Optimal entrainment of circadian clocks in the presence of noise

Michele Monti,[1] David K. Lubensky,[2] and Pieter Rein ten Wolde[1]
[1]*AMOLF, Science Park 104, 1098 XE Amsterdam, The Netherlands*
[2]*Department of Physics, University of Michigan, Ann Arbor, Michigan 48109-1040, USA*

(Received 7 June 2017; published 12 March 2018)

Circadian clocks are biochemical oscillators that allow organisms to estimate the time of the day. These oscillators are inherently noisy due to the discrete nature of the reactants and the stochastic character of their interactions. To keep these oscillators in sync with the daily day-night rhythm in the presence of noise, circadian clocks must be coupled to the dark-light cycle. In this paper, we study the entrainment of phase oscillators as a function of the intrinsic noise in the system. Using stochastic simulations, we compute the optimal coupling strength, intrinsic frequency, and shape of the phase-response curve, that maximize the mutual information between the phase of the clock and time. We show that the optimal coupling strength and intrinsic frequency increase with the noise, but that the shape of the phase-response curve varies nonmonotonically with the noise: in the low-noise regime, it features a dead zone that increases in width as the noise increases, while in the high-noise regime, the width decreases with the noise. These results arise from a tradeoff between maximizing stability—noise suppression—and maximizing linearity of the input-output, i.e., time-phase, relation. We also show that three analytic approximations—the linear-noise approximation, the phase-averaging method, and linear-response theory—accurately describe different regimes of the coupling strength and the noise.

DOI: 10.1103/PhysRevE.97.032405

I. INTRODUCTION

Many organisms possess a circadian clock to anticipate the changes between day and night. Circadian clocks are biochemical oscillators that can tick without any external driving with an intrinsic, free-running period of about 24 h. In unicellular organisms these oscillations are formed by chemical reactions and physical interactions between molecules inside the cell, while in multicellular organisms these oscillations are typically shaped by a combination of intra- and intercellular interactions, which are, however, both mediated by molecular interactions. Due to the discreteness of molecules and the stochastic nature of chemical and physical interactions, circadian oscillations are inherently stochastic, which means that they have an intrinsic tendency to run out of phase with the day-night cycle. To keep the circadian oscillations in phase with the day-night rhythm, the oscillations must be coupled to daily cues from the environment, such as daily changes in light-intensity or temperature. This coupling makes it possible to lock the clock to, i.e., synchronize with, the daily rhythm. However, how the circadian clock should be coupled to entrainment cues is a question that is still wide open. It is neither clear what the natural performance measure for entrainment is, nor is it fully understood how this depends on the strength and form of the coupling, the characteristics of the entrainment signal, and the properties of the clock.

The function that is most commonly used to describe the coupling of the clock to the entrainment signal is called the phase-response curve [1]. It gives the shift of the phase of the clock as induced by a perturbation (a small change in, e.g., light intensity), as a function of the phase at which the perturbation was given. The phase-response curve has been measured for a wide variety of organisms, ranging from cyanobacteria, to fungi, plants, flies, and mammals [2]. Interestingly, these phase-response curves share a number of characteristic features: they typically consist of a positive and a negative lobe, and often possess a dead zone of no coupling during the subjective day (see Fig. 1). Yet, the width of the dead zone can vary significantly, and also the negative and positive lobes are not always equal in magnitude.

These observations naturally raise the question of what the best shape is for a phase-response curve. To answer this, a measure that quantifies the performance of the system is needed. Several measures have been put forward. A key characteristic of any locking scheme is the Arnold tongue [1], which describes the range of system parameters over which the deterministic system is locked to the driving signal. In general, this range tends to increase with the strength of the driving signal, and one performance measure that has been presented is how the range—the width of the Arnold tongue—increases with the magnitude of the driving; this derivative has been called the "entrainability" of the clock [3,4]. Another hallmark of any stochastic system is its robustness against noise, and, in general, the stability of an entrained clock depends not only on its intrinsic noise, but also on the strength and shape of the coupling function; one way to quantify clock stability is the so-called "regularity," which is defined as the variance of the clock period [3,4]. Another important property of any locked system is its sensitivity to fluctuations in the driving signal. To quantify this, Pfeuty *et al.* have defined two sensitivity measures, one that describes the change in the phase difference between the signal and the clock due to a change in the input, and another that quantifies the change in the stability of the fixed point (the slope of the phase-response curve) in response to a change in the input signal [2].

These performance measures make it possible to make predictions on the optimal shape of the phase-response curve. Pfeuty *et al.* argued that the shape of the phase-response

