Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Self-regulation in a stochastic model of chemical self-replication

Alessandro Borri[1] | Massimiliano d'Angelo*[2] | Pasquale Palumbo[2]

[1]Istituto di Analisi dei Sistemi e Informatica "A.Ruberti", Italian National Research Council (IASI-CNR), Rome, Italy
[2]Department of Biotechnologies and Biosciences, University of Milano-Bicocca, Milan, Italy

Correspondence
*Email: massimiliano.dangelo@unimib.it

Summary

In this paper we provide an analysis of noise propagation in a stochastic minimal model of chemical self-replication, where a given species can duplicate itself normally. A feedback from the end product on the source, acting as an inhibitor transcription factor, is considered. Stochasticity involves the intrinsic noise affecting gene expression, which is assumed to happen in bursts. The use of a stochastic approach is a novelty within such a framework. The investigation involves the role of the feedback: how it impacts noise attenuation with respect to different modeling choices of stochastic transcription, and with respect to different strengths of the feedback action. The quantification of noise propagation is measured by means of the so called metabolic noise, i.e. the coefficient of variation of the end product. Computations are carried out numerically, according to the Stochastic Simulation Algorithm (SSA) properly adapted for the proposed Stochastic Hybrid Systems, as well as analytically: the latter have been achieved by exploiting the linear approximation of the nonlinear terms involved, since otherwise there are no closed loop solutions for the first- and second-order moments. In such a way, noise propagation may be linked to the model parameters, with the SSA aiming at validating the approximated formulas. Results confirm the noise reduction paradigm with feedback.

KEYWORDS:
Self-regulation, chemical self-replication, stochastic hybrid systems, stochastic analysis, mRNA burst production.

1 | INTRODUCTION

Autocatalytic reactions relate to self-replication of a chemical player. Such possibility to duplicate without an external catalytic action has been suggested to be involved as a motor in species evolution. It can emerge to cope with low abundance of a required product at the beginning of prebiotic evolution[1], as well as at later stages, working as a driving force for the natural selection of information carriers[2] (see also[3] and references therein). In 1995, it has been also advanced that life originally arose as autocatalytic chemical networks[4], with auto-catalysis as a potential explanation for abiogenesis.

From a macroscopic point of view, autocatalytic reactions can be observed in abundance. Few examples are the *BZ reaction* (a class of reactions producing a chemical oscillator[5]), any system driven by light coupled to photo-polymerization reactions[6], the haloform reaction (which is one of the oldest organic reactions known[7]), and the so-called *vinegar syndrome*[8].

Fundamental classes of regulation contemplate the integration of biochemical networks[9]. It frequently happens in transcription networks, where transcription factors exert their regulation role in gene expression by properly acting as activators or

inhibitors[10]. While activation, characterized by positive feedback, is less common but still crucial for the onset of many phenomenons (such as multi-stability), the repression case, specified by a negative feedback, is more prevalent in biological systems. In fact, the latter has an intrinsic propensity to stabilize the concentrations of the chemical players involved[11].

Within this framework, different interconnections reveal to play an active role regards to the qualitative behavior of the autocatalytic reaction networks. In[12] it has been shown how the feedback of the final product onto the source of the self-replication system enhances stability of the stationary equilibria when acting as inhibitor of the source transcription, whilst it promotes bistability when acting as activator. These results have been formalized according to the standard deterministic approach, describing the system in terms of species concentrations, by means of Reaction Rate Equations (i.e. Ordinary Differential Equations (ODE) systems). However, the deterministic approach fails to account for the intrinsic noise that usually arises in biological systems at many levels, spanning from enzymatic reactions[13] to metabolism[14] and gene expression[15,16]. In many cases, cells have to guarantee the desired average level of specific proteins and to attenuate as much as possible noise fluctuations. To this end, typical feedback control strategies have been selected by Nature to make robust the required protein production. Straightforwardly, mathematical control theory has recently gained an increasing interest within the synthetic biology framework whenever cells are designed to produce prescribed homeostatic levels robustly with respect to noisy uncertainties[17].

The stochastic approach, usually developed according to Chemical Master Equations (CME),[18] allows to faithfully understand how noise propagates through the network: noise propagation can be attenuated (or enhanced) by properly exploiting negative feedbacks, and the correct setting of the model parameters may be exploited to control the end-product fluctuations. These aspects have recently shown counter-intuitive results, highlighting that a negative feedback may also increase the noise according to a specific setting of the reactions kinetics[13,19].

In this work, we investigate how noise impacts in a chemical self-replication model, and we exploit both Systems Theory mathematical formalism and numerical simulations to provide design hints regards to both the feedback role and model parameters setting. The use of the stochastic approach for chemical self-replication is the novelty of this contribution. Similarly to[12] we consider a feedback from the end-product on the source, acting as an inhibitor transcription factor. A continuous-time Markov chain is exploited to model transcription, by assuming it happens in bursts[15,16]. Between any two bursts, the chemical players kinetics evolve according to ODEs and, therefore, the whole system is modeled by means of a Stochastic Hybrid System (SHS)[20]. Noise propagation is investigated in terms of the end-product *metabolic noise*[13], i.e. the square of the Coefficient of Variation (CV^2), namely

$$CV^2 = \frac{\sigma^2}{\alpha^2}, \tag{1}$$

where σ^2 and α refer to the stationary variance and average value of an underlying process, respectively. Computations are carried out according to a twofold approach. On the one hand, analytic computations are provided, on the ground of the linearization of the nonlinear terms involved in the continuous-time Markov chain as well as in the ODE (we remark that, because of the nonlinearities, it is not possible to write, in general, the equations in closed form[21]) and these computations will be exploited to infer information on how noise propagates according to different settings of the model parameters. On the other hand, these results are validated by means of numerical Monte Carlo simulations provided by the Gillespie Stochastic Simulation Algorithm (SSA)[22]. In particular, we highlight the following contributions and results.

- The use of the stochastic approach highlights how a given qualitative behavior may be lost because of high fluctuations in gene transcription. The use of a negative feedback mitigates such a drawback.
- The feedback "noise reduction" paradigm is confirmed, i.e. the feedback decreases the impact of the noise on the system.
- A useful closed-form approximated expression of the variance of the end-product of the self-replication model is provided, according to which the metabolic noise is readily constrained to the model parameters.

The paper is organized as follows. Section 2 describes the stochastic chemical self-replication model in terms of a Stochastic Hybrid Systems. In Section 3 we derive the approximated dynamics, in terms of ODEs, of the first- and second-order moments of the underlying state variables. In particular, we show that it is possible to find a solution for a stable equilibrium of the first-order moments (the expected values), and we provide an approximated closed-form solution for the second-order moments (actually, the variances) of the underlying state variables. Section 4 motivates the choice of the model parameters. Finally, in Section 5 we analyze and validate the approximated closed-form expressions through the Gillespie SSA. Conclusions follow.

Notation. On a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$, we denote by $\mathbb{E}[\cdot]$ the expected value, and for compactness, we shall also use $\langle\cdot\rangle$. If z_t is an adapted stochastic process with value in $\mathbb{R}$ and $\phi : \mathbb{R} \to \mathbb{R}$ is a measurable function, then we

denote $\overline{\langle\phi(z)\rangle} = \lim_{t\to+\infty} \mathbb{E}\left[\phi(z_t)\right]$ when it exists. We shall indicate with capital letters the chemical players (e.g. A), by small letters the copy numbers of the player (e.g. a) and by $[A]$ the concentration of player A.

2 | STOCHASTIC CHEMICAL SELF-REPLICATION MODEL

An established reaction framework to model chemical self-replication is the following [23,12]

$$\emptyset \xrightarrow{r(B)} A \tag{2}$$

$$\emptyset \xrightarrow{r(B)} C \tag{3}$$

$$A + C \xrightarrow{k_1} B \tag{4}$$

$$A + C + B \xrightarrow{k_2} 2\,B \tag{5}$$

$$B + E \underset{k_4}{\overset{k_3}{\rightleftharpoons}} BE \tag{6}$$

$$BE \xrightarrow{k_5} E \tag{7}$$

Reactions (2) and (3) represent the production of the sources A and C with reaction rates r regulated in feedback by the product B. In other words, B is a transcription factor for the production of proteins A and C. Besides, A and C contribute to produce B according to both an uncatalyzed reaction (4) and to the autocatalytic replication of B, reaction (5). Finally, B degrades according to the action of an enzyme E, reactions (6) and (7). The working assumptions are that

i) the system of reactions has no limiting reagents so that, because both sources are produced according to the same reaction rate $r(B)$, one may assume that the concentrations of A and C are always the same, i.e. $[A] = [C]$ (see [24,12] for details). That means, reactions (4) and (5) simplify into

$$2\,A \xrightarrow{k_1} B \tag{8}$$

$$2\,A + B \xrightarrow{k_2} 2\,B. \tag{9}$$

ii) the binding/unbinding reactions providing precursor complex BE to produce B degradation occur at a faster rate, so that the Quasi-Steady State Approximation applies [25] and the classical Michaelis-Menten (MM) saturating function models B degradation rate, so that reactions (6)-(7) simplify as

$$B \xrightarrow{k_5\Lambda(B)} \emptyset, \qquad \Lambda(b) = \frac{b}{b+\zeta} \tag{10}$$

where ζ is the Michaelis-Menten constant.

