Precision of Genetic Oscillators and Clocks

Luis G. Morelli* and Frank Jülicher†

Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, 01187 Dresden, Germany

(Received 29 September 2006; published 31 May 2007)

We develop a stochastic description of feedback oscillators in which functional molecules are produced by an assembly line consisting of many identical steps. The initiation rate of this assembly is regulated by its products via a negative feedback. This model is motivated by genetic oscillators such as circadian clocks. We show that precise oscillations of high quality are possible even when the number of product molecules is low and the fluctuations of amplitude are large. We discuss parameter values which can account for high quality oscillations as observed in single cells. Furthermore, we discuss effects of stochastic amplification steps on precision to account for translational bursting.

DOI: 10.1103/PhysRevLett.98.228101 PACS numbers: 87.10.+e, 05.40.−a, 87.19.Jj

Many organisms possess circadian clocks which permit them to anticipate the daily changes of environmental conditions. Such clocks are implemented on the cellular level by oscillators which generate periodic variations in the levels of certain proteins with a period of approximately 24 hours [1]. Such oscillations can be achieved by the transcriptional regulation of gene expression via a delayed negative feedback [1,2]. Genetic oscillators play an important role in various other cellular processes such as in the segmentation oscillator involved in the patterning of vertebrates [3] and oscillations in the tumor suppressor gene p53, which have been observed in the cellular response to DNA damage [4].

Molecular processes such as the regulation of gene expression, transcription and translation, as well as protein degradation and post-translational modifications are inherently stochastic [5–8]. Both intrinsic and extrinsic noise of gene expression have been identified and quantified [9,10]. Translational bursting has been implicated as an important source of intrinsic fluctuations in protein numbers [11,12]. Fluctuations induce variations of the oscillation period T of a genetic oscillator [13,14]. During subsequent periods these variations accumulate, resulting in a finite time τ during which oscillations are phase coherent. This raises the question of what limits the precision of genetic oscillators in the presence of fluctuations. This precision can be quantified by the quality of oscillations $\mathcal{Q} = \tau/T$, which is equal to $\mathcal{Q} = \pi^{-1} f/\Delta f$, where $f = 1/T$ is frequency and Δf describes the width of the spectral peak [15–17].

Genetic oscillators are often described by the feedback regulation of the transcription of a single or of multiple genes [18]; see Fig. 1(a). Spontaneous oscillations can be obtained using kinetic models which involve mRNA levels and protein numbers [19]. It has been shown that time delays which can be attributed to translation, transcription, intracellular transport, and other post-translational steps, play a key role for the properties of oscillations [1,20–22]. Theoretical approaches to noisy genetic oscillators include methods from stochastic chemical reaction kinetics with Poissonian waiting time distributions [23]. Such descriptions share many features with mesoscopic chemical oscillators [24,25], and lead to high quality of oscillations in the limit of large system size or molecule numbers [23,24]. Descriptions based on feedback between an activator and a repressor have been shown to operate with good quality at low molecule numbers [13,26]. More recently, deterministic time delays have been introduced in stochastic descriptions based on chemical kinetics [27].

In this Letter we introduce a simple but general description of genetic feedback oscillators where the kinetic processes between initiation of gene expression and the completion of the final product introduce stochastic time delays. Our description is motivated by a single gene with negative feedback, see Fig. 1(a). We represent this system by an assembly line which leads to the production of functional proteins via a stochastic, multistep process. Functional products have a finite life time, their number regulates the initiation of assembly, Fig. 1(b). The initiation rate decreases with increasing molecule numbers, introducing a negative feedback. We also consider the case of a stochastic amplification step which could account

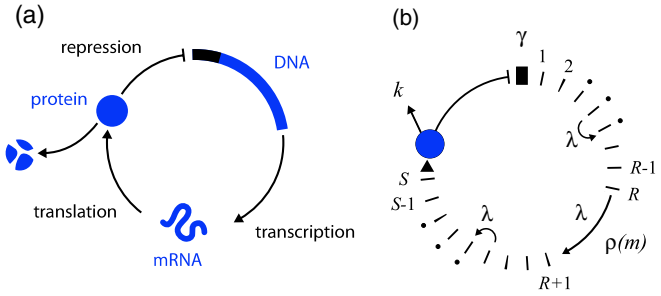

FIG. 1 (color online). (a) Schematic representation of a genetic oscillator based on a single gene and negative feedback. Expression of the gene involves transcription to mRNA and subsequent translation into a functional product (protein), which then represses the gene. (b) Simplified representation by an assembly line with negative feedback. Assembly is initiated at a rate γ which depends on the number N of functional products. It requires S steps at a rate λ until the product is complete. Product molecules decay at a rate k. Our model can include a stochastic amplification step R during assembly at which a single molecule gives rise to m identical copies with probability $\rho(m)$.

for translational bursting where a single mRNA gives rise to several or many proteins.