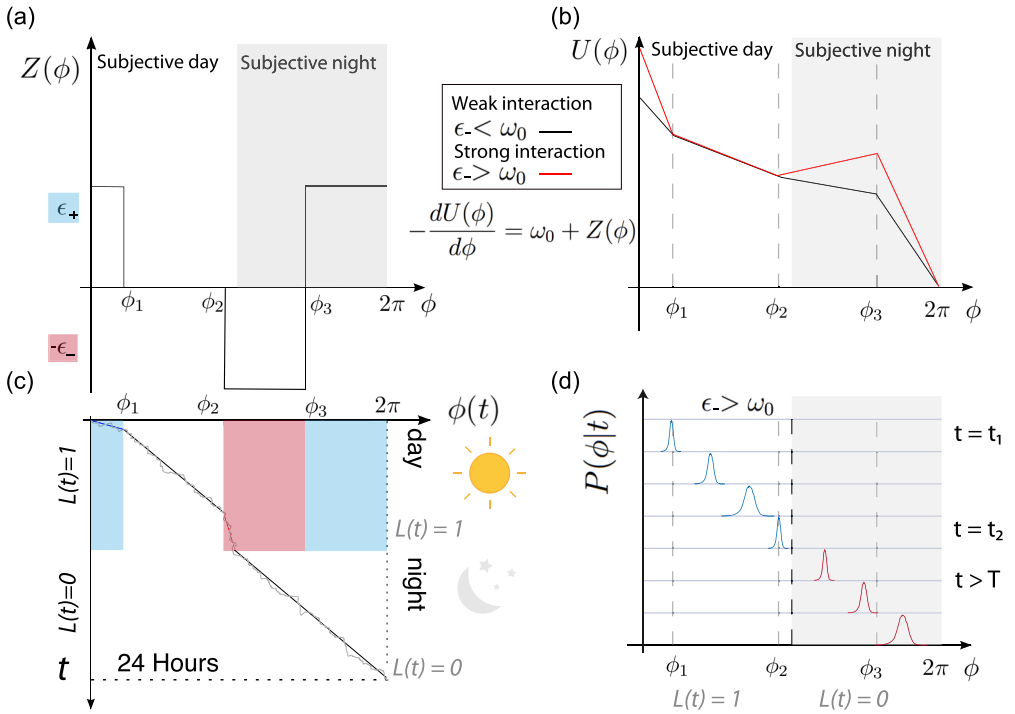

FIG. 1. Cartoon of the system. (a) The instantaneous phase-response curve $Z(\phi)$, characterized by the five parameters ϵ_+,ϵ_- and ϕ_1,ϕ_2,ϕ_3. The driving signal is given by $L(t) = 1$ during the day and $L(t) = 0$ during the night. (b) The phase evolution of the system, $d\phi/dt$, can be interpreted as that of a particle in a potential $U(\phi)$, with a force $-dU(\phi)/d\phi = \omega_0 + Z(\phi)L(t)$. Note that the particle only experiences a force during the day, when $L(t) = 1$, and not during the night, when $L(t) = 0$. (c) The phase evolution of the system, in the limit of small noise. During the night the deterministic system always evolves with its intrinsic frequency ω_0. During the day, it evolves with its intrinsic frequency ω_0 when the phase is between ϕ_1 and ϕ_2; between $\phi_3 - 2\pi$ and ϕ_1, the system is "pushed," moving with a frequency $\omega_0 + \epsilon_+$, while between ϕ_2 and ϕ_3 it is slowed down, moving at frequency $\omega_0 - \epsilon_-$. (d) Illustration of how $P(\phi)$ evolves in time, in the regime of strong coupling. At dawn, the system is pushed, narrowing the distribution; during the dead zone in which $Z(\phi) = 0$, the distribution tends to widen; near dusk, the system is slowed down, narrowing the distribution; during the night, the system evolves freely, widening the distribution again.

curve is determined by the requirement that the clock should respond to changes in light intensity that are informative on the day-night rhythm, namely light-intensity changes during dawn and dusk, but should ignore uninformative fluctuations in light intensity during the day, arising, e.g., from clouds [2]. This naturally gives rise to a dead zone in the phase-response curve, which allows the clock to ignore the input fluctuations during the day. Hasegawa and Arita argued that the shape of the phase-response curve is determined by a tradeoff between regularity (stability) and entrainability [3,4]. Entrainability requires not only changes in light intensity, but also that a change in the copy number n_i of a component i, as induced by the changing light signal, leads to a change in the phase ϕ of the clock: the gain $d\phi/dn_i$ should be large. However, a higher gain also means that the evolution of the phase becomes more susceptible to noise in n_i. Maximizing entrainability for a given total noise strength integrated over 24 h then yields a phase-response curve with a dead zone: During the day, when informative variations in light intensity are low, a high gain will not significantly enhance entrainability but will increase the integrated noise, implying that the gain should be as low as possible during the middle of the day.

In this paper, we introduce another measure to quantify the performance of the system, the mutual information [5]. The mutual information quantifies the number of signals that can be transmitted uniquely through a communication channel. As such it is arguably the most powerful measure for quantifying information transmission, and in recent years the mutual information has indeed been used increasingly to quantify the quality of information transmission in cellular signaling systems [6–24]. In the context studied here, the central idea is that the cell needs to infer from a variable of the clock, e.g., its phase ϕ, the time of the day t. The mutual information then makes it possible to quantify the number of distinct time points that can be inferred uniquely from the phase of the clock. Importantly, how many time states can be inferred reliably depends not only on the noise in the system, but also on the shape of the input-output curve, $\overline{\phi}(t)$, i.e., the average phase $\overline{\phi}(t)$ at time t.

We study how the mutual information between the clock phase and the time depends on the shape and magnitude of the phase-response curve in the presence of intrinsic noise in the system; we thus do not consider fluctuations in the input signal. The clock is modelled as a phase oscillator and the phase-response curve is described via a piecewise linear function (see Fig. 1), which allows for optimization and analytical results. We find that for a given amount of noise in the system there exists an optimal coupling strength that maximizes the mutual information: Increasing the coupling strength too much will decrease the mutual information. However, as the noise in the system increases, the optimal coupling strength increases. Moreover, for a given shape of the phase-response curve

featuring a dead zone, the optimal intrinsic (free running) period of the clock is nonmonotonic: as the noise is increased, the optimal period first becomes larger than 24 h, but then decreases to become smaller than 24 h. Optimizing over not only the coupling strength and the intrinsic period but also over the shape of the phase-response curve reveals that the optimal width of the dead zone is also nonmonotonic. As the noise is increased, the width first increases, but then decreases. We show that all of these results can be understood as a tradeoff between linearity and stability. At low noise, it is paramount to make the input-output relation $\overline{\phi}(t)$ as linear as possible, because this maximizes the mutual information; this is enhanced by a large dead zone and weak coupling. However, for large noise strengths, stability becomes key, which favors a small dead zone, a stronger coupling, and a smaller intrinsic period.

In the next section, we first briefly present the chemical Langevin description of a biochemical network, because this is important for understanding not only the phase-reduction method that reduces the system to a phase-oscillator model, but also for understanding some important characteristics of the mutual information. In the subsequent section, we then introduce the mutual information. We emphasize that the mutual information is insensitive to a coordinate transformation and that the mutual information between all degrees of freedom of the system (i.e., copy numbers of all components) and the input (i.e., time t) is always larger than that between one degree of freedom and the input. This means that the mutual information that we will compute between the phase of the clock and the time will provide a firm lower bound on the actual mutual information. We then briefly describe our phase-oscillator model and how we model the phase-response curve.

In the results section, we first present the results of stochastic simulations of our phase-oscillator model. By performing very extensive simulations we find the system parameters that maximize the mutual information, and by explicitly computing the linearity and stability as a function of parameters, we show that the optimal design as a function of the noise arises from the tradeoff mentioned above between linearity and stability.

Finally, we present and apply three different analytic approximations (or "theories"), and show that each recapitulates the simulations in a different parameter regime. The linear-noise approximation accurately describes the regime of low noise and strong coupling. The phase-averaging method [1] captures the regime of low noise and weak coupling. Finally, the linear-response theory accurately describes the mutual information in the regime of high noise and weak coupling. Whereas the first two approximations are valid in the vicinity of the optimal coupling for an appropriate range of noise strengths, the third turns out to hold only far from optimality.

II. MODEL

Below we derive the phase-oscillator model for a biochemical system starting from the chemical Langevin equation. This description is generic: the biochemical system can either be a single bacterial cell such as a cyanobacterium or a higher eukaryotic organism.

A. Chemical Langevin description

We consider a self-sustained oscillator of M components with copy numbers $n_1, n_2, \ldots, n_M$, denoted by the vector $\mathbf{n}$. Its dynamics is given by

$$\frac{d\mathbf{n}}{dt} = \mathbf{A}(\mathbf{n}), \tag{1}$$

where $\mathbf{A}(\mathbf{n})$ is determined by the propensity functions of the chemical reactions that constitute the network. The limit cycle of the free-running oscillator is the stable periodic solution of this equation, $\mathbf{n}(t) = \mathbf{n}(t + T_0)$, where T_0 is the intrinsic period of the oscillator.

Due to the stochasticity of the chemical reactions and the discreteness of the molecules, the evolution of the network is stochastic. When the copy numbers are sufficiently large, then the dynamics can be described by the chemical Langevin equation [25],

$$\frac{d\mathbf{n}}{dt} = \mathbf{A}(\mathbf{n}) + \underline{\eta}(\mathbf{n}), \tag{2}$$

where the vector $\underline{\eta}(t)$ describes the Gaussian white noise, characterized by the noise matrix with elements $\langle \eta_i(\mathbf{n}(t))\eta_j(\mathbf{n}(t'))\rangle = D_{ij}(\mathbf{n})\delta(t - t')$.

