Entrainment of a cellular circadian oscillator by light in the presence of molecular noise

Guanyu Wang* and Charles S. Peskin
Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, USA

(Received 4 July 2017; revised manuscript received 24 February 2018; published 25 June 2018)

In this paper, we consider a stochastic molecular circadian oscillator described by a sequence of biological reactions and its deterministic kinetics governed by a system of ordinary differential equations in the limit of large numbers of molecules. The oscillations in the model are generated by negative feedback regulation of a gene. The focus of this paper is the entrainment of the oscillator by a periodic light signal that affects the maximal transcription rate of the gene. We introduce two scalings of the model parameters that provide independent control over the natural frequency of the oscillator and the relative noise level. We study entrainment in two ways: by visualizing the stochastic limit cycle in various projections of the discrete phase space of the system and by evaluating the maximum of the normalized cross correlation of the light signal with the number of protein molecules in the cell. The visualization method ignores the phase of the oscillator, and we find in this way that entrainment has a subtle organizing effect on the limit cycle as a whole. The cross correlation results reveal an interval of natural frequencies of the oscillator surrounding the frequency of the light signal within which maximal entrainment occurs with rather sharp drops in entrainment at the edges of this interval. The width of the interval of maximal entrainment increases with the amplitude of the light signal. These statements are applicable both to the stochastic oscillator and to its deterministic limit, but the results are most clear-cut in the deterministic case and degrade from there as the relative noise level increases.

DOI: 10.1103/PhysRevE.97.062416

I. INTRODUCTION

This paper continues our study of entrainment of stochastic oscillators. In a previous paper [1], we considered the stochastic analog of a phase oscillator, i.e., a system that performs a random walk around a circle of states, and we studied the entrainment of the oscillator by a stochastic phase resetting signal. Here we consider a schematic model of a cellular circadian oscillator entrained by light. The model is fully stochastic so that the integer numbers of molecules of all molecular species are tracked, and reactions occur at random times with specified probabilities per unit time. For comparison, we also consider the deterministic limit in which the system is described by ordinary differential equations involving the concentrations of the different molecular species.

The model considered here involves one gene and four molecular species: the messenger RNA (mRNA) and the protein encoded by the gene, which are separately tracked in the nucleus and in the cytoplasm of the cell. The protein product inhibits the transcription of the gene that encodes it thus providing negative feedback that may lead to oscillations if the parameters of the system are properly chosen. We consider two different feedback mechanisms in this paper. In the first one, the protein is itself an inhibitory transcription factor [2–8], and in the second mechanism, the protein binds to an activating transcription factor thus sequestering the activator and thereby reducing the rate of transcription [8–14]. In both cases, we assume that the effect of light is to modulate the maximal rate of transcription.

The outline of this paper is as follows. First, we use the deterministic version of the model to find parameters that yield spontaneous limit-cycle oscillation. Next, we introduce two different kinds of scaling: One that allows us to tune the natural frequency of oscillation, and another that controls the level of molecular noise. Finally, we study entrainment by different depths of modulation of the ambient light while holding the mean light level constant throughout our studies. We find that the model can be entrained by a given periodic light signal over a range of natural frequencies of the oscillator that includes natural frequencies both above and below those of the entraining light signal. The width of the interval of entrainment grows with the amplitude of the light signal and is surprisingly well defined especially in the deterministic limit.

The models used in this paper are very much simplified in comparison to other cellular circadian clock models [6,7,9,15–24], but we have attempted here to capture the essential features of such models in their simplest form so that the interaction of intrinsic molecular noise with the phenomenon of entrainment can be elucidated.

II. CELLULAR CIRCADIAN OSCILLATOR

A. Model

We study a four-variable schematic cellular model, which includes the negative feedback regulation that is characteristic of the mammalian circadian clock. The circadian clock in mammals resembles that in *Drosophila* and in *Neurospora*. The clock gene period (*Per*) is a sequence homolog of the clock gene *per* in *Drosophila*, but its response to light is more similar to the clock gene *frq* in *Neurospora* [25–31]. The *Per*

*Corresponding author: guanyu@cims.nyu.edu

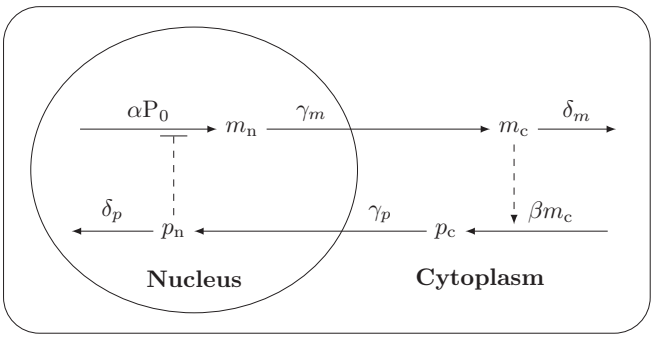

FIG. 1. A schematic cellular circadian clock.

gene is transcribed into *Per* mRNAs in the nucleus. The *Per* mRNAs are exported from the nucleus to the cytoplasm to be translated into PER protein and to degrade. The PER proteins then enter the nucleus where they inhibit the transcription of the *Per* gene and degrade in proteasomes in the nucleus.

The model is schematized in Fig. 1, and the reactions are defined in Table I, where m represents the number of *Per* mRNA molecules and p represents the number of PER protein molecules, the subscripts n and c are short for nucleus and cytoplasm, respectively, and P_0 is the probability that transcription happens. The influence of light will be to modulate the maximal transcription rate α as discussed later.

A key assumption of the model is that the PER protein inhibits the transcription of the *Per* gene [31–35]. Here, we describe a model in which this inhibition is direct, and later, in Sec. IV, we will consider instead an indirect inhibitory mechanism known as sequestration. We assume that there are r protein-binding sites at which this inhibition occurs and that occupation of any one of them is sufficient to block transcription. We further assume that the binding and unbinding reactions are fast [19] so that it is sufficient even in a stochastic model to consider only the probability that a protein-binding site is unoccupied. Let p_n be the total number of PER protein molecules in the nucleus at some particular time, counting both those that are free and those that are bound to one of the r protein-binding sites. Let ξ be the rate that any one protein molecule binds to any one site, let η be the corresponding unbinding rate, and let L be the number of occupied sites $L \leqslant p_n$ and $L \leqslant r$. The probability that l sites are occupied is denoted by $\Pr(L=l)$. We assume that the net rate of transition from $L=l-1$ to $L=l$ balances the net rate of transition from $L=l$ to $L=l-1$. Therefore,

$$\xi[p_n-(l-1)][r-(l-1)]\Pr(L=l-1)=\eta l\Pr(L=l) \tag{1}$$

$$\Rightarrow \quad \frac{\Pr(L=l)}{\Pr(L=l-1)}=\frac{\xi[p_n-(l-1)][r-(l-1)]}{\eta l}. \tag{2}$$

Iterating this, we find

$$\frac{\Pr(L=l)}{\Pr(L=0)}=\left(\frac{\xi}{\eta}\right)^l\frac{p_n!}{(p_n-l)!}\frac{r!}{(r-l)!}\frac{1}{l!}, \tag{3}$$

from which it follows that

$$1=\Pr(L=0)\sum_{l=0}^{\min(p_n,r)}\left(\frac{\xi}{\eta}\right)^l\frac{p_n!}{(p_n-l)!}\frac{r!}{(r-l)!}\frac{1}{l!} \tag{4}$$

$$\Rightarrow P_0=\Pr(L=0)=\frac{1}{\sum_{l=0}^{\min(p_n,r)}\left(\frac{\xi}{\eta}\right)^l\frac{p_n!}{(p_n-l)!}\frac{r!}{(r-l)!}\frac{1}{l!}}=\frac{1}{\sum_{l=0}^{\min(p_n,r)}\left(\frac{1}{KV_n}\right)^l\frac{p_n!}{(p_n-l)!}\frac{r!}{(r-l)!}\frac{1}{l!}}, \tag{5}$$

in which K with units of concentration is the equilibrium constant of the unbinding and binding reactions. Equation (5) may also be derived by noting that

$$\frac{p_n!}{(p_n-l)!}\frac{r!}{(r-l)!}\frac{1}{l!}=\binom{p_n}{l}\binom{r}{l}l!, \tag{6}$$

which is the number of distinct ways of choosing l of the p_n protein molecules and l of the r sites and then making a 1-1 correspondence between the chosen molecules and the chosen sites.

