1 + \frac{\kappa_1^- v_p^*}{\kappa_1^- v_p^* + \kappa_1^+} \right), \end{aligned}$$

showing (23).

#### 4.4 Negative feedback, Dessalles et al. (2017)

Our result for  $[neg]$  has Theorem 2 of Dessalles et al. (2017) as a limit result. They study a similar model (with  $M = 1$ ), but with a scaling such that  $\lambda_3, \mu_3 = O(1)$  and  $\lambda_1^\ominus = O(N)$ , leading to low (i.e.  $O(1)$ ) abundance of protein in the system. Assuming the same time-scale separation as in our model, i.e. gene switching and mRNA processes evolve on a fast time-scale, the resulting birth–death process 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,

$$\frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} \stackrel{\rho \rightarrow \infty}{\approx} \frac{1}{2}. \quad (24)$$

For large  $\rho$ , our model simplifies to  $v_p^* \approx \sqrt{\rho}$ , and then (14) gives the same limit.

#### 4.5 Negative feedback in a simpler model, Ramos et al. (2015)

In Ramos et al. (2015), the authors derive in their equation (11) the Fano factor for the simplified model, which we introduced in Sect. 3.5. (We note that their model slightly differs since the protein binds to the gene and therefore cannot degrade in this state.) Their results connect the Fano factor to the covariance of the number of active genes and the number of proteins in equilibrium. Since they also give the limiting distribution (in terms of a confluent hypergeometric function), they can evaluate this covariance and also the Fano factor numerically. From their Figure 2, one can see that—if proteins are somewhat abundant—the Fano factor stabilizes around 1/2 for various parameter combinations; a result reminiscent of Dessalles et al. (2017) as discussed above.

#### 4.6 Negative feedback for small amounts of protein

Theorems 1 and 2—and all subsequent calculations—only hold under the scaling described in  $\square_*$  or in Appendix D. In these scalings, we have that  $X_P = O(N)$ , i.e. there the protein is abundant. In this case, we see from (14) that the Fano factor is at least 1/2. In Sect. 4.4, we discussed the results by Dessalles et al. (2017), where  $X_P = O(1)$  is used, but the limiting result for  $\rho \rightarrow \infty$  implies that proteins are abundant and leads to a Fano factor of at least 1/2.

However, the scaling of Dessalles et al. (2017) also allows for smaller values for the Fano factor, which is called the infra-Fano regime in Ramos et al. (2015). Precisely, consider the model from Dessalles et al. (2017), where the scaling

$$\lambda_1^+, \lambda_1^\ominus, \lambda_2, \mu_2 = O(N), \quad \lambda_3, \mu_3 = O(1)$$

is used. It is shown that  $X_P$  converges towards a birth–death process with

$$\text{death rates } \delta_n := \mu_3 n \text{ and birth rates } \beta_n := \frac{\lambda_3}{\mu_2} \frac{\lambda_2 \lambda_1^+}{\lambda_1^+ + \lambda_1^\ominus n},$$

if there are  $n$  proteins. For this process, they compute the equilibrium distribution

$$\pi(n) = \frac{1}{Z} \prod_{i=0}^{n-1} \frac{\rho}{(i+1)(i + \lambda_1^+ / \lambda_1^\ominus)}$$

with  $\rho = \lambda_1^+ \lambda_2 \lambda_3 / (\lambda_1^\ominus \mu_2 \mu_3)$  and  $Z$  as a normalizing constant. In this case, one can see that for  $\lambda_1^\ominus \gg \lambda_1^+$ ,  $\pi$  is concentrated around 1, i.e. there is a single molecule of the protein and hence, the Fano factor becomes arbitrarily small. Since this in particular means a Fano factor below 1/2, Ramos et al. (2015) call this the infra-Fano regime.

## 5 Conclusion

Quantifying noise in gene expression is essential for understanding regulatory networks in cells (Thattai and Oudenaarden 2001). Our results capture most of the previously derived results. While negative feedback is known to reduce noise under auto-regulated gene expression, we improve on the quantification of this effect, i.e. our results account for all possible sources of noise due to gene activation, mRNA fluctuations and the protein processes itself. We note that our results require that proteins are abundant. Since the infra-Fano regime described in Sect. 4.6 relies on small amounts of protein, our results do not recover this regime.

In addition, we provide the same quantification of noise also for positive feedback, where noise is increased. In particular, (14) shows that the average time the gene is *off* determines the reduction of noise in all cases relative to unregulated genes; see also Grönlund et al. (2013). As we saw earlier, for both, negative and positive feedback, noise difference between the non-regulated and the model with feedback is largest if the gene is *off* most of the time. This can be interpreted by the burstiness of gene expression. It is largest for genes which are *off* for long times and then turned *on* for a short time in which mRNA is produced. Interestingly, previous approaches mostly gave approximations for noise for negative feedback if switching the gene on and off is very fast (Thattai and Oudenaarden 2001; Swain 2004) and if the gene is *on* most of the time (Thattai and Oudenaarden 2001) or in a simplified model (Ramos et al. 2015). Hence, all previous papers could not have seen the effects of gene activation switching on protein noise. As in previous results also obtained in Dessalles et al. (2017), 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. Additionally we find that 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 Kim et al. (2015), some special cases are studied when a straight-forward approximation of the fluctuations works. In our analysis, we use a new approach by Kang et al. (2014) and can also interpret all terms arising in (14), see Sect. 3.3.

Due to taking into account all potential sources of fluctuations the fit of simulations and theory (see e.g. Fig. 2) is excellent and improves on previous studies. There, noise arising from the gene switching its state has been averaged out, and only the recent approach of Kang et al. (2014) reveals the impact of these stochastic processes on the noise in protein numbers.

In their paper, Kang et al. (2014) 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 division.

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 (Becksei and Serrano 2000; Cardelli et al. 2016). Experimentally, gene expression noise can be used to understand the dynamics of gene regulation (Munsky et al. 2012). Our analysis should provide an approach for distinguishing between different models of gene regulation based on measurements of noise levels.