According to a deterministic approach, an Ordinary Differential Equations (ODE) system can be written, describing the time course of the chemical players concentrations, providing different qualitative behaviors according to different setting of the feedback action of B onto the source production (see [12,26] and references therein). The feedback action exerted by B has been modeled in [12] by dealing with B as a transcription factor regulating the expression of the source. Deterministic approaches may be found in [24,27,28]

In this work, we extend the model to account also for the intrinsic noise affecting gene expression [15,16]. The working hypothesis is that the production of mRNA, encoding for A expression, occurs in bursts. Transcriptional bursting in gene expression has been proven for both eukaryotic [29] and prokaryotic [15] cells, according to single-cell experiments. Gene transcription (i.e. mRNA production) is the only source of stochasticity introduced, so that the whole system is modeled according to a Stochastic Hybrid System (SHS) framework, where an Ordinary Differential Equation continuously describes the evolution of the state variables and specific stochastic discrete events reset the state vector whenever they occur. Such a SHS framework has been often exploited in the recent years (see, among the others, [19,20,30,31,32]).

The SHS approach neglects other (than bursty transcription) noise sources and loses validity if the intrinsic noise in reactions is large enough (e.g. when the copy numbers are not that large). In these cases, the Effective Mesoscopic Rate Equation (EMRE)

approach, derived from a systematic expansion method and proposed in [33] (see also [34]), provides a trustworthy route to estimate the average concentrations as coming from the master equations.

More in details, let a, b denote the copy number of A and B, respectively, and m denotes the copy number of mRNA transcript of A. m production represents the unique source of noise, occurring in bursts and modeled according to the following reset maps

$$m(t) = \varphi_j(m(t^-)) = m(t^-) + j \qquad j \in \mathbb{N}, \tag{11}$$

with j denoting the size of the burst occurring according to the transition intensities

$$\lambda_j(b) = \Psi(b)p(j), \quad \text{with} \quad \Psi(b) = r_0 + (\eta - r_0)\frac{\theta^\nu}{\theta^\nu + b^\nu}, \qquad j \in \mathbb{N}, \tag{12}$$

where $\Psi(b)$ describes the feature of the feedback exerted by the transcription factor, here assumed to act as an inhibitor. r_0 is a leak transcription rate (occurring even when there is abundance of inhibitor), $\eta > r_0$ is the maximal transcription rate occurring in absence of inhibitor, θ is the end-product copy number according to which the feedback exerts half of its maximal strength over r_0, and $p(\cdot)$ is a probability mass function with positive support of a random variable with finite moments of any order (we shall see later the possible choices of such distribution). In other words, the probability that a burst production (11) takes place in an infinitesimal interval $[t, t + \mathrm{d}t)$ is given by $\lambda_j(b)\mathrm{d}t$. Within any two bursts, the state variables m, a, b evolve according to the following ODE

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -\gamma\, m \tag{13}$$

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \kappa m - k_1 a^2 - k_2 a^2 b \tag{14}$$

$$\frac{\mathrm{d}b}{\mathrm{d}t} = k_1 a^2 + k_2 a^2 b - k_5 \Lambda(b), \qquad \Lambda(b) = \frac{b}{b + \zeta}, \tag{15}$$

where a and b dynamics comprise the reaction rates from reactions (8) and (9) (according to standard *mass action law*) and (10). As regards to the mRNA dynamics, a linear degradation is considered (γ in (13) stands for mRNA degradation rate) and the translation rate is also assumed to be a linear function of the mRNA (κ in (14) stands for the source translational rate per mRNA).

3 | FIRST AND SECOND ORDER MOMENTS APPROXIMATIONS

The computation of the metabolic noise (1) strongly relies on the first- and second-order moments of the state variables involved in the SHS. In general, if $x \in \mathbb{R}^n$ describes the ODE $\dot{x} = f(x)$ associated to the SHS, $\varphi_j(x^-)$ refer to the reset maps and $\lambda_j(x)$ refer to the transition intensities, given a continuously differentiable function $\psi : \mathbb{R}^n \to \mathbb{R}$, it is possible to write the dynamics of the expectation of $\psi(x)$ as [20]

$$\frac{\mathrm{d}\mathbb{E}\left[\psi(x(t))\right]}{\mathrm{d}t} = \mathbb{E}\left[\frac{\mathrm{d}\psi(x(t))}{\mathrm{d}x} f(x(t))\right] + \sum_{j=1}^{\infty} \mathbb{E}\left[\left(\psi(x(t) + \Delta_j) - \psi(x(t))\right) \lambda_j(x(t))\right], \tag{16}$$

where Δ_j, with $j \in \mathbb{N}$, is the size of the jump of x because of the reset $\varphi_j(x^-)$.

3.1 | First-order moments

Let us write the ODE (13)–(15) in the compact form

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(x(t)), \qquad \text{with} \quad x(t) = \begin{bmatrix} m(t) \\ a(t) \\ b(t) \end{bmatrix}, \qquad f(x(t)) = \begin{bmatrix} f_1(x(t)) \\ f_2(x(t)) \\ f_3(x(t)) \end{bmatrix} = \begin{bmatrix} -\gamma m(t) \\ \kappa m(t) - k_1 a^2(t) - k_2 a^2(t) b(t) \\ k_1 a^2(t) + k_2 a^2(t) b(t) - k_5 \Lambda(b(t)) \end{bmatrix} \tag{17}$$

By virtue of (16), we can compute the dynamics of the moments of the proposed self-replication model. In particular, taking into account (11), (12) and the ODE (13)–(15), we can write the equations of the first-order moment[1]

$$\frac{\mathrm{d}\langle m(t)\rangle}{\mathrm{d}t} = -\gamma\langle m(t)\rangle + \mu r_0 + \mu(\eta - r_0)\left\langle \frac{\theta^\nu}{\theta^\nu + b^\nu(t)}\right\rangle, \tag{18}$$

$$\frac{\mathrm{d}\langle a(t)\rangle}{\mathrm{d}t} = \kappa\langle m(t)\rangle - k_1\left\langle a^2(t)\right\rangle - k_2\left\langle a^2(t)b(t)\right\rangle, \tag{19}$$

$$\frac{\mathrm{d}\langle b(t)\rangle}{\mathrm{d}t} = k_1\left\langle a^2(t)\right\rangle + k_2\left\langle a^2(t)b(t)\right\rangle - k_5\left\langle \frac{b(t)}{\zeta + b(t)}\right\rangle, \tag{20}$$

where we set the average burst size $\mu = \sum_{j=1}^{\infty} jp(j)$ which is finite because of the assumptions on p. Due to the nonlinear terms, it is not possible to write the expression of the first order (any order in general) moments in closed form. For, it is common, according to the van Kampen's system-size expansion for Chemical Master Equations [18] (see e.g. [30,19]) to make use of the linear approximation of the nonlinear terms around a fixed point $\bar{x} = (\bar{m}, \bar{a}, \bar{b})^\top$. In this spirit, (18)–(20) can be approximated by

$$\frac{\mathrm{d}\langle m(t)\rangle}{\mathrm{d}t} \approx -\gamma\langle m(t)\rangle + \mu\Psi(\bar{b}) + \mu\Psi'(\bar{b})\left(\langle b(t)\rangle - \bar{b}\right), \tag{21}$$

$$\frac{\mathrm{d}\langle a(t)\rangle}{\mathrm{d}t} \approx \kappa\langle m(t)\rangle - k_1\bar{a}^2 - 2k_1\bar{a}\left(\langle a(t)\rangle - \bar{a}\right) - k_2\bar{a}^2\bar{b} - 2k_2\bar{a}\bar{b}\left(\langle a(t)\rangle - \bar{a}\right) - k_2\bar{a}^2\left(\langle b(t)\rangle - \bar{b}\right), \tag{22}$$

$$\frac{\mathrm{d}\langle b(t)\rangle}{\mathrm{d}t} \approx k_1\bar{a}^2 + 2k_1\bar{a}\left(\langle a(t)\rangle - \bar{a}\right) + k_2\bar{a}^2\bar{b} + 2k_2\bar{a}\bar{b}\left(\langle a(t)\rangle - \bar{a}\right) + k_2\bar{a}^2\left(\langle b(t)\rangle - \bar{b}\right) - k_5\Lambda(\bar{b}) - k_5\Lambda'(\bar{b})\left(\langle b(t)\rangle - \bar{b}\right). \tag{23}$$

If the point $\bar{x} = (\bar{m}, \bar{a}, \bar{b})^\top$ is set such that

$$-\gamma\bar{m} + \mu\Psi(\bar{b}) = 0 \tag{24}$$

$$\kappa\bar{m} - k_1\bar{a}^2 - k_2\bar{a}^2\bar{b} = 0 \tag{25}$$

$$k_1\bar{a}^2 + k_2\bar{a}^2\bar{b} - k_5\Lambda(\bar{b}) = 0, \tag{26}$$

then equations (21)–(23) admit a unique stationary point exactly at $\bar{x} = (\bar{m}, \bar{a}, \bar{b})^\top$. Clearly, the validity of the proposed approximation is stronger in case of small noise fluctuations around the stationary point.