Equation (5) is exact in the limit that $\xi\to\infty$ and $\eta\to\infty$ with $\frac{\xi}{\eta}$ held constant. Thus, we are assuming that the binding and unbinding reactions are in rapid equilibrium, but we are still taking into account the discreteness of the individual molecules and binding sites.

Although we use Eq. (5) in our stochastic simulations, we also need its macroscopic limit in order to formulate the corresponding system of ordinary differential equations. If $p_n \gg r$, then $\frac{p_n!}{(p_n-l)!}\sim p_n^l$, and $\min(p_n,r)=r$. Therefore,

$$\Pr(L=0)\sim\frac{1}{\sum_{l=0}^{r}\binom{r}{l}\left(\frac{p_n}{KV_n}\right)^l} \tag{7}$$

$$=\frac{1}{\left(1+\frac{p_n}{KV_n}\right)^r} \tag{8}$$

$$=\left(\frac{K}{K+\frac{p_n}{V_n}}\right)^r. \tag{9}$$

TABLE I. Reaction table.

Reaction number	Reaction name	Rate (probability per unit time)	Result
1	Transcription of the *Per* gene	αP_0	$m_n \to m_n+1$
2	Export of mRNA from the nucleus	$\gamma_m m_n$	$m_n \to m_n-1$, $m_c \to m_c+1$
3	Degradation of mRNA (in the cytoplasm)	$\delta_m m_c$	$m_c \to m_c-1$
4	Translation of *Per* mRNA	βm_c	$p_c \to p_c+1$
5	Import of protein to the nucleus	$\gamma_p p_c$	$p_c \to p_c-1$, $p_n \to p_n+1$
6	Degradation of protein (in the nucleus)	$\delta_p p_n$	$p_n \to p_n-1$

Thus in the limit of large numbers of molecules, P_0 is given by $(\frac{K}{K+\frac{P_n}{V_n}})^r$, and the dynamics of the model are approximately governed by the following kinetic equations:

$$\frac{dM_n}{dt} = \frac{\alpha}{V_n}\left(\frac{K}{K+P_n}\right)^r - \gamma_m M_n, \tag{10}$$

$$\frac{dM_c}{dt} = \gamma_m\left(\frac{V_n}{V_c}\right)M_n - \delta_m M_c, \tag{11}$$

$$\frac{dP_c}{dt} = \beta M_c - \gamma_p P_c, \tag{12}$$

$$\frac{dP_n}{dt} = \gamma_p\left(\frac{V_c}{V_n}\right)P_c - \delta_p P_n, \tag{13}$$

in which the capitalized variables are concentrations of molecules, V_n is the volume of the nucleus, and V_c is the cytoplasmic volume. Note that Eqs. (10)–(13) are a special case of the Goodwin oscillator [36,37].

B. Stability analysis

To determine the range of parameter values in which the system (10)–(13) can produce sustained periodic oscillations, we perform a stability analysis of the system.

The steady-state equations of Eqs. (10)–(13) are

$$\frac{\alpha}{V_n}\left(\frac{K}{K+P_n^0}\right)^r = \gamma_m M_n^0, \tag{14}$$

$$\gamma_m\left(\frac{V_n}{V_c}\right)M_n^0 = \delta_m M_c^0, \tag{15}$$

$$\beta M_c^0 = \gamma_p P_c^0, \tag{16}$$

$$\gamma_p\left(\frac{V_c}{V_n}\right)P_c^0 = \delta_p P_n^0, \tag{17}$$

where M_n^0, M_c^0, P_c^0, and P_n^0 are steady-state values of the corresponding concentrations. Multiplying Eqs. (14)–(17) together gives

$$\frac{\alpha}{V_n}\beta\left(\frac{K}{K+P_n^0}\right)^r = \delta_m\delta_p P_n^0. \tag{18}$$

The graphical solution of Eq. (18) verifies that there is a unique positive steady state. Linearization around the steady state gives us the following equations:

$$\frac{d\widetilde{M}_n}{dt} = -a\widetilde{P}_n - \gamma_m\widetilde{M}_n, \tag{19}$$

$$\frac{d\widetilde{M}_c}{dt} = \gamma_m\left(\frac{V_n}{V_c}\right)\widetilde{M}_n - \delta_m\widetilde{M}_c, \tag{20}$$

$$\frac{d\widetilde{P}_c}{dt} = \beta\widetilde{M}_c - \gamma_p\widetilde{P}_c, \tag{21}$$

$$\frac{d\widetilde{P}_n}{dt} = \gamma_p\left(\frac{V_c}{V_n}\right)\widetilde{P}_c - \delta_p\widetilde{P}_n, \tag{22}$$

where $a = r\frac{\delta_m\delta_p}{\beta}\frac{P_n^0}{K+P_n^0}$ and the variables with tildes are the deviations from the steady-state values.

Let $\tilde{x} = \begin{pmatrix}\widetilde{M}_n\\ \widetilde{M}_c\\ \widetilde{P}_c\\ \widetilde{P}_n\end{pmatrix}$, $A = \begin{pmatrix}-\gamma_m & 0 & 0 & -a\\ \gamma_m\left(\frac{V_n}{V_c}\right) & -\delta_m & 0 & 0\\ 0 & \beta & -\gamma_p & 0\\ 0 & 0 & \gamma_p\left(\frac{V_c}{V_n}\right) & -\delta_p\end{pmatrix}$, then

$$\frac{d\tilde{x}}{dt} = A\tilde{x}. \tag{23}$$

The eigenvalues of A satisfy

$$0 = \det(\lambda I - A) \tag{24}$$

$$= \det\begin{pmatrix}\lambda+\gamma_m & 0 & 0 & a\\ -\gamma_m\left(\frac{V_n}{V_c}\right) & \lambda+\delta_m & 0 & 0\\ 0 & -\beta & \lambda+\gamma_p & 0\\ 0 & 0 & -\gamma_p\left(\frac{V_c}{V_n}\right) & \lambda+\delta_p\end{pmatrix} \tag{25}$$

$$= (\lambda+\gamma_m)(\lambda+\delta_m)(\lambda+\gamma_p)(\lambda+\delta_p) + a\beta\gamma_m\gamma_p \tag{26}$$

$$= (\lambda+\gamma_m)(\lambda+\delta_m)(\lambda+\gamma_p)(\lambda+\delta_p) + G\gamma_m\delta_m\gamma_p\delta_p \tag{27}$$

where $G = (\frac{P_n^0}{K+P_n^0})r$. Now consider the special case[1] in which $\gamma_m = \gamma_p = \delta_m = \delta_p = \nu$, then Eq. (27) becomes

$$0 = (\lambda+\nu)^4 + G\nu^4, \tag{28}$$

$$\lambda + \nu = (-1)^{1/4}G^{1/4}\nu, \tag{29}$$

$$\lambda = \nu[-1 + (-1)^{1/4}G^{1/4}]. \tag{30}$$

There will be a root with $\mathrm{Re}(\lambda) > 0$ if and only if

$$G > 4. \tag{31}$$

Since r is an integer, to get instability we need

$$r \geqslant 5. \tag{32}$$

This power is larger than the one found in the secant condition of the Goodwin oscillator with four variables [11,38]. The reason is that we have a different nonlinearity in Eq. (10). With $r = 5$, Eq. (31) requires

$$P_n^0 > 4K. \tag{33}$$

To achieve this, by Eq. (18), we also need

$$\frac{\alpha}{V_n}\frac{\beta}{\delta_m\delta_p} > K\frac{4}{\left(\frac{1}{1+4}\right)^5} = K4(1+4)^5. \tag{34}$$

Substitute $\delta_m = \delta_p = \nu$ into Eq. (34); this is then equivalent to

$$\frac{\alpha}{V_n}\beta > \nu^2 K4(1+4)^5. \tag{35}$$

When this inequality is reversed, the system evolves towards a steady state via damped oscillations. Our model parameters are chosen to satisfy Eq. (35) and yield periodic oscillations with a period close to 24 h in continuous dark. We use[2] $\frac{\alpha}{V_n} = 1800\,000/(\mathrm{pL\,h})$, $\beta = 10/\mathrm{h}$, $\nu = \frac{2\pi}{22}/\mathrm{h}$, and $K =$

[1]Although we do not prove it here, it can be shown that this is the most unstable case, i.e., the case in which it is easiest to obtain oscillations.