**Acknowledgements** Open access funding provided by Max Planck Society. We thank Jens Timmer, Freiburg, for discussion. Additionally, we thank three anonymous referees for carefully reading the manuscript and helpful comments which improved the manuscript.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## A 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



where  $v_k = (v_{ks})_{s \in \mathcal{S}}$  and  $v'_k = (v'_{ks})_{s \in \mathcal{S}}$  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

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

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

$$\zeta_k := v'_k - v_k,$$

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

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

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

$$X(t) = X(0) + \sum_{k \in \mathcal{K}} \zeta_k R_k(t) = X(0) + \sum_{k \in \mathcal{K}} \zeta_k Y_k \left( \int_0^t \Lambda_k(X(s)) ds \right). \tag{26}$$

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 infinitesimal generator of the full process.

For some scaling parameter  $N$ , assume that  $X = X^N = (X_s^N, X_f^N)$  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  $L^N$  of  $X^N$  has the form

$$L^N = L_1^N + N L_2^N, \tag{27}$$

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

$$V_s^N \xrightarrow{N \rightarrow \infty} V_s, \tag{28}$$

$$U^N := \sqrt{N} (V_s^N - V_s) \xrightarrow{N \rightarrow \infty} U \tag{29}$$

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

1. We have that (with the projection  $\pi_s$  on the slow species and  $F^N := L^N \pi_s = L_1^N \pi_s$ )

$$M_1^N(t) := V_s^N(t) - V_s^N(0) - \int_0^t F^N (V_s^N(s), V_f^N(s)) ds$$

is a (local) martingale. For the convergence (28), we assume that

$$\int_0^t F^N (V_s^N(s), V_f^N(s)) \xrightarrow{N \rightarrow \infty} \int_0^t F (V_s(s)) ds$$

for some unique process  $V_s$ , which holds for

$$F(v_s) = \lim_{N \rightarrow \infty} \mathbb{E}_{\rho_N} \left[ L^N v_s \right], \tag{1}$$

where  $\rho_N$  is the equilibrium of  $V_f^N$  for fixed slow species  $v_s$ . Thus, the convergence  $V_s^N \xrightarrow{N \rightarrow \infty} V_s$  holds with

$$V_s(t) = V_s(0) + \int_0^t F(V_s(s)) ds$$

and we have shown (28).

2. Note that

$$\begin{aligned} U^N(t) - U^N(0) &= \sqrt{N} \left( V_s^N(t) - V_s(t) \right) - \sqrt{N} \left( V_s^N(0) - V_s(0) \right) \\ &= \sqrt{N} \left( M_1^N(t) + \int_0^t \left( F^N \left( V_s^N(s), V_f^N(s) \right) - F \left( V_s^N(s) \right) \right) ds \right. \\ &\quad \left. + \int_0^t \left( F \left( V_s^N \right) - F \left( V_s \right) \right) ds \right). \end{aligned}$$

Assume that we

$$\text{find } h^N \text{ such that } L_2^N h^N \approx F^N - F. \tag{2}$$

(The ‘ $\approx$ ’ is controlled by  $\varepsilon_2^N$  below. Note that this is a Poisson equation.) With

$$\begin{aligned} \varepsilon_1^N(t) &:= \frac{1}{N} \left( h^N \left( V_s^N(t), V_f^N(t) \right) - h^N \left( V_s^N(0), V_f^N(0) \right) \right. \\ &\quad \left. - \int_0^t L_1^N h^N \left( V_s^N(s), V_f^N(s) \right) ds \right), \\ \varepsilon_2^N(t) &:= - \int_0^t \left( L_2^N h^N \left( V_s^N(s), V_f^N(s) \right) \right. \\ &\quad \left. - \left( F^N \left( V_s^N(s), V_f^N(s) \right) - F \left( V_s^N(s) \right) \right) \right) ds, \end{aligned} \tag{30}$$

we obtain that

$$M_2^N(t) := \varepsilon_1^N(t) + \varepsilon_2^N(t) - \int_0^t \left( F^N \left( V_s^N(s), V_f^N(s) \right) - F \left( V_s^N(s) \right) \right) ds$$

is a (local) martingale. Hence

$$\sqrt{N} \left( V_s^N(t) - V_s(t) \right) - \sqrt{N} \left( V_s^N(0) - V_s(0) \right)$$



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

We assume for  $\varepsilon_1^N, \varepsilon_2^N$  from (30) that

$$\sqrt{N}\varepsilon_1^N, \sqrt{N}\varepsilon_2^N \xrightarrow{N \rightarrow \infty} 0. \tag{3}$$

- In order to show convergence of  $M_1^N - M_2^N$ , we use the Martingale Central Limit Theorem; see Appendix A.1 in Kang et al. (2014). Note that since the quadratic variation of all integrals  $\int dt$  vanishes, we find that (recall the notation for Chemical Reaction Networks from (25)–(26), which we now equip with a superscript  $N$  to account for the scaling constant), with  $z^{\otimes 2} = zz^\top$

$$\begin{aligned} \left[ \sqrt{N} \left( M_1^N - M_2^N \right) \right]_t &= N \left[ V_s^N - \frac{1}{N} \left( h^N \left( V_s^N, V_f^N \right) \right) \right]_t \\ &= \frac{1}{N} \sum_{k \in \mathcal{K}} \int_0^t \left( \zeta_{ks} + h^N \left( V_s^N(s-), V_f^N(s-) \right) \right. \\ &\quad \left. - h^N \left( V_s^N(s-) + \frac{1}{N} \zeta_{ks}, V_f^N(s-) + \zeta_{kf} \right) \right)^{\otimes 2} dR_k^N(s), \end{aligned}$$

where  $\zeta_{ks}$  and  $\zeta_{kf}$  are the stoichiometric 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

$$\begin{aligned} \lim_{N \rightarrow \infty} \left[ \sqrt{N} \left( M_1^N - M_2^N \right) \right]_t &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k \in \mathcal{K}} \int_0^t \left( \zeta_{ks} + h^N \left( V_s^N(s), V_f^N(s-) \right) \right. \\ &\quad \left. - h^N \left( V_s^N(s) + \frac{1}{N} \zeta_{ks}, V_f^N(s-) + \zeta_{kf} \right) \right)^{\otimes 2} \Lambda_k^N(s) ds. \end{aligned}$$