A clock is only a useful timing device if it has a stable and precise phase relationship with the daily rhythm. Biochemical noise tends to disrupt this relationship. To keep the clock in sync with the day-night rhythm in the presence of noise, the clock must be coupled to the light signal:

$$\frac{d\mathbf{n}}{dt} = \mathbf{A}(\mathbf{n}) + \epsilon\mathbf{p}(\mathbf{n},t) + \underline{\eta}(\mathbf{n}). \tag{3}$$

Here $\mathbf{p}(\mathbf{n},t)$ describes the coupling to the light signal and ϵ the strength of the coupling. The coupling force $\mathbf{p}(\mathbf{n},t) = \mathbf{p}(\mathbf{n},t + T)$ has a period T and frequency $\omega = 2\pi/T$, which in general is different from the intrinsic period T_0 and intrinsic frequency $\omega_0 = 2\pi/T_0$, respectively, of the free-running oscillator. In this paper, we will assume that the light signal is deterministic. We thus only consider the biochemical noise in the clock.

B. Mutual information

The organism needs to infer the time t from the concentrations of the clock components. This inference will be imprecise, because of the noise in the clock. We will quantify the accuracy of information transmission via the mutual information, which is a measure for how many distinct time states can be resolved from the concentrations of the clock components [5].

The mutual information $I(\mathbf{n};t) = I(\{n_1, \ldots, n_M\};t)$ between the copy numbers of all components and the time is given by

$$I(\mathbf{n};t) = \int d\mathbf{n} \int dt\, P(\mathbf{n};t) \log_2 \frac{P(\mathbf{n};t)}{P(\mathbf{n})P(t)}, \tag{4}$$

where $P(\mathbf{n};t)$ is the probability that copy numbers $\mathbf{n}$ are found at time t. $I(\mathbf{n};t)$ measures the reduction in uncertainty about t upon measuring $\{n_1, \ldots, n_M\}$, or vice versa. The quantity is indeed symmetric in $\mathbf{n}$ and t:

$$I\mathbf{n};t) = H(t) - \langle H(t|\mathbf{n})\rangle_{\mathbf{n}} \tag{5}$$

$$= H(\mathbf{n}) - \langle H(\mathbf{n}|t)\rangle_t, \tag{6}$$

where $H(\mathbf{a}) = -\int d\mathbf{a} P(\mathbf{a}) \log_2 P(\mathbf{a})$, with $P(\mathbf{a})$ the probability distribution of $\mathbf{a}$, is the entropy of $\mathbf{a}$; $H(\mathbf{a}|\mathbf{b}) =$

$-\int d\mathbf{a} P(\mathbf{a}|\mathbf{b}) \log_2 P(\mathbf{a}|\mathbf{b})$ is the information entropy of $\mathbf{a}$ given $\mathbf{b}$, with $P(\mathbf{a}|\mathbf{b})$ the conditional probability distribution of $\mathbf{a}$ given $\mathbf{b}$; $\langle f(c)\rangle_c$ denotes an average of $f(c)$ over the distribution $P(c)$.

A key point worthy of note is that the mutual information is invariant under a coordinate transformation, which allows us to put a firm lower bound on the mutual information between time and the clock components. Specifically, we can first make a nonlinear transformation from $\mathbf{n}$ to some other set of variables $\mathbf{x}$, of which two components are the amplitude R of the clock and its phase ϕ. Because the mutual information is invariant under this transformation,

$$I(\mathbf{n},t) = I(\mathbf{x},t). \tag{7}$$

Second, if the time is inferred not from all the components of $\mathbf{x}$, but rather from R and ϕ, then, in general,

$$I(R,\phi;t) \leqslant I(\mathbf{x};t). \tag{8}$$

By combining this expression with Eq. (7), we find that

$$I(\mathbf{n};t) \geqslant I(R,\phi;t). \tag{9}$$

Hence, once we have defined a mapping between $\mathbf{n}$ and $\mathbf{x}$ and hence (R,ϕ), the mutual information $I(R,\phi;t)$ between the combination of the amplitude and phase of the clock (R,ϕ) and time t, puts a lower bound on the mutual information $I(\mathbf{n};t)$. A weaker lower bound is provided by the mutual information between the phase of the clock and time:

$$I(\mathbf{n};t) \geqslant I(R,\phi;t) \geqslant I(\phi;t). \tag{10}$$

However, we expect this bound to be rather tight, since a reasonable, natural, mapping between $\mathbf{n}$ and (R,ϕ) should put the information on time in the phase of the clock.

C. Phase oscillator

The bound of Eq. (10) makes it natural to develop a description of the clock in terms of the phase. Here, we review the derivation of such a description, largely following the standard arguments in [1], but paying special attention to the appropriate form of the effective noise on the phase variable. In the absence of any coupling and noise, the temporal evolution of the phase is given by

$$\frac{d\phi(\mathbf{n})}{dt} = \omega_0, \tag{11}$$

where $\omega_0 = 2\pi/T_0$ is the intrinsic frequency of the clock, with T_0 the intrinsic period. As the phase is a smooth function of $\mathbf{n}$, the evolution of ϕ is also given by

$$\frac{d\phi(\mathbf{n})}{dt} = \sum_i \frac{\partial\phi}{\partial n_i}\frac{dn_i}{dt}. \tag{12}$$

Combining the above two equations with Eq. (1) yields the following expression for the intrinsic frequency:

$$\omega_0 = \sum_i \frac{\partial\phi}{\partial n_i} A_i(\mathbf{n}). \tag{13}$$

This equation *defines* a mapping $\phi(\mathbf{n})$. This mapping is defined such that for each point $\mathbf{n}$ in state space, the time derivative $d\phi(\mathbf{n})/dt = d\phi/dt$ of the phase is constant and equal to ω_0. The surfaces of constant $\phi(\mathbf{n})$, defined according to this mapping, are called isochrones.

In the presence of noise, the phase dynamics is, combining Eqs. (2) and (12),

$$\frac{d\phi(\mathbf{n})}{dt} = \sum_i \frac{\partial\phi}{\partial n_i}[A_i(\mathbf{n}) + \eta_i(\mathbf{n})] \tag{14}$$

$$= \omega_0 + \xi(\mathbf{n}), \tag{15}$$

which yields for the noise on the phase variable

$$\xi(\mathbf{n}) = \sum_i \frac{\partial\phi}{\partial n_i}\eta_i(\mathbf{n}). \tag{16}$$