[2]In this paper, concentration is represented by the number of molecules per unit volume. pL is short for picoliter, and h is short for hour.

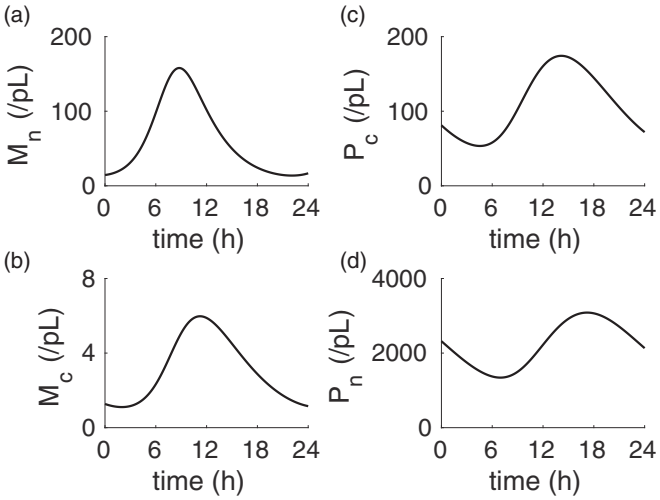

FIG. 2. Deterministic circadian oscillations in continuous dark after the system has reached a periodic steady state. The result shown here is the numerical solution of Eqs. (10)–(13).

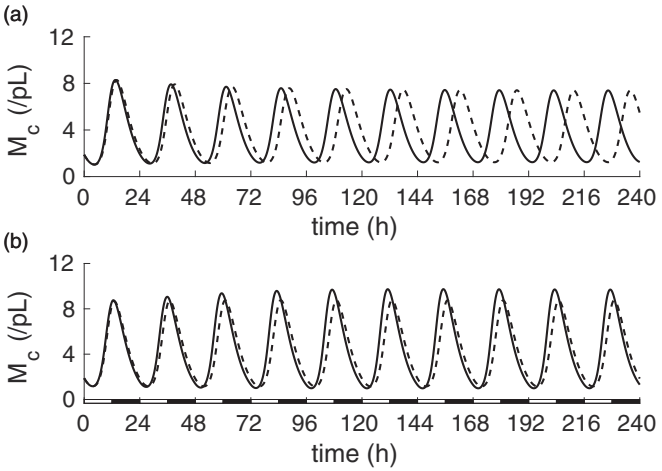

FIG. 3. (a) Deterministic circadian oscillators with free-running period $T = 23.5$ h (solid curve) and 24.5 h (dashed curve) in constant light. (b) Both oscillators are entrained by the 12:12 LD cycles (white and black bars along the horizontal axis).

200/pL, which produces an oscillator with autonomous period of $T_0 = 23.2$ h in continuous dark. Although in this case, the overall phase is arbitrary, we fix the phase by specifying the following initial conditions: $M_n(0) = 10.9$/pL, $M_c(0) = 1.88$/pL, $P_c(0) = 129.21$/pL, and $P_n(0) = 3531.75$/pL. As Fig. 2 shows, during continuous dark, the concentration of PER protein peaks about 6 h after the concentration of *Per* mRNA, in agreement with the observed phase relationship in mouse liver *in vivo* [6,25].

C. Analysis and simulation of the oscillator

In order to vary the free-running period of the oscillator before exposing it to light, we scale all of the rate constants (α, β, and ν) by a factor of $\theta = \frac{T_0}{T}$, where $T_0 = 23.2$ h is the autonomous period of the oscillator. This has the effect of setting the period of the oscillator to T.

It is known that light-induced resetting of the mammalian circadian clock is associated with rapid induction of *Per* mRNA transcription [25,39–42]. We incorporate the effect of light into the model by making the maximal transcription rate α of the *Per* gene be a function of the light level. We set

$$\alpha(t) = \alpha_0 + \bar{\alpha}[1 + \epsilon(t)]. \tag{36}$$

In this equation, α_0 denotes the maximum rate of transcription in the dark, and the second term on the right-hand side represents the effect of light on the maximum rate of transcription. In the studies presented in this paper, we hold α_0 and $\bar{\alpha}$ constant, and we choose $\epsilon(t)$ to be a periodic square wave with a period of 24 h such that $\epsilon(t) = +\epsilon_0$ for 12 h and then $\epsilon(t) = -\epsilon_0$ for 12 h, where $\epsilon_0 \in [0,1]$. The function $\epsilon(t)$ therefore has mean zero, and ϵ_0 is a dimensionless measure of the depth of modulation of the light signal. The parameter ϵ_0 can also be called the relative amplitude of the light signal. The case $\epsilon_0 = 1$ (which we consider in this section) corresponds to 12 h of light and 12 h of dark in each 24-h period, and the case $\epsilon_0 = 0$ corresponds to a constant intermediate light level. We emphasize that the mean light level remains constant as ϵ_0 varies.

1. Entrainment under 12:12 light-dark cycles

We solve the system (10)–(13) with free-running period $T = 23.5$ and 24.5 h using Euler's method under constant light conditions and under 12:12 light-dark (LD) cycles (i.e., 12 h of light followed by 12 h of darkness each day). Note that one of the oscillators has a shorter free-running period and the other a longer free-running period than that of the entraining signal. We plot the concentrations of *Per* mRNA in cytoplasm in Fig. 3. We see from the bottom plot that both oscillators are entrained by the 12:12 LD cycles.

In addition to looking for a deterministic solution, we want to investigate the entrainment behavior of the oscillator by simulating it stochastically, so we simulate the oscillator with the corresponding initial condition as used in the deterministic case by the Gillespie method [43]. To get the corresponding initial condition, we multiply each concentration by the appropriate volume to get a number of molecules, rounding to the nearest integer if necessary. To compare the stochastic simulation result with the numerical solution of the differential equations for the deterministic case, we draw a figure of one representative run with the deterministic solution superimposed. As Fig. 4 shows there is good agreement between the two cases under 12:12 LD cycles. In the unentrained case, there is good agreement for the first few cycles, but eventually the phase of the stochastic simulation drifts, so the two plots are no longer comparable at corresponding times, even though the stochastic and deterministic oscillators are still showing similar behaviors. This is unavoidable and is an inherent difficulty in the study of unentrained stochastic oscillators.