Now, for the equilibrium  $\rho_N$  of the fast species for given concentration of slow species,  $v_s$ , if

$$\begin{aligned} \frac{1}{N} \sum_{k \in \mathcal{K}} \mathbb{E}_{\rho^N} \left[ \left( \zeta_{ks} + h^N \left( v_s, V_f^N \right) - h^N \left( v_s + \frac{1}{N} \zeta_{ks}, V_f^N + \zeta_{kf} \right) \right)^{\otimes 2} \Lambda_k^N \left( V_s \right) \right] \\ \xrightarrow{N \rightarrow \infty} c(v_s), \end{aligned} \tag{4}$$

we have that

$$\begin{aligned} & \frac{1}{N} \sum_{k \in \mathcal{K}} \int_0^t \left( \zeta_{ks} + h^N \left( V_s^N(s), V_f^N(s-) \right) \right. \\ & \left. - h^N \left( V_s^N(s) + \frac{1}{N} \zeta_{ks}, V_f^N(s-) + \zeta_{kf} \right) \right)^{\otimes 2} \bar{\Lambda}_k^N \left( V_s^N \right) ds \xrightarrow{N \rightarrow \infty} \int_0^t c(V(s)) ds, \end{aligned}$$

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 the martingale central limit theorem (see Appendix A.1 in Kang et al. 2014) that

$$\sqrt{N}(M^{N,1} - M^{N,2}) \xrightarrow{N \rightarrow \infty} M$$

where  $M$  satisfies

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

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

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

we find that, if  $U^N \xrightarrow{N \rightarrow \infty} U$ , 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 (29).

## B Proof of Theorem 2

*Proof* We will make use of notation from Appendix A and have to show (1)–(4) from Appendix A in all cases. Note that the function  $F$  from Theorem 1 already satisfies (1). In all cases, the system  $(V_{on}, V_R, V_P)$  is a Markov process with a generator of the form (27) with

$$\begin{aligned} L_1^N f(u, r, v_P) &= \kappa_3 r N \left( f \left( u, r, v_P + \frac{1}{N} \right) - f(u, r, v_P) \right) \\ &\quad + v_P v_3 N \left( f \left( u, r, v_P - \frac{1}{N} \right) - f(u, r, v_P) \right) \\ &= (\kappa_3 r - v_P v_3) \frac{\partial f}{\partial v_P}(u, r, v_P) + o(1) \end{aligned}$$

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

$$F^N(r, v_P) = \kappa_3 r - v_3 v_P.$$

For [neu],

$$\begin{aligned} L_2^N f(u, r, v_P) &= (M - u)\kappa_1^+(f(u + 1, r, v_P) - f(u, r, v_P)) \\ &\quad + u\kappa_1^-(f(u - 1, r, v_P) - f(u, r, v_P)) \\ &\quad + \kappa_2 u(f(u, r + 1, v_P) - f(u, r, v_P)) \\ &\quad + v_2 r(f(u, r - 1, v_P) - f(u, r, v_P)). \end{aligned}$$

From (2) and (4), we see that we need to solve

$$L_2^N h^N = \kappa_3 r - \frac{M\kappa_1^+\kappa_2\kappa_3}{v_2(\kappa_1^- + \kappa_1^+)}.$$

Choosing the Ansatz

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

we obtain the following equation which we need to solve

$$(M\kappa_1^+ - u(\kappa_1^+ + \kappa_1^-))g_1(v_P) + (\kappa_2 u - v_2 r)g_2(v_P) \stackrel{!}{=} \kappa_3 r - \frac{M\kappa_1^+\kappa_2\kappa_3}{v_2(\kappa_1^- + \kappa_1^+)}.$$

Solving for  $g_2$  and then for  $g_1$ , we obtain

$$g_2(v_P) = -\frac{\kappa_3}{v_2}, \quad g_1(v_P) = -\frac{\kappa_2\kappa_3}{v_2(\kappa_1^- + \kappa_1^+)}.$$

Then,  $\sqrt{N}\varepsilon_1^N \xrightarrow{N \rightarrow \infty} 0$  since  $h^N$  is bounded in  $N$  and  $\varepsilon_2^N = 0$  by construction. Hence, we have shown (3). For (4), if  $\pi$  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

$$\begin{aligned} &\frac{1}{N} \sum_{k=1}^6 \mathbb{E}_\pi \left[ \left( \zeta_{kP} + h^N(U, R, v_P) - h^N \left( U + \zeta_{kon}, R + \zeta_{kr}, v_P + \frac{1}{N} \zeta_{kP} \right) \right)^2 \Lambda_k(v_P) \right] \\ &= 2(g_1(v_P))^2 M \frac{\kappa_1^- \kappa_1^+}{\kappa_1^- + \kappa_1^+} + (g_2(v_P))^2 \left( \frac{M\kappa_1^+\kappa_2}{\kappa_1^- + \kappa_1^+} + \frac{M\kappa_1^+\kappa_2}{v_2(\kappa_1^- + \kappa_1^+)} v_2 \right) \\ &\quad + \left( \frac{M\kappa_1^+\kappa_2\kappa_3}{v_2(\kappa_1^- + \kappa_1^+)} + v_3 v_P \right) \\ &= \frac{M\kappa_1^+}{\kappa_1^- + \kappa_1^+} \left( \frac{2\kappa_1^- \kappa_2^2 \kappa_3^2}{v_2^2(\kappa_1^- + \kappa_2^+)^2} + \frac{2\kappa_2\kappa_3^2}{v_2^2} + \frac{\kappa_2\kappa_3}{v_2} \right) + v_3 v_P. \end{aligned}$$

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

### C The two-dimensional Ornstein–Uhlenbeck process

We recall results for the two-dimensional Ornstein–Uhlenbeck process; see e.g. Gardiner (2009). They are needed when refining the Fano factor in Appendix D.

