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

A. 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

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

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

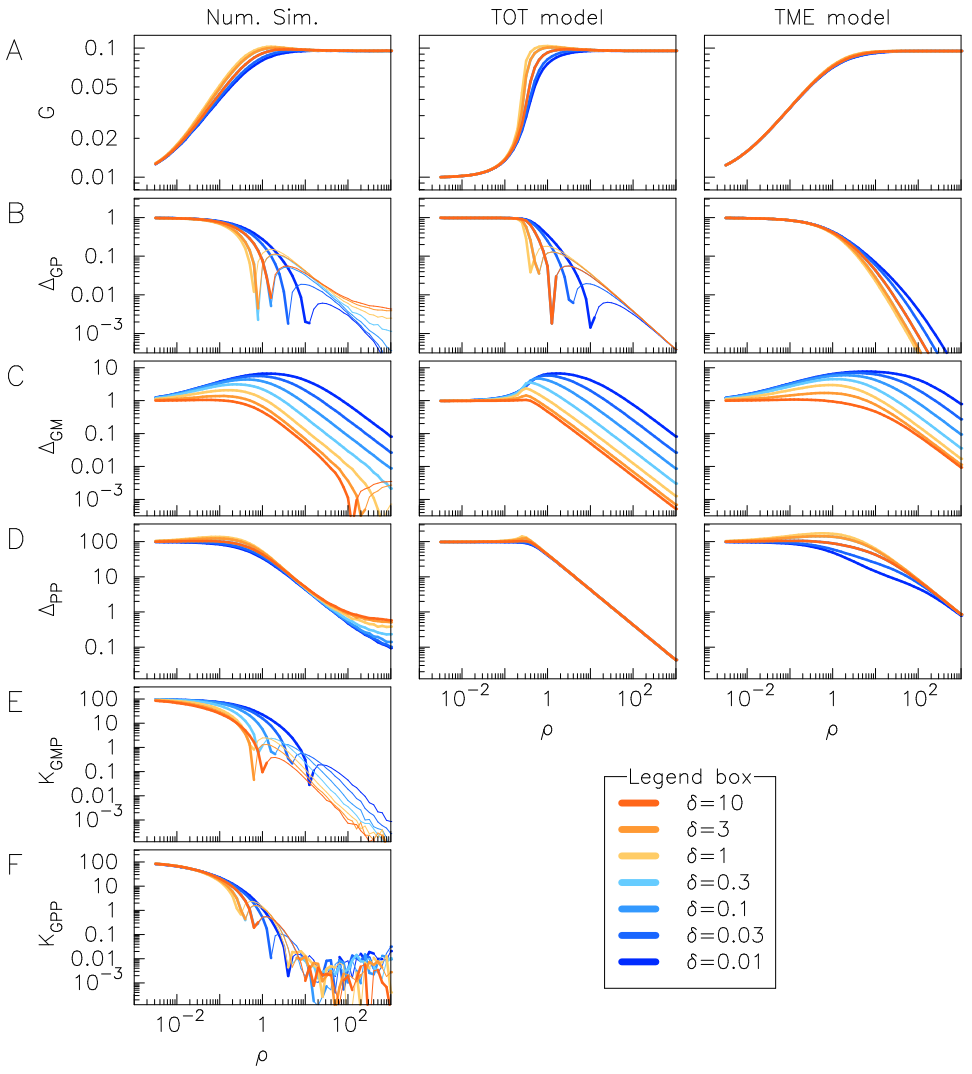

FIG. S1. **Comparison of averages and covariances obtained from stochastic simulations and from the fixed points of ODE models derived using the TOT and TME truncation schemes.** (A) Average gene activity G; (B), (C), (D) covariances Δ_{GP}, Δ_{GM} and Δ_{PP}; (E), (F) third-order cumulants computed 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:

$$U = P = \Lambda G; \quad \Delta_{P,P} = \Delta_{M,P} = \Delta_{U,P} = \Lambda\Delta_{G,U}; \tag{S14a}$$

$$\eta\Delta_{U,U} = (1+\eta)\,\Lambda\Delta_{G,U} - \Lambda\Delta_{G,P}; \tag{S14b}$$

$$\Delta_{G,P} = 1 - G - \Lambda G^2; \tag{S14c}$$

$$(\rho + \rho\Lambda G + \eta)\,\Delta_{G,P} = [\eta - \rho\Lambda G]\,\Delta_{G,U}; \tag{S14d}$$

$$\Delta_{G,P} + \rho\,(1 + 2\Lambda G)\,\Delta_{G,U} = \Lambda G\,(1 - G)\,; \tag{S14e}$$

$$\Delta_{M,M} = \Lambda\ \Delta_{G,M} = \Lambda\frac{(1+\delta)\,\Delta_{G,U} - \Delta_{G,P}}{\delta}. \tag{S14f}$$

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.