In general, the variance of ξ thus depends on all of the state variables $\mathbf{n}$, not just on the phase ϕ, and Eq. (15) does not give a closed description in terms only of ϕ. However, when the deviations from the limit cycle are small compared to the scale over which the noise strength changes as a function of distance from the limit cycle, we can estimate the noise by evaluating it at the limit cycle, $\mathbf{n}_0$:

$$\xi(\phi) = \sum_i \frac{\partial\phi(\mathbf{n}_0)}{\partial n_i}\eta_i(\mathbf{n}_0), \tag{17}$$

with Gaussian white noise statistics

$$\langle \xi(\phi(t))\xi(\phi(t'))\rangle = \sum_{i,j} \frac{\partial\phi}{\partial n_i}\frac{\partial\phi}{\partial n_j} D_{ij}(\mathbf{n}_0)\delta(t-t'), \tag{18}$$

$$\equiv 2D(\phi)\delta(t-t'). \tag{19}$$

When the system is coupled to light, the phase evolution becomes, from Eqs. (3) and (12),

$$\frac{d\phi(\mathbf{n})}{dt} = \sum_i \frac{\partial\phi}{\partial n_i}[A_i(\mathbf{n}) + \epsilon p_i(\mathbf{n},t) + \eta_i(\mathbf{n})]. \tag{20}$$

The force depends explicitly on time. This impedes a unique definition of the isochrones $\phi(\mathbf{n})$, because how the phase evolves at a particular point in phase space depends not only on $\mathbf{n}$ but also on t. Of course, one could still adopt the mapping of the free-running system, in which case the evolution of the phase is given by

$$\frac{d\phi(\mathbf{n})}{dt} = \omega_0 + \epsilon \sum_i \frac{\partial\phi}{\partial n_i} p_i(\mathbf{n},t) + \xi(\phi). \tag{21}$$

The problem is that, because along the surface $\phi(\mathbf{n})$ the light-coupling term is not constant, $d\phi(\mathbf{n})/dt$ will depend on $\mathbf{n}$. One can then not reduce the dynamics to that of a single phase variable.

However, if ϵ is small and the force only leads to small deviations from the limit cycle of the free-running system, then one may approximate the effect of the forcing by evaluating the corresponding term at the limit cycle, $\mathbf{n}_0$. We then have

$$\frac{d\phi(\mathbf{n})}{dt} = \omega_0 + \epsilon \sum_i \frac{\partial\phi(\mathbf{n}_0)}{\partial n_i} p_i(\mathbf{n}_0,t) + \xi(\phi). \tag{22}$$

In this case the evolution of the phase no longer explicitly depends on $\mathbf{n}$:

$$\frac{d\phi}{dt} = \omega_0 + Q(\phi,t) + \xi(\phi), \tag{23}$$

with

$$Q(\phi,t) = \epsilon \sum_i \frac{\partial \phi(\mathbf{n}_0(\phi))}{\partial n_i} p_i(\mathbf{n}_0(\phi),t). \tag{24}$$

How a circadian clock responds to a given light signal $L(t)$ depends on its phase ϕ; it does not explicitly depend on time. The coupling term can then be written as $Q(\phi,t) = Z(\phi)L(t)$, where $Z(\phi)$ is the instantaneous phase-response curve, which describes how the clock responds to the light signal as a function of its phase ϕ. In addition, while in general the noise strength depends on the phase, we will, motivated by the experimental observations of Mihalecescu and Leibler on the clock of the cyanobacterium *Synechococcus elongatus* [26], assume it is constant. We then finally arrive at the equation that describes the evolution of the phase in our model:

$$\frac{d\phi}{dt} = \omega_0 + Z(\phi)L(t) + \xi(t), \tag{25}$$

with $\langle \xi(t)\xi(t')\rangle = 2D\delta(t-t')$.

In what follows, we will study entrainment using the above equation not only when $Z(\phi)L(t)$ and D are much smaller than ω_0, so that the weak coupling assumptions necessary for the reduction to a phase oscillator clearly hold, but also when $Z(\phi)L(t)$ or D are of order ω_0 or larger. As we discuss in more detail in Sec. V, however, this does not present any contradiction, because it is perfectly possible for the noise and the external driving to be small compared to restoring forces orthogonal to the limit cycle, so that the system always stays near the limit cycle and the phase is the only relevant variable, while simultaneously strongly perturbing motion along the limit cycle. Hence, our description not only applies to organisms with highly stable rhythms, such as the cyanobacterium *S. elongatus* [26], but also to cells with more noisy clocks [27,28].

While it is clear that ϵ can be varied independently of the noise strength, it is perhaps less obvious that $Z(\phi)$ and D can be varied independently. When the size of the system, e.g., the volume of a living cell, is changed, then the noise strength D will change, but the coupling strength $Z(\phi)$ will, to first order, not change because the concentrations remain constant. This shows that at least fundamentally $Z(\phi)$ and D can be varied independently by changing the volume. In fact, even experimentally this might be possible: to study the effect of noise on the differentiation dynamics of the bacterium *Bacillus subtilus* (which does not have a circadian clock), Süel and co-workers varied cell length by inducing filamentation [29] and similar experiments could potentially be performed for cyanobacteria [30]. In addition, typically the system is coupled to light only via a relatively small number of reactions, while the noise is determined by all reactions. Also in this case, it seems natural to assume that $Z(\phi)$ and D can be varied independently. We note that the arguments of Hasegawa and Arita do not contradict our arguments that $Z(\phi)$ and D can be varied independently: the fact that changing the gain $\partial\phi/\partial n_i$ affects both the coupling to light (entrainability) and the phase noise [3,4], does not mean that the noise and the coupling cannot be varied independently if other parameters are changed (and vice versa). We thus imagine that $p_i(\mathbf{n})$ can be tuned (by evolution) independently of the $D_{ij}(\mathbf{n})$. We do not change the mapping $\phi(\mathbf{n})$, determined by the properties of the uncoupled system.

D. System

We will approximate $Z(\phi)$ and $L(t)$ as step functions, shown in Fig. 1. This makes it possible to analytically obtain the Arnold tongue, i.e., the range of parameters for which the deterministic system locks to the day-night rhythm in the absence of noise. The light-dark function $L(t)$ is unity for $0 < t < T/2$ and zero for $T/2 < t < T$. The shape of the instantaneous phase-response curve $Z(\phi)$ is inspired by experimentally characterized response curves, featuring a positive lobe, a dead zone in which $Z(\phi)$ is essentially zero, a negative lobe, followed by a positive lobe again [2]. It is characterized by five variables, the coupling strengths ϵ_+ and ϵ_-, and the phases ϕ_1,ϕ_2,ϕ_3:

$$Z(\phi) = \begin{cases} \epsilon_+ & 0 < \phi < \phi_1 \\ 0 & \phi_1 < \phi < \phi_2 \\ -\epsilon_- & \phi_2 < \phi < \phi_3 \\ \epsilon_+ & \phi_3 < \phi < 2\pi \end{cases}, \tag{26}$$

where ϵ_+ and ϵ_- are greater than 0. With these five variables, a wide range of experimentally characterized phase-response curves can be described.