**Theorem 3** (Stationary variance of two-dimensional Ornstein–Uhlenbeck process)  
 Let  $X = (X_1, X_2)$  solve

$$dX = -AXdt + BdW,$$

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

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

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

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

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  $2 \times 2$ -matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , we have that (for the unit matrix  $I_2$ )

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

Hence, we can write

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

i.e.

$$\begin{aligned} \mathbb{E}[X_0 X_0^\top] &= \int_{-\infty}^0 (\alpha_s + \beta_s A) B B^\top (\alpha_s + \beta_s A^\top) ds \\ &= \alpha B B^\top + \beta (A B B^\top + B B^\top A^\top) + \gamma A B B^\top A^\top \\ &= \gamma \left( A + \frac{\beta}{\gamma} \right) (B B^\top) \left( A^\top + \frac{\beta}{\gamma} \right) + \left( \alpha - \frac{\beta^2}{\gamma} \right) B B^\top \end{aligned} \tag{31}$$

for

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

We then write

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

which is only possible if

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

i.e.

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

Combining this with (31) then gives the result. □

**Corollary 1** (Diagonal matrix  $B$ ) *Note that if*

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

then

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

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

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

□

### D Refining the Fano factor

Here, we study the case

$$\lambda_3 = \tilde{\kappa}_3, \quad \mu_2 = \tilde{\nu}_2$$

which leads to  $X_R = O(N)$ , such that we have the scaling  $\alpha_{\text{off}} = \alpha_{\text{on}} = 0$  and  $\alpha_R = \alpha_P = 1$ ; compare with (1). In particular, we use that—see (2)— $V_R = X_R/N$ ,  $V_P = X_P/N$ . Here,  $V_{\text{on}} = X_{\text{on}}$  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

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

We will denote the unique solution of  $F(V_R, V_P) = 0$  by  $(v_R^*, v_P^*)$ .

**Theorem 4** (Central Limit Theorem) *Let  $V_R^N, V_R, V_P^N, V_P$  and  $F$  be as above and  $\sqrt{N}((V_R^N(0), V_P^N(0)) - (V_R(0), V_P(0))) \xrightarrow{N \rightarrow \infty} (U_R(0), U_P(0))$ . Then, for the models [neu], [neg] and [pos],  $\sqrt{N}((V_R^N, V_P^N) - (V_R, V_P)) \xrightarrow{N \rightarrow \infty} (U_R, U_P)$ , where*

$(U_R, U_P)$  solves

$$\begin{aligned}
 U_R(t) &= U_R(0) + \int_0^t \sqrt{c_R(V_P(s))} dW(s) \\
 &\quad - \int_0^t (\tilde{v}_2 U_R(s) + d_R(V_P(s)) U_P(s)) ds, \\
 U_P(t) &= U_P(0) + \int_0^t \sqrt{\tilde{\kappa}_3 V_R(s) + v_3 V_P(s)} dW'(s) \\
 &\quad + \int_0^t (\tilde{\kappa}_3 U_R(s) - v_3 V_P(s) U_P(s)) ds
 \end{aligned} \tag{32}$$

with  $W, W'$  independent Brownian motions,

$$c_R(v_R, v_P) = \begin{cases} \frac{M\kappa_1^+ \kappa_2}{\kappa_1^- + \kappa_1^+} \left( \frac{2\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2} + 1 \right) + \tilde{v}_2 v_R & \text{for [neu]}, \\ \frac{M\kappa_1^+ \kappa_2}{\kappa_1^\ominus v_P + \kappa_1^+} \left( \frac{2\kappa_1^\ominus v_P \kappa_2}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} + 1 \right) + \tilde{v}_2 v_R & \text{for [neg]}, \\ \frac{M\kappa_1^\oplus v_P \kappa_2}{\kappa_1^- + \kappa_1^\oplus v_P} \left( \frac{2\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^\oplus v_P)^2} + 1 \right) + \tilde{v}_2 v_R & \text{for [pos]} \end{cases} \tag{33}$$

and

$$d_R(v_P) = \begin{cases} 0 & \text{for [neu]}, \\ \frac{M\kappa_1^\ominus \kappa_1^+ \kappa_2}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} & \text{for [neg]}, \\ -\frac{M\kappa_1^- \kappa_1^\oplus \kappa_2}{(\kappa_1^- + \kappa_1^\oplus v_P)^2} & \text{for [pos]}. \end{cases} \tag{34}$$

*Remark 2* In equilibrium, we have

$$c_R(v_R^*, v_P^*) = \begin{cases} \frac{2\tilde{v}_2 v_3 v_P^*}{\tilde{\kappa}_3} \left( \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2} + 1 \right) & \text{for [neu]}, \\ \frac{2\tilde{v}_2 v_3 v_P^*}{\tilde{\kappa}_3} \left( \frac{\kappa_1^\ominus v_P \kappa_2}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} + 1 \right) & \text{for [neg]}, \\ \frac{2\tilde{v}_2 v_3 v_P^*}{\tilde{\kappa}_3} \left( \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^\oplus v_P)^2} + 1 \right) & \text{for [pos]}. \end{cases} \tag{35}$$

