Stochastic oscillations induced by intrinsic fluctuations in a self-repressing gene

Jingkui Wang, Marc Lefranc, Quentin Thommen

Stochastic oscillations induced by intrinsic fluctuations in a self-repressing gene

Jingkui Wang
Laboratoire de Physique des Lasers, Atomes, et Molécules, CNRS, UMR8523
Université Lille 1, F-59655 Villeneuve d'Ascq, France

Marc Lefranc
Laboratoire de Physique des Lasers, Atomes, et Molécules, CNRS, UMR8523
Université Lille 1, F-59655 Villeneuve d'Ascq, France

Quentin Thommen[1]
Laboratoire de Physique des Lasers, Atomes, et Molécules, CNRS, UMR8523
Université Lille 1, F-59655

September 20, 2014

[1]Corresponding author. Address: quentin.thommen@univ-lille1.fr

Abstract

Biochemical reaction networks are subjected to large fluctuations attributable to small molecule numbers, yet underlie reliable biological functions. Thus, it is important to understand how regularity can emerge from noise. Here, we study the stochastic dynamics of a self-repressing gene with arbitrarily long or short response time. We find that when the mRNA and protein half-lives are approximately equal to the gene response time, fluctuations can induce relative regular oscillations in the protein concentration. To gain insight into this phenomenon at the crossroads of determinism and stochasticity, we use an intermediate theoretical approach, based on a moment-closure approximation of the master equation, which allows us to take into account the binary character of gene activity. We thereby obtain differential equations that describe how nonlinearity can feed back fluctuations into the mean-field equations to trigger oscillations. Finally, our results suggest that the self-repressing *Hes1* gene circuit exploits this phenomenon to generate robust oscillations, as its time constants satisfy precisely the conditions we have identified.

Key words: Stochastic gene expression; Self-repressing gene; Genetic oscillations; Master equation; Moment closure; *Hes1*

Introduction

Most cellular functions are controlled by molecular networks involving genes and proteins which regulate each other so as to generate the adequate dynamical behavior. A major goal of systems biology is to understand how sophisticated functional modules emerge from the combination of elementary processes such as transcriptional regulation, complex degradation, active transport,... and how each of these processes influences the collective dynamics (1).

A specificity of regulatory networks viewed as dynamical systems is that they are both strongly nonlinear and inherently stochastic, which considerably complicates the mathematical analysis. In a cell, protein and mRNA molecules are often found in low abundance so that variations of their copy numbers by one unit represent significant fluctuations. Furthermore, there are generally very few copies of a gene-carrying DNA fragment, with only a few possible configurations depending on promoter occupancy. When its transcription is regulated by a single protein, a gene can essentially be in two states: free, or bound to its transcription factor. Gene activity is then described mathematically by a binary variable, which more generally can also account for the transcriptional pulsing that has been observed both in prokaryotes (2) and eukaryotes (3–6). The stochastic dynamics of the gene randomly flipping between the bound and free states with probabilities depending on transcription factor abundance is a major source of intrinsic fluctuations, all the more at it was shown that this flipping can occur at time scales which are comparable to other biochemical processes (2). While stochasticity in gene networks has been often viewed as an undesirable perturbation blurring deterministic behavior, it is increasingly recognized that noise can in fact be harnessed to become a functional component of a regulatory network and make its dynamics richer (7–11). It is thus important to understand how the deterministic and stochastic aspects of cellular processes interact and contribute to the same global dynamics, all the more as they are intimately coupled in nonlinear systems.

However, even moderately complex regulatory networks resist mathematical analysis and require formidable computational resources. A natural strategy to study such general questions as the interplay of dynamics and noise is to focus on small genetic networks comprising only a few elementary components, whose analysis can identify the key mechanisms and parameters and cast light on the dynamics of more complex networks. This approach is all the more valuable as the recent developments of synthetic biology allow experimental tests of the theoretical analyses (12).

Here, we study how stochastic fluctuations in gene activity feed back into the deterministic dynamics of the smallest genetic network, which consists of a single gene repressed by its own protein product. This system is an ideal workbench to investigate how the dynamics of the network emerges from the properties of its elementary components. In fact, this motif is very common in transcriptional networks and is thus biologically relevant (around 40% of *Escherichia coli* transcription factors are self-repressing (13–15)). Self-repression is known to be an important ingredient for generating oscillatory behavior (16). For instance, Hirata *et al* proposed that the somite clock network is governed by the self-repressing gene *Hes1* (17).

Accordingly, the dynamics of the self-repressing gene has been actively investigated throughout mathematical biology (18–26).

Most theoretical analyses of the self-repressing gene based on a deterministic description assume that gene state flipping occurs on much faster time scales than other processes such as transcription, translation, and degradation. The flipping dynamics can then be taken into account through an average activity, which adapts to protein concentration either instantaneously or after a time delay. If intrinsic fluctuations are neglected, the analysis of the rate equations reveals that oscillatory behavior can only be found by either (1) introducing an explicit time delay in the equations (e.g., to take into account the transcriptional dynamics (16, 23–25, 27, 28)); (2) inducing an implicit time delay via a reactional step, which can be intrinsic (20) or describe transport between two compartments (22); (3) incorporating complex degradation mechanisms (16, 26, 28, 29). However, recent experiments have shown that gene activity may display an intrinsic dynamics on time scales comparable to that of other cellular processes (2, 4–6). This may be taken into account in a deterministic model by introducing an average gene activity variable, which reacts gradually to protein concentration (30). In particular, how such a transcriptional delay and a nonlinear degradation mechanism conspire to generate oscillations has been studied in detail by Morant *et al.* (26), who obtained analytical expressions for the instability thresholds.

To take into account the binary nature of gene state and its stochasticity, the most general approach to study the dynamics of the self-repressing gene is to use the chemical master equation (CME) (31). The steady-state solution of the CME provides the probability distribution of molecular copy numbers, characterizing both the averages and the fluctuations around them. An analytical solution of the CME for the self-repressing gene can be obtained when the mRNA variable is considered to be fast and can be eliminated adiabatically (32, 33), but this assumption is unrealistic for the *Hes1* feedback network, where mRNA and protein have similar lifetimes (17). A classical strategy for approximating the CME is the system-size expansion also known as Van Kampen's Ω-expansion (31). Assuming that the system size is large but not infinite, the solution is expanded in powers of the inverse system size. The deterministic mean-field equations are obtained at lowest order while next-to-leading order corrections determine finite-size fluctuations in the so-called linear noise approximation (LNA). This approach can be used to estimate the amplitude of fluctuations (34) but also to determine their spectrum. In particular, the LNA has been useful to characterize the appearance of stochastic oscillations in parameter regions where the mean-field equations predict stable steady behavior (35, 36) or to verify that oscillations predicted by a deterministic modeling persist in presence of fluctuations (37), two problems which have been actively studied (38–41). To overcome the fact that LNA does not allow one to determine precisely when the steady state loses stability, Scott, Ingalls and Hwa proposed an extension of this method which takes into account how fluctuations modify the linearized dynamics around steady state and allows one to study how bifurcation diagrams are modified by noise (42). However, all these methods based on system size expansion assume that fluctuations vanish in the infinite size limit, without affecting the average values. This assumption clearly does not hold

when the gene state is a binary variable which fluctuates between two discrete values, regardless of system size. A different approach must then be taken.

In this paper, we propose a strategy to describe the stochastic dynamics of a basic self-repressing gene circuit, with no cooperativity in the transcriptional regulation and a linear degradation mechanism (Fig. 1A). The gene switches stochastically between the active and inactive state, so that this circuit can be viewed as a random telegraph signal generator, whose output is sent through a low-pass filter before being fed back to itself (Fig. 1B). It is well known that a mean-field model of this system is unconditionally stable (see, e.g., (26)). Our main result is that a low-dimensional model taking fluctuations into account predicts oscillatory behavior in a region of parameter space where we observe relatively regular spiking in protein concentration.

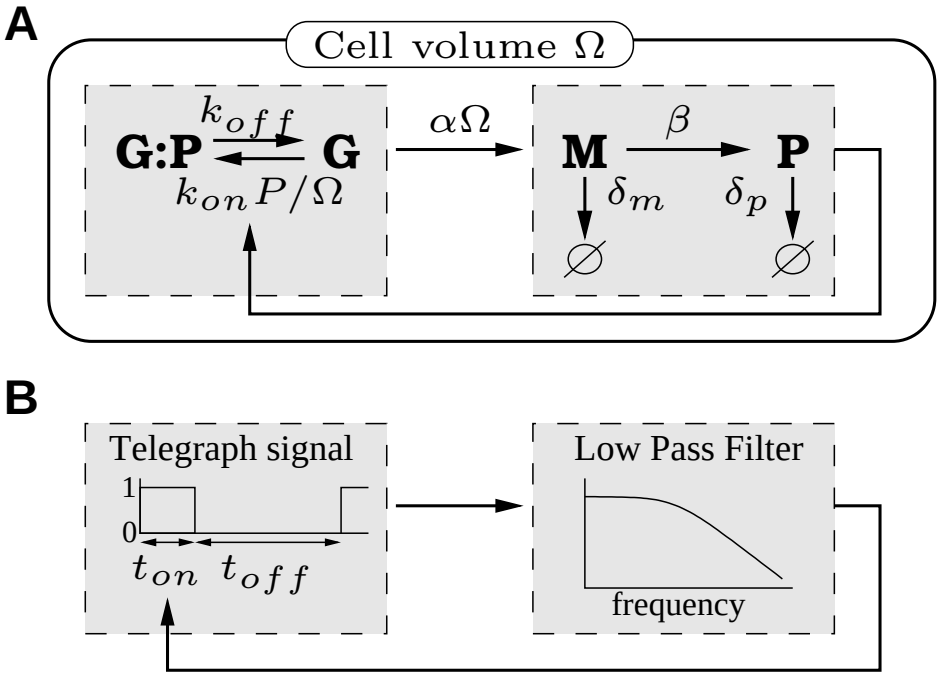

Figure 1: **Schematic view of the self-repressing gene network.** (A) Biochemical reactions composing the network. P, M, G, and $G:P$ denote protein, mRNA, free gene and bound gene chemical species, respectively. The kinetic constants of the reactions are indicated, with Ω denoting cell volume. In the limit where Ω is large, the mRNA and protein copy numbers become macroscopic variables, with decreasing fluctuations, while the gene state remains microscopic and displays full-scale variations. (B) Block diagram representation of the network, consisting of a random telegraph signal generator representing the gene state-flip dynamics, and of a low-pass filter of cut-off frequency ω_c representing proteins and mRNA dynamics. The telegraph signal regulates its frequency and duty cycle through feedback from the low pass filter.

To derive this model, we use a moment-closure approximation of the master equation (33, 43–45), and derive a set of ordinary differential equations which generalize the usual mean-field description while taking into account the binary nature of the gene state variable. These equations describe the combined time evolution of average quantities and of fluctuations around them. They reproduce accurately the stationary state values of the dynamical variables and predict the appearance of oscillations, without any assumption

on the gene switch time scale nor on the statistical distribution of random variables. We then explain the appearance of stochastic oscillations by a resonance effect between the characteristic times scales of the stochastic network and derive an analytical criterion for their appearance. Finally, the parameter values relevant for the Hes1 network suggests that the mechanism we describe may be exploited to generate robust oscillations in *Hes1* expression. Our findings highlight the functional role of intrinsic fluctuations arising from the gene state-flip dynamics as an important ingredient for shaping the dynamics of genetic networks.

Methods

To assess the validity of moment-closure approximations of the chemical master equations, we performed numerical stochastic simulations of the chemical network of Fig. 1-A for various values of the reduced parameters ρ, Λ, and δ (see Results). The stochastic simulations were performed using an implementation of the next reaction method (Gibson-Bruck algorithm (46)). The integration time used for numerical estimation of moments was chosen to ensure a relative error of the average gene activity estimator smaller than 10^{-4}, by monitoring the convergence of the estimator and its fluctuations. To estimate the Fano factor quantifying the regularity of protein spikes, we recorded 4000 interspike intervals, following a transient whose duration was chosen by monitoring the convergence of the estimator for the gene average activity.

To obtain a one-to-one correspondence between the original parameter space $\{k_{on}, k_{off}, \beta, \alpha, \delta_m, \delta_p\}$ and the reduced parameter space $\{\rho, \Lambda, \eta\}$, three constraints are required. Thus, we fixed (1) the ratio $\beta/\delta_p = 10$ to enforce a protein to mRNA concentration ratio of 10, which is a realistic assumption for a biological network, (2) the gene repression threshold $\Omega\, k_{off}/k_{on} = 100$ to keep computation time within reasonable limits while being consistent with the assumption of infinite cell volume, and (3) $\delta_m = 1$ to set the time scale to the mRNA half-life. Since stochastic simulations deal with copy numbers instead of concentration, the cell volume has no influence and we fixed $\Omega = 1$. The validity of the truncation schemes investigated can then be assessed by comparing the values of the averages in the stochastic simulation with the fixed point values of the ODE models obtained by truncating the moment expansion.

Results

Corrections to the rate equation

Three stochastic variables characterize the network dynamical state: the gene state g, the mRNA copy number m and the protein copy number p. The time evolution of the probabilities $P_{g,m,p}$ of being in a state

with given values of g, m and p is given by the following chemical master equation:

$$\frac{d}{dt}P_{g,m,p} = (-1)^g \left[\frac{k_{on}}{\Omega}\,(p+1-g)\,P_{1,m,p+1-g} - k_{off}P_{0,m,p-g}\right] \tag{1a}$$

$$+\delta_{g,1}\alpha\Omega\,[P_{g,m-1,p} - P_{g,m,p}] + \beta m\,[P_{g,m,p-1} - P_{g,m,p}] \tag{1b}$$

$$+\delta_m\,[(m+1)P_{g,m+1,p} - mP_{g,m,p}] + \delta_p\,[(p+1)\,P_{g,m,p+1} - p\,P_{g,m,p}] \tag{1c}$$

which can be read from Fig. 1A and provides the most general description of the dynamics. The parameters $k_{\rm on}$ and $k_{\rm off}$ characterize the kinetics of protein-DNA binding and unbinding, respectively. The transcription rate and translation rate are α/Ω and β, where Ω represents the cell volume, and δ_m and δ_p are the mRNA and protein degradation rates. The equations are normalized so that in the large volume limit, the average gene activity $\langle g\rangle$ and average concentrations $\langle m\rangle/\Omega$ and $\langle p\rangle/\Omega$ become independent of Ω.

Unfortunately, the master equation has generally no analytical solution. Contrary to the mRNA and protein copy numbers, which become much larger than one in the large volume limit and have then negligible fluctuations when a single molecule is created or destroyed, the gene state is a binary variable and its relative jump size does not decrease. Therefore, the standard approximation method based on the large-volume expansion of the master equation with the van Kampen ansatz fails (31). Alternatively, the chemical master equation can be reformulated as an infinite hierarchy of coupled differential equations whose variables are the moments of the random variables g, m, and p (31). This strategy leads to deterministic differential equations taking the fluctuations into account and having the mean-field rate equations as a limiting case.