2. Scaling to vary the amount of noise

By the law of large numbers, we expect that the relative noise level of the system will be smaller when the numbers of molecules are larger. This can be achieved by increasing the volume of the cell. We use the following scaling under which the deterministic model Eqs. (10)–(13) is invariant. Recall that V_n is the volume of the nucleus, V_c is the volume of the

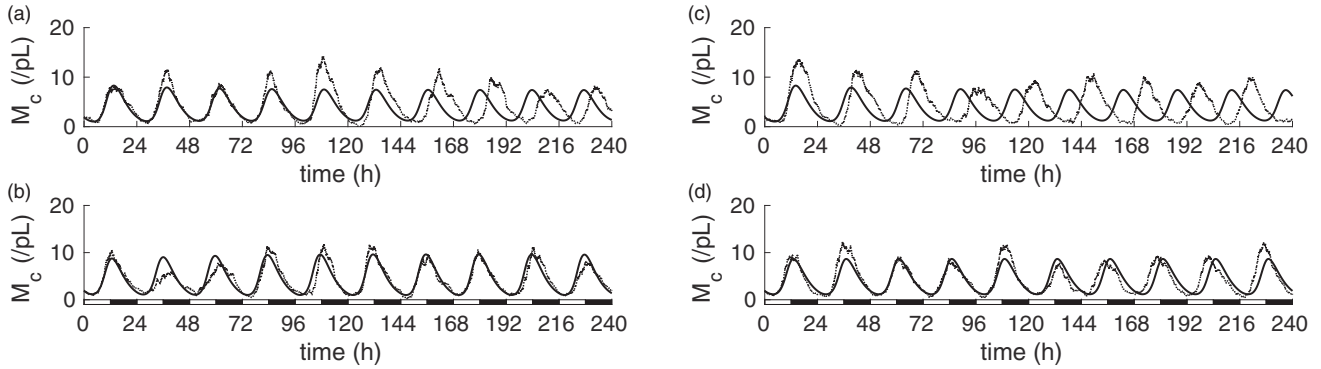

FIG. 4. Top: Stochastic simulation and deterministic solution for oscillators with free-running period (a) $T = 23.5$ h and (c) 24.5 h in constant light. Bottom: Both oscillators with free-running period (b) $T = 23.5$ h and (d) 24.5 h are entrained by the 12:12 LD cycles, and the stochastic model shows good agreement with the deterministic model. In this figure the numbers of molecules from the stochastic simulation have been expressed as concentrations for comparison with the deterministic result.

cytoplasm, and α is the parameter that governs the transcription rate of the *Per* gene. In the scaling used here, $\frac{\alpha}{V_n}$, $\frac{V_n}{V_c}$, and the initial concentrations are all held constant. To describe the results we use the cell volume $V = V_n + V_c$ as a parameter. In particular, we show results for $V = V_0$ and $V = 4V_0$, where V_0 is a realistic mammalian cell volume of 2.1 pL ($V_n = 0.1$ pL and $V_c = 2$ pL).

Figure 5 gives only a qualitative impression of the noise level in the system by comparing different realizations. To make this quantitative we consider the autocorrelation of the circadian rhythm under constant light. We let $f(t) = p_n(t) + p_c(t)$, that is, the total number of PER protein molecules in the cell at time t and define the normalized autocorrelation $\phi(\tau)$ as

$$\phi(\tau) = \frac{\int_0^{t_{\max}-\tau}[f(t) - \bar{f}_1(\tau)][f(t+\tau) - \bar{f}_2(\tau)]dt}{\int_0^{t_{\max}}[f(t) - \bar{f}]^2 dt}, \quad (37)$$

where

$$\bar{f} = \frac{1}{t_{\max}}\int_0^{t_{\max}} f(t)dt, \quad (38)$$

$$\bar{f}_1(\tau) = \frac{1}{t_{\max}-\tau}\int_0^{t_{\max}-\tau} f(t)dt, \quad (39)$$

$$\bar{f}_2(\tau) = \frac{1}{t_{\max}-\tau}\int_0^{t_{\max}-\tau} f(t+\tau)dt, \quad (40)$$

and where $t_{\max} = 600$ days $= 14\,400$ h is the total duration of a simulation. Thus $\phi(\tau)$ is a function that takes values in $[-1,1]$. For each oscillator with cell volumes V_0 and $4V_0$, we perform seven independent simulations under constant light and calculate their autocorrelations. Then we plot the averaged autocorrelation with shaded standard deviation. We expect that the autocorrelation will have a damped oscillatory character with a period close to 24 h; and, moreover, we expect that in a smaller cell, the stochastic process will lose correlation faster

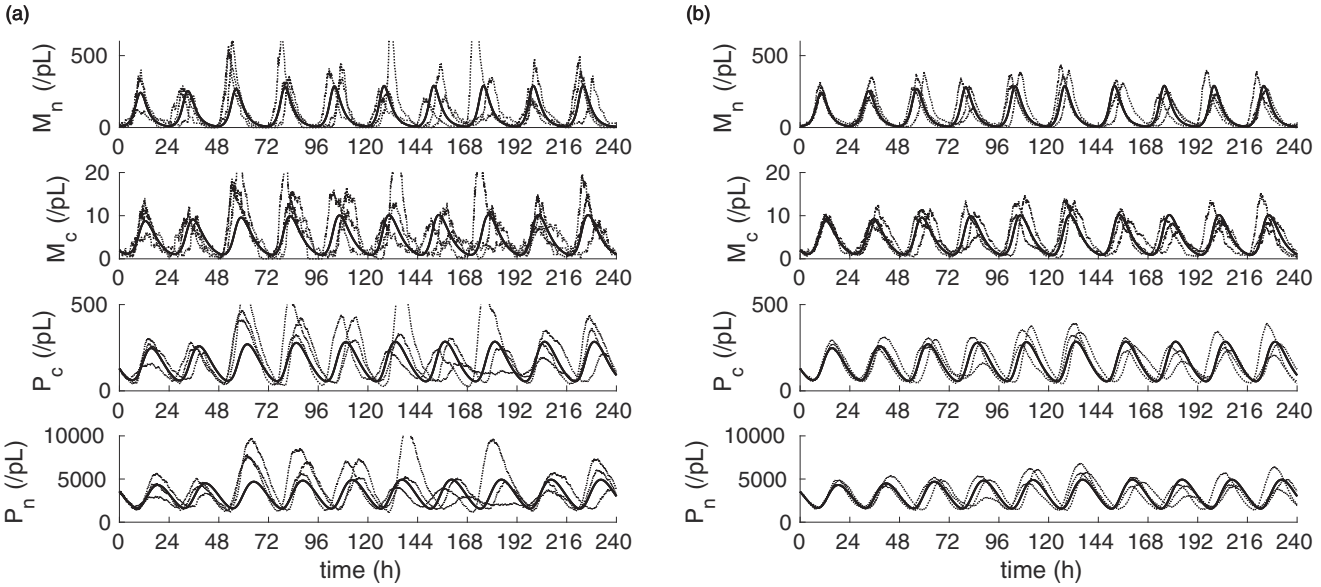

FIG. 5. Independent realizations (dotted curve) of concentrations of *Per* mRNAs and PER proteins under 12:12 LD cycles with the deterministic solution (solid curve) superimposed. Cell volume in (b) is four times that in (a). As in the previous figure, the numbers of molecules have been expressed as concentrations for comparison with the deterministic result.

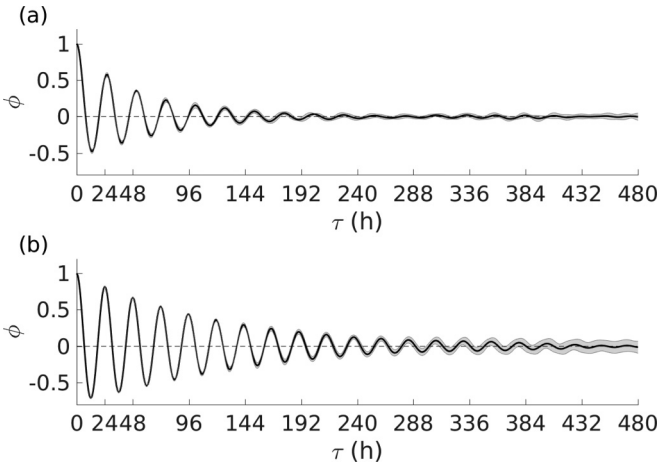

FIG. 6. Averaged autocorrelation $\phi(\tau)$ over seven independent realizations each of 14 400-h (600-day) duration of the total number of PER proteins with shaded standard deviation under constant light. Cell volume in (b) is four times that in (a).

than in a larger cell because of the relatively larger amount of molecular noise in the smaller cell. These phenomena are captured in Fig. 6.