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

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

The 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, 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 (S2) 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.

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 $\langle GUP\rangle_{\rho\to\infty} = \langle GP^2\rangle_{\rho\to\infty} =$

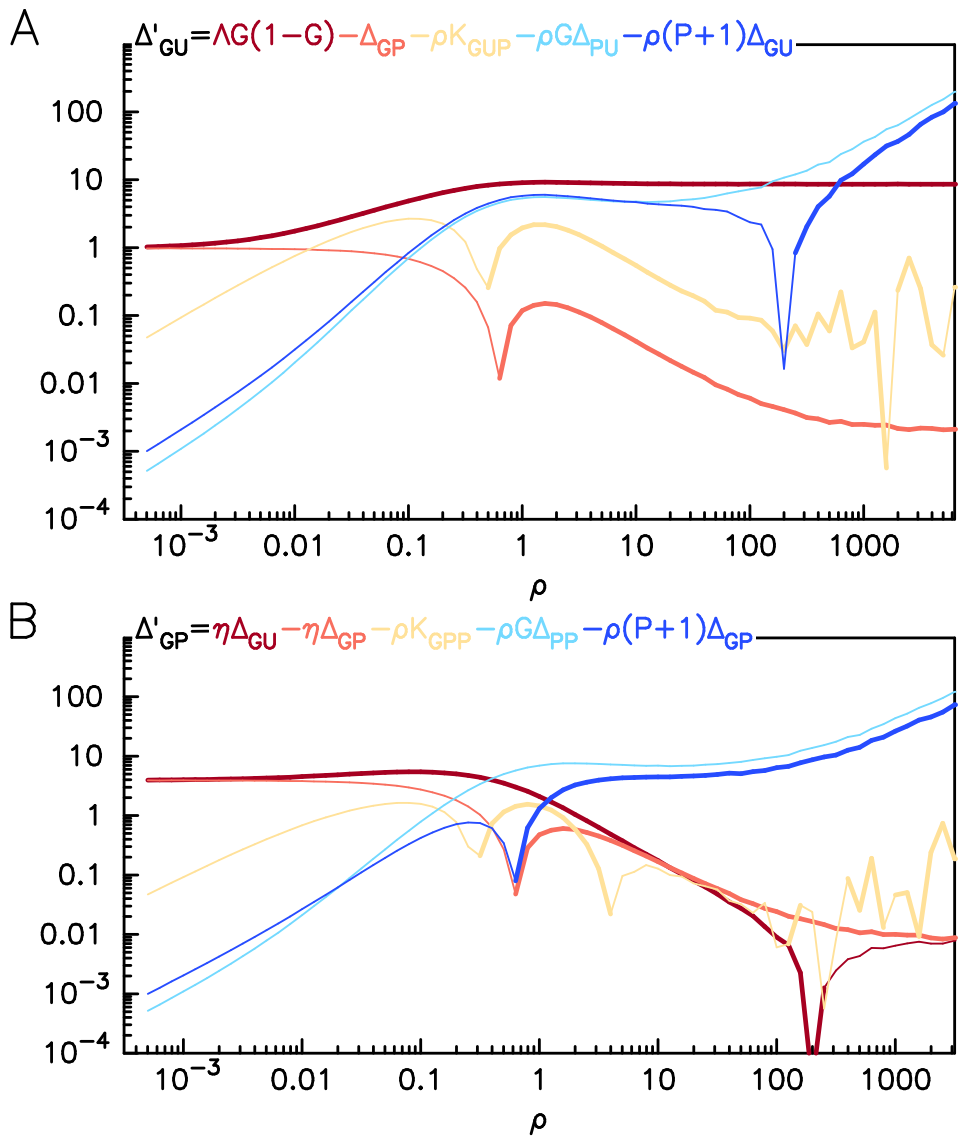

FIG. S2. **Dynamical influence of the third order cumulants** The magnitudes of the different terms appearing in Eqs. (S15) are numerically computed using stochastic simulations for $\Lambda = 100$, $\delta = 1$ (equal degradation rates), and various ρ. The differential equation at the top of each panel indicates the color code. In each panel, thick lines (resp., thin) lines indicate positive (resp., negative) values.

$\Lambda\langle G\rangle^* = \sqrt{\Lambda}$, 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 (S3) 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.