From a mathematical viewpoint, the search for the solution of (24)–(26) may be seen as a particular case of the wider investigation carried out in [12]. After straightforward manipulation of (24)–(26), the solution $\bar{b}$ satisfies

$$\Psi(\bar{b}) = \frac{k_5\gamma}{\kappa\mu}\Lambda(\bar{b}), \tag{27}$$

and we can give a graphical interpretation of it. Indeed, the left-hand side of (27) is the monotonically decreasing function defined in (12), starting from η for $\bar{b} = 0$ and asymptotically approaching $r_0 < \eta$ for $\bar{b} \mapsto +\infty$. On the other hand, the right-hand side of (27) is a monotonically increasing function starting from 0 for $\bar{b} = 0$ and asymptotically approaching $k_5\gamma/(\kappa\mu)$ for $\bar{b} \mapsto +\infty$. Therefore, a unique solution occurs if, and only if, the lower bound of the left-hand side is lower than the upper bound of the right-hand side, see Fig. 1 for the curves at different values of parameter μ. That means

$$r_0 < \frac{k_5\gamma}{\kappa\mu}. \tag{28}$$

It is worth noticing that the average size of the burst μ plays a crucial role, since, if it increases enough to violate condition (28), the equilibrium is lost. On the other hand, for a trivial leak transcription rate $r_0 = 0$, condition (28) is always satisfied, whatever is the burst average size.Such a result is coherent with the one reported in [35] where a variant of the classical Michaelis-Menten enzymatic framework is considered, accounting for substrate replenishment. In that case, steady-state solutions were proved to occur if the input rate of replenishment of the substrate was smaller than the maximum rate of production of the product, i.e. the maximum clearance rate for B in the present case. Indeed, by increasing the transcription leakage rate r_0 we increase the production rate for B, namely $\frac{\kappa\mu\Psi(\bar{b})}{\gamma}$ from (24)-(26), that may exceed the maximum rate of degradation k_5, thus losing stability according to (28): the enzyme will not be able to degrade B fast enough and hence the stationary state is lost.

In Figure 1 one can infers the graphical solution of (27) for the parameters setting of Section 4 for different values of the average burst size μ. It is apparent that, by increasing μ, the average copy number of species B definitely increases up to infinity for $\mu \mapsto k_5\gamma/(\kappa r_0)$.

[1] we recall the notation $\mathbb{E}[\cdot] = \langle\cdot\rangle$.

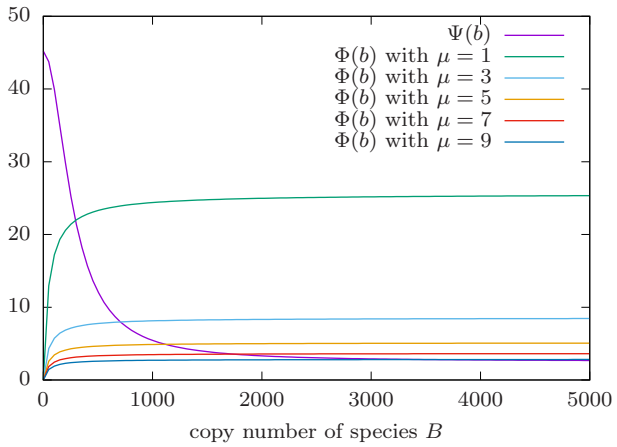

FIGURE 1 Graphical interpretation of (27), where $\Psi(b)$ is the transition intensity defined in (12) and $\Phi(b) := \frac{k_5\gamma}{\kappa\mu}\Lambda(\bar{b})$, for different values of the average burst size μ. The model parameters have been set according to Section 4.

With respect to the role of the feedback, by assuming to increase parameter θ enough to make negligible the inhibitor action of B, constraint (27) becomes

$$\eta = \frac{k_5\gamma}{\kappa\mu}\Lambda(\bar{b}) \tag{29}$$

that means

$$\eta < \frac{k_5\gamma}{\kappa\mu} \tag{30}$$

an inequality much difficult to obtain, since $\eta > r_0$. In summary, the greater is the average burst size μ, the more a negative feedback would be required to ensure the existence of an equilibrium point. More in general, by denoting with

$$\bar{\lambda}_j = \bar{\eta}p(j), \qquad j \in \mathbb{N}, \tag{31}$$

the transition intensities for the no-feedback case, if

$$\bar{\eta} = \Psi(\bar{b}) \tag{32}$$

with $\bar{b}$ standing for the stationary value achieved according to the feedback case, then parameters are set in order to keep fixed the same equilibrium point with or without the feedback.

In what follows, we choose $\bar{x} = (\bar{m}, \bar{a}, \bar{b})^\top$ satisfying (24)–(26) and thus we have $\lim_{t\to+\infty}\langle x(t)\rangle = \bar{x}$ in our approximation.

3.2 | Second-order moments

Similarly to the first-order moments computations, by applying (16) for $\psi(x) = x_i x_j$, $i, j = 1, 2, 3$, second-order moments equations are achieved. According to the same linear approximations exploited in the ODE nonlinearities (13)-(15) as well as in the nonlinear transition intensities (12), the following 6th-order linear algebraic system is achieved for the centered second-order moments (see Appendix for the details)

$$2\gamma\chi_1 - 2\tilde{\Psi}\chi_5 = \Gamma\Psi(\bar{b}) \tag{33}$$

$$-\kappa\bar{\delta}\chi_2 + \kappa\chi_4 - \bar{k}_2\chi_6 = 0 \tag{34}$$

$$\tilde{\Lambda}\chi_3 + \kappa\bar{\delta}\chi_6 = 0 \tag{35}$$

$$-\kappa\chi_1 + \left(\gamma + \kappa\bar{\delta}\right)\chi_4 + \bar{k}_2\chi_5 - \tilde{\Psi}\chi_6 = 0 \tag{36}$$

$$-\tilde{\Psi}\chi_3 - \kappa\bar{\delta}\chi_4 + (\gamma - \tilde{\Lambda})\chi_5 = 0 \tag{37}$$

$$\kappa\bar{\delta}\chi_2 - \bar{k}_2\chi_3 + \kappa\chi_5 + \left(\tilde{\Lambda} - \kappa\bar{\delta}\right)\chi_6 = 0 \tag{38}$$

in the unknown variables

$$\chi := \left(\left(\overline{\langle m^2\rangle} - \bar{m}^2\right), \left(\overline{\langle a^2\rangle} - \bar{a}^2\right), \left(\overline{\langle b^2\rangle} - \bar{b}^2\right), \left(\overline{\langle ma\rangle} - \bar{m}\bar{a}\right), \left(\overline{\langle mb\rangle} - \bar{m}\bar{b}\right), \left(\overline{\langle ab\rangle} - \bar{a}\bar{b}\right)\right)^\top \in \mathbb{R}^6,$$

where we recall the overline symbol stands for the limit as $t \to +\infty$, and

$$\Gamma := \sum_j j^2 p(j), \qquad \bar{\delta} = 2\frac{\bar{m}}{\bar{a}}, \qquad \bar{k}_2 = k_2 \bar{a}^2, \qquad \tilde{\Lambda} = \bar{k}_2 - k_5 \Lambda'(\bar{b}), \qquad \tilde{\Psi} = \mu \Psi'(\bar{b}).$$

Clearly, equations (33)–(38) can be written in matrix form as

$$\Pi \chi = \Upsilon, \tag{39}$$

where the expressions of Π and Υ can be readily derived (see equation (A7) in the Appendix). In Appendix the analytical expression is reported, providing the explicit solution to (39) (equation (A8)).