Assembly line with negative feedback.—We denote the number of products by N, which depends on time t. Assembly of new products occurs via S identical intermediate substeps $\mu = 1, \ldots, S$, Fig. 1(b). The stepping rate of the assembly line is denoted by λ. A stochastic amplification step $\mu = R$ can be included, for which a single molecule is replaced by m identical copies. Here we choose a geometric distribution $\rho(m) = \bar{m}^{-1}(1 - \bar{m}^{-1})^{m-1}$ for the amplification factor m [28,29]. It is characterized by a single parameter, the mean value $\bar{m}$; the variance of m is $\bar{m}(\bar{m} - 1)$. Note that for $\bar{m} = 1$ no amplification occurs. Assembly is initiated with a probability per unit time $\gamma = \alpha g(N)$, which we assume to depend only on the number N of products. Here α is the initiation rate for $N = 0$, and $g(N)$ is a monotonously decreasing function of N, with $g(0) = 1$. We choose $g(N) = 1/[1 + (N/N_0)^h]$, where N_0 is the number of products for which the rate of initiation of assembly is repressed by a factor of 2, and h is a Hill coefficient which characterizes the cooperativity of the feedback. Finally, we assume that the number N of functional products decays at a rate k.

This assembly line with negative feedback and stochastic amplification can be described as a Markov chain. The state of the system at any given time t is unambiguously specified by the number of products $N(t)$ and the numbers $M_\mu(t)$ of unfinished products that are at the intermediate stage $\mu = 0, \ldots, S$ in the assembly line. The stochastic process can be described by the probability distribution $\mathcal{P}(N, \mathbf{M}, t)$ to find the system in state $(N, \mathbf{M})$, as well as the conditional probability distribution $\mathcal{P}(N, \mathbf{M}, t|N', \mathbf{M}', t')$ to find the system in this state at time t under the conditions that it was in state $(N', \mathbf{M}')$ at earlier time t'. Here we have defined $\mathbf{M} = \{M_0, \ldots, M_S\}$. Both distributions obey the master equation

$$\begin{aligned}\partial_t \mathcal{P}(N, \mathbf{M}, t) = \{&\alpha g(N)(b_0^- - 1) \\ &+ \lambda \sum_{\mu=0, \mu\neq R}^{S-1} [(M_\mu + 1) b_\mu^+ b_{\mu+1}^- - M_\mu] \\ &+ \lambda \sum_m \rho(m)[(M_R + 1) b_R^+ b_{R+1}^{-m} - M_R] \\ &+ \lambda[(M_S + 1) a^- b_S^+ - M_S] \\ &+ k[(N + 1) a^+ - N]\}\mathcal{P}(N, \mathbf{M}, t), \end{aligned} \quad (1)$$

where we have introduced the step operators $a^\pm$ and $b_\mu^\pm$, which act according to $a^\pm \mathcal{P}(N, \mathbf{M}, t) = \mathcal{P}(N \pm 1, \mathbf{M}, t)$ and $b_\mu^\pm \mathcal{P}(N, \mathbf{M}, t) = \mathcal{P}(N, \ldots, M_\mu \pm 1, \ldots, t)$. The conditional distribution satisfies the initial condition $\mathcal{P}(N, \mathbf{M}, t|N', \mathbf{M}', t) = \delta_{N,N'}\delta_{\mathbf{M},\mathbf{M}'}$. The first line in Eq. (1) accounts for the initiation of assembly at rate $\alpha g(N)$. The second to fourth lines account for the sequence of steps in the assembly line including the amplification and final step. The last line describes degradation of products. Note that the generalization to situations where substeps are not identical is straightforward.