*Proof* First, note that

$$DF(v_R, v_P) = \begin{cases} \begin{pmatrix} -\tilde{v}_2 & 0 \\ \tilde{\kappa}_3 & -v_3 \end{pmatrix} & \text{for}[neu], \\ \begin{pmatrix} -\tilde{v}_2 & -\frac{M\kappa_1^+\kappa_1^\ominus\kappa_2}{(\kappa_1^\ominus v_P + \kappa_1^+)^2} \\ \tilde{\kappa}_3 & -v_3 \end{pmatrix} & \text{for}[neg], \\ \begin{pmatrix} -\tilde{v}_2 & \frac{M\kappa_1^-\kappa_1^\oplus\kappa_2}{(\kappa_1^- + \kappa_1^\oplus v_P)^2} \\ \tilde{\kappa}_3 & -v_3 \end{pmatrix} & \text{for}[pos]. \end{cases}$$

Again, we have to show **(1)–(4)** from Appendix A in all cases for  $F$  as above, which already satisfies **(1)**. We focus on  $[neu]$  first. The system  $(V_{on}, V_R, V_P)$  is a Markov process with a generator of the form (27) with

$$\begin{aligned} L_1^N f(u, v_R, v_P) &= (\kappa_2 u - v_R \tilde{v}_2) \frac{\partial f}{\partial v_R}(u, v_R, v_P) \\ &\quad + (\tilde{\kappa}_3 v_R - v_P v_3) \frac{\partial f}{\partial v_P}(u, v_R, v_P) + o(1), \\ L_2^N f(u, v_R, v_P) &= (M - u)\kappa_1^+(f(u + 1, v_R, v_P) - f(u, v_R, v_P)) \\ &\quad + u\kappa_1^-(f(u - 1, v_R, v_P) - f(u, v_R, v_P)). \end{aligned}$$

This implies that

$$F^N(u, v_R, v_P) = \begin{pmatrix} \kappa_2 u - \tilde{v}_2 v_R \\ \tilde{\kappa}_3 v_R - v_3 v_P \end{pmatrix}.$$

Hence, for **(2)**, we have to solve

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

which is

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

For the quadratic variation in **(4)**, we find that with  $z^{\otimes 2} = z z^\top$

$$\begin{aligned} &N \left[ (V_R, V_P)^\top - \frac{1}{N} h^N(V_{on}, V_R, V_P) \right]_t \\ &= \frac{1}{N} \sum_{k \in \mathcal{K}} \int_0^t \left( (\xi_{kR}, \zeta_{kP})^\top + h^N(V_{on}(s-), V_R(s-), V_P(s-)) \right) \end{aligned}$$



$$\begin{aligned}
 & -h^N \left( V_{\text{on}}(s-) + \zeta_{\text{kon}}, V_R(s-) + \frac{1}{N} \zeta_{kR}, V_P(s-) + \frac{1}{N} \zeta_{kP} \right) \otimes^2 \\
 & \Lambda_k^N (V_{\text{on}}(s), V_R(s), V_P(s)) ds \\
 = & \int_0^t \left[ \begin{pmatrix} -\frac{\kappa_2}{\kappa_1^- + \kappa_1^+} \\ 0 \end{pmatrix} \otimes^2 \kappa_1^+ (M - V_{\text{on}}(s)) + \begin{pmatrix} -\frac{\kappa_2}{\kappa_1^- + \kappa_1^+} \\ 0 \end{pmatrix} \otimes^2 \kappa_1^- V_{\text{on}}(s) \right. \\
 & + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes^2 \kappa_2 V_{\text{on}}(s) + \begin{pmatrix} -1 \\ 0 \end{pmatrix} \otimes^2 \tilde{v}_2 V_R(s) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes^2 \tilde{\kappa}_3 V_R(s) \\
 & \left. + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \otimes^2 v_3 V_P(s) \right] ds \\
 = & \int_0^t \left[ \begin{pmatrix} \frac{\kappa_2^2}{(\kappa_1^- + \kappa_1^+)^2} & 0 \\ 0 & 0 \end{pmatrix} (\kappa_1^+ (M - V_{\text{on}}(s)) + \kappa_1^- V_{\text{on}}(s)) \right. \\
 & + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} (\kappa_2 V_{\text{on}}(s) + \tilde{v}_2 V_R(s)) \\
 & \left. + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} (\tilde{\kappa}_3 V_R(s) + v_3 V_P(s)) \right] ds \\
 \approx & \int_0^t \begin{pmatrix} \frac{2M\kappa_1^- \kappa_1^+ \kappa_2^2}{(\kappa_1^- + \kappa_1^+)^3} + \frac{M\kappa_1^+ \kappa_2}{\kappa_1^+ + \kappa_1^-} & 0 \\ 0 & \tilde{\kappa}_3 V_R(s) + v_3 V_P(s) \end{pmatrix} ds.
 \end{aligned}$$

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

### D.1 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 (35), and obtain Ornstein–Uhlenbeck processes in all cases. Since they are two-dimensional, their equilibrium (normal) distribution can be computed (see Appendix C).

..for  $[neu]$

We obtain

$$\begin{aligned}
 d \begin{pmatrix} U_R \\ U_P \end{pmatrix} = & - \begin{pmatrix} \tilde{v}_2 & 0 \\ -\tilde{\kappa}_3 & v_3 \end{pmatrix} \begin{pmatrix} U_R \\ U_P \end{pmatrix} dt \\
 & + \sqrt{\frac{2v_P^* \tilde{v}_2 v_3}{\tilde{\kappa}_3}} \begin{pmatrix} \sqrt{\frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2} + 1} & 0 \\ 0 & \sqrt{\frac{\tilde{\kappa}_3}{v_2}} \end{pmatrix} \begin{pmatrix} dW \\ dW' \end{pmatrix}.
 \end{aligned}$$

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

$$\mathbf{E}_{neu}[UU^\top] = \frac{v_P^*}{(\tilde{v}_2 + v_3)\kappa_3} \times \begin{pmatrix} \left(1 + \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2}\right) v_3(\tilde{v}_2 + v_3) & \left(1 + \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2}\right) \tilde{\kappa}_3 v_3 \\ \left(1 + \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2}\right) \tilde{\kappa}_3 v_3 & \tilde{\kappa}_3(\tilde{v}_2 + v_3) + \tilde{\kappa}_3^2 \left(1 + \frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^+)^2}\right) \end{pmatrix}. \tag{36}$$