Finally, since we are interested in the noise propagation through the Coefficient of Variation (CV, see (1)), we pay particular attention for the variance of the end product $\boldsymbol{B}$, namely the third component of χ, i.e. $\sigma_b^2 = \left(\overline{\langle b^2 \rangle} - \bar{b}^2\right)$, according to which, the closed form expression for CV_b^2 is

$$CV_b^2 = \frac{1}{\Delta} \frac{\bar{\delta}\kappa^3 \Gamma \Psi(\bar{b})}{\bar{b}^2} (\tilde{\Lambda} - \gamma - \bar{\delta}\kappa), \tag{40}$$

where Δ is defined in (A15).

As regards the fruitfulness of the achieved analytical results obtained according to the aforementioned linear approximation, differently from classical Monte Carlo simulations provided by the Gillespie SSA (in this case the τ-leaping version is required because of the SHS framework), they are easy to handle and allow to infer relationships between noise fluctuations (e.g. measured by the CV) and model parameters. This point may assume paramount importance in the case of synthetic design of an autocatalytic reaction whenever a fine tuning of the model parameters is required. Anyhow, any possible hint/inference needs to be validated by the numerical simulations. Section 5 will show how to exploit such a virtuous cycle binding approximated analytical results to validation by Monte Carlo simulations.

4 | PARAMETERS SETTING

Model parameters have been set by properly exploiting the ones provided by [12]. Since in [12] parameters are given in terms of a dimensionless ODE, without loss of generality we have assumed μM and seconds wherever dealing with concentrations and time units. Finally, we passed from μM to molecules by multiplying the values by the volume $V = 10^{-15}$L (taken from *E. Coli* volume [13,36]) and the Avogadro number $N_0 = 6.022 \cdot 10^{23}$. This way, all but parameters γ and κ are set. With respect to γ we assume an mRNA half-life of 2 minutes [28], so that $\gamma = \log(2)/120 = 5.8 \cdot 10^{-3}\text{s}^{-1}$. Finally, κ, the translation rate per mRNA molecule, is taken from [37] and is set equal to 24.6 amino acids per seconds: by assuming an average number of 300 amino acid/protein [38], it comes that parameter κ is set equal to 0.0819 protein copy number per mRNA copy number per second. We summarize the parameters of the model in Table 1.

We point out that, with this setting of the parameters, the condition (28) for the existence of an equilibrium point of the first order moment in terms of the burst size μ is

$$\mu < \frac{k_5 \gamma}{\kappa r_0} \approx 10, \tag{41}$$

and we see in Fig. 1 the graphical solution (for $\bar{b}$) of the algebraic system (24)–(26) for some values of the burst size, namely $\mu \in \{1, 3, 5, 7, 9\}$.

TABLE 1 Nominal model parameters. ♯ denotes A or B the copy numbers, whilst ‡ denotes mRNA copy numbers.

$\gamma =$	$5.8 \cdot 10^{-3} s^{-1}$	$r_0 =$	$2.56 \ddagger\, s^{-1}$
$\kappa =$	$0.08 \sharp\, \ddagger^{-1}\, s^{-1}$	$\eta =$	$45.21 \sharp^2 s^{-1}$
$k_1 =$	$1.5 \cdot 10^{-5} \sharp^{-1} s^{-1}$	$\theta =$	$270 \sharp$
$k_2 =$	$3.04 \cdot 10^{-6} \sharp^{-2} s^{-1}$	$\nu =$	2
$k_5 =$	$361.32 \sharp s^{-1}$	$\zeta =$	$49.17 \sharp$

5 | ANALYSIS OF NOISE PROPAGATION FROM GENE EXPRESSION

Two main frameworks are here considered. One refers to a deterministic burst size, i.e. we assume that transcription occurs according to a burst of fixed size whenever it happens. It is worth noticing that, although the burst size is fixed (i.e. not stochastic), the sojourn time is a random variable [18]. The other refers to a geometric probability distribution for the burst size. The τ-leaping version of the Gillespie SSA [22] is adopted to make numerical Monte Carlo simulations of the Stochastic Hybrid System described by the discrete jumps (11)-(12) and by the continuous flow (13)-(15) between any two jumps. Indeed, because of the continuous-time variation of species b, the transition intensity (12) also varies continuously in time, which makes the underlying stochastic process non-homogeneous and the exact Gillespie algorithm inapplicable. As a consequence, the τ-leaping algorithm is exploited in our simulations, where the transition intensity (12) is considered to be changing with sampling time of duration $\tau > 0$, and is approximated as a constant in the inter-sampling period.

5.1 | Deterministic burst size

According to the deterministic (or fixed) burst size assumption [33], there is only one transition intensity function (12), i.e.

$$\lambda_j(b) = \Psi(b)\delta(j-\mu),$$

where $\Psi(\cdot)$ is defined in (12) and $\delta(\cdot)$ is the Kronecker function, assuming a nontrivial value (equal to 1) only for a null entry. Because of (41), in order to have an equilibrium point satisfying (24)–(26), the deterministic burst size $\mu \in \{1, 2, \dots, 9\}$.

Fig. 2 shows the probability density functions of the equilibrium point $\bar{x} = (\bar{m}, \bar{a}, \bar{b})$ for different values of the burst size. In particular, the purple curve indicates the probability density function of $\bar{m}$ (top left), $\bar{a}$ (top right) and $\bar{b}$ (bottom), for different values of the deterministic burst size ($\mu = 1, 2, 3, 4$), obtained with a τ-leaping Gillespie stochastic simulation with time horizon 10^7s and $\tau = 0.05$s of the underlying SHS (13)–(15). Instead, the pale green impulse indicates the solution $\bar{m}$ (top left), $\bar{a}$ (top right) and $\bar{b}$ (bottom) of (24)–(26), i.e. the deterministic equilibrium. We notice that, by increasing the burst size μ, the variance of such densities increases.

Fig. 3 summarizes the first-order moments (expected values) of the aforementioned densities and the solution of (24)–(26), i.e. the deterministic equilibrium, for burst sizes $\mu < 10$ (which is the constraint (41)). Finally, Fig. 4 presents the behaviour of the squared Coefficient of Variation CV^2 (40) computed through the analytical approximated closed form expressions and through the τ-leaping Gillespie SSA for the feedback case and the no-feedback case. The curves exhibit good agreement between the *a priori* approximated theoretical values and the *a posteriori* empirical values. The comparison feedback vs no-feedback is carried out by assigning to the no-feedback case the same equilibrium points obtained for the feedback case (and pictured in Fig. 3). To this end, parameter $\bar{\eta}$ is chosen according to (32). Fig. 4 confirms the feedback noise reduction paradigm, i.e. feedback works in order to reduce noise propagation.

5.2 | Stochastic burst size

As anticipated at the beginning of the Section, we adopt here a geometric probability distribution $p(j)$ for the transition intensities (12), namely:

$$p(j) = \left(1 - \frac{1}{\mu}\right)^{k-1} \frac{1}{\mu}, \qquad \text{with} \quad \sum_{j=1}^{+\infty} j p(j) = \mu. \tag{42}$$

Note that in this case μ is the *average* burst size, and we recall again that, even in this case, μ must satisfy equation (41) in order to let $\bar{x}$ be an equilibrium point satisfying (24)–(26). Figs. 5 and 6 are obtained according to the aforementioned philosophy to fix the same equilibrium points for both the feedback/no-feedback cases, by setting parameter $\bar{\eta}$ in (31) according to (32). Again, these simulations stress the improvement provided by the feedback in noise propagation.