To be specific, let us consider the equations describing the time evolution of the averages of gene activity and mRNA and protein concentrations in the infinite volume limit:

$$\frac{d}{dt}\langle\mathcal{P}\rangle = \beta\,\langle\mathcal{M}\rangle - \delta_p\,\langle\mathcal{P}\rangle; \tag{2a}$$

$$\frac{d}{dt}\langle\mathcal{M}\rangle = \alpha\,\langle g\rangle - \delta_m\,\langle\mathcal{M}\rangle; \tag{2b}$$

$$\frac{d}{dt}\langle g\rangle = k_{off}\,[1-\langle g\rangle] - k_{on}\,[\langle g\,\mathcal{P}\rangle]\,,$$

$$= k_{off}\,[1-\langle g\rangle] - k_{on}\,[\langle g\rangle\langle\mathcal{P}\rangle + \mathrm{cov}\,(g,\mathcal{P})]\,, \tag{2c}$$

where $\mathcal{M}$ (resp., $\mathcal{P}$) denotes the mRNA (resp., protein) concentration m/Ω (resp. p/Ω), $\langle x\rangle = \sum_{g,m,p} xP_{g,m,p}$ is the average of the stochastic variable x and $\mathrm{cov}\,(x,y) = \langle xy\rangle - \langle x\rangle\langle y\rangle$ is the covariance of x and y. These equations are derived following the approach described in the Supporting Material. Because of the nonlinear term associated with DNA-protein binding in Eq. (2c), this equation can only be reformulated in terms of the average values $\langle x\rangle$ by introducing the covariance term $\mathrm{cov}\,(g,P)$. This term does not appear in the usual rate equations describing the kinetics of the self-repressing gene. It describes the feedback from stochastic fluctuations into the dynamics of the average values and plays therefore a key role to model the influence of the gene state-flip dynamics. Eqs. (2) also indicate that the dynamics of mRNA and proteins behaves as

a low-pass filter whose input is the mean gene state $\langle g \rangle$ and output is the mean protein concentration $\langle P \rangle$. The cut-off frequency of this low-pass filter depends only on mRNA and protein degradation rates and is well approximated by $\omega_c = \frac{\delta_m \delta_p}{\delta_m + \delta_p}$ (see Supporting Material)

Eqs. (2) are only the first of an infinite hierarchy of equations where time derivatives of the first raw moments (the averages) are expressed in terms of the first and second raw moments, the time derivatives of second raw moments are expressed in terms of second and third raw moments, and so on (see Supporting Material). In order to truncate this infinite hierarchy to a finite set of equations, a closure approximation must be used. For instance, the usual rate equations are obtained when infinite cell volume and vanishing covariances are assumed (i.e., the $\mathrm{cov}\,(g, P)$ term in Eqs. (2) is set to 0). The approximation neglects all fluctuations and assumes that all variables have precise values, which conflicts with the binary nature of the gene state.

Here, we derive and analyze a higher-order model, by using a closure approximation of the moment expansion hierarchy in the limit of an infinite cell volume. In this limit, protein and mRNA copy numbers are also infinite and thus their variation by one unit is negligible, whereas the gene state is a binary variable, whose time evolution is similar to a random telegraph signal. Then the only remaining fluctuations in the models are those induced by the gene flipping dynamics.

The moment expansion equations up to order 2 are most conveniently expressed in terms of g, p and of a new variable $u = (\beta\, m + \delta_m\, p)/(\delta_p + \delta_m)$, after suitable rescaling (see Supporting Material for a detailed derivation). More precisely, the equations read

$$\frac{d}{dT}\langle P \rangle = \eta\,[\langle U \rangle - \langle P \rangle]\,; \tag{3a}$$

$$\frac{d}{dT}\langle U \rangle = \Lambda \langle G \rangle - \langle P \rangle; \tag{3b}$$

$$\frac{d}{dT}\langle G \rangle = \rho\,(1 - \langle G \rangle - \langle GP \rangle)\,; \tag{3c}$$

$$\frac{d}{dT}\langle GU \rangle = \Lambda \langle G \rangle - \langle GP \rangle - \rho\,[\langle GUP \rangle + \langle GU \rangle - \langle U \rangle]\,; \tag{3d}$$

$$\frac{d}{dT}\langle GP \rangle = \eta\,[\langle GU \rangle - \langle GP \rangle] - \rho\left[\langle GP^2 \rangle + \langle GP \rangle - \langle P \rangle\right]; \tag{3e}$$

$$\frac{d}{dT}\langle U^2 \rangle = 2\,[\Lambda \langle GU \rangle - \langle PU \rangle]\,; \tag{3f}$$

$$\frac{d}{dT}\langle P^2 \rangle = 2\eta\left[\langle PU \rangle - \langle P^2 \rangle\right]; \tag{3g}$$

$$\frac{d}{dT}\langle UP \rangle = \Lambda \langle GP \rangle - \langle P^2 \rangle + \eta\left[\langle U^2 \rangle - \langle PU \rangle\right]. \tag{3h}$$

where P, U, G are rescaled concentration of the random variables p, u and g, and T is a rescaled time. The three control parameters η, Λ and ρ are defined below. The key point is that Eqs. 3 are not closed, because Eqs. (3d) and (3e) depend on third-order moments $\langle GUP \rangle$ and $\langle GP^2 \rangle$, respectively, whose time evolution is unknown.

The dynamics is controlled by three key parameters

The biochemical reaction network of the self-repressing gene (Fig. 1-A) has six independent kinetic parameters. Three parameter combinations represent scales and thus can be taken out of the equations by rescaling time as well as mRNA and protein concentrations. These are : $\frac{k_{off}}{k_{on}}$, which is the protein concentration at which gene is half repressed; $\frac{\delta_p k_{off}}{\beta k_{on}}$ which is the mRNA concentration corresponding to half-repression in steady state; and $\frac{\delta_m+\delta_p}{\delta_m\delta_p}$, which is the response time of the low-pass filter. There remain three reduced parameters, denoted below by ρ, Λ, η, which control the dynamics and are discussed below.

The first reduced parameter $\Lambda = \frac{\alpha\beta k_{on}}{\delta_m\delta_p k_{off}}$ corresponds to the maximum possible protein concentration relative to the half-repression protein concentration threshold $\frac{k_{off}}{k_{on}}$. Dynamically, Λ characterizes the amplification of the gene telegraph signal sent to the low pass-filter. A low value of Λ ($\Lambda \ll 1$) indicates that the gene remains unbound most of the time; the average period of one gene on/off cycle is essentially the "on" state duration t_{on}. On the contrary, a high value of Λ ($\Lambda \gg 1$) indicates that the gene is repressed most of the time; the period of the gene on/off cycle is essentially the "off" state duration t_{off}, governed by k_{off}. Thus, Λ can also be viewed as characterizing the strength of the feedback from the gene to itself via its protein product.

The second parameter $\rho = k_{off}\frac{(\delta_m+\delta_p)}{\delta_p\delta_m}$ measures the gene unbinding rate relative to the cut-off frequency of the low-pass filter. A low value of ρ indicates that the low-pass filter transmits all the fluctuations of the gene state : the protein concentration time profile displays square waveforms enslaved to the gene flip. By contrast, a high value of ρ corresponds to the case where the low-pass filter averages out the gene flip dynamics : protein concentration evolves with small amplitude fluctuations around its mean value.

The third parameter $\eta = \frac{(\delta_m+\delta_p)^2}{\delta_p\delta_m}$ characterizes whether the protein and mRNA degradation rates are balanced or not. This indicator reaches a minimum value of 4 for equal degradation rates ($\delta_m = \delta_p$) and increases to infinity as one of the degradation rates becomes negligible compared to the other. It is worth noting that the expressions of all key parameters ρ, Λ, and η are symmetric with respect to exchange of δ_m and δ_p. As a consequence, the dynamical properties are unchanged if the mRNA and protein degradation rates are swapped, a fact which was already noted in (26). To distinguish the two regimes which have identical ρ, Λ and η parameter values but different values of δ_m and δ_p, we will later consider the ratio of protein and mRNA degradation rates $\delta = \delta_p/\delta_m$, with $\eta = (1+\delta)^2/\delta$. Obviously, the value of η is unchanged under the transformation $\delta \leftrightarrow 1/\delta$.

In the fast and slow gene limits, asympotic expressions of the averages are obtained from the fixed point of Eqs. (3), regardless of how they are closed. In particular $\langle U\rangle = \langle P\rangle = \Lambda\langle G\rangle$ in all cases. The value of $\langle G\rangle$ depends on the gene response time scale. In the fast gene regime ($\rho \to \infty$), it is determined by equating expression (3c) to zero, whose solution satisfies $\langle G\rangle \sim 1/\sqrt{\Lambda}$ in the limit of strong feedback (large Λ). In the slow gene limit ($\rho \to 0$), one has $\langle GP\rangle = \langle P\rangle$ (expressing the fact that protein concentration quickly relaxes to 0 when gene is off), so that equating expression (3c) to zero now leads to $\langle G\rangle \sim 1/\Lambda$ for strong feedback.

The dramatic decrease in gene average activity is related to the longer memory of the gene, which remains off for longer times after the repressor has disappeared.

Simple considerations also allow us to obtain the asymptotic behavior of the two third-order moments appearing in Eqs. (3). In the fast gene limit, the protein and mRNA can be considered as constant so that $\langle GUP\rangle = \langle G\rangle\langle U\rangle\langle P\rangle \sim \sqrt{\Lambda} \sim 1/\langle G\rangle$. In the slow gene limit, $\langle U\rangle$ and $\langle P\rangle$ quickly relax to their equilibrium value Λ when gene switches on, so that $\langle GUP\rangle = \langle UP\rangle_{G=1} \sim \Lambda^2\langle G\rangle \sim \Lambda$ and thus that $\langle GUP\rangle \sim 1/\langle G\rangle$ again. Similarly, one finds that $\langle GP^2\rangle \sim 1/\langle G\rangle$ in both limits.

Truncation of the moment equations

A natural closure approximation, which is described in detail in the Supporting Material, would to assume vanishing third-order central moments $\langle (G-\langle G\rangle)\,(U-\langle U\rangle)\,(P-\langle P\rangle)\rangle$ and $\langle (G-\langle G\rangle)\,(P-\langle P\rangle)^2\rangle$. The resulting model is eight-dimensional, incorporating three averages and five covariances as dynamical variables. Note that since not all third-order central moments are constrained to zero, this is a weaker requirement than assuming that variables are Gaussian-distributed. The predictions of this model are exact when the gene is either infinitely fast or slow.

Here, we focus on another closure approximation, which leads to a simpler yet accurate model. It assumes that the two unknown third-order moments are slaved to the gene state according to:

$$\langle GUP\rangle = \langle GP^2\rangle = \frac{(1-\langle G\rangle)^2}{\langle G\rangle}. \tag{4}$$

This relation is obtained by requiring that, in the limit of strong feedback (large Λ), (i) it matches the slow-gene and fast-gene asymptotic behaviors of the two moments, as obtained in previous section; (ii) the fixed point of the resulting equations agrees with that of the rate equations in the fast gene limit (see Supporting Material). Using stochastic numerical simulations, we also checked that it is relatively well satisfied for all intermediate gene response time scales, as will be shown in next section.

When rewritten using relation (4), Eqs. (3a-e) decouple from the others and form a closed system of only five differential equations, named thereafter the Truncated Moment Expansion (TME) model, which read:

$$\frac{d}{dT}\langle P\rangle = \eta\,[\langle U\rangle - \langle P\rangle]\,; \tag{5a}$$

$$\frac{d}{dT}\langle U\rangle = \Lambda\langle G\rangle - \langle P\rangle; \tag{5b}$$

$$\frac{d}{dT}\langle G\rangle = \rho\,(1-\langle G\rangle - \langle GP\rangle)\,; \tag{5c}$$

$$\frac{d}{dT}\langle GU\rangle = \Lambda\langle G\rangle - \langle GP\rangle - \rho\left[\frac{(1-\langle G\rangle)^2}{\langle G\rangle} + \langle GU\rangle - \langle U\rangle\right]; \tag{5d}$$

$$\frac{d}{dT}\langle GP\rangle = \eta\,[\langle GU\rangle - \langle GP\rangle] - \rho\left[\frac{(1-\langle G\rangle)^2}{\langle G\rangle} + \langle GP\rangle - \langle P\rangle\right]; \tag{5e}$$

This model predicts the time evolution of the three averages and of the two moments involving the gene state ($\langle GU\rangle$ and $\langle GP\rangle$). This is consistent with that the fact that only the gene state fluctuations survive in the infinite volume limit. The TME model thus provides a minimal extension of the rate equations, allowing us to describe the stochastic dynamics of the network, including the stationary state average values, and thus to study the impact of the gene state fluctuations.

The truncated moment expansion reproduces well the time averages of the stochastic dynamics

To assess the influence of stochastic fluctuations of gene state on the dynamics of the self-repressing gene, we first performed stochastic numerical simulations to determine the values of the time averages and covariances of the rescaled random variables G, M, and P (see Methods) as a function of the control parameters. These time averages were then compared to the fixed point values of two truncations of the moment equation hierarchy: the rate equation model, defined by Eqs. (2), with the covariance term set to zero, and the TME model defined by Eqs. (5). These models are defined by sets of ordinary differential equations, whose fixed points are specified by the values of the variables such that all time derivatives are zero. These fixed points are usually stable and thus reflect the stationary regime, however we shall see later that they may become unstable in some conditions, indicating the appearance of spontaneous oscillations.

Let us first examine how the average gene activity depends on ρ, which characterizes the gene response time scale, and Λ, which characterize feedback strength, when protein and mRNA lifetimes are identical ($\delta = 1$). Gene average activity as determined by stochastic simulations is shown in Fig. 2-A. The rate equation model correctly predicts the output of stochastic simulations only when gene dynamics is fast ($\rho \to \infty$) or when gene repression is small ($\Lambda \ll 1$) (Fig. 2-B). In contrast to this, the TME model predicts quantitatively gene average activity in the entire (ρ, Λ) plane (Fig. 2-C), in particular in regions where the rate equation approximation fails.

A more detailed assessment of the TME model accuracy is provided in Fig. 3, which shows how the time averages of G, U, P, and of their products evolve with ρ and δ, depending on whether they are computed from stochastic simulations (Fig. 3, left column) or from the fixed point values of the TME model (Fig. 3, right column). The computations are carried out in the strong feedback (i.e., high repression) limit ($\Lambda = 100$). Note that in the rate equation approximation, all averages would be constant and the covariances would vanish. Fig. 3 displays only a subset of components of the TME model fixed point, from which the other can be obtained using the relations $\langle M\rangle^* = \langle P\rangle^* = \Lambda\langle G\rangle^*$, $\langle GP\rangle^* = 1 - \langle G\rangle^*$, $\langle P^2\rangle^* = \langle PU\rangle^* = \Lambda\langle GU\rangle^*$. An important finding is that ρ is the main parameter controlling the averages of the stochastic variables G, U, and P, as the curves obtained for various values of δ superimpose remarkably well (Fig. 3-A, left column). As expected, using the variable U leads to numerical results which are symmetrical with respect to the $\delta \leftrightarrow 1/\delta$ inversion.

Fig. 3 shows that the fixed point values of the TME model are in very good quantitative agreement with

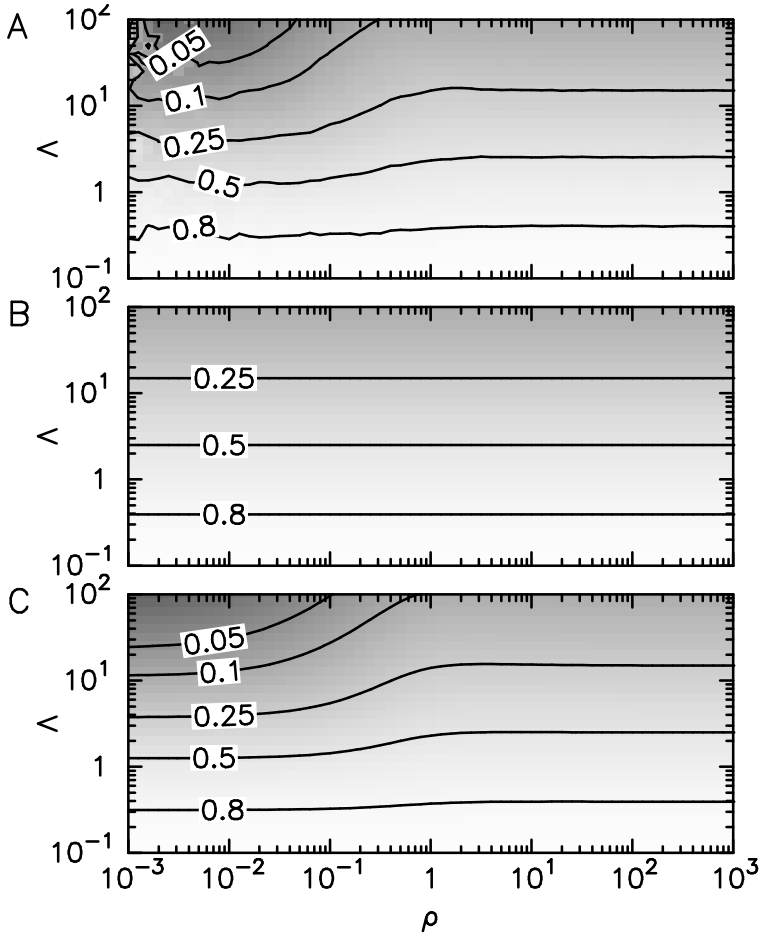

Figure 2: **Average gene activities as a function of the Λ and ρ parameter** (A) Numerical estimation of $\langle g \rangle$ using stochastic simulations with parameter values $k_{off}/k_{on} = 100$, $\delta = 1$, $\beta/\delta_p = 10$. (B) Average gene activity predicted by rate equation (C) Fixed point value of gene activity in the TME model.