III. RESULTS

A. Arnold tongue of the deterministic system

Motivated by the observation that circadian clocks typically lock 1:1 to the day-night rhythm, we will focus on this locking scenario, although we will also see that this system can exhibit higher-order locking, especially when the intrinsic period of the clock deviates markedly from that of the day-night rhythm. To derive the Arnold tongue, we first note that when the clock is locked to the light-dark cycle, it will have a characteristic phase ϕ_s at the beginning of the light-dark cycle, $t_s = 0$. In the case of 1:1 locking, the phase of the clock will then cross phase ϕ_1 at time t_1, ϕ_2 at time t_2, and ϕ_3 at time t_3. To obtain the Arnold tongue, we have to recognize that there are in total 12 possible locking scenarios: three for ϕ_s and four for t_1,t_2,t_3. The scenarios for ϕ_s are (1) $\phi_3 - 2\pi < \phi_s < \phi_1$; (2) $\phi_1 < \phi_s < \phi_2$; (3) $\phi_2 < \phi_s < \phi_3$. The four scenarios for t_1,t_2,t_3 are defined by where $T/2$ falls with respect to these times: (1) $T/2 < t_1 < t_2 < t_3$; (2) $t_1 < T/2 < t_2 < t_3$; (3) $t_1 < t_2 < T/2 < t_3$; (4) $t_1 < t_2 < t_3 < T/2$. For each of these 12 scenarios, we can analytically determine ϕ_s and t_1,t_2,t_3, which then uniquely specify $\phi(t)$. The four unknowns, ϕ_s,t_1,t_2,t_3, give each an inequality for T, and the range of T that satisfies all four inequalities determines the width of the Arnold tongue. For each of the 12 scenarios for the given ϵ_+,ϵ_-, we have an Arnold tongue, and those 12 tongues together give "the" Arnold tongue for those values of ϵ_+,ϵ_-. We now derive the tongue for scenario 1, which is also the most important one, as we will see: in this regime, the mutual information between time and the phase of the clock is the largest.

Scenario 1 is characterized by $\phi_3 - 2\pi < \phi_s < \phi_1$; $0 < t_1 < t_2 < T/2 < t_3$. The solution depends on whether ϵ_- is larger or smaller than ω_0. If $\epsilon_- < \omega_0$, then the deterministic system locks 1:1 to the driving signal when

$$\phi_s + (\epsilon_+ + \omega_0)t_1 + \omega_0(t_2 - t_1) + (-\epsilon_- + \omega_0)(T/2 - t_2) + \omega_0 T/2 = \phi_s + 2\pi. \tag{27}$$

To solve this, we note that $\phi_1 = \phi_s + (\omega_0 + \epsilon_+)t_1$, $\Delta\phi_{12} \equiv \phi_2 - \phi_1 = \omega_0(t_2 - t_1)$. The solution is

$$t_1 = \frac{2\pi - T(\omega_0 - \epsilon_-/2) - \epsilon_-\Delta\phi_{12}/\omega_0}{\epsilon_+ + \epsilon_-} \geqslant 0, \tag{28}$$

$$t_2 = \frac{\Delta\phi_{12}}{\omega_0} + t_1 < T/2, \tag{29}$$

$$t_3 = \frac{\Delta\phi_{23}}{\omega_0 - \epsilon_-} + t_2 > T/2, \tag{30}$$

$$\phi_s = \phi_1 - (\omega_0 + \epsilon_+)t_1 > \phi_3 - 2\pi, \tag{31}$$

where $\Delta\phi_{23} \equiv \phi_3 - \phi_2$. The above inequalities lead to the following inequalities for the period T, respectively:

$$T \leqslant \frac{2\pi - \epsilon_-\Delta\phi_{12}/\omega_0}{\omega_0 - \epsilon_-/2}, \tag{32}$$

$$T > \frac{2\pi + \epsilon_+\Delta\phi_{12}/\omega_0}{\epsilon_+/2 + \omega_0}, \tag{33}$$

$$T < \frac{2\pi + \epsilon_+\Delta\phi_{12}/\omega_0 + \Delta\phi_{23}(\epsilon_+ + \epsilon_-)/(\omega_0 - \epsilon_-)}{\epsilon_+/2 + \omega_0}, \tag{34}$$

$$T > \frac{(\Delta\phi_{13} - 2\pi)(\epsilon_+ + \epsilon_-)/(\omega_0 + \epsilon_+) + 2\pi - \epsilon_-\Delta\phi_{12}/\omega_0}{\omega_0 - \epsilon_-/2}, \tag{35}$$

where $\Delta\phi_{13} \equiv \phi_3 - \phi_1 = \Delta\phi_{12} + \Delta\phi_{23}$. The width of the Arnold tongue is given by the range of T that satisfies all inequalities.

If $\epsilon_- > \omega_0$, then the equation to solve is

$$\phi_s + (\epsilon_+ + \omega_0)t_1 + \omega_0(t_2 - t_1) + \omega_0 T/2 = \phi_s + 2\pi. \tag{36}$$

The solution is

$$t_1 = \frac{2\pi - \omega_0 T/2 - \Delta\phi_{12}}{\epsilon_+ + \omega_0} \geqslant 0, \tag{37}$$

$$t_2 = \frac{\Delta\phi_{12}}{\omega_0} + t_1 < T/2, \tag{38}$$

$$t_3 = \infty > T/,2 \tag{39}$$

$$\phi_s = \phi_1 - (\omega_0 + \epsilon_+)t_1 > \phi_3 - 2\pi. \tag{40}$$

The third inequality for t_3 does not contribute if the other inequalities are satisfied. We thus have three inequalities:

$$T \leqslant \frac{2(2\pi - \Delta\phi_{12})}{\omega_0}, \tag{41}$$

$$T > \frac{2\pi + \epsilon_+\Delta\phi_{12}/\omega_0}{\epsilon_+/2 + \omega_0}, \tag{42}$$

$$T > \frac{2\Delta\phi_{23}}{\omega_0}. \tag{43}$$

It is seen that the locking region does not depend on the absolute values of ϕ_1,ϕ_2,ϕ_3, but only on the separation between them, leaving only two independent parameters that are related to the phase: $\Delta\phi_{12} = \phi_2 - \phi_1$ and $\Delta\phi_{23} = \phi_3 - \phi_2$; the remaining interval is given by $2\pi - \Delta\phi_{13} = 2\pi - (\Delta\phi_{12} + \Delta\phi_{23})$. Shifting the absolute values of ϕ_1,ϕ_2,ϕ_3 only changes the definition of the phase of the clock, not the moments of the day—t_1,t_2,t_3—at which $Z(\phi)$ changes. The system thus has five independent parameters, four related to $Z(\phi)$—$\Delta\phi_{12},\Delta\phi_{23},\epsilon_+,\epsilon_-$—and one being the intrinsic frequency ω_0.

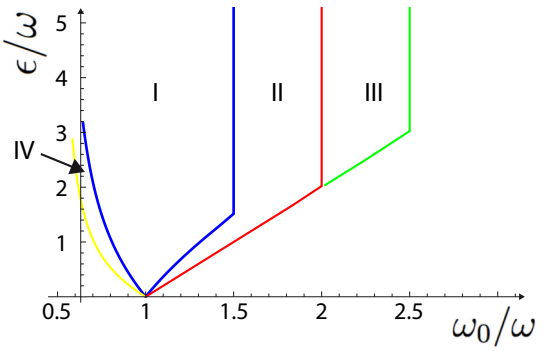

FIG. 2. The Arnold tongue for 1 : 1 locking in the deterministic model, with the coupling strength $\epsilon_+ = \epsilon_- = \epsilon$ in units of the (fixed) frequency of the day-night rhythm ω, plotted as a function of the intrinsic frequency of the clock, ω_0/ω. The different colors correspond to the different scenarios that yield a stable solution. The large region around $\omega_0/\omega = 1$, bounded by the blue lines, corresponds to the Arnold tongue of scenario (1). The adjoining region to the right, with the red boundaries, corresponds to scenario (2). The green lines bound the Arnold tongue of scenario (3), and the yellow lines on the far left yield the Arnold tongue of scenario (4). The other key parameters of $Z(\phi)$ are kept constant: $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$.