Dealing with both deterministic and stochastic burst sizes, it has to be stressed that the higher mismatch between analytical and numerical results is observed for large average burst sizes (see Figs. 4-6): indeed, by increasing μ, we get progressively closer to violating condition (41) on the existence of the equilibrium point (this is apparent for the simulations reported in Fig.4, where we come very close to the upper bound for μ). This fact might make both numerical and approximated analytical results less reliable. On the other hand, it is worth noticing that the feedback enhances results reliability, since the maximum discrepancy error between numerical and analytical results is strongly reduced (from a 9.88% for the feedback case to a 27.38% for the no-feedback case). A last set of simulations involves the following framework: (i) stochastic burst size, achieved with a geometric

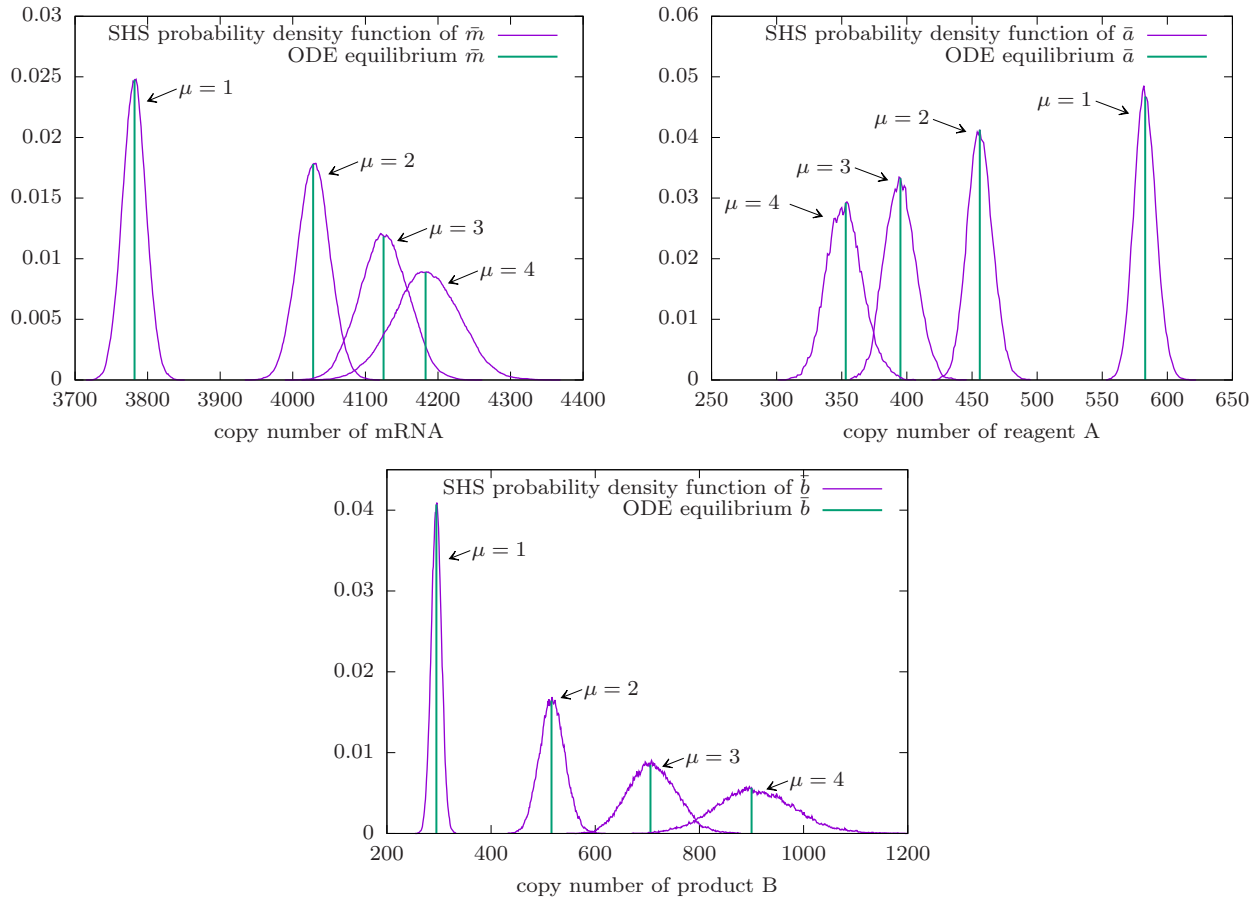

FIGURE 2 Case of deterministic burst size. Probability density functions of the equilibrium point $\bar{x} = (\bar{m}, \bar{a}, \bar{b})$ of the SHS for different values of the burst size ($\mu = 1, 2, 3, 4$) obtained with a τ-leaping Gillespie stochastic simulation with time horizon 10^7s and $\tau = 0.05$s.

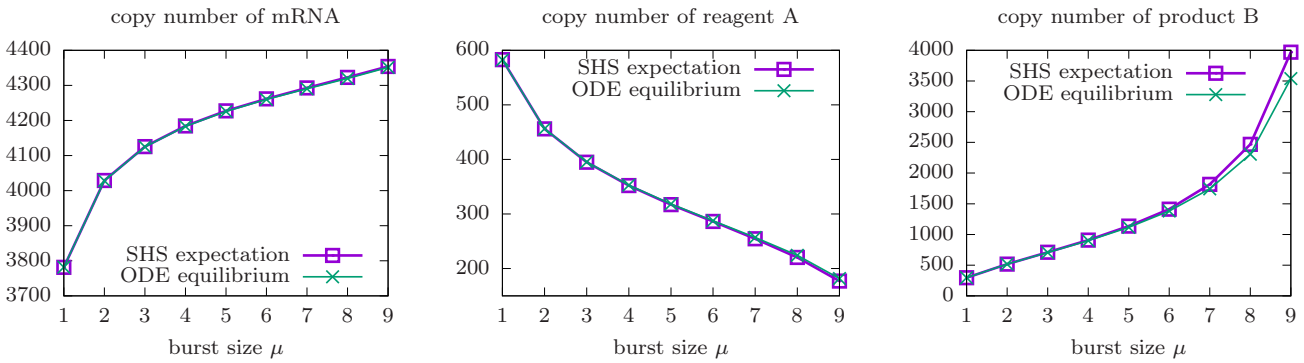

FIGURE 3 Case of deterministic burst size. Steady-state first-order moments of the SHS for $\mu = 1, 2, \ldots, 9$ obtained with a τ-leaping Gillespie stochastic simulation with time horizon 10^7s and $\tau = 0.05$s and the deterministic equilibrium.

distribution providing an average burst size $\mu = 4$; (ii) all but θ model parameters are kept fixed to the values of Table 1. Figs. 7, 8 and 9 analyze the behaviour of the steady-state first-order moments, standard deviations and CV^2 with respect to $1/\theta$ which corresponds to the repression strength[13]. As a matter of fact, when the repression strength vanishes, that is $1/\theta \mapsto 0$ (i.e. when

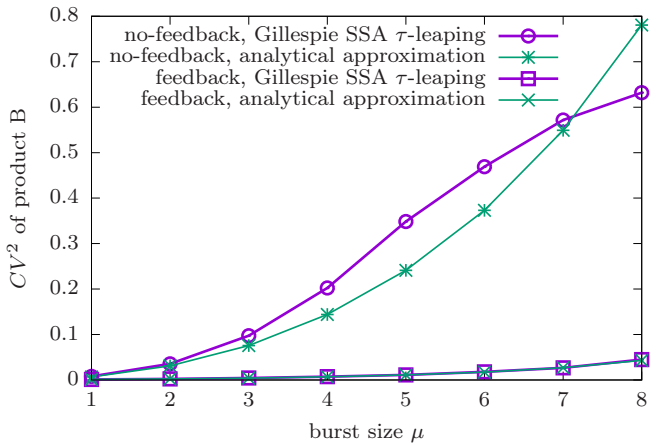

FIGURE 4 Case of deterministic burst size. Steady-state CV^2 of product B for different values of the deterministic burst size for the feedback case and no-feedback case. In particular, the CV^2 is obtained through a τ-leaping Gillespie stochastic simulation with time horizon 10^7s and $\tau = 0.05$s and the analytical closed form approximation (40).

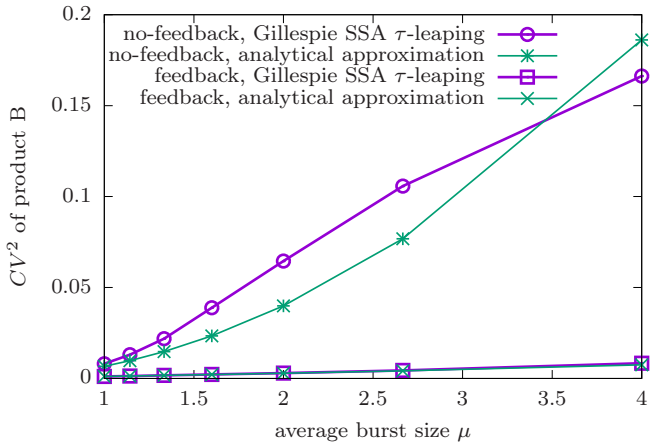

FIGURE 5 Case of stochastic burst size. Steady-state CV^2 of product B for different average values μ of the stochastic burst size for the feedback case and no-feedback case. In particular, the CV^2 is obtained through a τ-leaping Gillespie stochastic simulation with time horizon 10^7s and $\tau = 0.05$s and the analytical closed form approximation (40).

$\theta \mapsto +\infty$), the feedback does not exert its influence anymore, since $\Psi(b) \simeq \eta$. This kind of analysis may be of some help within the synthetic biology framework, in order to tune the model parameters to reduce noise propagation. Indeed, by looking at Fig. 9 it is apparent that there is not a monotonic behavior by varying the repression strength: there can be found a worst case when CV^2 reaches its maximum value for a given θ.

