Noise propagation in metabolic pathways: the role of growth-mediated feedback

A. Borri P. Palumbo A. Singh

***Abstract*— Metabolic networks are known to deal with the chemical reactions responsible to fuel cellular activities with energy and carbon source and, as a matter of fact, to set the growth rate of the cell. To this end, feedback and regulatory networks play a crucial role to handle adaptation to external perturbations and internal noise. In this work, a cellular resource is assumed to be activated at the end of a metabolic pathway, by means of a cascade of transformations. Such a cascade is triggered by the catalytic action of enzymes that promote the transformations. The final product is responsible for the cellular growth rate modulation. This mechanism acts in feedback at the enzymatic level, since the upstream enzyme (as well as all species) is subject to clearance, with the clearance rate proportional to growth. The upstream enzymatic production is modeled by the occurrence of noisy bursts: a Stochastic Hybrid System is exploited to model the network and to investigate how such noise propagates on growth fluctuations. A major biological finding is that the delay introduced by the cascade length helps in reducing the impact of enzymatic noise on to growth fluctuations. Further, if feedback is removed from the scheme, growth rate fluctuations increase, according to the same stationary values. Analytical results are supported by Monte Carlo simulations.**

***Index Terms*— Metabolic pathways, Enzymatic Reactions, Systems Biology, Feedback**

I. Introduction

Fluctuations in growth rate are known to be responsible for phenotypic heterogeneity, although the mechanisms behind them are still matter of investigation [20], [22]. A large effort has been spent to investigate such phenotypic heterogeneity since it is supposed to be involved in cellular growth control and cancer initiation (see [11] and references therein). Within this framework, metabolism has recently gained interest since the intrinsic noise in gene expression and enzymes accumulation has been shown to propagate to growth rate fluctuations through metabolic fluxes, according to single-cell experiments [8], [12], [21], [24].

In this note a metabolic pathway is considered (see the scheme in Fig. 1), dealing with a cellular resource X_0 required to undergo a set of p functional modifications before to exert cellular growth control as X_p. Such modifications are triggered by the catalytic action of enzymes. It can be thought of as a generic pathway, not formally constrained to any specific metabolic behavior (like respiration or fermentation, possibly working in a parallel way). The activated resource X_p indirectly controls, in turn, the degradation of all molecular players, according to the following feedback scheme: X_p accumulation proportionally regulates the growth rate, and the growth rate controls many cellular activities, including the ones related to molecules degradation.

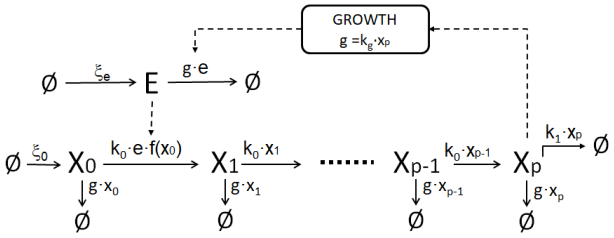

Fig. 1. Schematic figure of the metabolic pathway under investigation. Continuous lines refer to reactions; dashed lines refer to control actions.

Stochastic fluctuation in metabolic pathways has been matter of investigation in a wide range of literature, aiming at understanding the implications of gene expression stochasticity in cellular regulation and phenotypic diversity [14], [9], [3]. In [10] the analytical approach of a mathematical model exploiting Chemical Master Equations to deal with the cascade of chemical reactions showed that, for a linear metabolic pathway, fluctuations in different nodes of the pathway are uncorrelated and, consequently, fluctuations in enzyme levels have just local effects and do not propagate through the line of metabolic transformations.

The metabolic network here investigated differs from the framework adopted in [10], since we account for the feedback exerted by the final product that controls, by means of growth modulation, the clearance rate of all involved species. A similar interplay between enzymes, metabolites and growth rate has been investigated in [2], where the cellular metabolic resource X was not subject to a functional modification. Moreover, in [2] growth influenced only the enzyme clearance, whilst here all molecular players share the same clearance rate inherited from growth, according to a more biologically meaningful assumption.

A Stochastic Hybrid System (SHS) is exploited in order to model the network of biochemical reactions [7]. All molecular players are treated in terms of copy numbers. The upstream enzymatic production is supposed to occur in a

A. Borri is with the Istituto di Analisi dei Sistemi e Informatica "A. Ruberti", Italian National Research Council (IASI-CNR), Via dei Taurini 19, Roma, Italy. Email address: alessandro.borri@iasi.cnr.it

P. Palumbo is with Department of Biotechnologies and Biosciences, University of Milano-Bicocca, Milan, Italy 20126. E-Mail address: pasquale.palumbo@unimib.it, and with the Istituto di Analisi dei Sistemi e Informatica "A. Ruberti", Italian National Research Council (IASI-CNR), Via dei Taurini 19, Roma, Italy.

A. Singh is with Department of Electrical and Computer Engineering, Biomedical Engineering, Mathematical Sciences, Center for Bioinformatics and Computational Biology, University of Delaware, Newark, DE USA 19716. E-Mail address: absingh@udel.edu

discrete stochastic fashion: the enzyme copies are increased by a discrete amount each time a burst occurs. Within any two consecutive bursts of enzymatic production, all molecular players are supposed to vary continuously according to an Ordinary Differential Equation (ODE) system.

The goal of the manuscript is to investigate how noise propagates from the upstream enzyme to growth, by means of the metabolic transformations. To this end, the Coefficient of Variation is exploited in order to quantify the noise affecting molecular players fluctuations [10], [13], [1]; the analysis of the cross-correlation coefficient is also exploited to understand how the length of the pathway impact on noise propagation. In both cases, first- and second-order moments are required. Because of the ODE nonlinearities, finite-order moments equations are not provided in closed form [7], thus preventing any analytical computations except those involving moment closure techniques (e.g. the ones developed in [18]). To overcome such a problem, here we resort to linearizing the ODE nonlinearities around the steady-state solution that is proven to be the unique stationary equilibrium point.