Stochastic simulations.—We have performed numerical simulations of the system described by Eq. (1) using a standard Gillespie algorithm [30]. Since all events obey Poissonian statistics the Gillespie algorithm generates optimal time steps. Two examples of simulation results in the absence of an amplification step ($\bar{m} = 1$) are displayed in Fig. 2 for $Sk/\lambda = 11$ and $S = 440$. The period of oscillations in this case is $T \simeq 24$ h for $k \simeq 1$ h^{-1}. The top row corresponds to $\alpha/k = 20$, which leads to noisy oscillations of low quality $\mathcal{Q} \simeq 3$, Figs. 2(a) and 2(e), and a monomodal distribution $P_s(N)$ of product numbers in the steady state, Fig. 2(c). The coherence time τ can be determined from the two-point cumulant function $C(t) = \langle N(t)N(0)\rangle - \bar{N}^2$, where $\bar{N} = \langle N\rangle$, which decays for large t as $C(t) \sim e^{-t/\tau}\cos(\omega t)$, where $\omega = 2\pi/T$. The bottom row in Fig. 2 shows the corresponding simulation results for $\alpha/k = 60$. In this case, oscillations are of high quality with $\mathcal{Q} \simeq 160$. Note that the distribution $P_s(N)$ is now bimodal. Strong amplitude fluctuations remain, and the average number of products $\bar{N} \simeq 25$ is still small. The period of oscillations T and the corresponding quality values $\mathcal{Q}$ are displayed in Fig. 3 as a function of the step number S for different values of α/k. Note, that while T is approximately independent of α/k, the quality depends strongly on this ratio. For small $\alpha/k < 25$, $\mathcal{Q}$ saturates for large S at finite values. For large α/k, such a saturation cannot be seen in the data and the quality becomes very large for large S. The effects of the amplification step are shown in Fig. 4. Increasing the average burst size $\bar{m}$, the quality exhibits a maximum at an optimal value of $\bar{m}$. As

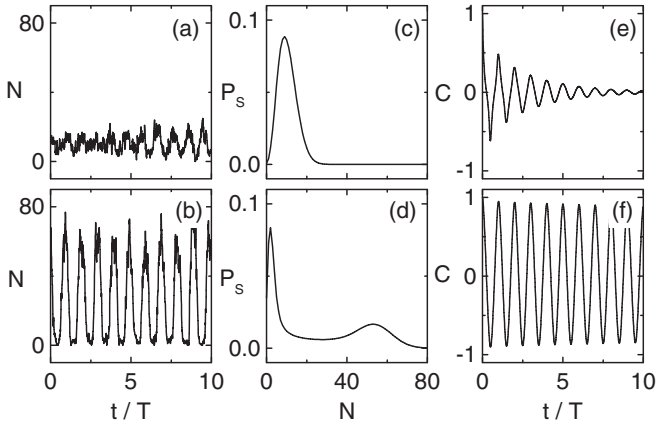

FIG. 2. Numeric simulations of the stochastic feedback oscillator for $T \simeq 24$ h and no amplification step. Low quality (top row, $\alpha/k = 20$, $\mathcal{Q} \simeq 3$) and high quality (bottom row, $\alpha/k = 60$, $\mathcal{Q} \simeq 160$) oscillations are shown. (a),(b) Number $N(t)$ of products as a function of time t, normalized to the average period T of oscillations. (c),(d) Histogram of product numbers N. The probability $P_s(N)$ to find N products at any given time during a long run is shown. (e),(f) Two-point cumulant function $C(t) = \langle N(t)N(t+\tau)\rangle - \langle N\rangle^2$ of the product number. Parameters are $k = 1$ h^{-1}, $\lambda/k = 40$, $S = 440$, $\bar{m} = 1$, $N_0 = 10$, and $h = 2$.

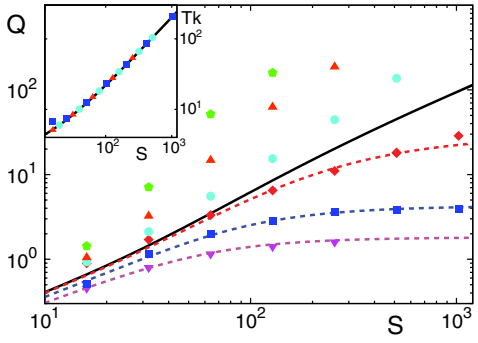

FIG. 3 (color online). The quality of the oscillators vs number of steps S. Dots stand for stochastic numerical simulations, and dotted lines are fits using the theory (numerical solution of Eq. (4)), with effective parameters $\bar{g}$ as indicated below. The solid line corresponds to $\bar{g}\alpha/k = 1$ in Eq. (4). When not shown, error bars are smaller than the dots. Inset: Period of the oscillators vs number of steps S in the assembly line, for several values of α/k. Dots correspond to numerical simulations, and the black line to Eq. (5). Symbols: down triangles $\alpha/k = 15$ ($\bar{g} = 0.0507$), squares $\alpha/k = 20$ ($\bar{g} = 0.0445$), diamonds $\alpha/k = 25$ ($\bar{g} = 0.0393$), circles $\alpha/k = 30$, up-triangles $\alpha/k = 40$, pentagons $\alpha/k = 60$. Other parameters: $k = 1\ \mathrm{h}^{-1}$, $\lambda/k = 10$, $\bar{m} = 1$, $N_0 = 10$ and $h = 2$.