Therefore,

$$\frac{N\mathbb{V}_{neu}[V_P]}{v_P^*} \approx 1 + \frac{\tilde{\kappa}_3}{\tilde{v}_2 + v_3} + \frac{\kappa_1^- \kappa_2 \tilde{\kappa}_3}{(\kappa_1^- + \kappa_1^+)^2 (\tilde{v}_2 + v_3)}. \tag{37}$$

...for [neg]

Here, we obtain

$$d \begin{pmatrix} U_R \\ U_P \end{pmatrix} = - \begin{pmatrix} \tilde{v}_2 & \frac{\kappa_1^\ominus v_P^* \tilde{v}_2 v_3}{(\kappa_1^\ominus v_P^* + \kappa_1^+) \tilde{\kappa}_3} \\ -\tilde{\kappa}_3 & v_3 \end{pmatrix} \begin{pmatrix} U_R \\ U_P \end{pmatrix} dt + \sqrt{\frac{2v_P^* \tilde{v}_2 v_3}{\tilde{\kappa}_3}} \begin{pmatrix} \sqrt{\frac{\kappa_1^\ominus v_P^* \kappa_2}{(\kappa_1^\ominus v_P^* + \kappa_1^+)^2} + 1} & 0 \\ 0 & \sqrt{\frac{\tilde{\kappa}_3}{\tilde{v}_2}} \end{pmatrix} \begin{pmatrix} dW \\ dW' \end{pmatrix}.$$

Hence, in equilibrium, and if ... denote the quantities from (36), and for  $b = \frac{\kappa_1^\ominus v_P^* \tilde{v}_2 v_3}{(\kappa_1^\ominus v_P^* + \kappa_1^+) \tilde{\kappa}_3}$ ,

$$\mathbb{E}_{neg}[UU^\top] = \frac{v_P^* \tilde{v}_2 v_3}{(\tilde{v}_2 v_3 + \tilde{\kappa}_3 b)(\tilde{v}_2 + v_3) \tilde{\kappa}_3} \begin{pmatrix} \dots + \frac{\tilde{\kappa}_3}{\tilde{v}_2} b^2 + \left(\frac{\kappa_1^\ominus v_P^* \kappa_2}{(\kappa_1^\ominus v_P^* + \kappa_1^+)^2} + 1\right) \tilde{\kappa}_3 b & \dots - \left(\frac{\kappa_1^\ominus v_P^* \kappa_2}{(\kappa_1^\ominus v_P^* + \kappa_1^+)^2} + 1\right) \tilde{\kappa}_3 b \\ \dots - \tilde{\kappa}_3 b & \dots + \frac{\tilde{\kappa}_3}{\tilde{v}_2} b \end{pmatrix}. \tag{38}$$

So,

$$\frac{N\mathbb{V}_{neg}[V_P]}{v_P^*} \approx \left(1 + \frac{\kappa_1^\ominus v_P^*}{\kappa_1^\ominus v_P^* + \kappa_1^+}\right)^{-1} \times \left(1 + \frac{\tilde{\kappa}_3}{\tilde{v}_2 + v_3} + \frac{\kappa_1^\ominus v_P^* \kappa_2 \tilde{\kappa}_3}{(\kappa_1^\ominus v_P^* + \kappa_1^+)^2 (\tilde{v}_2 + v_3)} + \frac{\kappa_1^\ominus v_P^* v_3}{(\kappa_1^\ominus v_P^* + \kappa_1^+) (\tilde{v}_2 + v_3)}\right). \tag{39}$$

...for [pos]

We obtain

$$d \begin{pmatrix} U_R \\ U_P \end{pmatrix} = - \begin{pmatrix} \tilde{v}_2 & -\frac{\kappa_1^- \tilde{v}_2 v_3}{(\kappa_1^- + \kappa_1^\oplus v_P^*) \tilde{\kappa}_3} \\ -\tilde{\kappa}_3 & v_3 \end{pmatrix} \begin{pmatrix} U_R \\ U_P \end{pmatrix} dt + \sqrt{\frac{2v_P^* \tilde{v}_2 v_3}{\tilde{\kappa}_3}} \begin{pmatrix} \sqrt{\frac{\kappa_1^- \kappa_2}{(\kappa_1^- + \kappa_1^\oplus v_P^*)^2} + 1} & 0 \\ 0 & \sqrt{\frac{\tilde{\kappa}_3}{\tilde{v}_2}} \end{pmatrix} \begin{pmatrix} dW \\ dW' \end{pmatrix}.$$

In equilibrium, we now have exactly (38) but with  $b = -\frac{\kappa_1^- \tilde{v}_2 v_3}{(\kappa_1^- + \kappa_1^\oplus v_P^*) \tilde{\kappa}_3}$ . Hence,

$$\frac{N \mathbb{V}_{pos}[V_P]}{v_P^*} \approx \left( 1 - \frac{\kappa_1^-}{\kappa_1^- + \kappa_1^\oplus v_P^*} \right)^{-1} \left( 1 + \frac{\tilde{\kappa}_3}{\tilde{v}_2 + v_3} + \frac{\kappa_1^- \kappa_2 \tilde{\kappa}_3}{(\kappa_1^- + \kappa_1^\oplus v_P^*)^2 (\tilde{v}_2 + v_3)} - \frac{\kappa_1^- v_3}{(\kappa_1^- + \kappa_1^\oplus v_P^*) (\tilde{v}_2 + v_3)} \right). \tag{40}$$

### E Results in unscaled parameter notation

In this section we state the formulas from Sect. 3 in terms of the unscaled parameters, i.e.  $\lambda$ 's and  $\mu$ 's. The deterministic approximation or law of large numbers for the number of protein is given in the following remark (compare with Theorem 1). Throughout, we have  $X_P \approx NV_P$ .