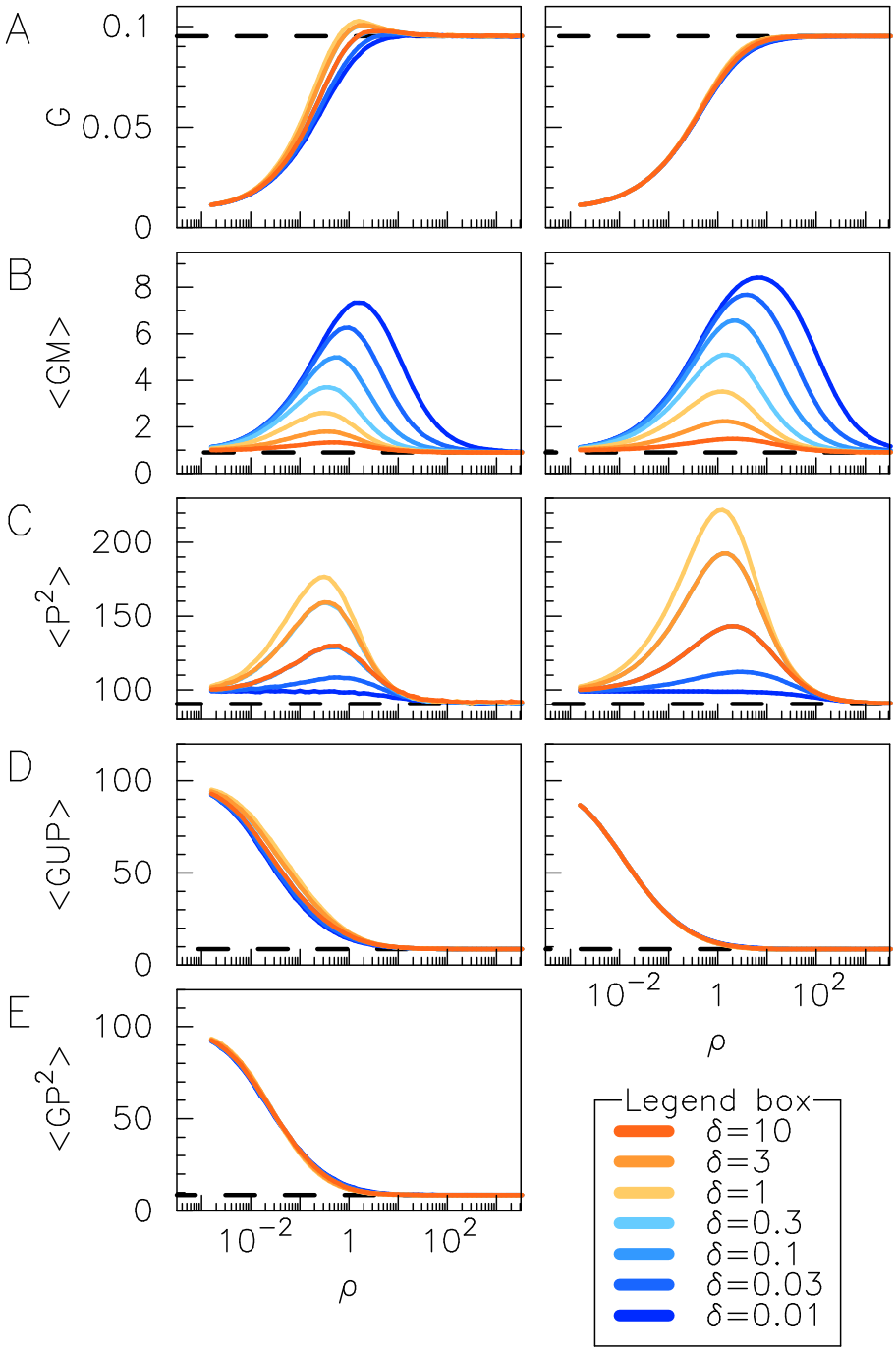

Figure 3: **Comparison of raw moments obtained from stochastic simulations (left column) and from the TME model (right column) as functions of parameter ρ, and for various values of δ.** (A) Average gene activity $\langle G \rangle$; (B), (C) second raw moments $\langle GM \rangle$ and $\langle P^2 \rangle$; (D), (E) third raw moments $\langle GUP \rangle$ and $\langle GP^2 \rangle$. Curves for different values of δ are color-coded according to legend box. The value of $\Lambda = 100$ used in the simulations corresponds to strong feedback (strong gene repression). Stochastic simulations are performed while constraining $k_{off}/k_{on} = 100$ and $\beta/\delta_p = 10$ (see Methods for details).

the numerical estimators. However, the stochastic estimators of the averages display an overshoot around $\rho = 1$ which is emphasized when η is small (Fig. 3-A). This behavior is not recovered by the TME model which predicts that averages increase monotonously with ρ. More precisely, the averages in the TME model are specified by

$$\langle G \rangle^* = \frac{2\kappa - 1 + \sqrt{1 + 4\Lambda\,(1-\kappa)}}{2\,(\Lambda + \kappa)}, \tag{6}$$

where $\kappa = \frac{\eta}{\rho(\eta+\rho)+\eta} \in [0,1]$. Presumably, this discrepancy could be resolved with a better closure approximation. The global evolution of the second raw moments is also well reproduced, even though the TME model overestimates the moments $\langle GM \rangle^*$ and $\langle P^2 \rangle^*$ (Fig. 3-B,C) for high values of ρ. The left column of Figs. 3-D,E display how $\langle GUP \rangle^*$ and $\langle GP^2 \rangle^*$ vary with ρ and δ, in good agreement with the closure assumption (4) (Fig. 3-D, right column). Although discrepancies are slightly more pronounced for $\langle GUP \rangle^*$, Figs. 3-D,E support the assumption that the two third-order raw moments have equal values. In comparison, the model obtained by assuming vanishing third-order central moments also predicts correctly stationnary values, captures more precisely the overshoot of the average near $\rho = 1$, but displays to a stiffer transition for the averages around $\rho = 1$ (see Supporting Material).

Dynamical considerations can explain the variation of averages with ρ observed. If $\rho \gg 1$, the gene flip dynamics is averaged by the low pass filter, and the stationary regime is correctly predicted by the fixed point values of the rate equations. In this limit, the gene remains bound or unbound for very short amounts of time, during which mRNA and proteins copy numbers can be considered as constant. RNA and protein levels keep a memory of many previous state switching cycles, and reach a stationary state with a probability distribution which is expected to be Gaussian. The coefficient of variation $CV = \frac{\sqrt{\langle P^2 \rangle - \langle P \rangle^2}}{\langle P \rangle}$ tends to zero as ρ increases, indicating that fluctuations in protein concentration become negligible compared to the average concentration in the limit of fast gene dynamics ($\rho \to \infty$).

Conversely, if $\rho \ll 1$, the gene reacts infinitely slowly. The dynamics is then driven by the gene jumping between two states according to a Poisson process. During the time where the gene is active (resp. inactive), protein and mRNA levels quickly converge to their maximum value Λ (resp. to zero); at the end of an gene switching state cycle, variables are always in the same state with no memory of previous cycles. Protein concentration temporal profiles feature a sequence of squared shape spikes, distributed in time according to a Poisson process, and characterized by a coefficient of variation $CV \approx \sqrt{\Lambda}$ increasing with the overall production rate Λ. Thus, fluctuations are enhanced by a slow gene and a high repression.

Then, a natural question is whether there exists between these two limit cases a dynamical regime which both behaves deterministically, as in the fast-gene limit, and displays strong variations of the protein concentration, as in the slow-gene limit. Such dynamical behavior would feature a sequence of protein concentration spikes, but with a time interval distribution more regular than a Poisson process. This intuition is based on the fact that when the gene flip frequency and the cut-off frequency of the low-pass filter are resonant ($\rho \approx 1$), the random fluctuations of gene flips generated by the Poisson process should

be partially buffered by the low-pass filter. This mechanism should prevent spike bunching, generating a more regular dynamical behavior that is the stochastic analogue of an oscillatory behavior (that we term thereafter stochastic oscillations). To assess the veracity of this idea, we developed a criterion to quantify the regularity of stochastic oscillations, described in next section.

Negative feedback induces protein spike antibunching

The regularity of a stochastic oscillatory behavior is often quantified using a temporal autocorrelation function (8, 37, 47). This measure is sensitive to reproducibility both in time and in amplitude. However, temporal regularity is certainly more relevant than amplitude regularity for biological protein signals. The highly nonlinear response of many signaling cascades can protect them against fluctuations in amplitude, for example by saturating output above an input threshold. A standard technique for assessing temporal regularity is to divide the state space into two regions I and II and to study the distribution of the times where the system leaves I to enter II. It is often useful to require a minimal excursion in region II to avoid spurious transitions induced by noise. Here, we detect events where the protein level crosses successively the mean protein level $\langle P \rangle^*$ and the $P'^* = \langle P \rangle^* + 0.25\,\mathrm{stdev}\,(P)$ level before falling back below the mean protein level, where stdev(P) denotes the standard deviation of P.

Given the list of times where the system transits from low to high protein levels, we compute the probability of detecting n transitions within a time interval of fixed duration. To be specific, we select a time interval equal to ten times the average time between two events, and characterize the probability distribution of the number of events by the variance to mean ratio, also known as the Fano factor (48). This method is inspired by how the temporal distribution of photons from a light source is generally characterized, with the event of interest being a photon detection. A Fano factor close to unity is obtained when time intervals between events follow a Poisson distribution. A Fano factor greater (less) than unity indicates super-Poissonian (resp., sub-Poissonian) behavior corresponding to a bunching (resp., anti-bunching) of protein spikes. Spike anti-bunching can be viewed as a stochastic counterpart of deterministic oscillations. While using the coefficient of variation of the interspike interval would give similar results, the method described above has the advantage to take into account correlations between the successive transitions.

Figure 4 displays stochastic simulations of the chemical reaction network of Fig. 1 for a slow, an intermediate and a fast gene, as well as the probability distribution of the number n of transitions within a given time window. As expected, protein spikes in the slow gene case (Fig. 4-A) are slaved to the switching process, leading to a Poisson probability distribution for n (Fig. 4-D) and accordingly a unit Fano factor. In the intermediate gene response time case (Fig. 4-B), protein spikes are mostly antibunched (see black circles). The probability distribution of spike number is Gaussian-like (Fig. 4-E), the Fano factor being around 0.25. This anti-bunching degrades in the case of a fast gene (Fig. 4-C) with the Fano factor rising to 0.9. Thus, we observe a resonance effect which results from the coincidence of the characteristic time of the gene response to protein variations with the time during which gene state history is remembered, which

is controlled by the protein and mRNA decay rates.

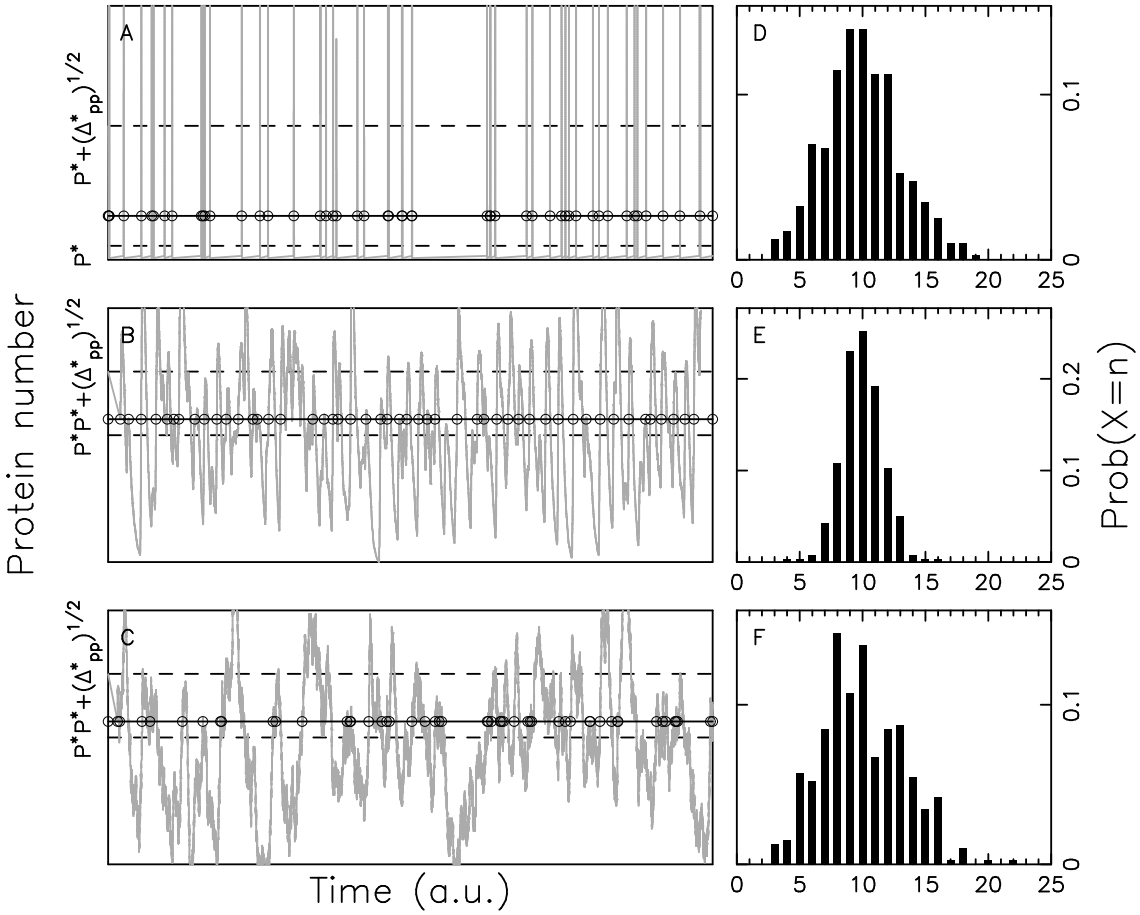

Figure 4: **Protein spike antibunching.** (A,B,C). Time evolution of protein copy number for $\Lambda = 100$, $\delta = 1$ and $\rho = 10^{-3},\ 1,\ 10^{3}$, respectively. Dashed lines indicate mean protein level and mean protein level plus standard deviation. Black lines correspond to the high trigger level and spiking events are indicated by black circles. (D, E, F) Probabilities of observing n spikes during a time window of 10 average transition times $\rho = 10^{-3},\ 1,\ 10^{3}$, respectively.

We studied systematically how the Fano Factor depends on the gene resonance parameter ρ and the relative protein decay rate δ in stochastic simulations (Fig. 5). We found that the regularity of protein spikes is reinforced by (1) similar mRNA and protein decay rates ($\delta \approx 1$), (2) a resonance parameter close to unity ($\rho \approx 1$), and (3) a sufficiently strong feedback ($\Lambda \gg 1$), as shown in Fig. 5-A,B. Thus, the most regular oscillations are observed when the gene cycling period resonates with the average mRNA/protein lifetime.