Results show that the longer is the length of the cascade, the less the noise in enzyme production impacts on the final product fluctuations, hence proving that the delay provided by the cascade length helps in limiting growth fluctuations. Moreover, the cross-correlation analysis highlights (as expected) the occurrence of an apparent delay from the upstream enzyme to growth fluctuations, with such delay increasing with the cascade length as well. Finally, the important role in noise attenuation of the feedback is highlighted by the investigation carried out also for the no-feedback case, modeled by assuming degradation occurring at a fixed clearance rate, independent of the end product accumulation. Indeed, keeping unchanged the stationary values, the growth rate fluctuations markedly increase with respect to the case when the feedback is active.

II. Model setting

Consider the metabolic pathway depicted in Fig. 1. Copy numbers of resources and upstream enzyme will be accounted for, denoting them with $\mathbf{x_0}$, $\mathbf{x_1}, \ldots,$ $\mathbf{x_1}$ and $\mathbf{e}$. The enzymatic production is the only noise source: according to [6], [16], it is assumed to occur in bursts, with the burst size γ a random variable taking values in $\{1, 2, \ldots\}$. Then, whenever a burst occurs, $\mathbf{e}$ is updated as follows

$$\mathbf{e} \mapsto \mathbf{e} + j, \qquad j = 1, 2, \ldots \tag{1}$$

with a propensity $a_j = \xi_e \mathbb{P}(\gamma = j)$. Analogously to past literature, [19], [1], such probability is chosen to comply with a geometric distribution

$$\mathbb{P}(\gamma = j) = \alpha(1-\alpha)^{j-1}, \qquad \alpha \in [0, 1], \quad j = 1, 2, \ldots \tag{2}$$

providing an average burst size $\overline{\gamma} = 1/\alpha$. Motivation for the modeling choice of letting the only source of noise in the enzymatic production stems from [8], [24], where single cell experiments showed that noise propagates from metabolic enzymes to growth rate (and not vice versa).

Between any two bursts, the state variables evolve according to an ODE describing all species dynamics by properly setting the reaction rates of all chemical reactions.

The reaction rate of the first transformation $X_0 \mapsto X_1$:

$$k_0 \mathbf{e} f(\mathbf{x_0}), \qquad f(\mathbf{x_0}) = \frac{\mathbf{x_0}}{\theta + \mathbf{x_0}} \tag{3}$$

is controlled by the upstream enzyme in a linear fashion, whilst X_0 enters according to a Michaelis-Menten saturating function, concerning fast equilibration between substrate and enzyme binding/unbinding reaction [15], [10]. All other transformations in the cascade ($X_i \mapsto X_{i+1}, i = 1, \ldots, p-1$) are assumed to be catalyzed by enzymes at saturation (zero-order kinetics), so that the enzymatic action is formally hidden by a constant parameter in the reaction rate that linearly depends only on $\mathbf{x_i}$:

$$k_0 \, \mathbf{x_i}. \tag{4}$$

All species are subject to degradation, a cellular activity supposed to be controlled by the growth rate, which in turn is proportional to X_p copy number. The modeling choice is to set each clearance rate proportional to the growth rate, according to the same coefficient λ:

$$\lambda g \, \mathbf{e}, \quad \lambda g \, \mathbf{x_i}, \quad i = 0, \ldots, p, \qquad\qquad g = k_g \mathbf{x_p}. \tag{5}$$

Indeed, X_p is supposed to modulate the growth rate by properly influencing the cell metabolism. The growth rate closes the loop on the upstream enzyme dynamics, since it rules the degradation of E. Actually, the growth rate rules the degradation of all other molecular players. Besides, X_p is also exported out according to a linear clearance rate

$$k_1 \mathbf{x_p}. \tag{6}$$

In summary, the ODE is written as follows

$$\begin{cases} \dfrac{d\mathbf{e}}{dt} = -\lambda k_g \mathbf{e} \mathbf{x_p} \\ \dfrac{d\mathbf{x_0}}{dt} = \xi_0 - k_0 \mathbf{e} f(\mathbf{x_0}) - \lambda k_g \mathbf{x_0} \mathbf{x_p} \\ \dfrac{d\mathbf{x_1}}{dt} = k_0 \mathbf{e} f(\mathbf{x_0}) - \lambda k_g \mathbf{x_1} \mathbf{x_p} - k_0 \mathbf{x_1} \\ \dfrac{d\mathbf{x_i}}{dt} = k_0 \mathbf{x_{i-1}} - k_0 \mathbf{x_i} - \lambda k_g \mathbf{x_i} \mathbf{x_p}, \quad i = 2, \ldots, p-1 \\ \dfrac{d\mathbf{x_p}}{dt} = k_0 \mathbf{x_{p-1}} - \lambda k_g \mathbf{x_p}^2 - k_1 \mathbf{x_p} \end{cases} \tag{7}$$

and it will be shortly denoted by

$$\frac{d\mathbf{z}}{dt} = h(\mathbf{z}), \qquad \mathbf{z} = [\mathbf{e} \;\; \mathbf{x_0} \;\; \mathbf{x_1} \;\; \cdots \;\; \mathbf{x_p}]^T \in \mathbb{R}^{p+2}. \tag{8}$$

In the following, without loss of generality, we will fix $\lambda = 1$, formally treating the product λk_g as a unique parameter.