III. ANALYSIS OF ENTRAINMENT

A. Visualization of entrainment

In a deterministic model of an oscillator with or without entrainment we can visualize the oscillation by choosing any two variables and by plotting the trajectory of the system on the plane of those two variables thus producing a limit cycle. In the stochastic case, the noise obscures such a plot, which in principle becomes dense on the phase plane if we run the system indefinitely. To overcome this difficulty, we keep track of the cumulative time that the system is in any of its discrete states, and then at the end of the run we divide by the total time of the run to get the fraction of time spent in any discrete state. This fraction is plotted in Fig. 7 on two different phase planes, the axes of which are the numbers of *Per* mRNA molecules in Fig. 7(a) and the numbers of PER protein molecules in Fig. 7(b) (note different scales). Both figures show that the limit cycle becomes more organized as the amplitude of the entrainment signal increases.

B. Measurement of entrainment

We now seek a quantitative measure of entrainment. For this purpose, we use the maximum of the cross correlation between $f(t) = p_n(t) + p_c(t)$ and the normalized light signal $s(t) = 1 + \epsilon(t)$. We define the normalized cross correlation between the stochastic output of the model and the input light signal as

$$\varphi(\tau) = \frac{\frac{1}{t_{\max}-\tau}\int_0^{t_{\max}-\tau} \frac{f(t)-\bar{f}_1(\tau)}{\bar{f}_1(\tau)} \frac{s(t+\tau)-\bar{s}_1(\tau)}{\bar{s}_1(\tau)}\,dt}{\sqrt{\frac{1}{t_{\max}}\int_0^{t_{\max}} \left(\frac{f(t)-\bar{f}}{\bar{f}}\right)^2 dt}\sqrt{\frac{1}{t_{\max}}\int_0^{t_{\max}} \left(\frac{s(t)-\bar{s}}{\bar{s}}\right)^2 dt}}, \tag{41}$$

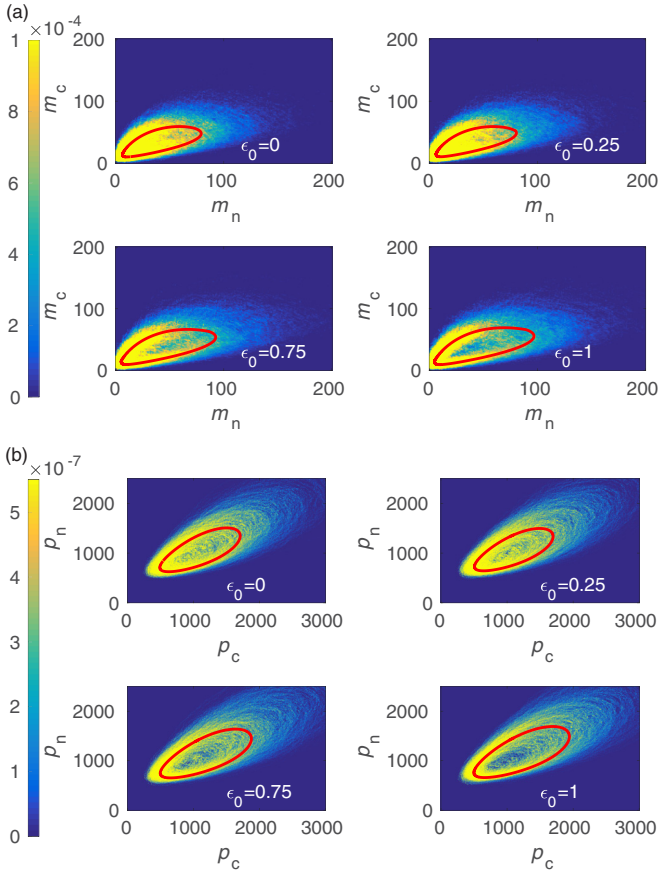

FIG. 7. Fraction of time spent in each system state projected onto the (m_n, m_c) plane in (a) and onto the (p_c, p_n) plane in (b), with the deterministic limit cycle superimposed (red curve). The label ϵ_0 on each plot indicates the depth of modulation of the entraining light signal, see Eq. (36) and the text that follows that equation. Note that the stochastic limit cycle becomes more organized as ϵ_0 increases. The bin width in these images is 1. That is, each bin corresponds to a unique pair of integers.

where

$$\bar{f} = \frac{1}{t_{\max}}\int_0^{t_{\max}} f(t)dt, \tag{42}$$

$$\bar{s} = \frac{1}{t_{\max}}\int_0^{t_{\max}} s(t)dt, \tag{43}$$

$$\bar{f}_1(\tau) = \frac{1}{t_{\max}-\tau}\int_0^{t_{\max}-\tau} f(t)dt, \tag{44}$$

$$\bar{s}_1(\tau) = \frac{1}{t_{\max}-\tau}\int_0^{t_{\max}-\tau} s(t)dt, \tag{45}$$

and where the duration of the simulation is $t_{\max} = 600$ days $= 14\,400$ h. Then $\varphi(\tau)$ is a function with values in $[-1,1]$, and the *maximal* value $\varphi^{\max}$ can serve as a good indication of how well the system is entrained by the signal. Note that the maximal value of the cross correlation $\varphi(\tau)$ can occur at any τ, not necessarily $\tau = 0$.

Using the numerical solution of system (10)–(13), we calculate $\varphi^{\max}$ of the oscillators with free-running period T ranging from 18 to 30 h for various relative intensities ϵ_0 of

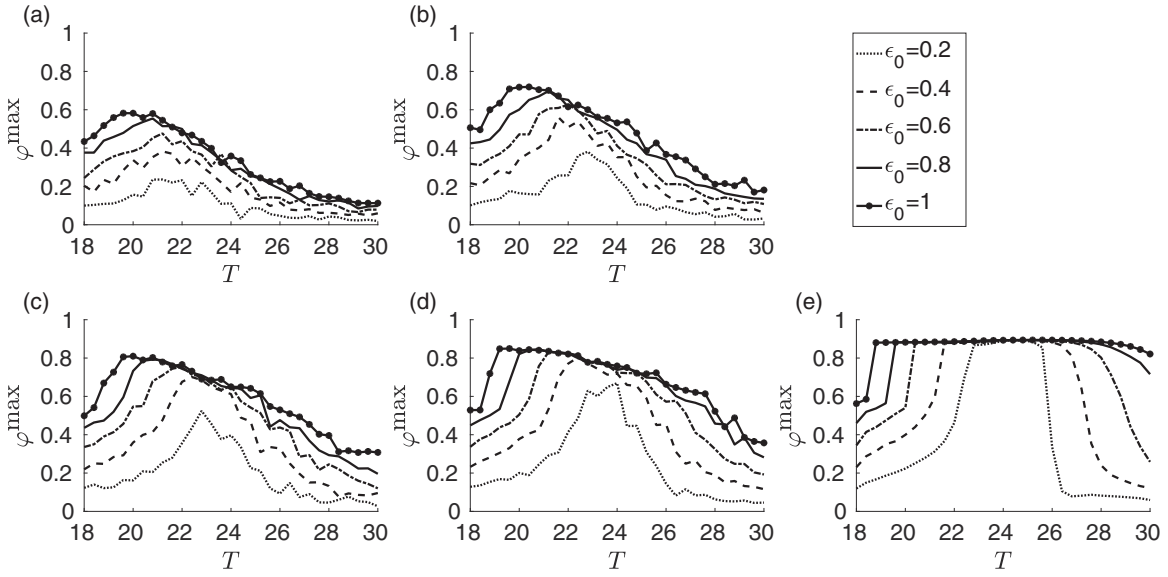

FIG. 8. Degree of entrainment $\varphi^{\max}$ for free-running period T from 18 to 30 h and depth of modulation ϵ_0 from 0.2 to 1. We show results for $V = V_0, 2V_0, 4V_0, 8V_0$ in (a)–(d) and for the deterministic system in (e). The roughness of curves in the stochastic cases is a result of the finite duration of the stochastic simulations ($t_{\max} = 14\,400$ h) and can be used to gauge the error that results from that finite duration.

the entraining light signal which has a fixed period of 24 h, and then we plot $\varphi^{\max}$ in Fig. 8(e). For each oscillator with increasing cell volumes, we also measure the entrainment level by simulating the oscillator stochastically and plot $\varphi^{\max}$ in Figs. 8(a)–8(d). Note that

(1) The circadian oscillator can be entrained regardless of whether the free-running period is shorter or longer than that of the entraining signal.