The lack of symmetry with respect to the transformation $\delta \leftrightarrow 1/\delta$ for low values of Λ (Fig. 5-B) results from numerical difficulties to reach the infinite cell volume limit for small δ_p ($\delta \ll 1$). As a control, we checked that the Fano Factor is almost independent of the ratio β/δ_p (see Fig. 5-C), which determines the protein to mRNA ratio.

In the large Λ limit, it is expected that the gene spends most of the time in the "off" state so that the

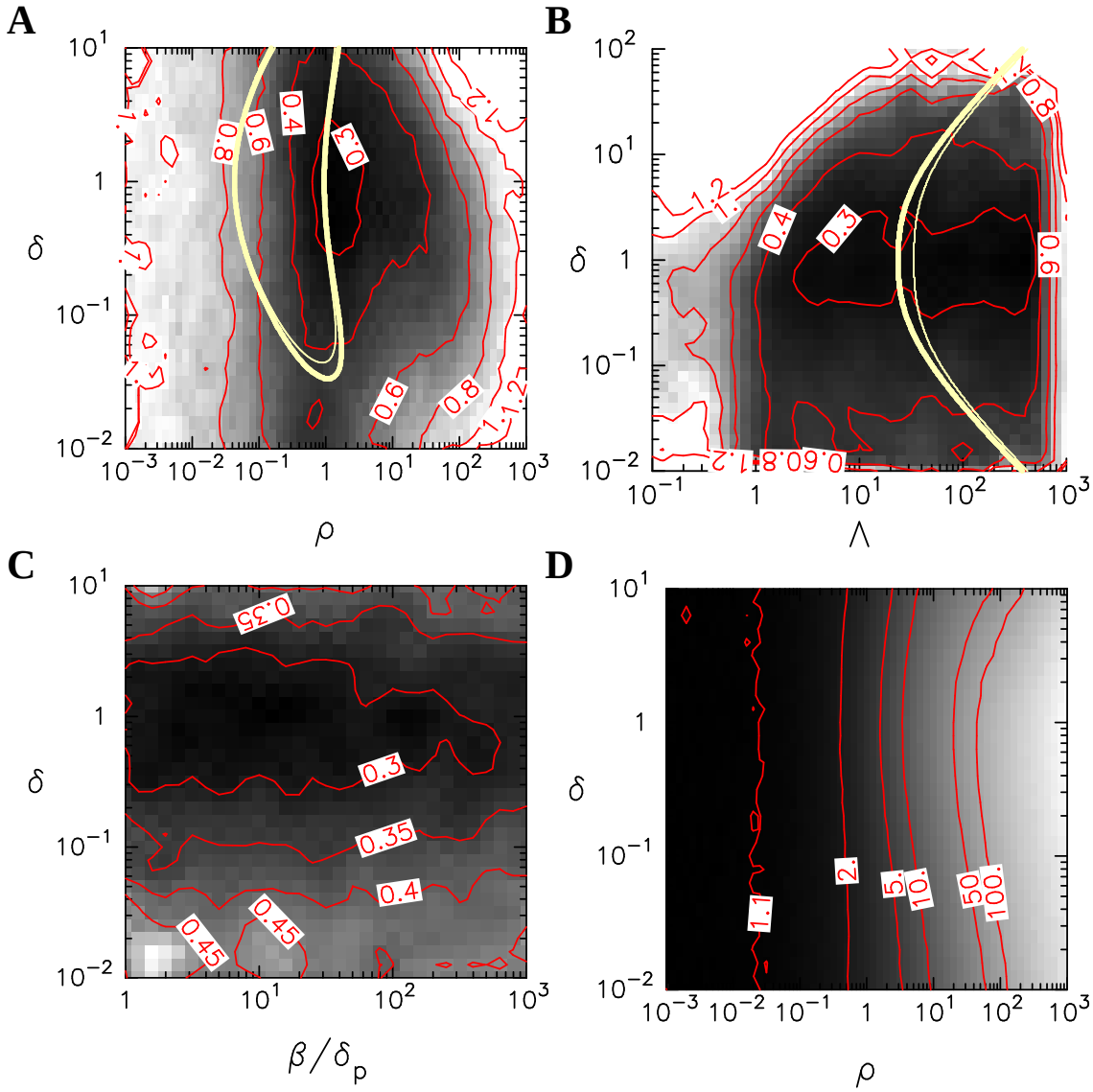

Figure 5: **Regularity of stochastic oscillations** In (A), (B), and (C), the value of the Fano factor F, which quantifies spiking regularity, is shown as a function of two parameters using a gray-scale color code with level lines displayed in red. The yellow (color online) lines in (A) and (B) enclose the region where the TME model predicts oscillations based on numerical analysis (thick lines) or analytical criterion Eq. (7) (thin lines). The regularity of stochastic oscillations is favored by balanced protein and mRNA degradation rates (corresponding to $\delta \simeq 1$) as well as (A) a resonance parameter ρ close to 1, (B) a high value of the overall production rate Λ. (C) The ratio β/δ controlling the relative mRNA to protein concentration has no effect on spike regularity. (D) The oscillation period (or protein average interspike time interval) is controlled by the resonance parameter ρ. The variation of χ, the average number of "on"/ "off" cycles in an interspike interval, is displayed using a gray color code as a function of the lifetime ratio δ and of the resonance parameter ρ. Level lines are displayed in red (color online). Stochastic simulations of the biochemical network have been carried out with $k_{off}/k_{on} = 100$; $\beta = 10\delta$ (A, B, and D); $\rho = 1$ (B and C); $\Lambda = 100$ (A, B, and C).

average duration of one "on"/"off" cycle is approximately given by $\tau_{\text{off}} = 1/k_{\text{off}}$ in original time units. To study the interplay between the gene state dynamics and the protein spike dynamics, a useful indicator is $\chi = k_{off}\langle T_s \rangle$ where T_s is the average time interval between two spikes. In the slow gene limit ($\rho \to 0$), χ tends to 1, indicating that protein and mRNA are slaved to the gene dynamics in a "fire and degrade" mode. Conversely, the high value of χ in the fast gene limit indicates that the gene dynamics is too fast to be relevant and justifies an adiabatic elimination of the gene state variable. In the parameter region where spikes are more regular, the intermediate values taken by χ (between 1 and 10) reveal that the gene dynamics and the mRNA/protein dynamics influence each other and generate together the stochastic oscillations observed.

The truncated moment expansion predicts the appearance of stochastic oscillations

If a moment-closure model such as the TME model (Eqs. (5)) is relevant to the dynamics of the self-repressing gene, it should be able to predict the stochastic oscillations evidenced in the previous section. While such models take fluctuations into account, they are deterministic ODE models, where the natural counterpart of the regular spiking observed in stochastic simulations would be the occurrence of self-sustained oscillations. A linear stability analysis of the TME model should then provide analytical insight into the key parameters controlling the stochastic oscillations.

Indeed, the TME model exhibits oscillatory behavior and the region in the parameter space where oscillation occurs (identified by numerical stability analysis) is consistent with the region of parameter space where regular stochastic oscillations are observed (Fig. 5-A,B), at least for small ρ (slow gene) or $\rho \approx 1$. In particular, the TME oscillation region contains the point ($\rho = 1$, $\delta = 1$) which is the organizing center of the stochastic oscillation region in parameter space. However, the TME model does not capture well the occurrence of regular stochastic oscillations when $\rho > 1$ (Fig. 5-A). The influence of feedback strength Λ on the appearance of oscillations depending on degradation rate balance is correctly captured (Fig. 5-B), although the TME model overestimates the value of Λ at which regular oscillations are first observed (Fig. 5-B).

It is interesting to note that while the mean-field model of the self-repressing gene is unconditionally stable, the simplest model incorporating the feedback from fluctuations, namely the TME model, predicts oscillations in good agreement with the observed regular stochastic oscillations. This strongly supports the idea that fluctuations play a functional role to promote oscillations. Thus, it is interesting to check whether an analytical criterion for the appearance of oscillations can be derived, so as to identify the role of each parameter in this dynamical behavior. In generally, oscillation criteria for systems with 4 or more variables are difficult to obtain. In the present case, however, a approximation in the stability analysis (see Supporting Material) leads to a relatively compact Routh–Hurwitz oscillation criterion (49), indicating the parameter space region in which oscillations originating from a Hopf bifurcation occur. In our case, the criterion states

that oscillations occurs in the TME model when the inequality

$$\mathcal{H}(\rho,\eta,\Lambda) = \rho^3\left(6+\frac{2}{\langle G\rangle^{*2}}\right) + 8\eta\rho^2 + \rho\eta\,(2\eta+2-\Lambda) + \eta^2 \;<0. \tag{7}$$

is satisfied, the average gene activity $\langle G\rangle^*$ being defined by Eq. (6). As Fig. 5-A,B show, the criterion (7) delimitates very accurately the parameter space regions where oscillatory behavior is observed in the TME model, as indicated by a numerical stability analysis. Therefore, it is tempting to use it to discuss the influence of the different parameters on the appearance of regular stochastic oscillations, at least for $\rho \leq 1$ where the TME model reflects the existence of these stochastic oscillations with oscillating averages and moments.

In particular, it can be seen that the only negative term in (7) is $-\rho\eta\Lambda$ and thus that a sufficiently large value of Λ is required for oscillations. On the other hand, η should not be too large, otherwise the η^2 term of expression (7) is dominant. Given that $\eta = (1+\delta)^2/\delta$, the lowest possible value for η is 4, which is reached when $\delta = 1$. This implies that the mRNA and protein degradation rates should not be too different, as is observed in the stochastic simulations. More precisely, a necessary condition for the occurrence of oscillatory behavior at $\rho = 1$ is $\Lambda \geq 3\eta + (13\eta+6)/(\eta-1)$, which shows that larger values of η require larger values of Λ (i.e., stronger feedback), as indicated by Fig. 5B.

Let us now consider how the value of ρ influences oscillatory behavior. When $\rho \to 0$ (i.e., the gene cycling period is much longer than average protein/mRNA lifetimes), the criterion is never satisfied because $\mathcal{H} = \eta^2 > 0$, and no oscillations occur. When $\rho \to \infty$, the dominant term is obviously positive, and no oscillations occur either. For intermediate values of ρ, the quantity $\mathcal{H}$ can become negative when Λ is sufficiently large and η is sufficiently close to its minimum value of 4, as discussed above.

The discussion is easier in the limit of large Λ, where the oscillation criterion simplifies considerably (see Sec. S.H in the Supporting Material). Two cases must be considered according to whether ρ is small or close to 1, because the gene average activity scales differently with Λ in these two cases. For $\rho = O(1)$, we find that the leading term of the exact Routh–Hurwitz criterion is

$$\mathcal{H} \approx -\frac{\rho\,\left(-2\,\eta\,\rho^2 - 2\,\rho^3 + \eta^2 - \rho\,\eta\right)}{(\eta+\rho)}\Lambda \tag{8}$$

so that oscillations occur when

$$2\,\rho^2\eta + 2\,\rho^3 - \eta^2 + \rho\,\eta < 0, \tag{9}$$

which describes well the rightmost boundary of the oscillation region in Fig. 5-A (see Fig. S3). The simplicity of the criterion allows us to discuss the relative influences of gene response time (characterized by ρ) and degradation rate balance (characterized by η) on the appearance of oscillations. We see that $\mathcal{H}$ changes

abruptly from large negative to large positive values at

$$\eta = \eta_c = (\rho + 1/2 + 1/2\,\sqrt{4\,\rho^2 + 12\,\rho + 1})\rho. \tag{10}$$

so that oscillations are lost for higher values of η. When ρ is small ($\rho = O(1/\Lambda)$), then the oscillation criterion simplifies to

$$\rho > \eta/\Lambda, \tag{11}$$

corresponding to the leftmost boundary of the oscillation region in Fig. 5-A (see Fig. S3).

Thus, oscillations are systematically found for $\rho \in [\eta/\Lambda, \rho_c]$, where ρ_c is the value of ρ statisfying (10) for a given value of η, with $\rho_c \approx 1.07$ for $\delta = 1$ ($\eta = 4$), and where $\mathcal{H}$ switches from negative infinity to positive infinity. This confirms unambigously the existence of a wide region of oscillation in parameter space. The singular behavior observed at ρ_c is presumably an artifact of the truncation scheme used, and is related to the fact that the choice of the moment-closure function affects the highest-order term in ρ in the criterion (7) (see Section S.H in Supplementary Information). It is plausible that the singularity would disappear at a higher truncation order, and that in this case, the oscillation region would extend further towards larger values of ρ, improving the agreement with stochastic simulations. However, the system would then perhaps be too complicated to obtain an analytical oscillation criterion.

Globally, the TME model captures well how the mean-field variables and fluctuations interact through nonlinearities to generate relatively regular stochastic oscillations, at least when the gene response time is not too small (i.e., when $\rho \leq 1$). The fact that an analytical criterion for its oscillation threshold, which becomes very simple in the limit of large Λ, can be obtained allows one to understand the role of the different parameters. In particular, it confirms that oscillations are favored when mRNA and protein degradation rates are close to each other (i.e., $\delta = 1$ and $\eta = 4$). It is also consistent with the fact that oscillations always occur in the neighborhood of $\rho = 1$.

Oscillations in Hes1 expression match the criterion for fluctuation-induced oscillations

The main result of this work is that stochastic fluctuations in a self-repressing gene can play a functional role in promoting the appearance of relatively regular oscillations in specific regions of the parameter space. It is then natural to ask whether oscillating self-repressing gene circuits found in biological systems operate in the parameter region we have identified. One such circuit that has been intensively studied is the *Hes1* gene, which is believed to be at the core of the somite clock (17). It is well known that a crucial ingredient of oscillations in *Hes1* expression is the presence of a time delay, associated to transcription, translation or transport. This time delay is often modeled as an explicit time delay (23–25, 40), however it can also result from a slow reactional step (21, 22, 26).

In our system, the time delay is due to the finite gene response time related to the binding/unbinding dynamics. This finite gene response time can also be viewed as taking into account phenomenologically other

sources of delay, if they arise from reactional steps and thus are exponentially distributed. More precisely, the gene can persist in the "off" state for some time after protein level goes down because of the characteristic time $\tau_g = k_{off}^{-1}$ (in original time units) needed to switch from the "off" to the "on" state. Therefore, this characteristic time can be viewed as the delay inducing oscillations, and large-scale variations of protein concentration will typically appear when it is not too small compared to protein half-life. We found that these variations are more regular when these two time scales are equal. Of course, the oscillations in our model remain less regular than those observed in *Hes1* because (1) the delay is exponentially distributed rather than constant and (2) there is no cooperativity. Yet, the models are sufficiently similar that if there is a parameter region where fluctuations promote oscillations in our stochastic self-repressing gene model, this should remain true for the *Hes1* circuit since oscillations would then be more robust to random variations of the delay. Such random variations could be due for instance to the presence of reactional steps. We should then expect this specific parameter region to be selected by evolution.

A first interesting observation is that in the *Hes1* oscillator, the protein and mRNA half-lives are approximately equal, with reported values of 22 and 24 minutes, respectively (17). This is fully consistent with both our observation that regular oscillations occur preferably for $\delta = 1$ (Fig. 5). Note that this contrasts with what is known for the mean-field model, where making degradation rates unbalanced while keeping their sum constant favors oscillations (26).

A crucial parameter for the regularity of the stochastic oscillations is the resonance parameter ρ, which depends on the time delay and on the mRNA and protein half-lives. However, the time delay in the *Hes1 circuit* is not known experimentally. In theoretical investigations (see, e.g., (40, 50)), it is generally assumed that the time delay ranges from 10 to 40 minutes. We assume here a value of 30 minutes, which is consistent with the fact that for large half-lives, the oscillation period of 120 minutes is approximately equal to four times the delay (23). With the known values for the mRNA and protein half-lives (which translate to $\delta_p \sim \delta_m \sim 0.03\,\mathrm{min}^{-1}$), this yields $\rho \sim 2$. Together with $\delta = 1$, this value corresponds precisely to the region of regular oscillations in Fig. 5A. Furthermore, note that Fig. 5D indicates that for $\rho = 2$, the ratio of the oscillation period to the delay $1/k_{off}$ is indeed close to 4.