In Appendix A, we derive the Arnold tongues for the other scenarios. It turns out that only scenarios (1)–(4) yield stable solutions; the solutions of the other scenarios are unstable.

Figure 2 shows the Arnold tongues for the four scenarios. Since we imagine that the period of the light-day cycle is fixed while the clock can adjust its intrinsic frequency ω_0, we plot the range of $\epsilon = \epsilon_+ = \epsilon_-$ over which the system exhibits a stable deterministic solution, as a function of ω_0/ω; $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$. The different colors correspond to the different scenarios. Clearly, the Arnold tongues of the respective scenarios are adjoining. The region in the middle, around $\omega_0 = \omega$, bounded by the blue lines, corresponds to our natural scenario, i.e., scenario 1, discussed above. The green lines bound the Arnold tongue of scenario 3. This is an unnatural scenario, because in this scenario the clock is driven backwards when the light comes up. Moreover, for $\omega_0/\omega > 2$, the system can also exhibit higher-order locking, which is biologically irrelevant. We will therefore focus on the regime $0.5 < \omega_0 < 2$.

Figure 2 shows that for $\epsilon < 1$ the Arnold tongue exhibits the characteristic increase in its width as the coupling strength is increased: coupling increases the range of frequencies over which the clock can be entrained. However, for $\epsilon > 1$, the width does not change significantly; in fact, it does not change at all when $\omega_0 > \omega$. This is because (a) during the day, for $\epsilon_- = \epsilon = \epsilon_+ > 1$, the phase evolution comes to a halt at ϕ_3—the particle sits in the potential well of Fig. 1(b), and (b) during the night the system evolves with a fixed speed ω_0, independent of ϵ.

Following earlier work [31], the relationship between the Arnold tongue and the phase of entrainment—the phase difference between the Zeitgeber and the clock—has recently been studied in considerable detail [32–34]. The phase difference depends on the frequency mismatch $\omega_0 - \omega$ and the coupling strength and was found to vary for a range of systems by approximately 180° within the Arnold tongue [32–34]. Our work underscores these observations: the phase ϕ_s (which

quantifies the phase difference between clock and input) varies with the intrinsic frequency ω_0 and coupling strengths ϵ_+ and ϵ_- within the Arnold tongue (Fig. 2), changing continuously from one region to the next [see Eqs. (37) and (40) for region I, and Appendix A for other regions]. Interestingly, within regions I, II, and IV, ϕ_s varies for a given coupling strength $\epsilon_+ = \epsilon_- = \epsilon$ from $\phi_3 - 2\pi$ (the left border of region IV, corresponding to the smallest intrinsic frequency ω_0 for which the system exhibits 1 : 1 locking) to ϕ_2 (the right border of region II, corresponding to the highest value of ω_0); ϕ_s is thus excluded from $\phi_2 < \phi_s < \phi_3$ (see Fig. 1). Hence, the range over which ϕ_s varies within I, II, IV is $\phi_2 - \phi_3 + 2\pi$: clearly, for ϕ_3 and ϕ_2 such that $\phi_3 - \phi_2 = \pi$, also in our model the phase difference varies within the Arnold tongue by 180°, supporting the earlier observations [32–34]. In fact, if, additionally, $\phi_2 = \phi_1$, such that there is no dead zone, the scenario with $\phi_3 - \phi_2 = \pi$ corresponds to a symmetric sinusoidal-like phase-response curve, for which the 180° rule has been reported [32].

B. Optimal coupling strength and intrinsic frequency in presence of noise

While the Arnold tongue shows the range of parameters over which the deterministic system can exhibit stable 1 : 1 locking, it does not tell us how reliably the time can be inferred from the phase in the presence of noise. To address this question, we have computed the mutual information $I(\phi;t)$ between the phase of the clock, $\phi(t)$, and the time t. The mutual information captures how the reliability to infer the time depends on the mean input-output relation $\bar{\phi}(t)$, the noise in the system, and the restoring force for deviations away from the mean input-output relation. We have computed $I(\phi;t)$ by performing long stochastic simulations of the system, i.e., stochastically propagating Eq. (25).

Figure 3(a) shows a heat map of the mutual information as a function $\epsilon_+ = \epsilon_- = \epsilon$ and ω_0/ω, for $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$ and $D = 0.1/T$. Superimposed over the heat map are the deterministic Arnold tongues for scenarios (1)–(4), which are also shown in Fig. 2. It is seen that the mutual information is highest in the region bounded by the Arnold tongue of 1 : 1 locking in scenario 1. Interestingly, however, the figure does also show that the mutual information can be large outside of the 1 : 1 locking regimes, especially when $\omega_0/\omega > 2$. This is the result of higher-order locking.

The results of Fig. 3(a) are further elucidated in Figs. 3(b)–3(d), which show the mutual information as a function of ω_0/ω for different values of the diffusion constant D, and for three different values of ϵ/ω, respectively; the results for $D = 0.1/T$ in Figs. 3(b)– 3(d) correspond to three different cuts through the heat map of Fig. 3(a). The following points are worthy of note. First, it can be seen that for each value of ϵ/ω and ω_0/ω the mutual information always increases with decreasing D. Decreasing the noise makes the mapping from the time to the phase of the clock more deterministic, which means that the time can be more accurately inferred from the phase of the clock. Second, it is seen that the mutual information exhibits very characteristic peaks, which result from higher-order locking. For example, the peak at $\omega_0/\omega \approx 2.3$ for $\epsilon = 1.5\omega$, corresponds to 2 : 1 locking.