Fig. 4(a) reveals, the value of $\bar{m}$ for which $\mathcal{Q}$ is maximal increases if the amplification step occurs earlier in the assembly line. If the initiation rate is increased, quality increases, but amplification noise is also enhanced such that high quality occurs for smaller $\bar{m}$, see Fig. 4(b).

Averages and correlation functions.—Using the master Eq. (1), simplified expressions for the time dependent averages and correlation functions of the product number N can be obtained. The average number of products $\langle N(t)\rangle$ at time t, and the average number $\langle M_\mu(t)\rangle$ of unfinished products at intermediate stage μ, where the brackets $\langle\ldots\rangle$ denote ensemble averages, can be expressed in terms of the distribution $P(N, \mathbf{M}, t)$. Similarly, the two-point functions $\langle N(t)N(t')\rangle$ and $\langle M_\mu(t)N(t')\rangle$ can be expressed using the

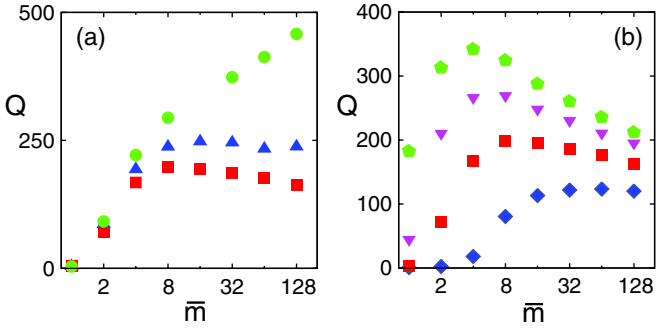

FIG. 4 (color online). Probing the effects of an amplification step. Quality factor as a function of burst size $\bar{m}$. (a) Symbols are: $R = S$ (squares), $R = S/2$ (triangles), $R = 1$ (circles), for $\alpha/k = 20$. (b) Symbols are: $\alpha/k = 10$ (diamonds), $\alpha/k = 20$ (squares), $\alpha/k = 30$ (down triangles), $\alpha/k = 40$ (pentagons), for $R = S$. Other parameters are $k = 1\ \mathrm{h}^{-1}$, $\lambda/k = 10$, $S = 256$, $N_0 = 10$ and $h = 2$.

joint probability distribution $\mathcal{P}(N, \mathbf{M}, t; N', \mathbf{M}', t') = \mathcal{P}(N, \mathbf{M}, t|N', \mathbf{M}', t')P(N', \mathbf{M}', t')$. The master Eq. (1) then defines dynamic relations between these functions. From these expressions for $\langle N(t)\rangle$ and $\langle N(t)N(t')\rangle$ all dependencies on the numbers M_μ of unfinished products can be eliminated because those variables effectively generate a distribution $G(t) = \lambda^{S+1}t^S e^{-\lambda t}/S!$ of delay times t between the start of assembly and the completion of the products. The average number of products satisfies for long times

$$\frac{d}{dt}\langle N(t)\rangle \simeq -k\langle N(t)\rangle + \bar{m}\alpha\int_0^t \langle g(N(s))\rangle G(t-s)ds. \quad (2)$$

On the right-hand side of Eq. (2), we have neglected terms which are proportional to $t^\mu e^{-\lambda t}$ for $\mu = 1, \ldots, S$. These terms vanish for times long compared to λ^{-1} and do not contribute to the long-time limit in which we are interested here. Note that the distribution $\rho(m)$ of the stochastic amplification enters explicitly only via its average $\bar{m}$. According to Eq. (2), the average $\langle N(t)\rangle$ relaxes after a relaxation time to a constant steady state value $\bar{N} = \langle N\rangle$ with $d\langle N\rangle/dt = 0$ and constant $\langle g(N)\rangle = \bar{N}k/\alpha\bar{m}$. By the same method, and again neglecting terms proportional to $t^\mu e^{-\lambda t}$, we obtain an equation for the two-point cumulant $C(t) = \langle N(t)N(0)\rangle - \bar{N}^2$ valid for long times, which reads

$$\frac{d}{dt}C(t) \simeq -kC(t) + \bar{m}\alpha\int_0^t C_g(s)G(t-s)ds. \quad (3)$$

Here $C_g(t) = \langle g(N(t))N(0)\rangle - \langle g(N)\rangle\bar{N}$, where averages are taken in the steady state. Interestingly, only the average $\bar{m}$ of the distribution $\rho(m)$ enters explicitly in Eq. (3). Note however that the variance $\langle N^2(t)\rangle$ as well as correlation functions of higher order depend on the variance and higher order moments of the distribution $\rho(m)$.