III. First-order moment equations

The expected value dynamics of any scalar transformation $\psi(\mathbf{z}) : \mathbb{R}^{p+2} \mapsto \mathbb{R}$ of the state of the SHS described by the

continuous ODE dynamics in (7)–(8) and endowed with the discrete resets (1) obeys to the following equation [7]:

$$\frac{d\langle\psi(\mathbf{z})\rangle}{dt} = \langle\frac{d\psi}{d\mathbf{z}}h(\mathbf{z})\rangle + \sum_{j=1}^{+\infty}\langle\big(\psi(\mathbf{z}+\Delta_j)-\psi(\mathbf{z})\big)a_j(\mathbf{z})\rangle \tag{9}$$

where Δ_j, $j = 1, 2, \ldots$ is the displacement on $\mathbf{z}$ because of reaction j in (1). This formula is exploited to compute the dynamics of any order moment. For instance, with respect to the first-order moments, we have:

$$\begin{aligned}
&\tfrac{d\langle\mathbf{e}\rangle}{dt} = \xi_e\overline{\gamma} - k_g\langle\mathbf{e}\mathbf{x_p}\rangle\\
&\tfrac{d\langle\mathbf{x_0}\rangle}{dt} = \xi_0 - k_0\langle\mathbf{e}f(\mathbf{x_0})\rangle - k_g\langle\mathbf{x_0}\mathbf{x_p}\rangle\\
&\tfrac{d\langle\mathbf{x_1}\rangle}{dt} = k_0\langle\mathbf{e}f(\mathbf{x_0})\rangle - k_g\langle\mathbf{x_1}\mathbf{x_p}\rangle - k_0\langle\mathbf{x_1}\rangle\\
&\tfrac{d\langle\mathbf{x_i}\rangle}{dt} = k_0\langle\mathbf{x_{i-1}}\rangle - k_0\langle\mathbf{x_i}\rangle - k_g\langle\mathbf{x_i}\mathbf{x_p}\rangle\\
&\qquad\qquad\qquad\qquad i = 2,\ldots,p-1\\
&\tfrac{d\langle\mathbf{x_p}\rangle}{dt} = k_0\langle\mathbf{x_{p-1}}\rangle - k_g\langle{\mathbf{x_p}}^2\rangle - k_1\langle\mathbf{x_p}\rangle
\end{aligned} \tag{10}$$

Because of the nonlinear terms, first-order moment equations cannot be written in a closed form [7], therefore we resort to the linear approximations of the ODE nonlinearities around the stationary point

$$\overline{\mathbf{z}} = (\overline{\mathbf{e}}, \overline{\mathbf{x}}_\mathbf{0}, \overline{\mathbf{x}}_\mathbf{1}, \ldots, \overline{\mathbf{x}}_\mathbf{p})^T, \qquad \overline{\mathbf{z}}_\mathbf{i} = \lim_{t\mapsto+\infty}\langle\mathbf{z_i}\rangle \tag{11}$$

provided by:

$$\begin{aligned}
&\mathbf{e}f(\mathbf{x_0}) \simeq \overline{\mathbf{e}}f(\overline{\mathbf{x}}_\mathbf{0}) + f(\overline{\mathbf{x}}_\mathbf{0})(\mathbf{e}-\overline{\mathbf{e}}) + \overline{\mathbf{e}}f'(\overline{\mathbf{x}}_\mathbf{0})(\mathbf{x_0}-\overline{\mathbf{x}}_\mathbf{0})\\
&\mathbf{e}\,\mathbf{x_p} \simeq \overline{\mathbf{e}}\;\overline{\mathbf{x}}_\mathbf{p} + \overline{\mathbf{x}}_\mathbf{p}(\mathbf{e}-\overline{\mathbf{e}}) + \overline{\mathbf{e}}(\mathbf{x_p}-\overline{\mathbf{x}}_\mathbf{p})\\
&\mathbf{x_i}\mathbf{x_p} \simeq \overline{\mathbf{x}}_\mathbf{i}\overline{\mathbf{x}}_\mathbf{p} + \overline{\mathbf{x}}_\mathbf{p}(\mathbf{x_i}-\overline{\mathbf{x}}_\mathbf{i}) + \overline{\mathbf{x}}_\mathbf{i}(\mathbf{x_p}-\overline{\mathbf{x}}_\mathbf{p})\\
&\qquad\qquad\qquad\qquad i = 0,1,\ldots,p
\end{aligned} \tag{12}$$

Thus, first-order moment equations are written, after computations, according to the following linear system

$$\frac{d}{dt}\langle\mathbf{z}\rangle = A(\overline{\mathbf{z}})\langle\mathbf{z}\rangle + b(\overline{\mathbf{z}}) \tag{13}$$

Matrix $A(\overline{\mathbf{z}})$ and vector $b(\overline{\mathbf{z}})$ are explicitly reported in Appendix.

By properly exploiting the fact that linearizations are achieved around the stationary point, this equilibrium has to comply with the following constraints

$$\begin{aligned}
&\xi_e\overline{\gamma} = k_g\overline{\mathbf{e}}\,\overline{\mathbf{x}}_\mathbf{p}\\
&\xi_0 = k_0\overline{\mathbf{e}}f(\overline{\mathbf{x}}_\mathbf{0}) + k_g\overline{\mathbf{x}}_\mathbf{0}\overline{\mathbf{x}}_\mathbf{p}\\
&k_0\overline{\mathbf{e}}f(\overline{\mathbf{x}}_\mathbf{0}) = k_g\overline{\mathbf{x}}_\mathbf{1}\overline{\mathbf{x}}_\mathbf{p} + k_0\overline{\mathbf{x}}_\mathbf{1}\\
&k_0\overline{\mathbf{x}}_\mathbf{i-1} = k_g\overline{\mathbf{x}}_\mathbf{i}\overline{\mathbf{x}}_\mathbf{p} + k_0\overline{\mathbf{x}}_\mathbf{i}, \qquad i = 2,\ldots,p-1\\
&k_0\overline{\mathbf{x}}_\mathbf{p-1} = k_g\overline{\mathbf{x}}_\mathbf{p}^2 + k_1\overline{\mathbf{x}}_\mathbf{p}
\end{aligned} \tag{14}$$

Theorem 1: There exists a unique, positive real solution for the nonlinear algebraic system (14), for any choice of positive model parameters.