(2) The entrainment level increases with increasing depth of modulation of the light.

(3) The deterministic case [Fig. 8(e)] shows an envelope of maximal entrainment that can be achieved over a wider range of free-running periods T when the amplitude of the light signal is larger. It is striking that the departure from the envelope of maximal entrainment is very abrupt as the free-running period becomes too large or too small for maximal entrainment to occur.

(4) The stochastic cases [Figs. 8(a)–8(d)] show some of the same trends as in the deterministic case but in a more diffuse form, and of course the resemblance to the deterministic case is stronger when the cell volume is larger.

IV. SEQUESTRATION MODEL

In the model considered in the foregoing, the PER protein directly inhibits transcription of the *Per* gene. Here we consider an alternative possibility known as sequestration [8–14] in which the PER protein instead acts indirectly by binding to an activator of *Per* transcription and thereby preventing the activator from carrying out its function.

In the sequestration model, we assume that the regulation of transcription involves two reactions: The PER protein molecules tightly bind activators to form an inactive 1:1 stoichiometric complex, and only free activators can activate transcription by binding to DNA. Let p_n be the total number of PER protein molecules in the nucleus at some particular time counting both those that are free and those that are bound to the activators. Let a be the total number of activator molecules in the nucleus counting both those that are free and those that are bound to the PER protein or to the regulatory site on the DNA where activation occurs. Let ξ be the rate that any particular free PER protein molecule binds to any particular free activator molecule, and let η be the corresponding unbinding rate. Let λ be the rate that any particular free activator molecule binds to the site that activates transcription of the *Per* gene on the DNA, provided that site is not already occupied, and let μ be the corresponding unbinding rate. Let $s = 1$ if an activator molecule is bound to DNA, allowing transcription; 0 otherwise. Let c be the total number of protein-activator complexes $c \leqslant \min(a - s, p_n)$. Let $P_{ij} = \Pr(s = i \text{ and } c = j)$. By the same analysis as in the previous model [Eq. (3)], we have

$$P_{0j} = P_{00}\left(\frac{\xi}{\eta}\right)^j \frac{p_n!}{(p_n - j)!}\frac{a!}{(a-j)!}\frac{1}{j!}, \tag{46}$$

$$P_{1j} = P_{10}\left(\frac{\xi}{\eta}\right)^j \frac{p_n!}{(p_n - j)!}\frac{(a-1)!}{(a-1-j)!}\frac{1}{j!}, \tag{47}$$

from which it follows that

$$P_0 = \Pr(s = 0) \tag{48}$$

$$= \sum_{j=0}^{\min(a,p_n)} P_{0j} \tag{49}$$

$$= \left(\sum_{j=0}^{\min(a,p_n)} \left(\frac{\xi}{\eta}\right)^j \frac{p_n!}{(p_n - j)!}\frac{a!}{(a-j)!}\frac{1}{j!}\right) P_{00} \tag{50}$$

$$= f(a, p_n) P_{00}. \tag{51}$$

Similarly,

$$P_1 = \Pr(s = 1) \tag{52}$$
$$= f(a-1, p_n) P_{10} \tag{53}$$
$$= f(a-1, p_n) \frac{\lambda a}{\mu} P_{00}. \tag{54}$$

$P_0 + P_1 = 1$, so

$$P_{00} = \frac{1}{f(a, p_n) + f(a-1, p_n)\frac{\lambda a}{\mu}}, \tag{55}$$

and therefore,

$$P_1 = \frac{f(a-1, p_n)\frac{\lambda a}{\mu}}{f(a, p_n) + f(a-1, p_n)\frac{\lambda a}{\mu}} \tag{56}$$
$$= \frac{\frac{f(a-1, p_n)\frac{\lambda a}{\mu}}{f(a, p_n)}}{1 + \frac{f(a-1, p_n)\frac{\lambda a}{\mu}}{f(a, p_n)}}. \tag{57}$$

Equation (57) is exact in the limit that $\xi \to \infty$ and $\eta \to \infty$ with $\frac{\xi}{\eta}$ held constant. We further assume tight binding between the proteins and the activators (i.e., $\frac{\xi}{\eta} \to \infty$) so that it is sufficient to consider only the probability that a free activator binds to DNA when the total number of activator molecules is greater than the total number of protein molecules, i.e., when $a > p_n$, and then we have

$$P_1 = \frac{\lambda(a - p_n)}{\lambda(a - p_n) + \mu}$$
$$= \frac{1}{1 + \frac{K' V_n}{a - p_n}}, \tag{58}$$

in which K' with units of concentration is the equilibrium constant of the unbinding and binding reaction from or to DNA. Therefore,

$$P_1 = \begin{cases} 0, & a \leqslant p_n, \\ \frac{1}{1 + \frac{K' V_n}{a - p_n}}, & a > p_n. \end{cases} \tag{59}$$

Note that in the sequestration model P_1 plays the same role as P_0 did in the previous model since now transcription can only proceed if the regulatory site is occupied, whereas previously, transcription could only proceed if all of the regulatory sites were unoccupied.

In the limit of large cell volume, the dynamics of the model are governed by

$$\frac{dM_n}{dt} = F(P_n) - \gamma_m M_n, \tag{60}$$
$$\frac{dM_c}{dt} = \gamma_m \left(\frac{V_n}{V_c}\right) M_n - \delta_m M_c, \tag{61}$$
$$\frac{dP_c}{dt} = \beta M_c - \gamma_p P_c, \tag{62}$$
$$\frac{dP_n}{dt} = \gamma_p \left(\frac{V_c}{V_n}\right) P_c - \delta_p P_n, \tag{63}$$

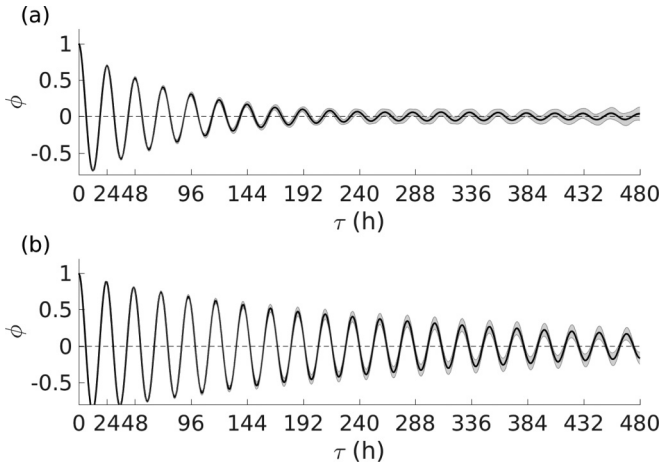

FIG. 9. Averaged autocorrelation $\phi(\tau)$ for the total number of PER proteins in the sequestration model. Compare Fig. 6, and see the legend of Fig. 6 for additional details.

where

$$F(P_n) = \begin{cases} 0, & A \leqslant P_n, \\ \frac{\alpha}{V_n}\left(\frac{1}{1 + \frac{K'}{A - P_n}}\right), & A > P_n, \end{cases} \tag{64}$$

the capitalized variables are concentrations of molecules, V_n is the volume of the nucleus, and V_c is the cytoplasmic volume. Note that the rate of transcription here has essentially the same form microscopically and macroscopically. This was not the case in the previous model. In the sequestration model with the tight-binding limit, it is necessary to scale up the number of activator molecules along with the cell volume in order to get a sensible limit, and this is fundamentally why the microscopic and macroscopic rates depend on the amount of PER protein in the nucleus in essentially the same way.