When one expresses the function $C_g(t)$ in terms of averages and correlation functions of N, the nonlinearities of $g(N)$ become important. We can expand $g(N)$ in a series around the average $\bar{N}$ as $g(N) = g_0 + \sum g_n(N-\bar{N})^n/n!$, where $g_0 = g(\bar{N})$, $g_n = d^n g/dN^n|_{N=\bar{N}}$. This expansion reveals that $C_g(t)$ depends on cumulants of all orders and thus also on the shape of the distribution $\rho(m)$ of the amplification step. We can obtain an approximative expression for $C_g(t)$ by neglecting all contributions of order higher than three. This approximation is valid in situations where the distribution $P_s(N)$ is monomodal, close to a Gaussian and sharply peaked, which typically occurs for small quality of oscillations. This simple approximation leads to $C_g(t) \approx -\bar{g}C(t)$, where we have introduced the effective feedback strength $\bar{g} \approx -g_1 - g_3C(0)/2$. The effective feedback strength $\bar{g}$ is not a fixed parameter, but depends itself on the variance $C(0) = \langle(N-\bar{N})^2\rangle$ of product numbers. Note, that $C(0)$ and $\bar{g}$ depend on the variance of the distribution $\rho(m)$.

Period and quality of oscillations.—In the limit of large step numbers S, the delay distribution approaches a Gaussian $G(t) \simeq \exp[-(t-t_0)^2/2\sigma^2]/(2\pi\sigma^2)^{1/2}$, where

the mean time delay is $t_0 = S/\lambda$ and the variance is $\sigma^2 = S/\lambda^2$. We obtain from Eq. (3) a simple characteristic equation for the complex number z which describes the long-time decay of the cumulant function $C(t) \simeq C_0 e^{-zt}$. For $t_0 \gg \sigma$ we find

$$z - k \simeq \alpha \bar{m}\, \bar{g}\, e^{zt_0 + z^2\sigma^2/2}. \quad (4)$$

For given $\bar{g}$ and set of parameters, the complex zeros $z = \tau^{-1} + i\omega$ of this equation determine the period $T = 2\pi/\omega$ and quality $\mathcal{Q} = \tau/T$ of noisy oscillations. Oscillations of high quality are typically found for $k^{-1} \ll t_0$. Using this separation of time scales, we find a simple expression for the oscillation period to first order in $(kt_0)^{-1}$

$$T \approx 2t_0(1 + 1/kt_0). \quad (5)$$

Note that T depends neither on α/k nor on the effective feedback $\bar{g}$. In the inset of Fig. 3, the period of oscillations is displayed for different values of α/k as a function of $S = t_0^2/\sigma^2$ with $\lambda = t_0/\sigma^2$ kept constant. The solid line corresponds to Eq. (5), which is in excellent agreement with the simulation results.

A simple expression for the quality $\mathcal{Q}$ can be found using Eq. (4) for $t_0 \gg \sigma$. Neglecting $(kt_0)^{-1}$ it is given by

$$\mathcal{Q} \approx [\pi^2\sigma^2/t_0^2 - 2\ln(\alpha\bar{m}\,\bar{g}\,/k)]^{-1}. \quad (6)$$

According to Eq. (6) the quality depends on the ratio α/k. This is indeed the case as discussed above; see Fig. 3. Furthermore, $\mathcal{Q}$ depends on the burst size $\bar{m}$ and the effective feedback strength $\bar{g}$, which is a function of all parameters.

Discussion.—We can choose parameter values which could correspond to oscillations with a period of $T \simeq 24$ h using, e.g., $t_0 k = 11$ and $k = 1\ \mathrm{h}^{-1}$; see Fig. 2. Our simulations show that high quality oscillations with $\mathcal{Q} \simeq 160$ are generated by negative feedback control with $\alpha/k = 60$ using an assembly line of 440 steps without the stochastic amplification step. This leads to precise oscillations with small phase fluctuations even though there are on average $N_0 \simeq 25$ product molecules and high amplitude noise, see Fig. 2(b). While the required delay time of about 11 hours seems long compared to typical times of protein synthesis, which are of the order of tens of minutes, it has been suggested that in the circadian clock of Drosophila longer delays of about 10 hours result largely from post-translational events [1].