Finally, other theoretical investigations (see, e.g., (40, 50)) assume that $\Lambda \gg 1$. This in fact a natural condition, which requires that the maximum protein level reached when the gene is fully active must be much larger than the half-repression threshold. This ensures that the protein level can go below and above this threshold in the course of oscillations, and that the gene is strongly repressed when protein level is high.

Taken together, these facts strongly suggest that the *Hes1* mRNA and protein half-lives have been tuned to be both close to the time delay in order to make oscillations in *Hes1* expression robust against stochastic fluctuations in delay.

Discussion and conclusion

In this paper, we have studied the stochastic dynamics of a self-repressing gene model, in which the gene switches randomly between the active and inactive state with a characteristic time which can be arbitrarily small or large compared to mRNA and protein lifetimes. The regularity of the protein spikes generated by the dynamics was characterized using a Fano-like indicator. This allowed us to evidence a dynamical resonance phenomenon, namely that a more regular time evolution of protein concentration is observed when the protein and mRNA degradation rates and the gene response time are nearly equal. It should be stressed that fluctuations are here the only factor triggering oscillations, since our model does not incoporate cooperativity nor nonlinear degradation. The regularity of the oscillations displayed in Fig. 4B would be significantly improved by using these two ingredients, as is done in most theoretical investigations, or by considering a fixed time delay in addition to the exponentially-distributed gene response time.

To understand the resonance phenomenon, we developed a deterministic ODE model using a moment-closure approximation of the master equation. The TME model describes the combined time evolution of the average gene activity, protein and mRNA concentrations and of the two raw moments linking gene activity with protein and mRNA concentrations, in accordance with the fact that the gene state remains a binary variable in the infinite volume limit and thus is the most stochastic variable. Such a model naturally describes how nonlinearity injects fluctuations into the average dynamics, which can be substantially modified. The steady state of the TME model predicts well how averages and covariances vary with the gene response time and the ratio of mRNA and protein degradation rates. In particular, it reproduces the fact that the average gene activity is significantly reduced in the slow gene limit. Unlike rate equations, which are unconditionally stable, the TME model displays a Hopf bifurcation and predicts oscillations in a parameter region where numerical simulations indicate that the protein spikes are indeed more regularly spaced. It remains that the agreement is less good when $\rho > 1$, presumably because the moment closure approximation affects the ρ^3 term in the oscillation criterion, making it sensitive to tiny variations. A direction for future research will thus be to better understand this effect and to derive a more comprehensive description of the self-repressing gene.

Globally, we believe that our results support the idea that deriving deterministic equations through a moment-closure approximation of the master equation is an interesting approach to describe the bifurcation diagram of stochastic dynamical systems, which is generally a difficult problem (see, e.g., (51, 52)). This approach is all the more interesting as computer software is available to derive the hierarchy of equations for the moments or cumulants of increasing order (43, 53). The approach describe here is also well fitted to problems where one variable remains microscopic, such as gene state, and where fluctuations dramatically affect the average values. It thus brings a distinctive advantage compared to other methods based on the linear noise approximation (42).

To check whether the resonance effect discussed here is relevant in real genetic oscillators, we examined

the time scales reported for the *Hes1* self-repressing gene (17, 23, 24). In this circuit, the mRNA and protein lifetimes are approximately equal to the time delay. We found that this situation is characterized by the reduced parameters $\rho = 2$, $\delta = 1$, which are located near the center of the parameter region where regular protein spiking is observed. This strongly suggests that the phenomenon of stochastic resonance we have unveiled plays an important role for generating robust genetic oscillations, independently of other oscillation-enhancing effects such as cooperativity in the transcriptional regulation (54) or nonlinear degradation (26, 29), which can be simultaneously harnessed. A possibly related observation by Murugan and Kreiman is that protein response times fluctuate less when mRNA and protein lifetimes are closer (55). Another interesting fact is that when the sum of half-lives is kept constant, balanced degradation leads to a longer delay in the feedback loop (see Section S.G. in Supporting Information). From a mathematical point of view, the fact that several important time scales coincide may favor the appearance of complex conjugate eigenvalues in the Jacobian matrix, a prerequisite for oscillations.

More generally, we believe that our findings provide a remarkable example of how stochastic fluctuations, which are unavoidable in genetic networks, may play a functional role to shape their dynamics (7).

Acknowledgments

This work has been supported by Ministry of Higher Education and Research, Nord-Pas de Calais Regional Council and FEDER through the Contrat de Projets État-Région (CPER) 2007–2013, as well as by LABEX CEMPI (ANR-11-LABX-0007). The authors thank Benjamin Pfeuty for careful reading of the manuscript.

References

[1] Hartwell, L., J. Hopfield, S. Leibler, and A. Murray, 1999. From molecular to modular cell biology. *Nature* 402:C47–C52.

[2] Golding, I., J. Paulsson, S. M. Zawilski, and E. C. Cox, 2005. Real-time kinetics of Gene activity in Individual bacteria. *Cell* 113:1025–1036.

[3] Darzacq, X., Y. Shav-Tal, V. de Turris, Y. Brody, S. M. Shenoy, R. D. Phair, and R. H. Singer, 2007. In vivo dynamics of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* 14:796–806.

[4] Chubb, J. R., T. Trcek, S. M. Shenoy, and R. H. Singer, 2006. Transcriptional Pulsing of a Developmental Gene. *Curr. Biol.* 16:1018–1025.

[5] Suter, D. M., N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef, 2011. Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science* 332:472–474. `http://www.sciencemag.org/content/332/6028/472.abstract`.

[6] Harper, C. V., B. Finkenstdt, D. J. Woodcock, S. Friedrichsen, S. Semprini, L. Ashall, D. G. Spiller, J. J. Mullins, D. A. Rand, J. R. E. Davis, and M. R. H. White, 2011. Dynamic Analysis of Stochastic Transcription Cycles. *PLoS Biol* 9:e1000607. http://dx.doi.org/10.1371%2Fjournal.pbio.1000607.

[7] Eldar, A., and M. Elowitz, 2010. Functional roles for noise in genetic circuits. *Nature* 467:167–273.

[8] Simpson, M. L., C. D. Cox, and G. S. Sayler, 2003. Frequency domain analysis of noise in autoregulated gene circuits. *Proceedings of the National Academy of Sciences* 100:4551–4556. http://www.pnas.org/content/100/8/4551.abstract.

[9] Warren, P. B., S. Tanase-Nicola, and P. R. ten Wolde, 2006. Exact results for noise power spectra in linear biochemical reaction ne tworks. *The Journal of Chemical Physics* 125:144904. http://link.aip.org/link/?JCP/125/144904/1.

[10] Lestas, I., J. Paulsson, N. Ross, and G. Vinnicombe, 2008. Noise in gene regulatory networks. *Automatic Control, IEEE Transactions on* 53:189–200.

[11] Aquino, T., E. Abranches, and A. Nunes, 2012. Stochastic single-gene autoregulation. *Phys. Rev. E* 85:061913. http://link.aps.org/doi/10.1103/PhysRevE.85.061913.

[12] Stricker, J., S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty, 2008. A fast, robust and tunable synthetic gene oscillator. *Nature* 456:516–520.

[13] Hermsen, R., B. Ursem, and P. R. ten Wolde, 2010. Combinatorial Gene Regulation Using Auto-Regulation. *PLoS Comput Biol* 6:e1000813. http://dx.doi.org/10.13712Fjournal.pcbi.1000813.

[14] Salgado, H., A. Santos-Zavaleta, S. Gama-Castro, D. Milln-Zrate, E. Daz-Peredo, F. Snchez-Solano, E. Prez-Rueda, C. Bonavides-Martnez, and J. Collado-Vides, 2001. RegulonDB (version 3.2): transcriptional regulation and operon organization in Escherichia coli K-12. *Nucleic Acids Research* 29:72–74. http://nar.oxfordjournals.org/content/29/1/72.abstract.

[15] Keseler, I. M., J. Collado-Vides, S. Gama-Castro, J. Ingraham, S. Paley, I. T. Paulsen, M. Peralta-Gil, and P. D. Karp, 2005. EcoCyc: a comprehensive database resource for Escherichia coli. *Nucleic Acids Research* 33:D334–D337. http://nar.oxfordjournals.org/content/33/suppl_1/D334.abstract.

[16] Novak, B., and J. J. Tyson, 2008. Design principles of biochemical oscillators. *Nat. Rev. Mol. Cell Biol.* 9:981–991.

[17] Hirata, H., S. Yoshiura, T. Ohtsuka, Y. Bessho, T. Harada, K. Yoshikawa, and R. Kageyama, 2002. Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop. *Science* 298:840–843.

[18] Goodwin, B. C., 1965. Oscillatory behavior of enzymatic control processes. *Adv. Enzyme Regul.* 3:425–439.

[19] Griffith, J. S., 1968. Mathematics of cellular control processes I. Negative feedback to one gene. *J. Theor. Biol.* 20:202–208.

[20] Bliss, R. D., P. R. Painter, and A. G. Marr, 1982. Role of feedback inhibition in stabilizing the classical operon. *J. Theor. Biol.* 97:177–193.

[21] Goldbeter, A., 1995. A model for circadian oscillations in the Drosophila period protein (PER). *Proc. R. Soc. Lond. B* 261:319–324.

[22] Leloup, J.-C., D. Gonze, and A. Goldbeter, 1999. Limit cycle models for circadian rhythms based on transcriptional regulation in Drosophila and Neurospora. *J. Biol. Rhythms* 14:433–448.

[23] Lewis, J., 2003. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. *Curr. Biol.* 13:1398–1408.

[24] Monk, N. A. M., 2003. Oscilatory expression of Hes1, p53 and NK-κB driven by transcriptional time delays. *Curr. Biol.* 13:1409–1413.

[25] Jensen, M. H., K. Sneppen, and G. Tiana, 2003. Sustained oscillations and time delays in gene expression of protein Hes1. *FEBS Lett.* 541:176–177.

[26] Morant, P.-E., Q. Thommen, F. Lemaire, C. Vandermoëre, B. Parent, and M. Lefranc, 2009. Oscillations in the Expression of a Self-Repressed Gene Induced by a Slow Transcriptional Dynamics. *Phys. Rev. Lett.* 102:068104. `http://link.aps.org/doi/10.1103/PhysRevLett.102.068104`.

[27] Tiana, G., S. Krishna, S. Pigolotti, M. H. Jensen, and K. Sneppen, 2007. Oscillations and temporal signalling in cells. *Phys. Biol.* 4:R1–R17.

[28] Mengel, B., A. Hunziker, L. Pedersen, A. Trusina, M. H. Jensen, and S. Krishna, 2010. Modelling oscillatory control in NF-kB, p53, and Wnt signaling. *Current Opinion in Genetics and Development* 20:656–664.

[29] Tyson, J. J., C. I. Hong, C. D.Thron, and B. Novak, 1999. A simple model of circadian rhythms based on dimerization and proteolysis of PER and TIM. *Biophys. J.* 77:2411–2417.

[30] François, P., and V. Hakim, 2005. Core genetic module: the mixed feedback loop. *Phys. Rev. E* 72:031908.

[31] van Kampen, N. G., 2007. Stochastic processes in physics and chemistry. Elsevier.

[32] Hornos, J. E. M., D. Schultz, G. C. P. Innocentini, J. Wang, A. M. Walczak, J. N. Onuchic, and P. G. Wolynes, 2005. Self-regulating gene: An exact solution. *Phys. Rev. E* 72:051907.

[33] Grima, R., D. Schmidt, and T. Newman, 2012. Steady-state fluctuations of a genetic feedback loop: An exact solution. *J. Chem. Phys.* 137:035104.

[34] Elf, J., and M. Ehrenberg, 2003. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Res.* 13:2475–2484.

[35] McKane, A. J., and T. J. Newman, 2005. Predator-prey cycles from resonant amplification of demographic stochasticity. *Phys. Rev. Lett.* 94:218102.

[36] McKane, A. J., J. D. Nagy, T. J. Newman, and M. O. Stefanini, 2007. Amplified biochemical oscillations in cellular systems. *J. Stat. Phys.* 128:165–191.

[37] Galla, T., 2009. Intrinsic fluctuations in stochastic delay systems: Theoretical description and application to a simple model of gene regulation. *Phys. Rev. E* 80:021909.

[38] Loinger, A., and O. Biham, 2007. Stochastic simulations of the repressilator circuit. *Phys. Rev. E* 76:051917.

[39] Blossey, R., L. Cardelli, and A. Phillips, 2008. Compositionality, stochasticity, and cooperativity in dynamic models of gene regulation. *HFSP Journal* 2:17–28.

[40] Barrio, M., K. Burrage, A. Leier, and T. Tian, 2006. Oscillatory regulation of hes1: Discrete stochastic delay modelling and simulation. *PLoS Comput. Biol.* 2:1017–1030.

[41] Lepzelter, D., H. Feng, and J. Wang, 2010. Oscillation, cooperativity, and intermediates in the self-repressing gene. *Chem. Phys. Lett.* 490:216.

[42] Scott, M., T. Hwa, and B. Ingalls, 2007. Deterministic characterization of stochastic genetic circuits. *Proc. Nat. Acad. Sci. USA* 104:7402–7407.

[43] Gillespie, C. S., 2009. Moment-closure approximations for mass action models. *IET Syst. Biol.* 3:52–58.

[44] Lee, C. H., K.-H. Kim, and P. Kim, 2009. A moment closure method for stochastic reaction networks. *J. Chem. Phys.* 130:134107. `http://link.aip.org/link/?JCP/130/134107/1`.

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

[46] Gibson, M. A., and J. Bruck, 2000. Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. *J. Phys. Chem. A* 104:1876–1889.

[47] Coulon, A., O. Gandrillon, and G. Beslon, 2010. On the spontaneous stochastic dynamics of a single gene: complexity of the molecular interplay at the promoter. *BMC systems biology* 4:2.

[48] Fano, U., 1947. Ionization yield of radiation .2. The fluctuations of the number of ions. *Phys. Rev.* 72:26–29.

[49] Gradshteyn, I. S., and I. M. Ryzhik, 2000. Tables of Integrals, Series, and Products. Academic Press, San Diego.

[50] Bernard, S., B. Cajavec, L. Pujo-Menjouet, M. Mackey, and H. Herzel, 2006. Modelling transcriptional feedback loops: the role of Gro/TLE1 in Hes1 oscillations. *Phil. Trans. R. Soc. A* 364:1155–1170.

[51] Bratsun, D., D. Volfson, L. S. Tsimring, and J. Hasty, 2005. Delay-induced stochastic oscillations in gene regulation. *Proc. Nat. Acad. Sci. USA* 102:14593–14598. `http://www.pnas.org/content/102/41/14593.abstract`.

[52] Song, C., H. Phenix, V. Abedi, M. Scott, B. P. Ingalls, M. Krn, and T. J. Perkins, 2010. Estimating the Stochastic Bifurcation Structure of Cellular Networks. *PLoS Comput. Biol.* 6:e1000699. `http://dx.doi.org/10.1371%2Fjournal.pcbi.1000699`.

[53] Vidal, S., M. Petitot, F. Boulier, F. Lemaire, and C. Kuttler, 2012. Models of Stochastic Gene Expression and Weyl Algebra. *In* K. Horimoto, M. Nakatsui, and N. Popov, editors, Algebraic and Numeric Biology, Springer Berlin Heidelberg, volume 6479 of *Lecture Notes in Computer Science*, 76–97. `http://dx.doi.org/10.1007/978-3-642-28067-2_5`.

[54] Purcell, O., N. J. Savery, C. S. Grierson, and M. di Bernardo, 2010. A comparative analysis of synthetic genetic oscillators. *J. R. Soc. Interface* 7:1503–1524. `http://rsif.royalsocietypublishing.org/content/7/52/1503.abstract`.

[55] Murugan, R., and G. Kreiman, 2011. On the Minimization of Fluctuations in the Response Times of Autoregulatory Gene Networks. *Biophys. J.* 101:1297–1306.

Supporting Material for Stochastic oscillations induced by intrinsic fluctuations in a self repressing gene: a deterministic approach.