We choose parameters such that the steady-state values are the same as in the original model. We use $A = 2300$/pL and $K' = 5.11 \times 10^7$/pL. Simulation results for the sequestration model are shown in Figs. 9–11, and for comparison with those of the previous model, see Figs. 6–8, respectively. Note that the noise level is reduced compared to that in the original model: The decay of correlation is slower in Fig. 9 than in Fig. 6; the limit cycle is more organized in each frame of Fig. 10 than in the corresponding frame of Fig. 7; and the degree of entrainment reached at maximal entrainment is somewhat higher for each plot in Figs. 11(a)–11(d) than in the corresponding plot of Figs. 8(a)–8(d). Another effect, which can be seen by comparing Figs. 11(e) and 8(e), is that the sequestration oscillator seems to be more sharply tuned so that the width of its entrainment region is smaller for each depth of modulation of the light signal. This sharper tuning in the case of the sequestration model may explain the relatively smaller impact of molecular noise in that case, but we do not have any explanation of what makes the sequestration oscillator more sharply tuned. In any case, these effects are not large, and the two oscillators in all respects show similar qualitative behavior.

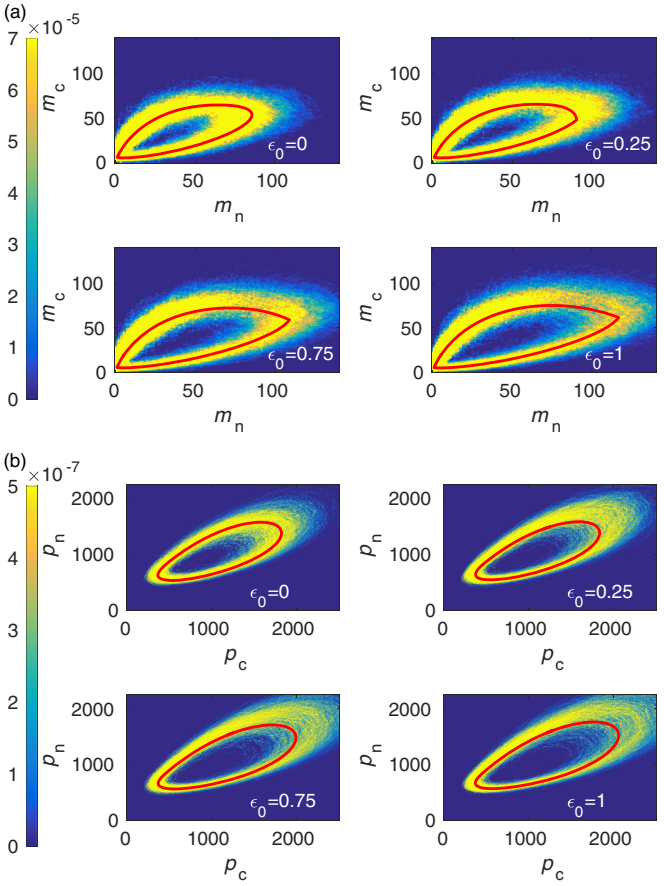

FIG. 10. Fraction of time spent in each system state for the sequestration model. Compare Fig. 7, and see the legend of Fig. 7 for additional details.

V. SUMMARY AND CONCLUSIONS

In this paper, we have introduced and studied a schematic model of a cellular circadian clock in which a single gene encodes a protein that inhibits the transcription of that gene. The model is fully stochastic with intrinsic noise that arises from the random nature of biochemical reactions. The focus of the paper is the entrainment of the oscillator by light, which we introduce as a periodic signal that affects the maximal rate at which the gene in question is transcribed.

In the study of entrainment, we have made use of two scalings of the model parameters, one of which allows us to control the relative noise level by varying the cell volume, and we have included in our paper the case of zero noise level, which is the large-volume deterministic limit of the model. The second scaling allows us to specify the natural frequency of the oscillator to study how entrainment depends on the difference between this natural frequency and the frequency of the entraining light signal. The reason that we vary the parameters of the oscillator rather than those of the light signal is that, in the natural world, the light signal always has the same 24-h period but the oscillators presumably have parameters that vary from species to species, from individual to individual, and perhaps even from cell to cell within an individual.

In all of our studies of entrainment, we have kept the mean light level constant while varying the relative amplitude, i.e., the depth of modulation, of the light signal. We have studied entrainment in two ways. The first method is to visualize the fraction of time spent by the system in each of its discrete states. This is performed for two variables at a time so that the result is a plot of intensity over a plane (actually, over the lattice of points with non-negative integer coordinates within the plane) in each case (Figs. 7 and 10). The pairs of variables chosen are the numbers of mRNA molecules in the cytoplasm and in the nucleus and the numbers of protein molecules in the cytoplasm and in the nucleus. This choice has the effect that within each pair we have comparable numbers of molecules. The conclusion of this visual study is that entrainment seems to *organize* the limit cycle of the stochastic oscillator. With increasing amplitude of the light signal, the limit cycle of the oscillator becomes increasingly well defined. This is somewhat surprising when we consider that the method we are using to

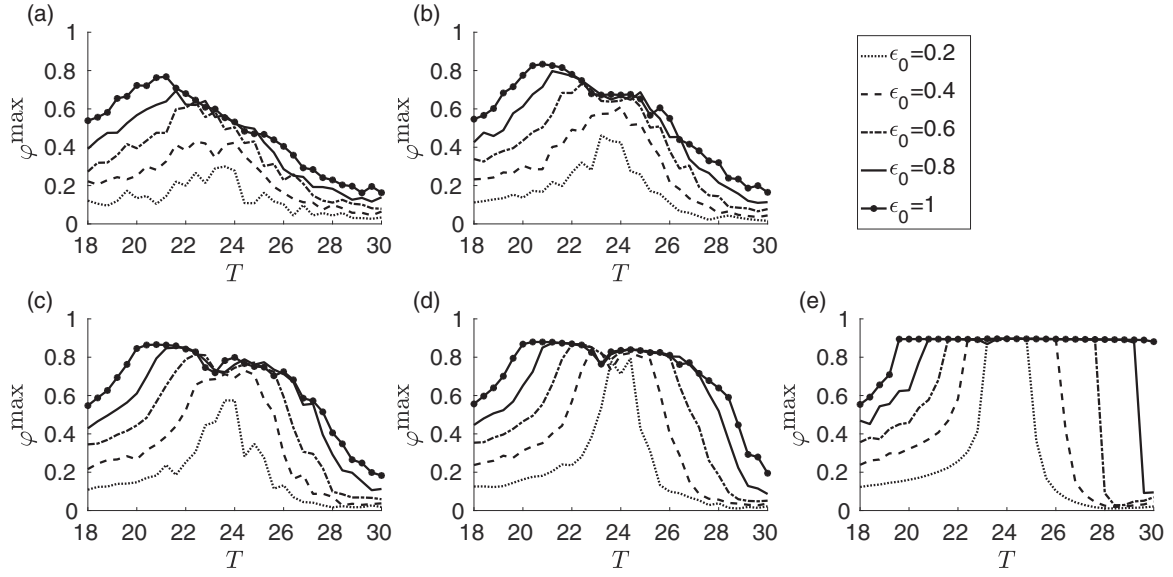

FIG. 11. Degree of entrainment $\varphi^{\max}$ for the sequestration model. Compare Fig. 8, and see the legend of Fig. 8 for additional details.

visualize the results is completely insensitive to the frequency and phase of the oscillator, which are the quantities that entrainment is known to affect. Thus, we are dealing here with a subtle consequence of entrainment that deserves further attention. We have not quantified this phenomenon, and we certainly have not explained it.