Stochastic amplification can enhance quality. For example, using $\alpha/k = 20$ and an amplification of $\bar{m} = 10$ the quality exceeds $\mathcal{Q} \simeq 200$ for only $S = 256$ steps. While amplification enhances the feedback strength, it also involves additional burst noise and noise amplification. As a consequence the quality can decrease for large amplification factors; see Fig. 4. Interestingly, if the amplification occurs for $R = 1$, the quality becomes large and we do not see a decrease for large $\bar{m}$. The mechanism presented here, which permits high precision of oscillations for small molecule numbers comes at a cost, namely, the relatively high turnover of the products. During one cycle, the total number of molecules produced and degraded is of the order of $\bar{m}\alpha t_0$. Cells might have additional mechanisms to reduce this cost, such as interlinked positive and negative feedback loops, post-translational modifications, and complex degradation pathways [1,13,31].

We thank A. Oates for introducing us to genetic oscillations, and S. Ares, N. Becker, T. Bollenbach, L. Herrgen, K. Kruse, F. Peruani, C. Schröter, and C. Svaneborg for helpful comments and fruitful discussions.

*Electronic address: morelli@mpipks-dresden.mpg.de
†Electronic address: julicher@mpipks-dresden.mpg.de

[1] M. W. Young and S. A. Kay, Nat. Rev. Genet. **2**, 702 (2001).
[2] M. B. Elowitz and S. Leibler, Nature (London) **403**, 335 (2000).
[3] O. Pourquié, Science **301**, 328 (2003).
[4] G. Lahav *et al.*, Nat. Genet. **36**, 147 (2004).
[5] O. G. Berg, J. Theor. Biol. **71**, 587 (1978).
[6] H. H. McAdams and A. Arkin, Proc. Natl. Acad. Sci. U.S.A. **94**, 814 (1997).
[7] M. Thattai and A. van Oudenaarden, Proc. Natl. Acad. Sci. U.S.A. **98**, 8614 (2001).
[8] A. Becskei and L. Serrano, Nature (London) **405**, 590 (2000).
[9] M. B. Elowitz *et al.*, Science **297**, 1183 (2002).
[10] J. Paulsson, Nature (London) **427**, 415 (2004).
[11] E. M. Ozbudak *et al.*, Nat. Genet. **31**, 69 (2002).
[12] M. Kærn *et al.*, Nat. Rev. Genet. **6**, 451 (2005).
[13] N. Barkai and S. Leibler, Nature (London) **403**, 267 (2000).
[14] I. Mihalcescu, W. Hsing, and S. Leibler, Nature (London) **430**, 81 (2004).
[15] R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, New York, 1967), Vol. 2.
[16] A. S. Pikovsky and J. Kurths, Phys. Rev. Lett. **78**, 775 (1997).
[17] B. Lindner *et al.*, Phys. Rep. **392**, 321 (2004).
[18] B. C. Goodwin, *Temporal Organization in Cells* (Academic, New York, 1963).
[19] A. Goldbeter, Nature (London) **420**, 238 (2002).
[20] M. A. Lema, D. A. Golombek, and J. Echave, J. Theor. Biol. **204**, 565 (2000).
[21] J. Lewis, Curr. Biol. **13**, 1398 (2003).
[22] N. MacDonald, *Biological Delay Systems: Linear Stability Theory* (Cambridge University Press, Cambridge, England, 1989).
[23] D. Gonze and A. Goldbeter, Chaos **16**, 026110 (2006).
[24] P. Gaspard, J. Chem. Phys. **117**, 8905 (2002).
[25] H. Qian, J. Phys. Chem. B **110**, 15 063 (2006).
[26] J. M. G. Vilar *et al.*, Proc. Natl. Acad. Sci. U.S.A. **99**, 5988 (2002).
[27] D. Bratsun, D. Volfson, L. S. Tsimring, and J. Hasty, Proc. Natl. Acad. Sci. U.S.A. **102**, 14 593 (2005).
[28] J. Yu *et al.*, Science **311**, 1600 (2006).
[29] I. Golding *et al.*, Cell **123**, 1025 (2005).
[30] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[31] J. Tomita *et al.*, Science **307**, 251 (2005).