Proof: As a preliminary step it will be proven that, for $i = 1,\ldots,p-1$:

$$\overline{\mathbf{x}}_\mathbf{i} = \frac{(k_0+k_g\overline{\mathbf{x}}_\mathbf{p})^{p-i-1}(k_1+k_g\overline{\mathbf{x}}_\mathbf{p})\overline{\mathbf{x}}_\mathbf{p}}{k_0^{p-i}} \tag{15}$$

The proof comes from mathematical induction. Indeed, (15) is true for $i = p-1$, since, by straightforward substitution, we readily obtain the last constraint in (14). Now, let (15) be true for a given i. Then from the penultimate constraint in (14), one gets:

$$k_0\overline{\mathbf{x}}_\mathbf{i-1} = (k_0+k_g\overline{\mathbf{x}}_\mathbf{p})\overline{\mathbf{x}}_\mathbf{i} = (k_0+k_g\overline{\mathbf{x}}_\mathbf{p})^{p-i}\cdot\frac{(k_1+k_g\overline{\mathbf{x}}_\mathbf{p})\overline{\mathbf{x}}_\mathbf{p}}{k_0^{p-i}} \tag{16}$$

from which it straightforwardly comes that (14) is true also for $i-1$.

By properly manipulating the second and third equations in (14) one has:

$$\xi_0 - k_g\overline{\mathbf{x}}_\mathbf{0}\overline{\mathbf{x}}_\mathbf{p} = k_g\overline{\mathbf{x}}_\mathbf{1}\overline{\mathbf{x}}_\mathbf{p} + k_0\overline{\mathbf{x}}_\mathbf{1} \tag{17}$$

from which, after substituting in (17) eq. (15) computed for $i = 1$, we have also $\overline{\mathbf{x}}_\mathbf{0}$ as a function of $\overline{\mathbf{x}}_\mathbf{p}$:

$$\overline{\mathbf{x}}_\mathbf{0} = \frac{\xi_0 k_0^{p-1} - (k_0+k_g\overline{\mathbf{x}}_\mathbf{p})^{p-1}(k_1+k_g\overline{\mathbf{x}}_\mathbf{p})\overline{\mathbf{x}}_\mathbf{p}}{k_g k_0^{p-1}\overline{\mathbf{x}}_\mathbf{p}} \tag{18}$$

This relationship will be referred to in the sequel as $\overline{\mathbf{x}}_\mathbf{0} = \varphi_1(\overline{\mathbf{x}}_\mathbf{p})$.

On the other hand, also $\overline{\mathbf{e}}$ can be easily written as a function of $\overline{\mathbf{x}}_\mathbf{p}$ from the first equation of (14):

$$\overline{\mathbf{e}} = \frac{\xi_e\overline{\gamma}}{k_g}\cdot\frac{1}{\overline{\mathbf{x}}_\mathbf{p}}. \tag{19}$$

Therefore, after substituting (19) and eq. (15) computed for $i = 1$ in the third constraint of (14), by suitably exploiting the saturating function (3), one has the following further relationship between $\overline{\mathbf{x}}_\mathbf{0}$ and $\overline{\mathbf{x}}_\mathbf{p}$:

$$\overline{\mathbf{x}}_\mathbf{0} = \frac{\theta k_g(k_0+k_g\overline{\mathbf{x}}_\mathbf{p})^{p-1}(k_1+k_g\overline{\mathbf{x}}_\mathbf{p}){\overline{\mathbf{x}}_\mathbf{p}}^2}{\xi_e\overline{\gamma}k_0^p - k_g(k_0+k_g\overline{\mathbf{x}}_\mathbf{p})^{p-1}(k_1+k_g\overline{\mathbf{x}}_\mathbf{p}){\overline{\mathbf{x}}_\mathbf{p}}^2}. \tag{20}$$

This relationship will be referred to in the sequel as $\overline{\mathbf{x}}_\mathbf{0} = \varphi_2(\overline{\mathbf{x}}_\mathbf{p})$.

In summary, the biologically meaningful stationary solutions are provided by the positive real intersections of the curves (18)-(20) in the $(\overline{\mathbf{x}}_\mathbf{0}, \overline{\mathbf{x}}_\mathbf{p})$ plane. Regards to $\varphi_1(\overline{\mathbf{x}}_\mathbf{p})$, it is monotonically decreasing in the positive orthant, with

$$\lim_{\overline{\mathbf{x}}_\mathbf{p}\mapsto 0^+}\varphi_1(\overline{\mathbf{x}}_\mathbf{p}) = +\infty, \qquad \lim_{\overline{\mathbf{x}}_\mathbf{p}\mapsto+\infty}\varphi_1(\overline{\mathbf{x}}_\mathbf{p}) = -\infty, \tag{21}$$

so that it has a unique positive value for which it vanishes. On the other hand, the second curve is monotonically increasing in the positive orthant, with

$$\varphi_2(0) = 0, \qquad \lim_{\overline{\mathbf{x}}_\mathbf{p}\mapsto+\infty}\varphi_2(\overline{\mathbf{x}}_\mathbf{p}) = -\theta, \tag{22}$$

with a vertical asymptote in the unique point that vanishes the denominator of $\varphi_2(\cdot)$. As a matter of fact, there exists a unique intersection of the two curves in the positive orthant. ■

IV. Second-order moment equations

According to what anticipated in the Introduction, we exploit the following definition of metabolic noise (e.g. for enzyme or growth fluctuations) provided by the squared Coefficient of Variation [10], [13], [1]:

$$CV_i^2 = \frac{\sigma_i^2}{\bar{i}^2}, \qquad i \in \{\mathbf{e}, g\}, \tag{23}$$

where σ_e, σ_g are the standard deviations of $\mathbf{e}$ and g.