Jingkui Wang, Marc Lefranc, and Quentin Thommen

S.A Master equation

The self-repressing gene reaction network involves four chemical species: the unbound gene G, mRNA M, protein P and the DNA-protein complex GP. These molecular actors interact via the following biochemical reactions:

$$\mathrm{G} + \mathrm{P} \underset{k_{off}}{\overset{k_{on}/\Omega}{\rightleftharpoons}} \mathrm{GP} \tag{S1a}$$

$$\mathrm{G} \xrightarrow{\alpha\Omega} \mathrm{G} + \mathrm{M} \tag{S1b}$$

$$\mathrm{M} \xrightarrow{\beta} \mathrm{M} + \mathrm{P} \tag{S1c}$$

$$\mathrm{M} \xrightarrow{\delta_M} \varnothing \tag{S1d}$$

$$\mathrm{P} \xrightarrow{\delta_P} \varnothing \tag{S1e}$$

The cell volume parameter Ω allows us to consider the limit where the protein and mRNA copy numbers are macroscopic variables and are not affected by a one-copy variation. Defining the DNA-protein binding rate as k_{on}/Ω and the transcription rate of the free gene as $\alpha\Omega$ ensures that in the infinite volume limit, the average amount of time spent by the gene in the active state as well as the mRNA and protein average concentrations m/Ω and p/Ω remain bounded. The unbinding rate is k_{off}. The parameter δ_m (resp., δ_p) is the linear mRNA (resp., protein) degradation rate and β is the translation rate.

If $P_{g,m,p}(t)$ denotes the probability to find the gene in stage g (where $g = 0$ represents the bound gene and $g = 1$ the unbound state), together with m mRNA and p protein copies at time t, its time evolution is governed by the following master equation :

$$\begin{aligned}\frac{d}{dt}P_{g,m,p} &= (-1)^g \left[\frac{k_{on}}{\Omega}\ (p+1-g)\ P_{1,m,p+1-g} - k_{off}P_{0,m,p-g}\right] \\ &+\delta_{g,1}\alpha\Omega\left[\mathbb{E}_m^- - 1\right]P_{g,m,p} + \beta m\left[\mathbb{E}_p^- - 1\right]P_{g,m,p} \\ &+\delta_m\left[\mathbb{E}_m^+ - 1\right]m\ P_{g,m,p} + \delta_p\left[\mathbb{E}_p^+ - 1\right]p\ P_{g,m,p}.\end{aligned} \tag{S2}$$

where $\mathbb{E}_x^{\pm}$ is the usual step operator (1) defined by $\mathbb{E}_x^{\pm} f(x,y) = f(x \pm 1, y)$.

S.B Moment expansion

The moments of the probability distribution $P_{g,m,p}$ are defined by:

$$M_{n_1,n_2,n_3} = \langle g^{n_1} m^{n_2} p^{n_3} \rangle = \sum_{g,m,p} g^{n_1} m^{n_2} p^{n_3} P_{g,m,p}. \tag{S3}$$

The idea of a moment expansion is to use the chemical master equation to derive equations describing the time evolution of these statistical quantities, taking into account that the $P_{g,m,p}$ generally evolves with time (1). More precisely, the time derivative of the moments defined by (S3) involves time derivatives of the $P_{g,m,p}$ probabilities, which may be expressed in terms of the $P_{g,m,p}$ themselves using the master equation (S2). The resulting expression can be rewritten in terms of moments (2).

It is well known that closed equations can only be obtained when the underlying dynamics is linear. When it is nonlinear, as is the case here, the time derivative of a cumulant of given order depends on higher-order cumulants, so that there is essentially an infinite number of equations to be considered. A common strategy to obtain a finite-dimensional set of equations approximating the chemical master equations is to truncate this infinite hierarchy in some way. In the present case, we will only consider the infinite cell volume limit, so that the variations of protein and mRNA copy numbers by one unit is negligible. The remaining fluctuations in the mRNA and protein concentrations are then only due to gene fluctuations.

The moment expansion that we derive below takes a simpler form if we replace the mRNA copy number by the weighted average

$$u = \frac{\beta\, m + \delta_m\, p}{\delta_p + \delta_m}, \tag{S4}$$

and by using the following rescaled variables

$$r_t = \tfrac{\delta_m+\delta_p}{\delta_p\delta_m}; \quad r_g = 1; \quad r_u = r_p = \tfrac{k_{on}}{k_{off}\Omega}; \quad r_m = \tfrac{\beta k_{on}}{\delta_p k_{off}\Omega}; \tag{S5a}$$

$$T = r_t\, t; \quad G = r_g\, g; \quad U = r_u\, u; \quad P = r_p\, p. \tag{S5b}$$

Note that since g is a binary variable, $\langle g^n \rangle = \langle g \rangle$, which simplifies the cumulant expansion.

Introducing the following rescaled parameters

$$\rho = \frac{k_{off}\,(\delta_m + \delta_p)}{\delta_p\delta_m}; \quad \Lambda = \frac{\alpha\beta k_{on}}{\delta_m\delta_p k_{off}}; \quad \eta = \frac{(\delta_m + \delta_p)^2}{\delta_m\delta_p}, \tag{S6}$$

the normalized time evolution equations for the averages in the infinite cell volume limit read:

$$\frac{d}{dT}\langle P\rangle = \eta\left[\langle U\rangle - \langle P\rangle\right]; \tag{S7a}$$
$$\frac{d}{dT}\langle U\rangle = \Lambda\langle G\rangle - \langle P\rangle; \tag{S7b}$$
$$\frac{d}{dT}\langle G\rangle = \rho\left(1 - \langle G\rangle - \langle GP\rangle\right); \tag{S7c}$$
$$\frac{d}{dT}\langle GU\rangle = \Lambda\langle G\rangle - \langle GP\rangle - \rho\left[\langle GUP\rangle + \langle GU\rangle - \langle U\rangle\right]; \tag{S7d}$$
$$\frac{d}{dT}\langle GP\rangle = \eta\left[\langle GU\rangle - \langle GP\rangle\right] - \rho\left[\langle GP^2\rangle + \langle GP\rangle - \langle P\rangle\right]; \tag{S7e}$$
$$\frac{d}{dT}\langle U^2\rangle = 2\left[\Lambda\langle GU\rangle - \langle PU\rangle\right]; \tag{S7f}$$
$$\frac{d}{dT}\langle P^2\rangle = 2\eta\left[\langle PU\rangle - \langle P^2\rangle\right]; \tag{S7g}$$
$$\frac{d}{dT}\langle UP\rangle = \Lambda\langle GP\rangle - \langle P^2\rangle + \eta\left[\langle U^2\rangle - \langle PU\rangle\right]. \tag{S7h}$$

Because of the binary gene binding reaction (S1a), the time derivatives of the second order moments $\langle GU\rangle$ and $\langle GP\rangle$ depend on the third-order moments $\langle GUP\rangle$ and $\langle GP^2\rangle$ which are unspecified at this stage. Thus Eqs. (S7) do not form a closed system of equations.

The moments involving the natural variables G, M, and P can be recovered by the relations

$$\langle M\rangle = \frac{(1+\delta)\langle U\rangle - \langle P\rangle}{\delta}, \tag{S8a}$$
$$\langle GM\rangle = \frac{(1+\delta)\langle GU\rangle - \langle GP\rangle}{\delta}, \tag{S8b}$$
$$\langle MP\rangle = \frac{(1+\delta)\langle PU\rangle - \langle P^2\rangle}{\delta}, \tag{S8c}$$
$$\langle M^2\rangle = \frac{(1+\delta)^2\langle U^2\rangle - 2(1+\delta)\langle PU\rangle + \langle P^2\rangle}{\delta^2}. \tag{S8d}$$

S.C First order truncation of the moment expansion

A first strategy to truncate the hierarchy of moment equations is to set all covariances (the second order centered moments) to zero (1)

$$\langle (X - \langle X\rangle)(Y - \langle Y\rangle)\rangle = 0,$$

which enslaves the covariances to the means $\langle XY\rangle = \langle X\rangle\langle Y\rangle$. Under this approximation, all fluctuations are neglected and the following deterministic rate equations for the averages are obtained:

$$\frac{d}{dT}\langle P\rangle = \eta\left[\langle U\rangle - \langle P\rangle\right]; \tag{S9a}$$

$$\frac{d}{dT}\langle U\rangle = \Lambda\langle G\rangle - \langle P\rangle; \tag{S9b}$$

$$\frac{d}{dT}\langle G\rangle = \rho\left(1 - \langle G\rangle - \langle G\rangle\langle P\rangle\right). \tag{S9c}$$

The steady state solution of (S9) is given by

$$\Lambda\langle G\rangle^* = \langle U\rangle^* = \langle P\rangle^* = \frac{1}{2}\left(\sqrt{1+4\Lambda} - 1\right), \tag{S10}$$

which does not depend on ρ, and is stable in the entire parameter space. Indeed, it was noted by Morant *et al.* (3) that besides the finite gene response time, a nonlinear degradation mechanism is needed to induce oscillations in this system.

Incorporating fluctuations in the dynamics of the average quantities requires truncating the hierarchy at a higher order. We discuss two different strategies in the following sections.

S.D Second order truncation, the TOT model

S.D.1 Derivation of the model

A natural extension of the previous developed truncation is to keep the second order moments and enslave the third order moments to the means and covariances by assuming vanishing third order centered moments. So assuming that

$$K_{GUP} = \langle\left(G - \langle G\rangle\right)\left(U - \langle U\rangle\right)\left(P - \langle P\rangle\right)\rangle = 0, \quad \text{and} \quad K_{GPP} = \langle\left(G - \langle G\rangle\right)\left(P - \langle P\rangle\right)^2\rangle = 0,$$

fixes the two following dependencies

$$\langle GUP\rangle = G\langle UP\rangle + U\langle GP\rangle + P\langle GU\rangle - 2\langle G\rangle\langle U\rangle\langle P\rangle, \tag{S11}$$

$$\langle GP^2\rangle = G\langle P^2\rangle + 2P\langle GP\rangle - 2\langle G\rangle\langle P\rangle^2. \tag{S12}$$

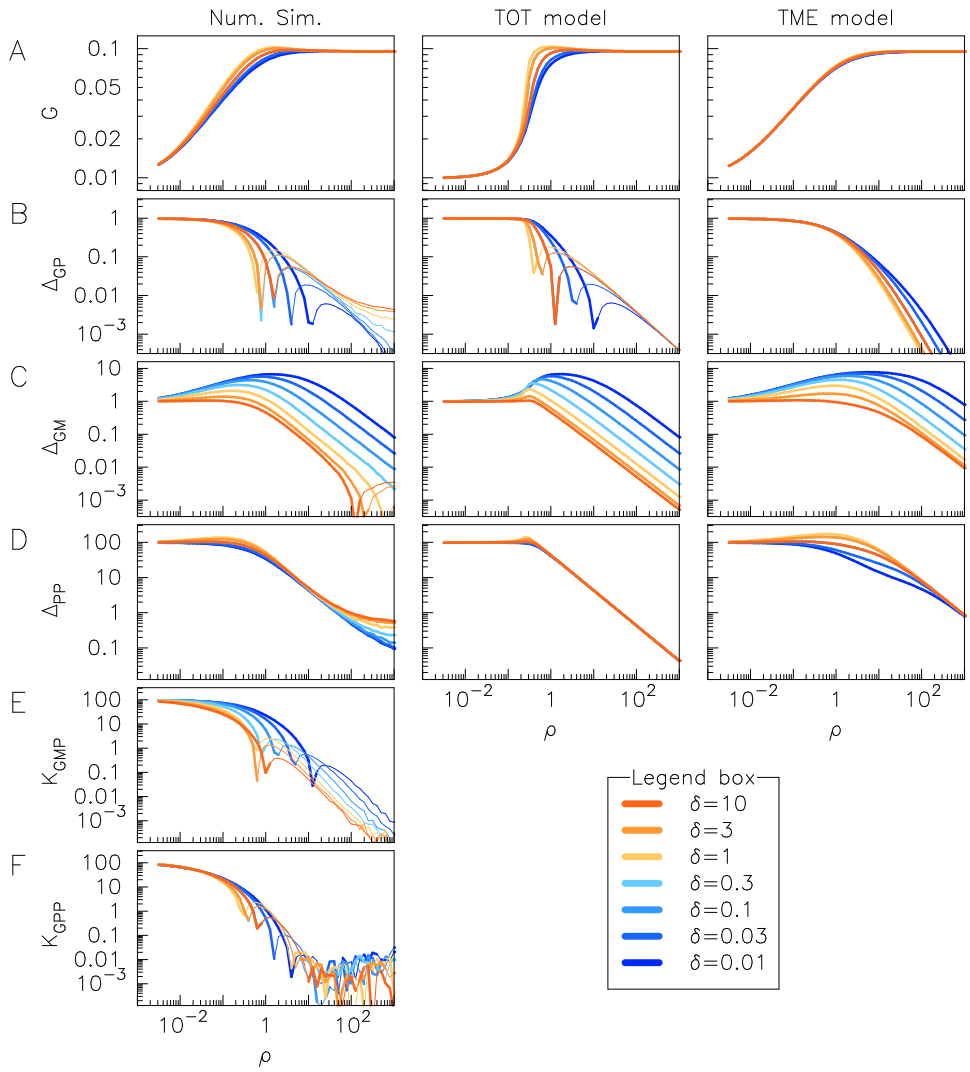

Figure S6: **Comparison of averages and covariances obtained from stochastic simulations and from the fixed points of ODE models derived using the TOT and TME truncation schemes.** (A) Average gene activity G; (B), (C), (D) covariances Δ_{GP}, Δ_{GM} and Δ_{PP}; (E), (F) third-order cumulants computed from numerical simulations. Curves for different values of δ are color-coded according to legend box. In each panel, thick lines (resp., thin) lines indicate positive (resp., negative) values.

Under this hypothesis, the time evolution of averages and covariances is described by the closed differential system :

$$\frac{d}{dT}P = \eta\,(U-P)\,; \tag{S13a}$$
$$\frac{d}{dT}U = \Lambda G - P; \tag{S13b}$$
$$\frac{d}{dT}G = \rho\,(1-G-GP-\Delta_{G,P})\,; \tag{S13c}$$
$$\frac{d}{dT}\Delta_{G,U} = \Lambda G\,(1-G) - \Delta_{G,P} - \rho\,[G\Delta_{P,U} + (P+1)\,\Delta_{G,U}]\,; \tag{S13d}$$
$$\frac{d}{dT}\Delta_{G,P} = \eta\,[\Delta_{G,U} - \Delta_{G,P}] - \rho\,[G\Delta_{P,P} + (P+1)\,\Delta_{G,P}] \tag{S13e}$$
$$\frac{d}{dT}\Delta_{U,U} = 2\,[\Lambda\Delta_{G,U} - \Delta_{P,U}]\,; \tag{S13f}$$
$$\frac{d}{dT}\Delta_{P,P} = 2\eta\,(\Delta_{P,U} - \Delta_{P,P})\,; \tag{S13g}$$
$$\frac{d}{dT}\Delta_{P,U} = \Lambda\Delta_{G,P} - \Delta_{P,P} + \eta\,[\Delta_{U,U} - \Delta_{P,U}]\,, \tag{S13h}$$

where $\Delta_{X,Y}$ stand for the covariance of random variables X and Y : $\Delta_{X,Y} = \langle XY\rangle - \langle X\rangle\langle Y\rangle$. We refer to model (S13) as the Third-Order Truncation (TOT) model.