As usual, we notice the good agreement between the *a priori* approximated theoretical values (obtained through Eqs. (40) and (A8)) and the *a posteriori* empirical values (provided by the τ-leaping Gillespie SSA with time horizon 10^7s and $\tau = 0.05$s).

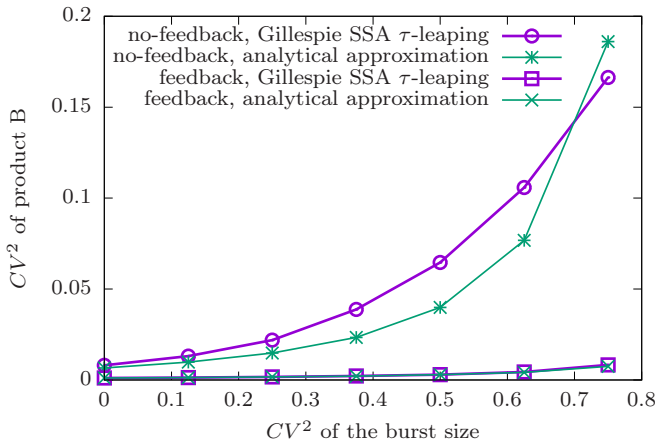

FIGURE 6 Case of stochastic burst size. Steady-state CV^2 of product B for different values of the burst size CV^2 for the feedback case and no-feedback case. In particular, the CV^2 of product B is obtained through a τ-leaping Gillespie stochastic simulation with time horizon 10^7s and $\tau = 0.05$s and the analytical closed-form approximation (40).

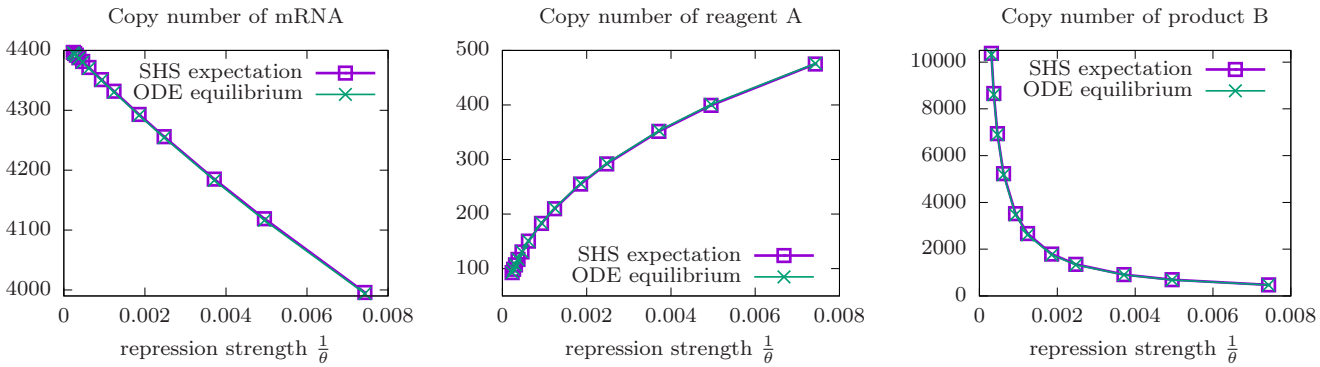

FIGURE 7 Case of stochastic burst size. Steady-state first-order moments of the SHS obtained with a τ-leaping Gillespie stochastic simulation with time horizon 10^7s and $\tau = 0.05$s and the deterministic equilibrium with respect to the repression strength, i.e. $\frac{1}{\theta}$. In particular, the average burst size has been set $\mu = 4$.

6 | CONCLUSIONS AND FUTURE DIRECTIONS

For the proposed self-regulation stochastic model, the noise reduction paradigm in the presence of negative feedback is confirmed. A closed form approximated expression of the variance of the product of the chemical network is provided and it allows to analyze the noise propagation from gene expression in an easier way than according to numerical simulations (like SSA), potentially cumbersome and time-consuming. Many other future directions can be taken in consideration. A positive feedback case has shown to provide bi-stability in a deterministic framework: it would be interesting to investigate whether a bimodal stationary probability distribution comes out in the stochastic framework. Further future directions could be the use of a more accurate model accounting for cell division and cell cycle duration variability that have profound effects on the shape of mRNA distributions[39], or even to account for a spatial dimension like in[26].

□

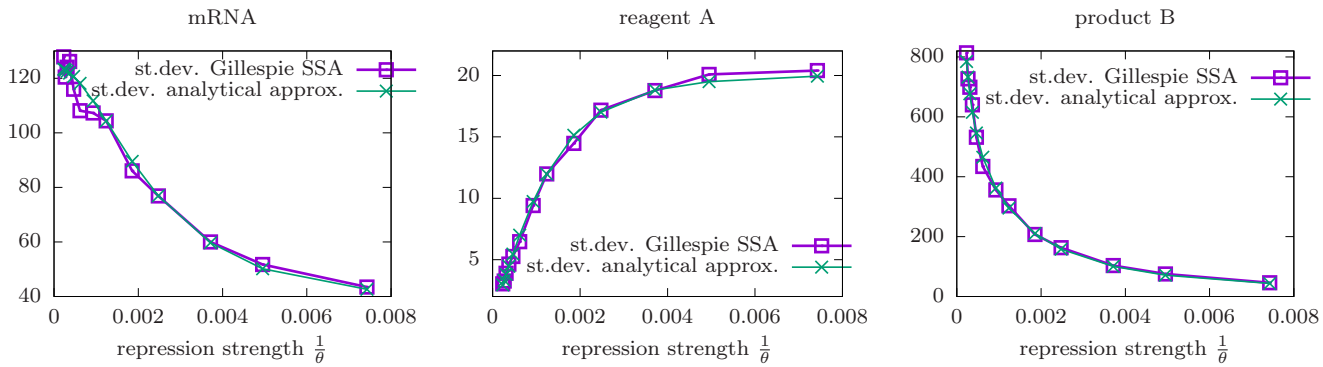

FIGURE 8 Case of stochastic burst size. Standard deviations of the steady-state first-order moments (τ-leaping Gillespie stochastic simulation with time horizon 10^7s and $\tau = 0.05$s) and the analytic approximated standard deviations computed through (A8) with respect to the repression strength, i.e. $\frac{1}{\theta}$. In particular, the average burst size has been set $\mu = 4$.

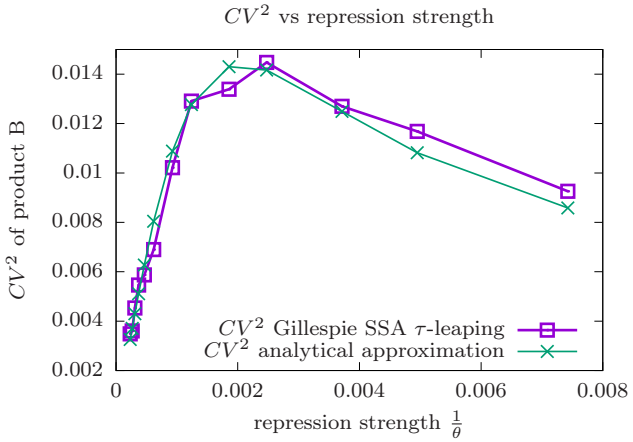

FIGURE 9 Case of stochastic burst size. Steady-state CV^2 of product B (τ-leaping Gillespie stochastic simulation with time horizon 10^7s and $\tau = 0.05$s) and its analytic approximation computed through (40) with respect to the repression strength, i.e. $\frac{1}{\theta}$. In particular, the average burst size has been set $\mu = 4$.