Second-order moments dynamics is written according to (9) [7]:

$$\begin{aligned}\frac{d}{dt}\langle \mathbf{z_i z_k}\rangle &= \langle \mathbf{z_k} h_i(\mathbf{z})\rangle + \langle \mathbf{z_i} h_k(\mathbf{z})\rangle \\ &+ \sum_{j=1}^{+\infty} \left\langle \big((\mathbf{z_i}+\Delta_{ij})(\mathbf{z_k}+\Delta_{kj}) - \mathbf{z_i z_k}\big) a_j(\mathbf{z})\right\rangle\end{aligned} \tag{24}$$

where Δ_{ij}, $i = 1, \ldots, p+2$, $j = 1, 2, \ldots$ is the displacement on $\mathbf{z_i}$ because of reaction j.

Similarly to what has been done for the first-order moments, to have closed equations, we exploit the linearized ODE to approximate $h_i(\mathbf{z})$ and $h_k(\mathbf{z})$ in (24), so that the stationary values of the second-order moments equations come out as the solutions of a linear system, whose formal expression is achieved according to straightforward, though cumbersome computations, and is not here reported. The order of the system, by properly exploiting the simmetries, is:

$$\sum_{i=1}^{p+2} i = \frac{(p+3)(p+2)}{2}. \tag{25}$$

In order to evaluate how noise fluctuations in the upstream enzyme impact in growth, we compute the stationary cross-correlation

$$\rho_{eg} = \frac{\overline{\langle \mathbf{e}(t) g(t+\tau)\rangle} - \overline{\mathbf{e}}\,\overline{\mathbf{g}}}{\sigma_e \sigma_g} \tag{26}$$

where $\tau \in \mathbb{R}$ refers to the noise propagation delay and the overbars denote the stationary average values so that, for example:

$$\overline{\langle \mathbf{e}(t) g(t+\tau)\rangle} = \lim_{t \mapsto +\infty} \langle \mathbf{e}(t) g(t+\tau)\rangle . \tag{27}$$

Because of the linear relationship between g and $\mathbf{x_p}$, (5), one has:

$$\rho_{eg} = \frac{\overline{\langle \mathbf{e}(t)\mathbf{x_p}(t+\tau)\rangle} - \overline{\mathbf{e}}\,\overline{\mathbf{x}}_\mathbf{p}}{\sigma_e \sigma_p} = \rho_{ex_p} \tag{28}$$

with σ_p being the standard deviation of $\mathbf{x_p}$. Correlation analysis is carried out according to the same approach developed in [17], [2] so that, for $\tau \geq 0$ we take advantage of

$$\langle \mathbf{e}(t)\mathbf{x_p}(t+\tau)\rangle = \langle \mathbf{e}(t)\, \langle \mathbf{x_p}(t+\tau)|\mathbf{z}(t)\rangle\rangle \tag{29}$$

Then, by exploiting the explicit solution of the linear equation in (13), we have, provided that A^{-1} exists,

$$\langle \mathbf{x_p}(t+\tau)|\mathbf{z}(t)\rangle = C_{xp} e^{A\tau}\mathbf{z}(t) + C_{xp} A^{-1}(e^{A\tau} - I_{p+2}) b \tag{30}$$

where $C_{xp} = [O_{1\times(p+1)} \;\; 1]$, I_{p+2} is the identity matrix in $\mathbb{R}^{(p+2)\times(p+2)}$. By substituting (30) in (29), we have

$$\begin{aligned}\langle \mathbf{e}(t)\mathbf{x_p}(t+\tau)\rangle &= C_{xp} e^{A\tau}\, \langle \mathbf{e}(t)\mathbf{z}(t)\rangle \\ &+ \langle \mathbf{e}(t)\rangle\, C_{xp} A^{-1}(e^{A\tau} - I_{p+2}) b\end{aligned} \tag{31}$$

and, at steady-state

$$\overline{\langle \mathbf{e}(t)\mathbf{x_p}(t+\tau)\rangle} = C_{xp} e^{A\tau}\overline{\langle \mathbf{ez}\rangle} + \overline{\mathbf{e}} C_{xp} A^{-1}(e^{A\tau} - I_{p+2}) b \tag{32}$$

where $\overline{\langle \mathbf{ez}\rangle} = \left[\overline{\langle \mathbf{e}^2\rangle} \;\; \overline{\langle \mathbf{e x_0}\rangle} \;\; \cdots \;\; \overline{\langle \mathbf{e x_p}\rangle}\right]^T$ are part of the stationary second-order moments dealt with at the beginning of the Section.

In case of negative τ:

$$\langle \mathbf{e}(t)\mathbf{x_p}(t+\tau)\rangle = \langle \mathbf{e}(t)\mathbf{x_p}(t-|\tau|)\rangle \tag{33}$$

so that, at steady-state, it is

$$\overline{\langle \mathbf{e}(t)\mathbf{x_p}(t+\tau)\rangle} = \overline{\langle \mathbf{x_p}(t)\mathbf{e}(t+|\tau|)\rangle} \tag{34}$$

The computation of $\langle \mathbf{x_p}(t)\mathbf{e}(t+|\tau|)\rangle$ follows the same steps developed before:

$$\langle \mathbf{x_p}(t)\mathbf{e}(t+|\tau|)\rangle = \langle \mathbf{x_p}(t)\, \langle \mathbf{e}(t+|\tau|)|\mathbf{z}(t)\rangle\rangle \tag{35}$$

with

$$\langle \mathbf{e}(t+|\tau|)|\mathbf{z}(t)\rangle = C_e e^{A|\tau|}\mathbf{z}(t) + C_e A^{-1}(e^{A|\tau|} - I_{p+2}) b \tag{36}$$

where $C_e = [1 \;\; O_{1\times(p+1)}]$ so that

$$\begin{aligned}\langle \mathbf{x_p}(t)\mathbf{e}(t+|\tau|)\rangle &= C_e e^{A|\tau|}\, \langle \mathbf{x_p}(t)\mathbf{z}(t)\rangle \\ &+ \langle \mathbf{x_p}(t)\rangle\, C_e A^{-1}(e^{A|\tau|} - I_{p+2}) b\end{aligned} \tag{37}$$