The steady state of model (S13) is obtained by solving the following equations:

$$U = P = \Lambda G; \qquad \Delta_{P,P} = \Delta_{M,P} = \Delta_{U,P} = \Lambda\Delta_{G,U}; \tag{S14a}$$
$$\eta\Delta_{U,U} = (1+\eta)\,\Lambda\Delta_{G,U} - \Lambda\Delta_{G,P}; \tag{S14b}$$
$$\Delta_{G,P} = 1 - G - \Lambda G^2; \tag{S14c}$$
$$(\rho + \rho\Lambda G + \eta)\,\Delta_{G,P} = [\eta - \rho\Lambda G]\,\Delta_{G,U}; \tag{S14d}$$
$$\Delta_{G,P} + \rho\,(1+2\Lambda G)\,\Delta_{G,U} = \Lambda G\,(1-G)\,; \tag{S14e}$$
$$\Delta_{M,M} = \Lambda\,\Delta_{G,M} = \Lambda\frac{(1+\delta)\,\Delta_{G,U} - \Delta_{G,P}}{\delta}. \tag{S14f}$$

The steady state value of Δ_{GP}, which is the joint correlation between the gene state and the protein copy number, vanishes when $\rho\Lambda G = \eta$. The steady state values of averages in the model (S13) then coincide with those derived from the rate equations (S9), given by (S10). Except in this particular case, equations (S14) do not admit analytical solutions. However, asymptotic expressions for the steady state values of averages and covariances can be obtained by a perturbative expansion when the resonance parameter ρ and feedback strength Λ are either very large or very small, as is summarized in Table S1. In this computation, the ratio δ is assumed to be neither very large nor very small. The expressions given in Table S1 allow us to characterize the effect of fluctuations in the different limiting cases considered.

The second column of Fig. (S1) shows that the fixed point values of the TOT model are in good quantitative agreement with the numerical estimators (first column of Fig. (S1)). Regarding the averages,

	$\rho \to 0$ $\Lambda \ll 1$	$\rho \to \infty$ $\Lambda \ll 1$	$\rho \to 0$ $\Lambda \gg 1$	$\rho \to \infty$ $\Lambda \gg 1$
$\Lambda G^* = U^* = P^* \simeq$	$\Lambda + \rho\Lambda$	$\Lambda + \frac{\Lambda^4}{\rho}$	$1 + 3\rho$	$\sqrt{\Lambda}\left(1 + \frac{1}{4\rho}\right)$
$\Delta^*_{GP} \simeq$	$\Lambda^2 - \rho$	$-\frac{\Lambda^3}{\rho}$	$1 - \frac{9\rho}{\Lambda}$	$-\frac{1}{2\rho}$
$\Delta^*_{M,P} = \Delta^*_{P,P} = \Delta^*_{U,P} = \Lambda\Delta^*_{G,U} \simeq$	$\Lambda^3 - 3\rho\Lambda^2$	$\frac{\Lambda^3}{\rho}$	$\Lambda + 3\rho\frac{\Lambda - 3\eta}{\eta}$	$\frac{\sqrt{\Lambda}}{2\rho}$
$\Delta^*_{U,U} \simeq$	$\Lambda^3 - \rho\Lambda$	$\frac{\Lambda^3}{\rho}\frac{\eta+2\Lambda}{\eta}$	$\Lambda + 3\rho\Lambda\frac{1+\eta}{\eta^2}$	$\frac{\sqrt{\Lambda}}{2\rho}\frac{2+\eta}{\eta}$
$\Delta^*_{M,M} = \Lambda\Delta^*_{G,M} \simeq$	$\Lambda^3 - \rho\Lambda$	$\frac{\Lambda^3}{\rho}\frac{\delta+2\Lambda}{\delta}$	$\Lambda + 3\rho\Lambda\frac{1+\delta}{\delta\eta}$	$\frac{\sqrt{\Lambda}}{2\rho}\frac{2+\delta}{\delta}$

Table S1: Asymptotic expressions of the steady state values of averages and covariances for Eqs. (S13).

the overall shapes of the curves, with a maximum around $\rho = 1$, are very similar and the evolution of this maximum with δ is reproduced (Fig. S1-A). The main discrepancy is that the transition from the fast to the slow gene regime is more abrupt in the TOT model than in stochastic simulations, presumably because higher-order contributions to the averages are neglected. The global evolution of the covariances is also well reproduced, and the values of ρ where Δ_{GP} becomes zero are also well predicted for the different values of δ (Fig. S1B). Similarly, the variation of Δ_{GM} with δ is captured. (Fig. S1-C). However, the TOT model fixed point values overestimate the covariances Δ_{GM} and Δ_{PP} (Fig. S1-C,D) in the fast gene limit. Still, the asymptotic values of the TOT model steady states (summarized in Table S1 in the Supporting Material) are correctly reproduced.

A key assumption of the TOT model is that the third centered moments $K_{G,U,P}$ and $K_{G,P,P}$ vanish, which is correct in the fast gene limit. However, Figs. S1-E,F show that they take rather large values in the stochastic simulations, of the order of Λ, in the slow gene limit. One may thus wonder why the TOT model is effective in this regime. Examining the structure of the equations solves this paradox.

Consider the dynamical equations for the covariance Δ_{GP} and Δ_{GU}:

$$\frac{d}{dT}\Delta_{G,U} = \Lambda G\left(1-G\right) - \Delta_{G,P} - \rho\left[K_{G,U,P} + G\Delta_{P,U} + \left(P+1\right)\Delta_{G,U}\right]; \tag{S15a}$$

$$\frac{d}{dT}\Delta_{G,P} = \eta\left[\Delta_{G,U} - \Delta_{G,P}\right] - \rho\left[K_{G,P,P} + G\Delta_{P,P} + \left(P+1\right)\Delta_{G,P}\right]. \tag{S15b}$$

The key point is that K_{GUP} and K_{GPP} are both weighted by ρ, so their dynamical influence vanishes in the slow-gene limit even thought their are non zero.

Figure (S7) displays the values of all terms in Eqs. (S15) for various ρ. In the slow gene regime, ($\rho \to 0$) the first two terms of each equation dominate (i.e., $\Lambda G\left(1-G\right)$ and Δ_{GP} for Eq. (S15a); $\eta\Delta_{GU}$ and $\eta\Delta_{GP}$ For eq. (S15b)) whereas in the fast gene limits, the last two terms dominate ($\rho\Delta_{PU}$ and $\rho\left(P+1\right)\Delta_{GU}$ for eq. (S15a)); $\rho\Delta_{PP}$ and $\rho\left(P+1\right)\Delta_{GP}$ for eq. (S15b)). The fact that third-order central moments do not converge to zero in numerical simulations when $\rho \to \infty$ is due to numerical cancellation errors in their computation, because two nearly equal numbers are being substracted, and should not be taken into account.

In both regimes, the influence of the third order cumulants vanish. It turns out that terms involving

third-order cumulants play a more important role in the intermediate regime where they are the dominant negative terms in the expression of the time derivative of Δ_{GU} for $\rho \simeq 0.1$. Therefore, the TOT model provides an excellent approximation for both fast and slow gene dynamics, and provides only a reasonable description of the dynamics in the intermediate regime.

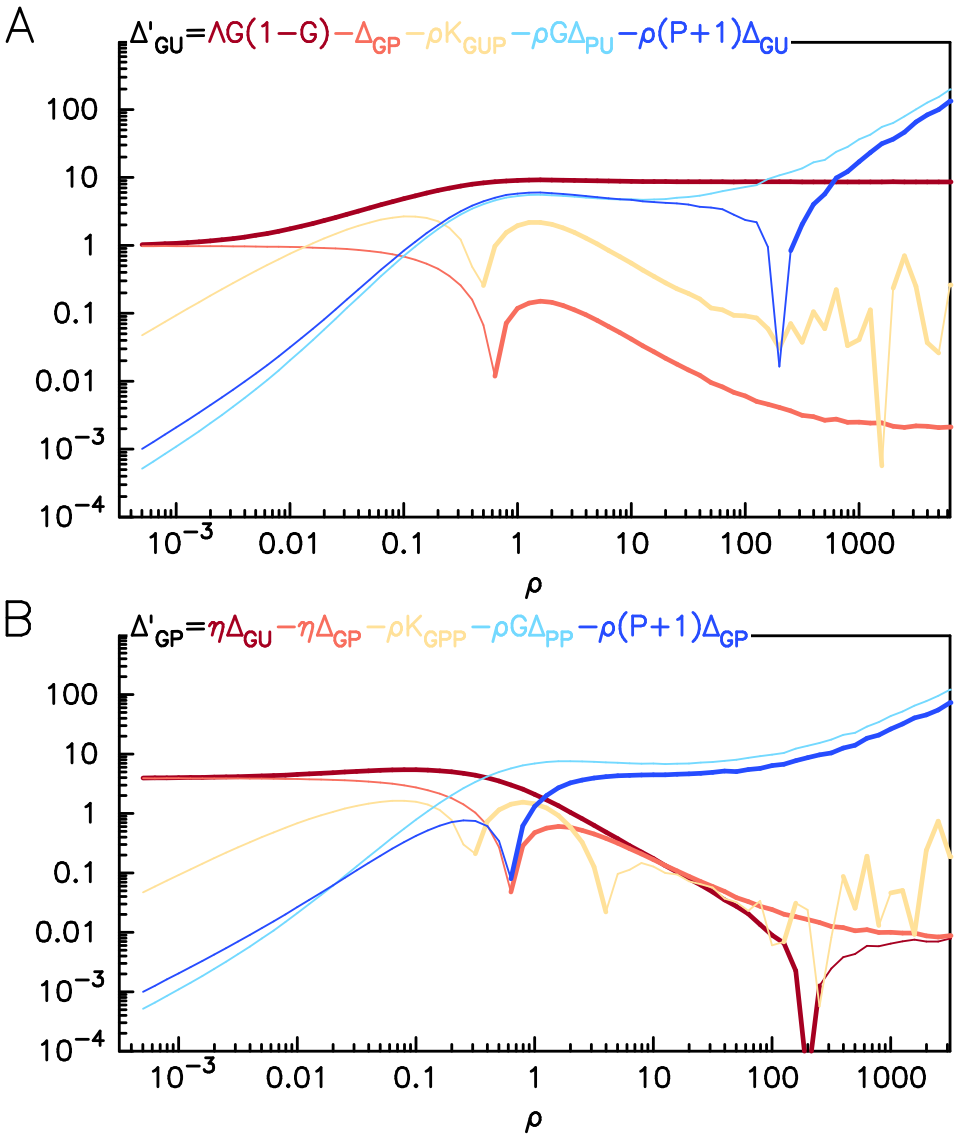

Figure S7: **Dynamical influence of the third order cumulants** The magnitudes of the different terms appearing in Eqs. (S15) are numerically computed using stochastic simulations for $\Lambda = 100$, $\delta = 1$ (equal degradation rates), and various ρ. The differential equation at the top of each panel indicates the color code. In each panel, thick lines (resp., thin) lines indicate positive (resp., negative) values.

S.E Alternative truncation, the TME model

In the moment expansion (S7), Eqs. (S7d-e) describing the time evolution of $\langle GP\rangle$ and $\langle GU\rangle$ are independent of $\langle U^2\rangle$, $\langle P^2\rangle$, and $\langle UP\rangle$. However, Eqs. (S7a-e) do not form a closed system due to the presence of the $\langle GUP\rangle$ and $\langle GP^2\rangle$ terms. Here, we use an another closure approximation by enslaving the third moment $\langle GUP\rangle$ and $\langle GP^2\rangle$ to the average gene activity $\langle G\rangle$ *via* a phenomenological function.

In the case of a strong repression (i.e., $\Lambda \gg 1$), the moments $\langle GUP\rangle$ and $\langle GP^2\rangle$ can be derived from considerations both in the slow and the fast gene limits. In the fast gene limit, the proteins and mRNA number of copies are almost constant over a gene switch and correspond to their stationary value $G \simeq U \simeq \sqrt{\Lambda}$ so $\langle GUP\rangle_{\rho\to\infty} = \langle GP^2\rangle_{\rho\to\infty} = \Lambda\langle G\rangle^* = \sqrt{\Lambda}$, as $\langle G\rangle^* \simeq 1/\sqrt{\Lambda}$. In the slow gene limit ($\rho \ll 1$) the averages, covariances and third-order joint cumulants can be computed because all variables are slaved to the gene state variable. In particular, the values of P and U alternate between 0 when the gene is off and Λ when the gene is on. In normalized time units, the gene is active during a time $t_{ON} = 1/\Lambda$ and inactive during a time $t_{OFF} = 1$ so that its average activity is $G^* = t_{ON}/(t_{ON} + t_{OFF}) = 1/(1+\Lambda)$ and $P^* = U^* = \Lambda G^* = \Lambda/(1+\Lambda) \approx 1$. Because P and U can be assumed to have a constant value of Λ during the phase where $G = 1$, it follows that $\langle GUP\rangle_{\rho\to 0} = \langle GP^2\rangle_{\rho\to 0} = \Lambda\langle G\rangle^* = \frac{1}{\Lambda}$. Finally, we get $\langle GUP\rangle = \langle GP^2\rangle = \frac{1}{\langle G\rangle}$ in the two limits. We then seek to express $\langle GUP\rangle$ and $\langle GP^2\rangle$ in terms of the same function depending on $\langle G\rangle$ only :

$$\langle GUP\rangle = \langle GP^2\rangle = F\left(\langle G\rangle\right).$$

The moment expansion (S7) then reduces to a five dimensional ODE system:

$$\frac{d}{dT}\langle P\rangle = \eta\left[\langle U\rangle - \langle P\rangle\right]; \tag{S16a}$$

$$\frac{d}{dT}\langle U\rangle = \Lambda\langle G\rangle - \langle P\rangle; \tag{S16b}$$

$$\frac{d}{dT}\langle G\rangle = \rho\left(1 - \langle G\rangle - \langle GP\rangle\right); \tag{S16c}$$

$$\frac{d}{dT}\langle GU\rangle = \Lambda\langle G\rangle - \langle GP\rangle - \rho\left[F\left(\langle G\rangle\right) + \langle GU\rangle - \langle U\rangle\right]; \tag{S16d}$$

$$\frac{d}{dT}\langle GP\rangle = \eta\left[\langle GU\rangle - \langle GP\rangle\right] - \rho\left[F\left(\langle G\rangle\right) + \langle GP\rangle - \langle P\rangle\right]. \tag{S16e}$$

The fixed point of Eqs. (S16) is obtained by solving

$$-\left(\Lambda+1\right)\langle G\rangle^* + 1 + \frac{\rho\left(\eta+\rho\right)}{\rho\left(\eta+\rho\right)+\eta}F\left(\langle G\rangle^*\right) = 0. \tag{S17}$$

Requesting that the solution of Eq. (S17) in the limit of fast gene ($\rho \gg 1$) coincides with the stationary state of the rate equation (S9) allows one to obtain the asymptotic form of the unknown function F:

$$\lim_{\rho\to\infty} F\left(\langle G\rangle\right) = F_\infty\left(\langle G\rangle\right) = \frac{\left(1-\langle G\rangle\right)^2}{\langle G\rangle}.$$

By extending this asymptotic form to the whole ρ axis and fixing

$$F\left(\langle G\rangle\right) = \frac{\left(1-\langle G\rangle\right)^2}{\langle G\rangle},$$

an alternate moment-closure model is obtained, which we term the Truncated Moment Expansion (TME) model. Its fixed point describe well the stationary values of the averages (see main text).

The third column of Fig S1 displays the fixed points of the TME model recast in terms of averages and covariances of the natural rescaled variables G, M, and P, so as to allow comparison with the fixed points of the TOT model and provide information which is complementary to that of Fig 3. The absence of an overshoot in the averages near $\rho = 1$ is correlated with the fact that Δ_{GP} does not change its sign as rho increases.

S.F Stability analysis

Figure (S8) compares the parameter space regions where TOT and TME models oscillate, as indicated by a numerical stability analysis. In the slow gene regime ($\rho \to 0$), the two models display similar behavior, as could be expected from the fact that the closure approximations are consistent in this case. Similarly, none of the two models displays oscillations in the large ρ limit.

However, the oscillation region of the TOT model is much narrower and moreover, is clearly disconnected from the ($\rho = 1$, $\delta = 1$) central point where the regular stochastic oscillations are preferentially observed. This suggests that the third-order cumulants play an important role in the dynamics, in accordance with their importance in the equations describing the time evolution of covariances involving the gene state (Fig. (S7)).