The second method for studying entrainment is to record the maximum of the normalized cross correlation between the periodic light signal and the number of protein molecules in the cell. We interpret this maximum value, which by construction lies in [0,1], as a measure of the degree of entrainment, and we study its behavior as a function of the natural period of the oscillator for various amplitudes of the entraining light signal (Figs. 8 and 11). It is especially striking in the deterministic case that there is an envelope of maximal entrainment that is independent of the depth of modulation of the light signal, although the interval of natural periods over which maximal entrainment is achieved does depend on the amplitude of the light signal, the interval being wider (as one would expect) for larger amplitudes. Also striking in the deterministic case is how sharply defined the interval of maximal entrainment is at any given depth of modulation of the light signal. Note, too, that the 24-h period of the entraining light falls roughly in the middle of the interval of maximal entrainment, which shows that entrainment is possible when the natural frequency of the oscillator is longer or shorter than the frequency of the light signal. This is in contrast to some other nonlinear oscillators, such as electrically active cells with pacemaker activity, that can only be entrained when the natural frequency of the oscillator is lower than the frequency of the entraining stimulus.

The properties of entrainment that we have just described in the deterministic limit seem to be reflected also in the corresponding stochastic results but in a somewhat degraded form. In particular, as the noise level increases, the envelope of maximal entrainment is pushed down to somewhat lower degrees of entrainment, and the minimum amplitude of the light signal that is needed to reach that envelope increases. Also, the interval of oscillator periods over which maximal entrainment occurs at any given amplitude of the light signal becomes less and less sharply defined.

We have also studied a model in which transcriptional repression occurs indirectly by sequestration of an activator instead of by direct binding of an inhibitor to the DNA. In this model, the noise level is somewhat reduced compared to that in the original model, but the qualitative picture remains the same.

The models of this paper are simplified and schematic ones, but they contain elements that are characteristic of actual cellular circadian clocks, including the stochasticity that arises from the random nature of biochemical reactions. We hope that the methods employed here will be useful in the future study of more complicated models of this kind and that the results obtained here will help to illuminate the phenomenology of entrainment in stochastic cellular circadian oscillators.

ACKNOWLEDGMENTS

We thank the anonymous reviewers of this paper for their invaluable suggestions. G.W. was supported by a MacCracken Fellowship from New York University.

[1] G. Wang and C. S. Peskin, Phys. Rev. E **92**, 052718 (2015).
[2] J. C. Leloup, D. Gonze, and A. Goldbeter, J. Biol. Rhythms **14**, 433 (1999).
[3] G. Kurosawa and A. Goldbeter, J. Theor. Biol. **242**, 478 (2006).
[4] T. Roenneberg, E. J. Chua, R. Bernardo, and E. Mendoza, Curr. Biol. **18**, R826 (2008).
[5] D. DeWoskin, W. H. Geng, A. R. Stinchcombe, and D. B. Forger, Interface Focus **4**, 20130076 (2014).
[6] J. C. Leloup and A. Goldbeter, Proc. Natl. Acad. Sci. USA **100**, 7051 (2003).
[7] J. Yan *et al.*, Nucleic Acids Res. **42**, 10278 (2014).
[8] B. Novak and J. J. Tyson, Nat. Rev. Mol. Cell. Biol. **9**, 981 (2008).
[9] J. K. Kim and D. B. Forger, Mol. Syst. Biol. **8**, 630 (2012).
[10] J. K. Kim, Z. P. Kilpatrick, M. R. Bennett, and J. Kresimir, Biophys. J. **106**, 2071 (2014).
[11] J. K. Kim, IET Syst. Biol. **10**, 125 (2016).
[12] N. Barkai and S. Leibler, Nature (London) **403**, 267 (2000).
[13] J. M. Vilar, H. Y. Kueh, N. Barkai, and L. Stanislas, Proc. Natl. Acad. Sci. USA **99**, 5988 (2002).
[14] P. Francois and V. Hakim, Phys. Rev. E **72**, 031908 (2005).
[15] J. C. Leloup and A. Goldbeter, IET Syst. Biol. **5**, 44 (2011).
[16] S. Becker-Weimann, J. Wolf, H. Herzel, and A. Kramer, Biophys. J. **87**, 3023 (2004).
[17] H. Zeng, Z. Qian, M. P. Myers, and M. Rosbash, Nature (London) **380**, 129 (1996).
[18] D. B. Forger and C. S. Peskin, Proc. Natl. Acad. Sci. USA **100**, 14806 (2003).
[19] D. B. Forger and C. S. Peskin, Proc. Natl. Acad. Sci. USA **102**, 321 (2005).
[20] A. Goldbeter, Proc. R. Soc. London, Ser. B. **261**, 319 (1995).
[21] D. Gonze, J. Halloy, and A. Goldbeter, J. Biol. Phys. **28**, 637 (2002).
[22] D. Gonze, J. Halloy, and A. Goldbeter, Proc. Natl. Acad. Sci. USA **99**, 673 (2002).
[23] J. C. Leloup and A. Goldbeter, J. Theor. Biol. **230**, 541 (2004).
[24] H. P. Mirsky, A. C. Liu, D. K. Welsh, S. A. Kay, and F. J. Doyle III, Proc. Natl. Acad. Sci. USA **106**, 11107 (2009).
[25] Y. Shigeyoshi *et al.*, Cell **91**, 1043 (1997).
[26] S. Sun, U. Alsbrecht, O. Zhuchenko, J. Bailey, G. Eichele, and C. Lee, Cell **90**, 1003 (1997).
[27] H. Tei, H. Okamura, Y. Shigeyoshi, C. Fukuhara, R. Ozawa, M. Hirose, and Y. Sakaki, Nature (London) **389**, 512 (1997).
[28] S. C. Crosthwaite, J. J. Loros, and J. C. Dunlap, Cell **81**, 1003 (1995).
[29] S. C. Crosthwaite, J. C. Dunlap, and J. J. Loros, Science **276**, 763 (1997).

[30] J. Qiu and P. Hardin, Mol. Cell. Biol. **16**, 4182 (1996).
[31] S. Reppert and D. Weaver, Nature (London) **418**, 935 (2002).
[32] P. E. Hardin, J. C. Hall, and M. Rosbash, Nature (London) **343**, 536 (1990).
[33] C. Lee, J. P. Etchegaray, F. Cagampang, A. Loudon, and S. M. Reppert, Cell **107**, 855 (2001).
[34] L. P. Shearman *et al.*, Science **288**, 1013 (2000).
[35] K. Bae, C. Lee, D. Sidote, K.-Y. Chuang, and I. Edery, Mol. Cell Biol. **18**, 6142 (1998).
[36] B. C. Goodwin, Adv. Enzyme Regul. **3**, 425 (1965).
[37] B. C. Goodwin, Nature (London) **209**, 479 (1966).
[38] C. D. Thron, Bull. Math. Biol. **53**, 383 (1991).
[39] M. Hunter-Ensor, A. Ousley, and A. Sehgal, Cell **84**, 677 (1996).
[40] C. Lee, V. Parikh, T. Itsukaichi, K. Bae, and I. Edery, Science **271**, 1740 (1996).
[41] M. Myers, K. Wager-Smith, A. Rothenfluh-Hilfiker, and M. Young, Science **271**, 1736 (1996).
[42] M. J. Zylka, L. P. Shearman, D. R. Weaver, and S. M. Reppert, Neuron **20**, 1103 (1998).
[43] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).