*Remark 3 (Law of large numbers in unscaled parameters)* Given the assumptions of Theorem 1 we find that approximately (see (4))

$$\dot{X}_P \approx NF(X_P/N) = \begin{cases} \frac{M\lambda_1^+ \lambda_2 \lambda_3}{\mu_2 (\lambda_1^- + \lambda_1^+)} - \mu_3 X_P & \text{for [neu],} \\ \frac{M\lambda_1^+ \lambda_2 \lambda_3}{\mu_2 (\lambda_1^\ominus X_P + \lambda_1^+)} - \mu_3 X_P & \text{for [neg],} \\ \frac{M\lambda_1^\oplus \lambda_2 \lambda_3 X_P}{\mu_2 (\lambda_1^- + \lambda_1^\oplus X_P)} - \mu_3 X_P & \text{for [pos]} \end{cases}$$

with equilibria (see (5))

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

Next, we give the unscaled version of the central limit theorem derived in the main text (see also Theorem 2).

*Remark 4 (Central limit theorem in unscaled parameters)* Assuming the same setting as in Theorem 2 and writing down the resulting formula (7) in unscaled parameters we find

$$dX_P = N dv_P + \sqrt{N} dU \approx \left( N F(v_P) + \sqrt{N} F'(v_P) U \right) dt + \sqrt{\tilde{b}(X_P)} dW \quad (41)$$

with

$$\tilde{b}(X_P) = \begin{cases} \frac{M \lambda_1^+ \lambda_2 \lambda_3}{\mu_2 (\lambda_1^- + \lambda_1^+)} \left( \frac{2 \lambda_1^- \lambda_2 \lambda_3}{\mu_2 (\lambda_1^- + \lambda_1^+)^2} + \frac{2 \lambda_3}{\mu_2} + 1 \right) + \mu_3 X_P & \text{for [neu],} \\ \frac{M \lambda_1^+ \lambda_2 \lambda_3}{\mu_2 (\lambda_1^\ominus X_P + \lambda_1^+)} \left( \frac{2 \lambda_1^\ominus X_P \lambda_2 \lambda_3}{\mu_2 (\lambda_1^\ominus X_P + \lambda_1^+)^2} + \frac{2 \lambda_3}{\mu_2} + 1 \right) + \mu_3 X_P & \text{for [neg],} \\ \frac{M \lambda_1^\oplus \lambda_2 \lambda_3 X_P}{\mu_2 (\lambda_1^- + \lambda_1^\oplus X_P)} \left( \frac{2 \lambda_1^- \lambda_2 \lambda_3}{\mu_2 (\lambda_1^- + \lambda_1^\oplus X_P)^2} + \frac{2 \lambda_3}{\mu_2} + 1 \right) + \mu_3 X_P & \text{for [pos].} \end{cases}$$

Using Theorem 2 we can now write down the Fano factor obtained in equilibrium:

*Remark 5 (Fano factor in unscaled parameters)* From (41), we find that, with  $X_P(0) = X_P^*$ , in equilibrium (see (14))

$$\frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} \approx \begin{cases} \frac{\lambda_1^- \lambda_2 \lambda_3}{\mu_2 (\lambda_1^- + \lambda_1^+)^2} + \frac{\lambda_3}{\mu_2} + 1 & \text{for [neu],} \\ \left( 1 + \frac{\lambda_1^\ominus X_P^*}{\lambda_1^\ominus X_P^* + \lambda_1^+} \right)^{-1} \left( \frac{\lambda_1^\ominus X_P^* \lambda_2 \lambda_3}{\mu_2 (\lambda_1^- X_P^* + \lambda_1^+)^2} + \frac{\lambda_3}{\mu_2} + 1 \right) & \text{for [neg],} \\ \left( 1 - \frac{\lambda_1^- \mu_2 \mu_3}{M \lambda_1^\oplus \lambda_2 \lambda_3} \right)^{-1} \left( \frac{\lambda_1^- \lambda_2 \lambda_3}{\mu_2 (\lambda_1^- + \lambda_1^\oplus X_P^*)^2} + \frac{\lambda_3}{\mu_2} + 1 \right) & \text{for [pos].} \end{cases} \quad (42)$$

The comparison of the fluctuations in neutral version negative and positive feedback reads for unscaled parameters (see (16))

$$\begin{aligned} \frac{\mathbb{V}_{neg}[X_P^N]}{\mathbb{V}_{neu}[X_P^N]} &\approx \left(1 + \frac{\lambda_1^\ominus X_P^*}{\lambda_1^\ominus X_P^* + \lambda_1^+}\right)^{-1}, \\ \frac{\mathbb{V}_{pos}[X_P^N]}{\mathbb{V}_{neu}[X_P^N]} &\approx \left(1 - \frac{\lambda_1^-}{\lambda_1^- + \lambda_1^+ X_P^*}\right)^{-1}. \end{aligned} \tag{43}$$

For completeness, we also give the unscaled version of the refined Fano factor derived in Sect. 3.6 and Appendix D. From (18),

$$\frac{\mathbb{V}[X_P]}{\mathbb{E}[X_P]} \approx \begin{cases} \frac{\lambda_1^- \lambda_2 \lambda_3}{(\lambda_1^- + \lambda_1^+)^2 (\mu_2 + \mu_3)} + \frac{\lambda_3}{\mu_2 + \mu_3} + 1 & \text{for [neu],} \\ \left( \alpha_{neg} \frac{\mu_3}{\mu_2 + \mu_3} + \frac{\alpha_{neg}^2 \lambda_2 \lambda_3}{\lambda_1^\ominus N v_P^* (\mu_2 + \mu_3)} + \frac{\lambda_3}{\mu_2 + \mu_3} + 1 \right) / (1 + \alpha_{neg}) & \text{for [neg],} \\ \left( -\alpha_{pos} \frac{\mu_3}{\mu_2 + \mu_3} + \frac{\alpha_{pos}^2 \lambda_2 \lambda_3}{\lambda_1^- (\mu_2 + \mu_3)} + \frac{\lambda_3}{\mu_2 + \mu_3} + 1 \right) / (1 - \alpha_{pos}) & \text{for [pos],} \end{cases} \tag{44}$$

with

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

## References

Anderson D, Kurtz TG (2015) Stochastic analysis of biochemical systems. Springer, Berlin