S.G Analysis of the low-pass filter : Cut-off frequency and Feedback delay

In the infinite volume limit, Eqs. (2a-b) describing the time evolution of the averages of mRNA and protein concentrations are linear and do not depend on higher-order moments. Assume that mRNA and protein concentrations respond to gene activity considered as an external signal. The two equations

$$\frac{d}{dt}\langle\mathcal{P}\rangle(t) = \beta\,\langle\mathcal{M}\rangle(t) - \delta_p\,\langle\mathcal{P}\rangle(t); \tag{S18a}$$

$$\frac{d}{dt}\langle\mathcal{M}\rangle(t) = \alpha\,\langle g\rangle(t) - \delta_m\,\langle\mathcal{M}\rangle(t). \tag{S18b}$$

can be viewed as describing a low-pass filter, whose dynamics is easily characterized. If we denote by $\langle g\rangle(\omega)$ and $\langle\mathcal{P}\rangle(\omega)$ the Fourier transforms of the input $\langle g\rangle(t)$ and output $\langle\mathcal{P}\rangle(t)$ of the low-pass filter, then the

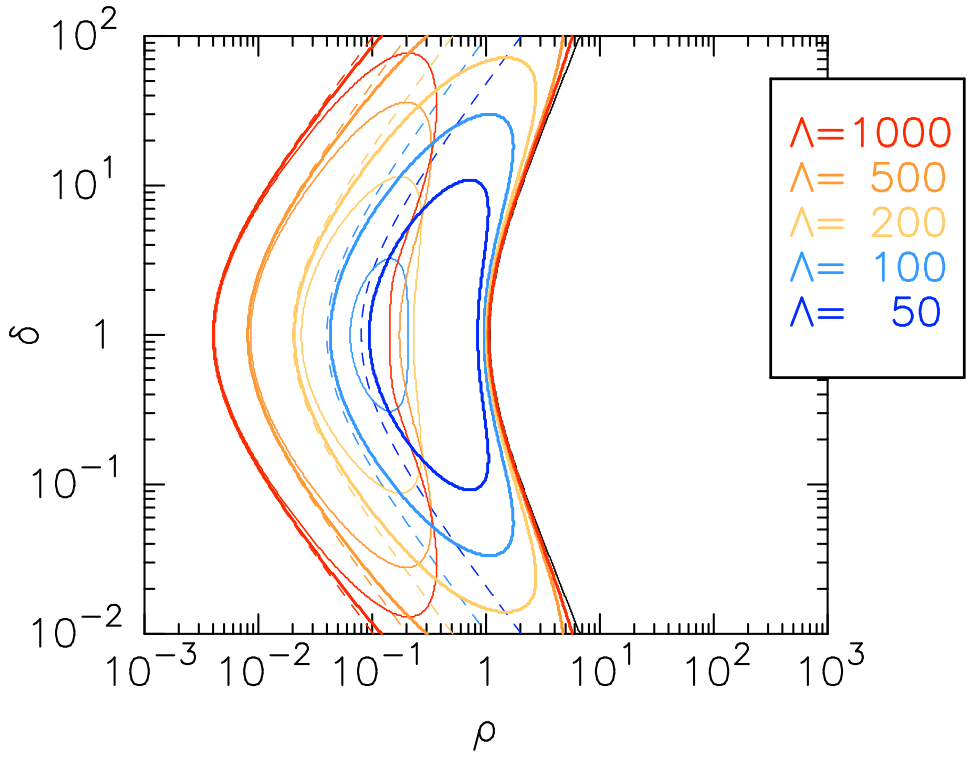

Figure S8: **Comparison between TOT and TME models oscillatory domain.** The lines enclose the region where the TOT model (thin lines) and the TME model (thick lines) oscillate for various Λ (see legend box for color code).. The analytical expression of the two boundaries derived for $\Lambda \gg 1$ are $\rho = \eta/\Lambda$ (dashed lines) and $\eta^2 - 2\rho^3 - 2\eta\rho^2 - \eta\rho = 0$ (black line), with $\eta = \frac{(1+\delta)^2}{\delta}$ (see Sec. S.H).

transfer function is given by

$$F(\omega) = \frac{\langle \mathcal{P} \rangle(\omega)}{\langle g \rangle(\omega)} = \frac{\alpha\beta}{\delta_p\delta_m - \omega^2 + i\omega(\delta_p + \delta_m)}. \tag{S19}$$

The cut-off frequency Ω_c, defined by $|F(\omega = \Omega_c)|^2 = \frac{1}{2}|F(\omega = 0)|^2$, characterizes the spectral interval in the input which is transmitted to output. More precisely, a sinusoidal input of frequency Ω_c and amplitude A induces a sinusoidal output of amplitude $A/\sqrt{2}$. The expression of the cut-off frequency Ω_c is

$$\Omega_c = \omega_c \sqrt{\frac{\eta^2 - 2\eta}{2}} \sqrt{\sqrt{1 + \frac{4}{(\eta-2)^2}} - 1}, \tag{S20}$$

where $\omega_c = \frac{\delta_p\delta_m}{\delta_p+\delta_m}$ and $\eta = \frac{(\delta_p+\delta_m)^2}{\delta_p\delta_m}$. Ω_c/ω_c varies between $2\sqrt{\sqrt{2}-1} \approx 1.29$ when $\eta = 4$ and 1 when η is large. Thus ω_c provides a good approximation of Ω_c (whose definition involves itself an arbitrary choice) and characterizes the relevant time scale.

If we rescale frequency with respect to ω_c by defining $\omega = \omega'\omega_c$, the transfer function reads

$$F(\omega') = \frac{K}{\eta - \omega'^2 + i\omega'\eta}, \tag{S21}$$

where K is a constant. This corresponds to the time rescaling that we use in model derivation and which are defined Eqs. (S5).

The low-pass filter, as any linear system, is fully characterized by its impulse response, computed as the inverse Fourier transform of the transfer function Eq. (S21). The impulse response represents the protein time profile created by an infinitely short pulse of gene activity at time 0 :

$$P_{IR}(T) \propto \frac{2\eta}{\sqrt{\eta^2 - 4\eta}} \sinh\left(\frac{1}{2}\sqrt{\eta^2 - 4\eta}\, T\right) e^{-\frac{1}{2}\eta T} \quad T \geq 0. \tag{S22}$$

The impulse response displays a maximum at $T = T_m$, where T_m depends on η only:

$$T_m = \frac{\log\left(\sqrt{\eta^2 - 4\eta} + \eta\right) - \log\left(-\sqrt{\eta^2 - 4\eta} + \eta\right)}{\sqrt{\eta^2 - 4\eta}}. \tag{S23}$$

The value of T_m, which corresponds to the delay between gene activity pulse and maximum protein concentration, decreases monotonously from its maximum value of 0.5 for $\eta = 4$ to 0 for large η (Fig. S9).

Assuming that the sum of mRNA and protein half-lives is fixed, the case of balanced half-lives ($\eta = 4$, $\delta = 1$) implies then a longer delay in the negative feedback loop.

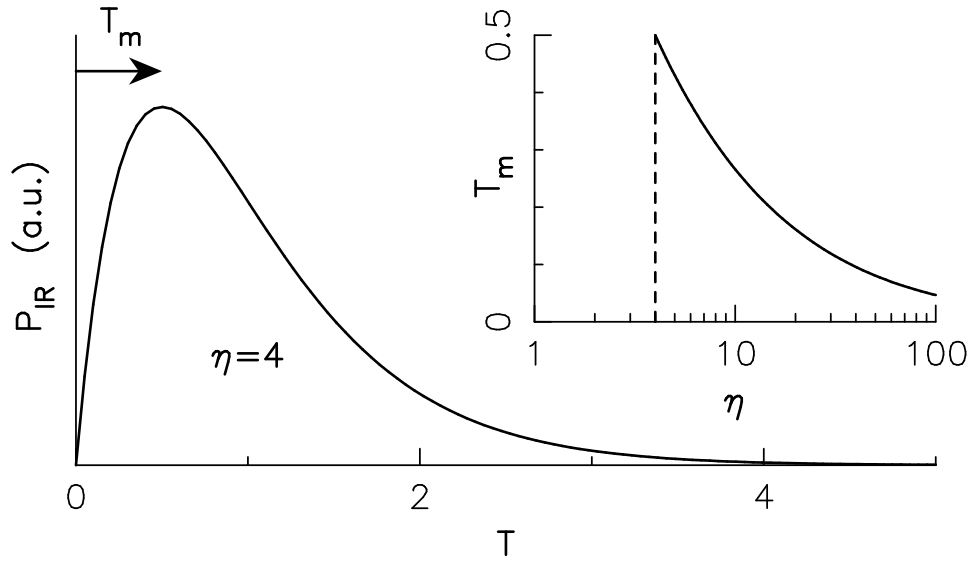

Figure S9: **Impulse response of the low pass filter.** The protein concentration time profile in response to an impulse gene signal displays a maximum around a time T_m, whose dependence on η is shown in the insert.

S.H Linear stability analysis of the TME model

The linear stability analysis characterizes the qualitative behavior of the trajectories of a dynamical system near a fixed point by examining the eigenvalues of the Jacobian matrix evaluated at the fixed point. If all eigenvalues have negative real parts, the fixed point is stable.

When the real part of a pair of complex conjugate eigenvalues crosses zero from negative to positive, the fixed point becomes unstable and generically gives birth to a limit cycle, associated with appearance of spontaneous oscillations (Hopf bifurcation) (4). The occurrence of such a bifurcation can be investigated using the Routh-Hurwitz criterion (5, 6) without having to compute the actual eigenvalues. The Routh-Hurwitz criterion provides one with a set of functions of the coefficients of the characteristic polynomial, which are all negative when the fixed point is stable. One of these functions go through zero at a Hopf bifurcation, and thus can be used as a criterion for the appearance of oscillations.

The dynamical properties of the TME model is governed by Eqs. (5a-e) of the main text. The Jacobian matrix evaluated at the fixed point reads

$$J = \begin{pmatrix} -\eta & \eta & 0 & 0 & 0 \\ -1 & 0 & \Lambda & 0 & 0 \\ 0 & 0 & -\rho & 0 & -\rho \\ 0 & \rho & \Lambda - \rho D & -\rho & -1 \\ \rho & 0 & -\rho D & \eta & -\eta - \rho \end{pmatrix}, \tag{S24}$$

where $D = D(\langle G \rangle^*) = D(\rho, \eta, \Lambda) = dF(X)/dX|_{X=\langle G \rangle^*}$ is the derivative of the function F used in the closure

approximation $\langle GUP\rangle = \langle GP^2\rangle = F(\langle G\rangle)$.

The analysis of the Routh Table computed using the characteristic polynomial of the Jacobian (S24) leads to an oscillation criterion $\mathcal{H}'(\rho,\eta,\Lambda)$ with a complicated expression, however the analysis of its structure reveals that

$$\mathcal{H}(\rho,\eta,\Lambda) = \rho^3\left(8-2D(\rho,\eta,\Lambda)\right) + 8\eta\rho^2 + \rho\eta\left(2\eta+2-\Lambda\right) + \eta^2 \ < 0 \tag{S25}$$

is a sufficient condition for the occurrence of spontaneous oscillations. Indeed, the Routh-Hurwitz criterion can be decomposed as $\mathcal{H}' = A\times\mathcal{H} - B < 0$ where A and B are two strictly positive functions of ρ, η, and Λ, and thus cannot become positive if H is not positive. In practice, numerical simulations show that $\mathcal{H}=0$ delimitates very accurately the oscillation region in parameter space (see Fig. 5 in main Text).

Interestingly, $\mathcal{H}<0$ corresponds to the stability criterion of the approximated Jacobian

$$J' = \begin{pmatrix} -\eta & \eta & 0 & 0 & 0 \\ -1 & 0 & \Lambda & 0 & 0 \\ 0 & 0 & -\rho & 0 & -\rho \\ 0 & \mathbf{0} & \Lambda-\rho D & -\rho & -1 \\ \mathbf{0} & 0 & -\rho D & \eta & -\eta-\rho \end{pmatrix}. \tag{S26}$$

where the leftmost entries on fourth and fifth row have been set to zero.

With the closure $F(X) = \frac{(1-X)^2}{X}$ used in the main Text, we have $D(X) = 1-\frac{1}{X^2}$ and the oscillation criterion reads

$$\mathcal{H}(\rho,\eta,\Lambda) = \rho^3\left(6+\frac{2}{\langle G\rangle^{*2}}\right) + 8\eta\rho^2 + \rho\eta\left(2\eta+2-\Lambda\right) + \eta^2 \ < 0. \tag{S27}$$

where $\langle G\rangle^* = \langle G\rangle^*(\rho,\eta,\Lambda)$ is given by expression (6) in the main Text.

Because the derivative of the closure function appears in the coefficient of ρ^3, the location of the oscillation region will typically be very sensitive to the choice of the closure function, especially in the region around $\rho=1$, where the more regular stochastic oscillations are observed, and even more for larger values of ρ. This probably explains why the agreement between the instability region of the TME model and the region where regular stochastic oscillations are observed is not very good for $\rho>1$.

A even simpler oscillation criterion can be obtained in the limit of strong feedback, when $\Lambda\to\infty$, without having to approximate the Routh-Hurwitz criterion. In this limit, we have to consider two cases depending on the value of ρ.

If ρ is $O(1)$, then the fixed point of the TME model is determined to leading order in $1/\Lambda$ by

$$< G >^* = \sqrt{\frac{\rho\ (\eta+\rho)}{\rho\ (\eta+\rho)+\eta}}\sqrt{\frac{1}{\Lambda}} \tag{S28}$$

If, however, ρ is sufficiently small that it can be written $\rho = \frac{K}{\Lambda}$ with $K = O(1)$, then the leading order

solution of the TME fixed point equations is

$$< G >^{*} = \frac{1+\sqrt{1+4\,K}}{2}\frac{1}{\Lambda} \tag{S29}$$

Note that the average gene activity scales differently with Λ in Eqs. (S28) and (S29).

To obtain the oscillation criterion in the limit of large Λ, we substitute expressions (S28) and (S29) in the Jacobian (S24) and compute the Hopf Routh-Hurwitz criterion to leading order in Λ, which considerably simplifies the expression.

We thus find that oscillations occur whenever

$$-2\,\rho^2\eta - 2\,\rho^3 + \eta^2 - \rho\,\eta > 0, \quad [\rho = O(1)] \tag{S30}$$

if the gene response time is similar to degradation rates, or when

$$\rho > \frac{\eta}{\Lambda} \quad [\rho = O(1/\Lambda)]. \tag{S31}$$

when the gene response time is large. Note that Eq. (S31) confirms that oscillations appear for very small ρ in the limit of large Λ, and also that it is consistent with the assumed scaling.

In spite of their simplicity, the two expressions provide excellent approximations of the two boundaries of the instability region when Λ is large, as can be seen in Fig. S8. This allows one to discuss the relative influences of gene response time (described by ρ) and degradation rate balance (described by η) on the appearance of oscillations.

Interestingly, the conditions (S31) and (S30) can also be recovered by injecting expressions (S28) and (S29) in the approximate criterion (S27), showing that the latter is all the more accurate as Λ is large.

References

[1] van Kampen, N. G., 2007. Stochastic processes in physics and chemistry. Elsevier.

[2] Gillespie, C. S., 2009. Moment-closure approximations for mass action models. *IET Syst. Biol.* 3:52–58.

[3] Morant, P.-E., Q. Thommen, F. Lemaire, C. Vandermoëre, B. Parent, and M. Lefranc, 2009. Oscillations in the Expression of a Self-Repressed Gene Induced by a Slow Transcriptional Dynamics. *Phys. Rev. Lett.* 102:068104. `http://link.aps.org/doi/10.1103/PhysRevLett.102.068104`.

[4] Kuznetsov, Y. A., 2006. Andronov-Hopf bifurcation. *Scholarpedia* 1:1858. revision #90964.

[5] Gradshteyn, I. S., and I. M. Ryzhik, 2000. Tables of Integrals, Series, and Products. Academic Press, San Diego.

[6] Hairer, E., S. P. Nørsett, and G. Wanner, 1991. Solving ordinary differential equations I. Springer.