and, at steady-state

$$\overline{\langle \mathbf{x_p}(t)\mathbf{e}(t+\tau)\rangle} = C_e e^{A|\tau|}\overline{\langle \mathbf{x_p z}\rangle} + \overline{\mathbf{x}}_\mathbf{p} C_e A^{-1}(e^{A|\tau|} - I_{p+2}) b \tag{38}$$

where $\overline{\langle \mathbf{x_p z}\rangle} = \left[\overline{\langle \mathbf{e}\,\mathbf{x_p}\rangle} \;\; \overline{\langle \mathbf{x_0 x_p}\rangle} \;\; \cdots \;\; \overline{\langle \mathbf{x_p}^2\rangle}\right]^T$ are part of the stationary second-order moments dealt with at the beginning of the Section.

V. Simulation results and discussion

Analytical solutions concerning both first- and second-order moments are achieved according to linear approximations of the nonlinear terms of the ODE. Therefore, results here reported are validated by Monte Carlo stochastic simulations for the SHS. These simulations are performed by means of the tau-leaping algorithm [5], by also exploiting the ergodic properties of the underlying stochastic process [23]. In this setting, the step selection has been chosen equal to 0.001 seconds within an overall simulation time of 5000 seconds. Part of the model parameters (Table I) have been taken or derived from [21] (e.g. k_1, θ); others have been tuned in order to obtain biologically meaningful outcomes, such as, for $p = 10$, a Mass Duplication Time (MDT) of about 46min (coherent with the MDT of *E. coli*), and species copy numbers complying with a realistic scale separation ($\simeq 100$ copies for enzyme and $\simeq 1000$ copies for the total metabolic resources), [21].

TABLE I

MODEL PARAMETERS (♯ DENOTES THE COPY NUMBERS)

$\xi_e = 60h^{-1}$	$\xi_0 = 4000\sharp h^{-1}$	$k_0 = 70h^{-1}$	$\alpha = 0.5$
$k_g = 0.02\sharp^{-1}h^{-1}$	$\theta = 1000\sharp$	$k_1 = 70h^{-1}$	

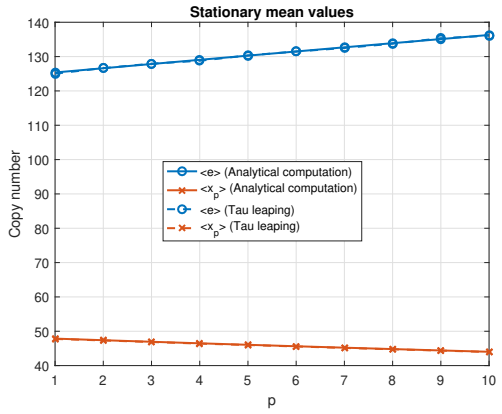

Fig. 2. Stationary mean values for variable cascade length (up to $p = 10$).

A first set of simulations aims at showing how the cascade length impacts on noise propagation. To this end, all model parameters have been fixed as the ones reported in Table I. Both analytical results and Monte Carlo simulations show that by increasing p, the steady-state average values of enzyme and final product copy numbers have opposite effects: stationary enzyme values increase, whilst stationary final product values decrease, see Fig. 2. The wiring of the network helps in understanding this behavior: by increasing the length of the cascade, nontrivial amount of substrate is stuck on intermediate values, thus reducing the final product accumulation; on the other hand, a reduced final product accumulation weakens the degradation of all molecular players, including the enzyme, thus allowing a greater enzyme accumulation. Fig. 3 reports how the noise associated to growth fluctuations varies for varying values of p, and we find out that such CVs decrease with p. That means, the delay provided by the cascade length helps in limiting growth fluctuations.

When plotting the cross-correlation $\rho_{eg}(\tau)$ (see Fig. 4), we find out that there exists an apparent delay showing that current enzyme expression correlates better with growth at some time later. This fact is coherent with experimental results reported in [24], thus explaining that growth fluctuations occur because of the noise in the enzyme expression. The present model allows to investigate also how the length of the cascade impacts on such delay. Fig. 4 shows that the delay increases with p.

A second set of simulations has been carried out for a fixed value of the cascade length ($p = 10$), by properly varying the enzymatic noisy production. Indeed, by varying the parameter α of the enzyme burst geometric distribution in (2) (with respect to the nominal value in Table I), it is possible to tune the variance of the incoming noise for the

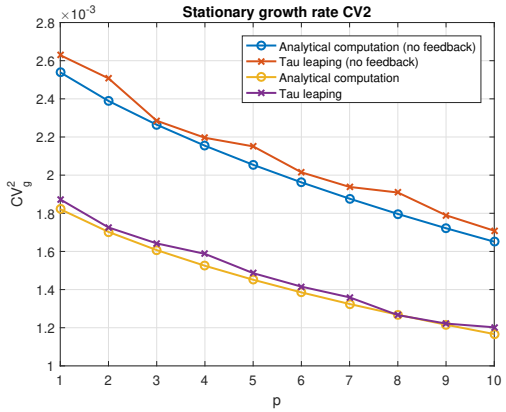

Fig. 3. Stationary values for the CV^2 of the growth rate, for variable cascade length (up to $p = 10$). The no-feedback case is also reported.

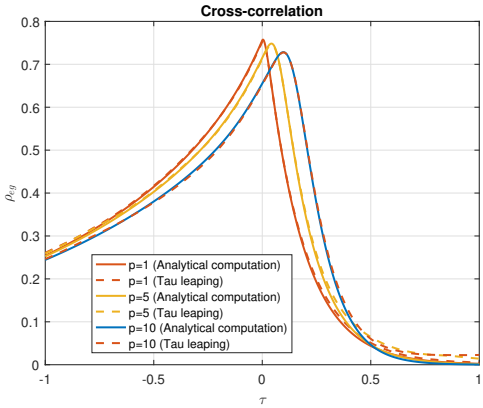

Fig. 4. Cross-correlation ρ_{eg} for different cascade lengths.

enzyme. It is worth to notice that, by varying $\alpha \in [0, 1]$, one obtains a wide range of burst size fluctuations and, as a matter of fact, different levels of enzymatic noise.