Balázsi G, van Oudenaarden A, Collins J (2011) Cellular decision making and biological noise: from microbes to mammals. *Cell* 144(6):910–925

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

Bar-Even A, Paulsson J, Maheshri N, Carmi M, O’Shea E, Pilpel Y, Barkai N (2006) Noise in protein expression scales with natural protein abundance. *Nat Genet* 38(6):636–643

Beckskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. *Nature* 405(6786):590–593

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

Cardelli L, Csikász-Nagy A, Dalchau N, Tribastone M, Tschaikowski M (2016) Noise reduction in complex biological switches. *Sci Rep* 6:20214

Darling RWR (2002) Fluid limits of pure jump Markov processes: a practical guide, 1–16. arxiv preprint [arXiv:math/0210109](https://arxiv.org/abs/math/0210109)

Dessalles R, Fromion V, Robert P (2017) A stochastic analysis of autoregulation of gene expression. *J Math Biol* 75:1253–1283

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

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

Ethier SN, Kurtz TG (1986) Markov processes: characterization and convergence. Wiley series in probability and mathematical statistics. Wiley, New York

- Fraser D, Kærn M (2009) A chance at survival: gene expression noise and phenotypic diversification strategies. *Mol Microbiol* 71(6):1333–1340
- Gardiner C (2009) Stochastic methods. A handbook for the natural and social sciences. Springer, Berlin
- Gillespie D (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81:2340–2361
- Grönlund A, Lötstedt P, Elf J (2013) Transcription factor binding kinetics constrain noise suppression via negative feedback. *Nat Commun* 4:1864
- Hornos JEM, Schultz D, Innocentini G, Wang J, Walczak A, Onuchic J, Wolynes PG (2005) Self-regulating gene: an exact solution. *Phys Rev E* 72(5 Pt 1):051907
- Hornung G, Barkai N (2008) Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. *PLoS Comput Biol* 4(1):e8
- Kaern M, Elston TC, Blake WJ, Collins JJ (2005) Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet* 6(6):451–464
- Kang HW, Kurtz T, Popovic L (2014) Central limit theorems and diffusion approximations for multiscale Markov chain models. *Ann Appl Probab* 24:721–759
- Kim JK, Josić K, Bennett MR (2015) The relationship between stochastic and deterministic quasi-steady state approximations. *BMC Syst Biol* 9:87
- Kuehn C (2015) Multiple time scale dynamics. Springer, Berlin
- Kumar N, Singh A, Kulkarni RV (2015) Transcriptional bursting in gene expression: analytical results for general stochastic models. *PLoS Comput Biol* 11(10):e1004292
- Kurtz T (1970a) Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J Appl Probab* 8:344–356
- Kurtz T (1970b) Solutions of ordinary differential equations as limits of pure jump Markov processes. *J Appl Probab* 7:49–58
- Lestas I, Vinnicombe G, Paulsson J (2010) Fundamental limits on the suppression of molecular fluctuations. *Nature* 467(7312):174–178
- Li GW, Xie XS (2011) Central dogma at the single-molecule level in living cells. *Nature* 475:308–315
- Maamar H, Raj A, Dubnau D (2007) Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science* 317(5837):526–529
- Mitosch K, Rieckh G, Bollenbach T (2017) Noisy response to antibiotic stress predicts subsequent single-cell survival in an acidic environment. *Cell Syst* 4(4):393–403.e5
- Munsky B, Neuert G, van Oudenaarden A (2012) Using gene expression noise to understand gene regulation. *Science* 336(6078):183–187
- Pardoux E, Veretennikov AY (2001) On Poisson equation and diffusion approximation 1. *Ann Probab* 29:1061–1085
- Pardoux E, Veretennikov AY (2003) On Poisson equation and diffusion approximation 2. *Ann Probab* 31:1166–1192
- Paulsson J (2005) Models of stochastic gene expression. *Phys Life Rev* 2:157–175
- Raj A, van Oudenaarden A (2008) Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 135(2):216–226
- Ramos A, Innocentini G, Hornos J (2011) Exact time-dependent solutions for a self-regulating gene. *Phys Rev E* 83(6 Pt 1):062902
- Ramos A, Hornos J, Reinitz J (2015) Gene regulation and noise reduction by coupling of stochastic processes. *Phys Rev E* 91(2):020701
- Raser JM, O’Shea EK (2005) Noise in gene expression: origins, consequences, and control. *Science* 309(5743):2010–2013
- Seegel LA, Slemrod M (1989) The quasi-steady-state assumption: a case study in perturbation. *SIAM Rev* 31(3):446–477
- Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. *Proc Natl Acad Sci USA* 105(45):17256–17261
- Silva-Rocha R, de Lorenzo V (2010) Noise and robustness in prokaryotic regulatory networks. *Annu Rev Microbiol* 64:257–275
- Singh A (2014) Transient changes in intercellular protein variability identify sources of noise in gene expression. *Biophys J* 107(9):2214–2220
- Singh A, Soltani M (2013) Quantifying intrinsic and extrinsic variability in stochastic gene expression models. *PLoS ONE* 8(12):e84301
- Swain PS (2004) Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *J Mol Biol* 344(4):965–976

- Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci USA* 99(20):12795–12800
- Thattai M, van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci* 98(15):8614–8619
- Wang Z, Zhang J (2011) Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *Proc Natl Acad Sci* 108(16):E67–E76