APPENDIX

A EXPRESSIONS OF THE APPROXIMATED SECOND-ORDER MOMENT DYNAMICS

We provide here the expressions of the approximated ODEs describing the dynamics of the second-order moments. Define $\Gamma = \sum_j j^2 p(j)$. By taking into account equation (16) and according to the linear approximations of the ODE nonlinearities

(13)-(15) as well as of the nonlinear transition intensities (12) we get

$$\frac{\mathrm{d}\left\langle m^2(t)\right\rangle}{\mathrm{d}t} \approx -2\gamma\left\langle m^2(t)\right\rangle + 2\mu\Psi'(\bar{b})\left\langle m(t)b(t)\right\rangle + 2\mu\left(\Psi(\bar{b}) - \bar{b}\Psi'(\bar{b})\right)\left\langle m(t)\right\rangle + \Gamma\Psi'(\bar{b})\left(\left\langle b(t)\right\rangle - \bar{b}\right) + \Gamma\Psi(\bar{b}), \tag{A1}$$

$$\frac{\mathrm{d}\left\langle a^2(t)\right\rangle}{\mathrm{d}t} \approx -4\bar{a}\left(k_1 + k_2\bar{b}\right)\left\langle a^2(t)\right\rangle + 2\kappa\left\langle m(t)a(t)\right\rangle - 2k_2\bar{a}^2\left\langle a(t)b(t)\right\rangle + 2\bar{a}^2\left(k_1 + 2k_2\bar{b}\right)\left\langle a(t)\right\rangle, \tag{A2}$$

$$\frac{\mathrm{d}\left\langle b^2(t)\right\rangle}{\mathrm{d}t} \approx 2\left(k_2\bar{a}^2 - k_5\Lambda'(\bar{b})\right)\left\langle b^2(t)\right\rangle + 4\bar{a}\left(k_1 + k_2\bar{b}\right)\left\langle a(t)b(t)\right\rangle - 2\left(k_1\bar{a}^2 + 2k_2\bar{a}^2\bar{b} + k_5\Lambda(\bar{b}) - k_5\bar{b}\Lambda'(\bar{b})\right)\left\langle b(t)\right\rangle, \tag{A3}$$

$$\begin{aligned}\frac{\mathrm{d}\left\langle m(t)a(t)\right\rangle}{\mathrm{d}t} \approx{}& \kappa\left\langle m^2(t)\right\rangle - \left(\gamma + 2k_1\bar{a} + 2k_2\bar{a}\bar{b}\right)\left\langle m(t)a(t)\right\rangle - k_2\bar{a}^2\left\langle m(t)b(t)\right\rangle + \mu\Psi'(\bar{b})\left\langle a(t)b(t)\right\rangle \\ &+ \mu\left(\Psi(\bar{b}) - \bar{b}\Psi'(\bar{b})\right)\left\langle a(t)\right\rangle - \bar{a}\left(k_1 - 2k_1\bar{a} - 2k_2\bar{a}\bar{b}\right)\left\langle m(t)\right\rangle,\end{aligned} \tag{A4}$$

$$\begin{aligned}\frac{\mathrm{d}\left\langle m(t)b(t)\right\rangle}{\mathrm{d}t} \approx{}& 2\bar{a}\left(k_1 + k_2\bar{b}\right)\left\langle m(t)a(t)\right\rangle - \left(\gamma - k_2\bar{a}^2 + k_5\Lambda'(\bar{b})\right)\left\langle m(t)b(t)\right\rangle + \mu\Psi'(\bar{b})\left\langle b^2(t)\right\rangle \\ &+ \left(k_1\bar{a} - 2k_1\bar{a}^2 - 2k_2\bar{a}^2\bar{b} - k_5\Lambda(\bar{b}) + k_5\bar{b}\Lambda'(\bar{b})\right)\left\langle m(t)\right\rangle + \mu\left(\Psi(\bar{b}) - \bar{b}\Psi'(\bar{b})\right)\left\langle b(t)\right\rangle,\end{aligned} \tag{A5}$$

$$\begin{aligned}\frac{\mathrm{d}\left\langle a(t)b(t)\right\rangle}{\mathrm{d}t} \approx{}& 2\bar{a}\left(k_1 + k_2\bar{b}\right)\left\langle a^2(t)\right\rangle - k_2\bar{a}^2\left\langle b^2(t)\right\rangle - \left(2k_1\bar{a} + 2k_2\bar{a}\bar{b} - k_2\bar{a}^2 + k_5\Lambda'(\bar{b})\right)\left\langle a(t)b(t)\right\rangle + \kappa\left\langle m(t)b(t)\right\rangle \\ &+ \left(k_1\bar{a} - 2k_1\bar{a}^2 - 2k_2\bar{a}^2\bar{b} - k_5\Lambda(\bar{b}) + k_5\bar{b}\Lambda'(\bar{b})\right)\left\langle a(t)\right\rangle + \bar{a}^2\left(k_1 + 2k_2\bar{b}\right)\left\langle b(t)\right\rangle,\end{aligned} \tag{A6}$$

where $\langle m(t)\rangle$,$\langle a(t)\rangle$ and $\langle b(t)\rangle$ are the dynamics of the first-order moments approximated by (21)–(23), and $\bar{x} = (\bar{m}, \bar{a}, \bar{b})^\top$ is a point in the state space. In particular, in order to obtain (A1)-(A6) we use the first-order approximation of the functions $\Psi(b)$ around $\bar{b}$, namely $\Psi(b) \approx \Psi(\bar{b}) + \Psi'(\bar{b})(b - \bar{b})$, of the functions a^2 and a^2b, around $\bar{a}$ and $\bar{b}$, namely $a^2 \approx \bar{a}^2 + 2\bar{a}(a - \bar{a})$ and $a^2b \approx \bar{a}^2\bar{b} + 2\bar{a}\bar{b}(a - \bar{a}) + \bar{a}^2(b - \bar{b})$, and of the function $\Lambda(b)$ around $\bar{b}$, namely $\Lambda(b) = \Lambda(\bar{b}) + \Lambda'(\bar{b})(b - \bar{b})$.

When $\bar{x} = (\bar{m}, \bar{a}, \bar{b})^\top$ satisfies (24)–(26), the approximated stationary points of the second order moments can be written through the linear algebraic matrix equation (39), where Π and Υ are given by

$$\Pi = \begin{bmatrix} 2\gamma & 0 & 0 & 0 & -2\tilde{\Psi} & 0 \\ 0 & -\kappa\bar{\delta} & 0 & \kappa & 0 & -\bar{k}_2 \\ 0 & 0 & \tilde{\Lambda} & 0 & 0 & \kappa\bar{\delta} \\ -\kappa & 0 & 0 & \gamma + \kappa\bar{\delta} & \bar{k}_2 & -\tilde{\Psi} \\ 0 & 0 & \tilde{\Psi} & \kappa\bar{\delta} & \tilde{\Lambda} - \gamma & 0 \\ 0 & \kappa\bar{\delta} & -\bar{k}_2 & 0 & \kappa & \tilde{\Lambda} - \kappa\bar{\delta} \end{bmatrix}, \qquad \Upsilon = \begin{bmatrix} \Gamma\Psi(\bar{b}) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{A7}$$

where we recall $\bar{\delta} = 2\frac{\bar{m}}{\bar{a}}$, $\bar{k}_2 = k_2\bar{a}^2$, $\tilde{\Lambda} = \bar{k}_2 - k_5\Lambda'(\bar{b})$ and $\tilde{\Psi} = \mu\Psi'(\bar{b})$.

Finally, the analytical expression of the solution to (39) can be written as

$$\chi_j = \pi_j^\star\Gamma\Psi(\bar{b}), \tag{A8}$$

where

$$\begin{aligned}\pi_1^\star = \frac{1}{\Delta}\big(&\tilde{\Lambda}^3\bar{\delta}\kappa + \tilde{\Lambda}^3\gamma - \tilde{\Lambda}^2\bar{\delta}^2\kappa^2 - 2\tilde{\Lambda}^2\bar{\delta}\gamma\kappa - 2\tilde{\Lambda}^2\bar{\delta}\kappa\bar{k}_2 - \tilde{\Lambda}^2\gamma^2 - \tilde{\Lambda}^2\gamma\bar{k}_2 + \tilde{\Psi}\tilde{\Lambda}^2\kappa + \tilde{\Lambda}\bar{\delta}^2\gamma\kappa^2 + 2\tilde{\Lambda}\bar{\delta}^2\kappa^2\bar{k}_2 + \tilde{\Lambda}\bar{\delta}\gamma^2\kappa \\ &+ 2\tilde{\Lambda}\bar{\delta}\gamma\kappa\bar{k}_2 - \tilde{\Psi}\tilde{\Lambda}\bar{\delta}\kappa^2 + \tilde{\Lambda}\bar{\delta}\kappa\bar{k}_2^2 + \tilde{\Lambda}\gamma^2\bar{k}_2 - \tilde{\Psi}\tilde{\Lambda}\gamma\kappa - \bar{\delta}^2\gamma\kappa^2\bar{k}_2 + \tilde{\Psi}\bar{\delta}^2\kappa^3 - \bar{\delta}^2\kappa^2\bar{k}_2^2 - \bar{\delta}\gamma^2\kappa\bar{k}_2 + \tilde{\Psi}\bar{\delta}\gamma\kappa^2 - \tilde{\Psi}\bar{\delta}\kappa^2\bar{k}_2\big),\end{aligned} \tag{A9}$$