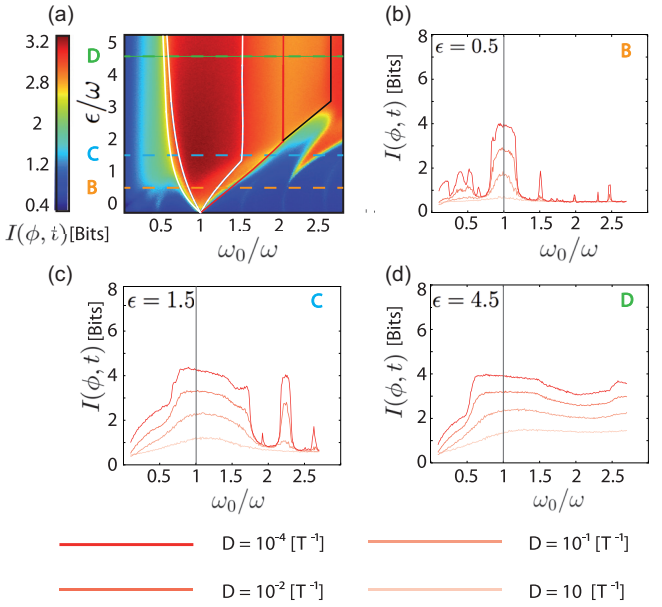

FIG. 3. The mutual information as a function of ϵ, D, and ω_0, keeping $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$. (a) Heat map of the mutual information as a function of ϵ/ω and ω_0/ω for $D = 0.1/T$, respectively. Superimposed are the Arnold tongues for 1 : 1 locking in scenarios (1)–(4). It is seen that the mutual information is high inside the Arnold tongues, with the region corresponding to scenario (1) being the most stable one. The mutual information can, however, also be high outside the 1 : 1 locking regions, because of higher-order locking, especially when $\omega_0/\omega > 2$. (b)–(d) The mutual information as a function of ω_0/ω for different values of the diffusion constant D, and for three values of the coupling strength ϵ/ω, as indicated by the dashed lines in panel (a): $\epsilon/\omega = 0.5$ (b), $\epsilon/\omega = 1.5$ (c), and $\epsilon/\omega = 4.5$ (d). For all values of ϵ, the mutual information increases as D decreases. The peaks outside the main locking region around $\omega_0 \approx \omega$ correspond to higher-order locking.

Figure 3 also shows that, for a given ω_0 and D, the mutual information initially increases with ϵ. This is not surprising, and is consistent with the observation that increasing the coupling strength ϵ tends to widen the Arnold tongue; locking is enhanced by increasing the coupling strength. However, a closer examination of the different panels of Fig. 3 suggests that the mutual information not only saturates as ϵ is increased further, but even goes down. The second surprising observation is that the optimal intrinsic frequency ω_0 that maximizes the mutual information is not equal to ω. In fact, it seems to be smaller than ω when D is small, but then becomes larger than ω as D is increased [Fig. 3(d)].

To elucidate the optimal design of the clock that maximizes the mutual information further, we show in Fig. 4(a) the mutual information $I_{\omega_0^{\rm opt}}(\phi;t)$ that has been obtained by maximizing $I(\phi;t)$ over ω_0 as a function of ϵ, for different values of D. It is seen that for all values of D, $I_{\omega_0^{\rm opt}}(\phi;t)$ first rises with ϵ, as expected. However, $I_{\omega_0^{\rm opt}}(\phi;t)$ then reaches a maximum, after which it comes down: there exists an optimal coupling strength $\epsilon_{\rm opt}$ that maximizes $I_{\omega_0^{\rm opt}}(\phi;t)$; increasing the coupling too much will actually *decrease* the mutual information. Figure 4(a) also shows, however, that the optimal coupling $\epsilon_{\rm opt}$ does increase with the diffusion constant. This is more clearly shown in Fig. 4(b): $\epsilon_{\rm opt}$ increases monotonically with D. This

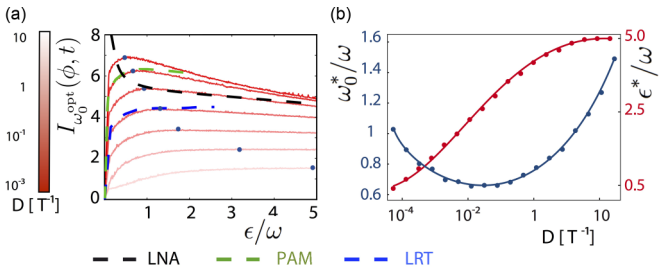

FIG. 4. Optimal design of the clock: parameters $\epsilon = \epsilon_+ = \epsilon_-$ and ω_0 of the phase-response curve $Z(\phi)$ that maximize the mutual information $I(\phi,t)$ as a function of the intrinsic clock noise D, keeping the shape of $Z(\phi)$ constant [see Fig. 1(a)]. (a) The mutual information $I_{\omega_0^{\rm opt}}(\phi;t)$ obtained by maximizing $I(\phi;t)$ over ω_0 as a function of ϵ, for different values of D. It is seen that there is an optimal coupling strength $\epsilon_{\rm opt}$ that maximizes the mutual information, which depends on the magnitude of the diffusion constant D; the blue dot denotes the maximum for each value of D. The figure also shows the predictions of three theories, each for their own regime of validity: the linear-noise approximation (LNA), which captures the regime of strong coupling ϵ and low diffusion D (result shown for $D = 10^{-2}/T$); the phase-averaging method (PAM), which describes the regime of weak coupling and weak noise (result shown for $D = 10^{-3}/T$); and linear-response theory (LRT), which describes the regime of high diffusion and weak coupling (result shown for $D = 1/T$). For a more detailed comparison of the accuracies of the respective theories, see Fig. 8. (b) The optimal coupling strength $\epsilon_{\rm opt}$ (red dots) and the optimal intrinsic frequency $\omega_0^{\rm opt}$ (blue dots), both obtained by maximizing $I(\phi;t)$ over *both* ϵ and ω_0, as a function of D. While $\epsilon_{\rm opt}$ increases with D monotonically, $\omega_0^{\rm opt}$ first decreases from $\omega_0 = \omega$, but then rises again to become larger than ω for higher D. The lines are a guide to the eye. Other parameters: $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$.

panel also shows the optimal intrinsic frequency $\omega_0^{\rm opt}$ obtained by maximizing the mutual information over both ω_0 and ϵ, as a function of D. For $D \to 0$, $\epsilon_{\rm opt}$ goes to zero, and ω_0 to ω—this is the free-running clock. As D is increased, however, ω_0 first decreases, but then increases again to become larger than ω for higher diffusion constants. The optimal intrinsic period that maximizes the mutual information depends in a nontrivial, nonmonotonic manner on the noise in the clock.

C. Optimal design arises from tradeoff between linearity and stability

To understand the optimal design of the clock, we have to recognize that, in general, the amount of information that is transmitted through a communication channel depends on the input distribution, the input-output relation, and on the noise that is propagated to the output. For a given amount of noise, the optimal shape of the input-output relation that maximizes the mutual information is determined by the shape of the input distribution. However, the shape that optimally matches the input-output curve to the input distribution is not necessarily the design that minimizes the noise in the output. Our system provides a clear demonstration of this general principle, and, as we will see, the optimal design of the clock can be understood as arising from a tradeoff between stability, i.e., noise minimization, and linearity, i.e., optimally matching the input-output curve to the statistics of the input.

When the noise is very weak, noise minimization is not important, and optimally matching the input-output curve to the input distribution is paramount. Since the input distribution $p(t)$ is flat, the optimal input-output curve is linear: the average phase $\overline{\phi}(t)$ should increase linearly with time t. This is indeed the solution of the free-running clock, $\phi(t) = \omega_0 t$, and it explains why in the low-noise limit the optimal design is that of an essentially free-running system that is only very weakly coupled to the input.

However, as the noise level is increased, the reliability by which each input signal is relayed becomes increasingly important. Here, a tradeoff could emerge: while increasing the coupling strength ϵ could reduce the noise at the output, which tends to enhance information transmission, it may also distort the input-output curve, pushing it away from its optimal linear shape, decreasing information transmission. Can we capture this tradeoff quantitatively?