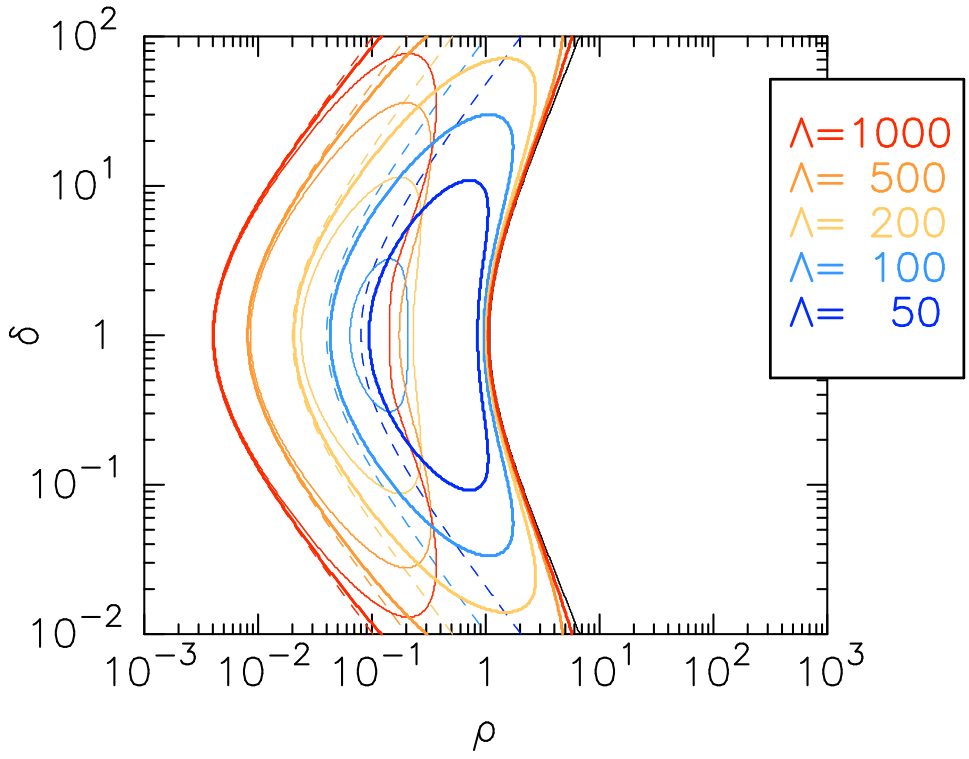

FIG. S3. **Comparison between TOT and TME models oscillatory domain.** The lines enclose the region where the TOT model (thin lines) and the TME model (thick lines) oscillate for various Λ (see legend box for color code).. The analytical expression of the two boundaries derived for $\Lambda \gg 1$ are $\rho = \eta/\Lambda$ (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).

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. (S2)).

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 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. S4).

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.

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

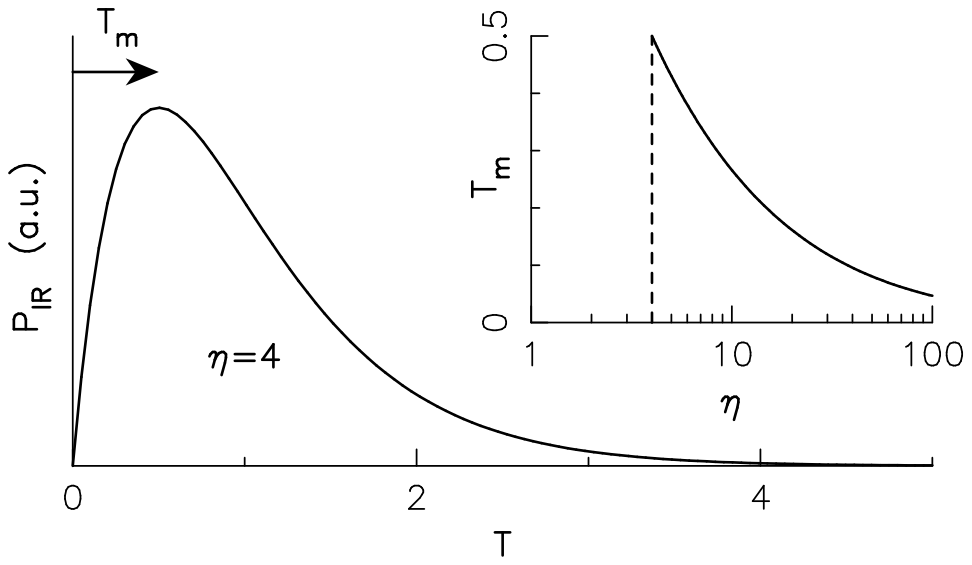

FIG. S4. **Impulse response of the low pass filter.** The protein concentration time profile in response to an impulse gene signal displays a maximum around a time T_m, whose dependence on η is shown in the insert.

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. S3. 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.

---

[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.