$$\pi_2^\star = \frac{1}{\Delta}\big(k(\tilde{\Lambda}^3 - \bar{\delta}\tilde{\Lambda}^2\kappa - g\tilde{\Lambda}^2 + \bar{\delta}\gamma\tilde{\Lambda}\kappa + \tilde{\Psi}\bar{\delta}\kappa^2 - \bar{\delta}\gamma\bar{k}_2\kappa)\big), \tag{A10}$$

$$\pi_3^\star = \frac{1}{\Delta}\big(\bar{\delta}\kappa^3(\tilde{\Lambda} - \gamma - \bar{\delta}\kappa)\big), \tag{A11}$$

$$\pi_4^\star = \frac{1}{\Delta}\big(\kappa(\tilde{\Lambda}^3 - \tilde{\Lambda}^2\bar{k}_2 - \tilde{\Lambda}^2\gamma + \tilde{\Lambda}\gamma\bar{k}_2 - \tilde{\Lambda}^2\bar{\delta}\kappa + \tilde{\Psi}\bar{\delta}\kappa^2 - \bar{\delta}\gamma\kappa\bar{k}_2 + \tilde{\Lambda}\bar{\delta}\gamma\kappa + \tilde{\Lambda}\bar{\delta}\kappa\bar{k}_2)\big), \tag{A12}$$

$$\pi_5^\star = \frac{1}{\Delta}\big(\bar{\delta}\kappa^2(\tilde{\Lambda}\bar{k}_2 - \tilde{\Psi}\kappa - \tilde{\Lambda}^2 + \tilde{\Lambda}\bar{\delta}\kappa - \bar{\delta}\kappa\bar{k}_2)\big), \tag{A13}$$

$$\pi_6^\star = \frac{1}{\Delta}\big(\tilde{\Lambda}\kappa^2(\gamma - \tilde{\Lambda} + \bar{\delta}\kappa)\big), \tag{A14}$$

with

$$\Delta = 2(\tilde{\Lambda}\gamma + \tilde{\Psi}\kappa - \gamma\bar{k}_2)(\tilde{\Lambda}^2\bar{\delta}\kappa + \tilde{\Lambda}^2\gamma - \tilde{\Lambda}\bar{\delta}^2\kappa^2 - 2\tilde{\Lambda}\bar{\delta}\gamma\kappa - \bar{k}_2\tilde{\Lambda}\bar{\delta}\kappa - \tilde{\Lambda}\gamma^2 + \bar{\delta}^2\gamma\kappa^2 + \bar{k}_2\bar{\delta}^2\kappa^2 + \bar{\delta}\gamma^2\kappa + \tilde{\Psi}\bar{\delta}\kappa^2). \tag{A15}$$

References

1. Hordijk W, Steel M. Chasing the tail: The emergence of autocatalytic networks. *Biosystems* 2017; 152: 1–10.
2. Markovitch O, Lancet D. Excess Mutual Catalysis Is Required for Effective Evolvability. *Life* 2012; 18: 243–266.
3. Skorb EV, Semenov SN. Mathematical Analysis of a Prototypical Autocatalytic Reaction Network. *Life* 2019; 42: 1–10.
4. Kauffman S. *At home in the universe: The search for the laws of self-organization and complexity*. Oxford university press . 1996.
5. Mikhailov AS, Ertl G. The Belousov–Zhabotinsky reaction. In: Springer. 2017 (pp. 89–103).
6. Biria S, Malley PP, Kahan TF, Hosein ID. Optical Autocatalysis Establishes Novel Spatial Dynamics in Phase Separation of Polymer Blends during Photocuring. *ACS Macro Letters* 2016; 5(11): 1237–1241.
7. Fuson RC, Tullock CW. The Haloform Reaction. XIV. An improved iodoform test. *Journal of the American Chemical Society* 1934; 56(7): 1638–1640.
8. Adelstein PZ, Reilly JM, Nishimura DW, Erbland C. Stability of cellulose ester base photographic film: Part I—Laboratory testing procedures. *SMPTE journal* 1992; 101(5): 336–346.
9. Aguda B, Friedman A. *Models of cellular regulation*. Oxford University Press . 2008.
10. Alon U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Boca Raton, FL, USA: CRC Press . 2006.
11. Rosenfeld N, Elowitz MB, Alon U. Negative autoregulation speeds the response times of transcription networks. *Journal of molecular biology* 2002; 323(5): 785–793.
12. Lou SJ, Peacock-López E. Self-regulation in a minimal model of chemical self-replication. *Journal of biological physics* 2012; 38(2): 349–364.
13. Oyarzun DA, Lugagne JP, Stan GBV. Noise Propagation in Synthetic Gene Circuits for Metabolic Control. *ACS Synth. Biol.* 2015; 4: 116–125.
14. Levine E, Hwa T. Stochastic fluctuations in metabolic pathways. *PNAS* 2007; 104(22): 9224–9229.
15. Golding I, Paulsson J, Zawilski S, Cox E. Real-time kinetics of gene activity in individual bacteria. *Cell* 2005; 123: 1025–1036.
16. Shahrezaei V, Swain PS. Analytical distributions for stochastic gene expression. *Proc. Nat. Acad. Sci.* 2008; 105: 17256–17261.
17. Del Vecchio D, Sontag E. Synthetic biology: A systems engineering perspective. *Control Theory and Systems Biology* 2009: 101–124.
18. VanKampen NG. *Stochastic processes in physics and chemistry*. North Holland, 2007, 3rd ed. . 2007.
19. Borri A, Palumbo P, Singh A. Impact of negative feedback in metabolic noise propagation. *IET Syst. Biol.* 2016; 10(5): 179–186.
20. Hespanha JP, Singh A. Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems. *International Journal of Robust and Nonlinear Control* 2005; 15: 669–689.
21. Singh A, Hespanha JP. Approximate moment dynamics for chemically reacting systems. *IEEE Trans, Aut. Contr.* 2011; 56: 414–418.
22. Gillespie DT. Exact Stochastic Simulation of Coupled Chemical Reactions. *Journal of Physical Chemistry* 1977; 81(25): 2340–2361.

23. Rebek J. Synthetic self-replicating molecules. *Scientific American* 1994; 271(1): 48–55.

24. Peacock-López E. Chemical oscillations: The templator model. *The Chemical Educator* 2001; 6(4): 202–209.

25. Segel L. On the validity of the steady state assumption of enzyme kinetics. *Bull. Math. Biol.* 1988; 50: 579–593.

26. Shi Y, Cao Q, Wu J, Jia Y. Steady-state solutions of the templator model in chemical self-replication. *Journal of Mathematical Chemistry* 2021; 59: 1068–1097.

27. Gray P, Scott SK. *Chemical oscillations and instabilities: non-linear chemical kinetics* . 1990.

28. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature* 2000; 403(6767): 335–338.

29. Kaern M, Elston TC, Blake WJ, Collins J. Stochasticity in gene expression: from theories to phenotypes. *Nature reviews, Genetics* 2005; 6: 451–464.

30. Singh R, Hespanha JP. Optimal feedback strength for noise suppression in autoregulatory gene networks. *Biophysical Journal* 2009; 96: 4013–4023.

31. Bortolussi L, Policriti A. *Hybrid systems and biology: continuous and discrete modeling for systems biology. In Formal methods for computational systems biology*. Berlin, Germany: Springer . 2008.

32. Singh A, Hespanha JP. Stochastic hybrid systems for studying biochemical processes. *Phil. Trans. R. Soc. A* 2010; 368: 4995–5011.

33. Grima R. An effective rate equation approach to reaction kinetics in small volumes: Theory and application to biochemical reactions in nonequilibrium steady-state conditions. *The Journal of chemical physics* 2010; 133(3): 07B604.

34. Thomas P, Matuschek H, Grima R. Computation of biochemical pathway fluctuations beyond the linear noise approximation using iNA. In: IEEE. ; 2012: 1–5.

35. Gima R. Noise-induced breakdown of the Michaelis-Menten equation in steady-state conditions. *Physical Review Letters* 2009; 102(21): 218103.

36. Kubitschek HE, Friske JA. Determination of bacterial cell volume with the Coulter Counter. *J. Bacteriol.* 1986; 168: 1466–1467.

37. Guet CC, Bruneaux L, Min TL, et al. Minimally invasive determination of mRNA concentration in single living bacteria. *Nucleic Acids Res.* 2008; 36(12): 1–8.

38. Milo R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *Bioessays* 2013; 35(12): 1050–1055.

39. Perez-Carrasco R, Beentjes C, Grima R. Effects of cell cycle variability on lineage and population measurements of messenger RNA abundance. *Journal of the Royal Society Interface* 2020; 17(168): 20200360.