In order to make a fair comparison, we properly tuned the model parameters to keep fixed stationary average values in spite of different burst sizes. Looking at the constraints in (14), it comes out that α affects only the first equation (by means of $\overline{\gamma}$), therefore we have properly modified ξ_e to have the product $\xi_e\overline{\gamma}$ fixed for any value of α. More in details:

$$\xi_e\overline{\gamma} = \xi_e \cdot \frac{1}{\alpha} = \xi_e^{\text{ref}} \cdot \frac{1}{\alpha^{\text{ref}}} = \frac{50}{0.5} = 100, \qquad (39)$$

so that $\xi_e = 100\alpha$, with reference values for ξ_e^{ref} and α^{ref} from Table I. Fig. 5 reports how growth fluctuations levels vary according to enzyme noise levels (in both cases quantified by their CV^2). It is apparent that, by increasing the stochastic enzymatic fluctuations, also the growth noise increases.

Finally, the system has been investigated with respect to the no-feedback case, by assuming a fixed clearance rate for all species, namely, $\eta = \lambda k_g \tilde{\mathbf{x}}_{\mathbf{p}}$, where $\tilde{\mathbf{x}}_{\mathbf{p}}$ is the stationary average value of $\mathbf{x}_{\mathbf{p}}$ when the feedback is active. This way we obtain the same stationary values achieved for the

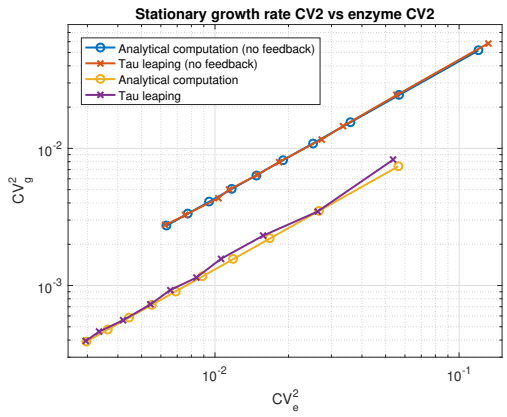

Fig. 5. Stationary growth rate CV_g^2 vs enzyme CV_e^2. The no-feedback case is also reported.

feedback case. Results are reported in Figs. 3 and 5, and highlight the role of noise attenuation for the feedback, since the CV_g^2 increases for the no-feedback case. Instead, the delay highlighted by the cross-correlation functions does not vary with respect to the no-feedback case, figure not shown.

VI. Conclusions

We have investigated how noise propagates in a metabolic pathway where a cellular resource undergoes a cascade of functional modifications, catalyzed by enzymes, before exerting cellular growth control. In a biological context, understanding the fluctuations in growth rate is of topic importance, since noise is known to be responsible for phenotypic heterogeneity, which is supposed to be involved, for example, in cancer initiation and growth. The analytical computations are based on a SHS model, and further validated via approximate stochastic simulations. Results show how the length of the cascade impacts in noise propagation, by introducing a delay from the noise source (upstream enzyme production in bursts) to the end product, and highlight the crucial role of noise attenuation played by the feedback.

Appendix

With respect to Eq. (13), the nontrivial entries of matrix $A(\overline{\mathbf{z}}) \in \mathbb{R}^{(p+2)\times(p+2)}$ are

$$
\begin{aligned}
&A_{1,1}(\overline{\mathbf{z}}) = -k_g \overline{\mathbf{z}}_{\mathbf{p+2}} \qquad A_{2,1}(\overline{\mathbf{z}}) = -k_0 f(\overline{\mathbf{z}}_{\mathbf{2}}) \\
&A_{2,2}(\overline{\mathbf{z}}) = -\left(k_g \overline{\mathbf{z}}_{\mathbf{p+2}} + k_0 \overline{\mathbf{z}}_{\mathbf{1}} f'(\overline{\mathbf{z}}_{\mathbf{2}})\right) \\
&A_{3,1}(\overline{\mathbf{z}}) = k_0 f(\overline{\mathbf{z}}_{\mathbf{2}}) \qquad A_{3,2}(\overline{\mathbf{z}}) = k_0 \overline{\mathbf{z}}_{\mathbf{1}} f'(\overline{\mathbf{z}}_{\mathbf{2}}) \\
&A_{j,p+2}(\overline{\mathbf{z}}) = -k_g \overline{\mathbf{z}}_{\mathbf{j}}, \qquad\qquad j = 1, \ldots, p+1 \\
&A_{p+2,p+2}(\overline{\mathbf{z}}) = -(2k_g \overline{\mathbf{z}}_{\mathbf{p+2}} + k_1) \\
&\text{IF} \quad p > 1 \\
&\quad A_{k,k-1}(\overline{\mathbf{z}}) = k_0, \qquad\qquad k = 4, \ldots, p+2 \\
&\quad A_{i,i}(\overline{\mathbf{z}}) = -(k_g \overline{\mathbf{z}}_{\mathbf{p+2}} + k_0), \qquad i = 3, \ldots, p+1
\end{aligned}
\tag{40}
$$

whilst, regards to $b(\overline{\mathbf{z}}) \in \mathbb{R}^{p+2}$, we have:

$$
\begin{aligned}
&b_1(\overline{\mathbf{z}}) = 2\xi_e \overline{\gamma} \\
&b_2(\overline{\mathbf{z}}) = \xi_0 + k_g \overline{\mathbf{z}}_{\mathbf{2}} \overline{\mathbf{z}}_{\mathbf{p+2}} + k_0 \overline{\mathbf{z}}_{\mathbf{1}} \overline{\mathbf{z}}_{\mathbf{2}} f'(\overline{\mathbf{z}}_{\mathbf{2}}) \\
&\text{IF} \quad p = 1 \\
&\quad b_3(\overline{\mathbf{z}}) = -k_0 \overline{\mathbf{z}}_{\mathbf{1}} \overline{\mathbf{z}}_{\mathbf{2}} f'(\overline{\mathbf{z}}_{\mathbf{2}}) + k_g \overline{\mathbf{z}}_{\mathbf{p+2}}^2 \\
&\text{ELSE} \\
&\quad b_3(\overline{\mathbf{z}}) = -k_0 \overline{\mathbf{z}}_{\mathbf{1}} \overline{\mathbf{z}}_{\mathbf{2}} f'(\overline{\mathbf{z}}_{\mathbf{2}}) + k_g \overline{\mathbf{z}}_{\mathbf{3}} \overline{\mathbf{z}}_{\mathbf{p+2}} \\
&\quad b_j(\overline{\mathbf{z}}) = k_g \overline{\mathbf{z}}_{\mathbf{j}} \overline{\mathbf{z}}_{\mathbf{p+2}}, \qquad j = 4, \ldots, p+2
\end{aligned}
\tag{41}
$$

References

[1] A. Borri, P. Palumbo, A. Singh, Impact of negative feedback in metabolic noise propagation, *IET Syst. Biol.*, 1-8, 2016.

[2] A. Borri, P. Palumbo, A. Singh, Noise propagation in feedback coupling between cell growth and metabolic activity, *Proceedings of the 57th IEEE Conference on Decision and Control (CDC 2018)*, Miami Beach, FL, USA, pp. 2679–2684, 2018.

[3] A. Borri, P. Palumbo, A. Singh, Time Delays in a Genetic Positive-Feedback Circuit, IEEE Control Systems Letters, 4(1), 163-168, 2020.

[4] D. T. Gillespie, Exact Stochastic Simulation of Coupled Chemical Reactions, *Journal of Physical Chemistry* 81(25), 23402361, 1977.

[5] D. T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems, *J. Chem. Phys.* 115(4): 1716, 2001.

[6] I. Golding, J. Paulsson, S. Zawilski, E. Cox, Real-time kinetics of gene activity in individual bacteria, *Cell* 123, 1025-1036, 2005.

[7] J.P. Hespanha, A. Singh, "Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems," *Int. J. of Robust and Nonlinear Control*, 15, 669–689. 2005

[8] D.J. Kiviet, P. Nghe, N. Walker, S. Boulineau, V. Sunderlikova, S.J. Tans, Stochasticity of metabolism and growth at the single-cell level. *Nature*, 514, 376–379, 2014.

[9] E. Kussell, S. Leibler, Phenotypic Diversity, population growth, and information in fluctuating environments, *Science* 309, 2075-78, 2005.

[10] E. Levine, T. Hwa, Stochastic fluctuations in metabolic pathways, *Proc. Natl. Acad. Sci.* 104(22), 9224-9229, 2007.

[11] A. Mahdipour-Shirayeh, K. Kaveh, M. Kohandel, S. Sivaloganathan, Phenotypic heterogeneity in modeling cancer evolution, *PLoS ONE* 12(10): e0187000, 2017.

[12] N. Nordholt, J. van Heerden, R. Kort, F.J. Bruggeman, Effects of growth rate and promoter activity on single-cell protein expression, *Scientific Reports*, 7:6299, DOI:10.1038/s41598-017-05871-3, 2018

[13] Oyarzun, D.A, Lugagne, J.-B., Stan, G.-B.V.: Noise propagation in synthetic gene circuits for metabolic control, ACS Synth. Biol., 2014.

[14] J.M. Raser, E.K. O'Shea, Noise in Gene Expression: Origins, Consequences, and Control, *Science* 309, 2010-2013, 2005.

[15] L.A. Segel, *Modeling Dynamic Phenomena in Molecular and Cellular Biology*, Cambridge Univ. Press, 1984

[16] V. Shahrezaei, P. S. Swain, Analytical distributions for stochastic gene expression, *Proc. Nat. Acad. Sci.*, 105, 17256–17261, 2008.

[17] A. Singh, B. Bokes, Consequences of mRNA Transport on Stochastic Variability in Protein Levels, *Biophys. J.* 103, 1087-1096, 2012.

[18] A. Singh, J.P. Hespanha, "Approximate moment dynamics for chemically reacting systems," *IEEE Transactions on Automatic Control.* 56, 414–418, 2011

[19] M. Soltani, C. Vargas, A. Singh, Conditional moment closure schemes for studying stochastic dynamics of genetic circuits. *IEEE Transactions on Biomedical Circuits and Systems*, 9(4), 518–526, 2015.

[20] P. Thomas, G. Terradot, V. Danos, A.Y. Weisse, Sources, propagation and consequences of stochasticity in cellular growth, *Nature Comm.* 9:4528, 2018. DOI: 10.1038/s41467-018-06912-9

[21] M.K. Tonn, P. Thomas, M. Barahona, D.A. Oyarzun, Stochastic modelling reveals mechanisms of metabolic heterogeneity, *Commun. Biol.* 2:108, 2009.

[22] S. Tsuru, J. Ichinose, A. Kashiwagi, B.-W. Ying, K. Kaneko, T. Yomo, Noisy cell growth rate leads to fluctuating protein concentration in bacteria, *Phys. Biol.* 6, 036015, 2009.

[23] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North Holland, third edition, 2007.

[24] M. Wehrens, F. Buke, P. Nghe, S.J. Tans, Stochasticity in cellular metabolism and growth: Approaches and consequences, *Current Opinion in Systems Biology*, 8, 131-136, 2018.