To study the tradeoff between linearity and stability, we have computed for each value of ϵ the value of ω_0 that makes the average input-output relation $\overline{\phi}(t)$ most linear, i.e., minimizes $\int_0^T dt[\overline{\phi}(t) - \omega t]^2$. The result is the blue line in Fig. 5(a), which lies in the Arnold tongue of scenario (1). Along this line of maximal linearity, ω_0 decreases as ϵ increases, which can be understood intuitively by noting that increasing ϵ introduces a curvature in the input-output relation, leading to a deviation away from the straight line ωt: at the beginning of the day, till the time t_1 at which the system crosses ϕ_1, the phase evolves with a speed $\omega_0 + \epsilon$, whereas between the time t_2 at which the system crosses ϕ_2 and the end of the day at $T/2$, the phase evolves either following ϕ_2 when $\epsilon = \epsilon_- > \omega_0$ or evolves with a speed $\omega_0 - \epsilon$ when $\epsilon = \epsilon_- < \omega_0$. While increasing ϵ tends to increase the curvature, this effect can be counteracted by decreasing ω_0.

To quantify the stability, we define the return map $F_t(\phi)$:

$$\phi(t+T) = F(\phi(t)) = F_t(\phi), \tag{44}$$

where the subscript t for F indicates that the return map depends on time; this subscript will be suppressed in what follows below when there is no ambiguity, in order to simplify notation. The deterministic solution $\phi^*(t)$ is given by $\phi^*(t) = \phi^*(t+T) = F(\phi^*(t))$. We now expand $F(\phi)$ around $\phi^*(t)$:

$$F(\phi^* + \delta\phi) = F(\phi^*) + F'(\phi^*)\delta\phi, \tag{45}$$

where $\delta\phi = \phi - \phi^*$ and we have dropped the subscript t because $F'(\phi)$, which gives the rate of exponential relaxation back to the limit cycle over many cycles, must be independent of t. Indeed, by exploiting that $F(\phi^*(t)) = \phi^*(t+T)$, we find that

$$\delta\phi(t+T) = F'(\phi^*)\delta\phi(t). \tag{46}$$

The quantity $F'(\phi^*) \equiv \partial F(\phi)/\partial\phi|_{\phi^*} = \partial\phi(t+T)/\partial\phi(t)|_{\phi^*(t)}$ determines the linear stability of the system, with $F' < 1$ meaning that the system is stable. The quantity can be directly obtained from the deterministic solutions. We first note that, since $L(t) = 0$ during the dark, $F'(\phi^*(t=0)) = \partial\phi(T)/\partial\phi(0) = \partial\phi(T/2)/\partial\phi(0)$. For scenario (1), when $\epsilon_- < \omega_0$, $\phi(T/2) = \phi_2 + (\omega_0 - \epsilon_-)(T/2 - t_2)$. We then find that, exploiting Eqs. (38) and (40), $F'(\phi^*(t=0)) = \partial\phi(T/2)/\partial\phi(0) = \partial\phi(T/2)/\partial\phi_s = (\partial\phi(T/2)/\partial t_2)(\partial t_2/\partial t_1)(\partial t_1/\partial\phi_s) = (\omega_0 - \epsilon_-)/(\omega_0 + \epsilon_+)$. Similarly, for scenario (2) we find

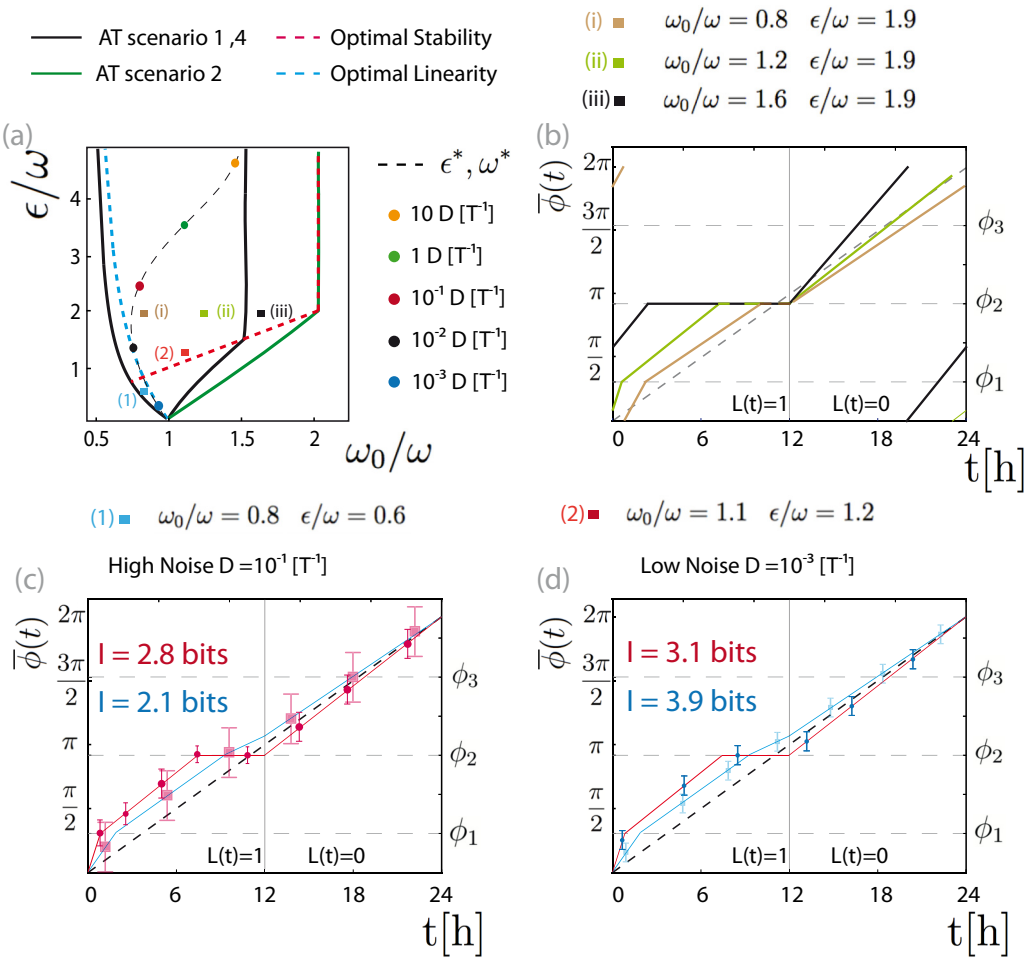

FIG. 5. The optimal design arises from a tradeoff between linearity and stability. (a) The black line shows the Arnold tongue (AT) for scenarios (1) and (4) while the green line shows the Arnold tongue of scenario (2) (see also Fig. 2). The dashed blue line shows for each value of ϵ the value of ω_0 that makes the input-output curve $\overline{\phi}(t)$ most linear, i.e., minimizes $\int_0^T dt[\overline{\phi}(t) - \omega t]^2$. The dashed red line shows for each value of ϵ the value of ω_0 that maximizes the stability. For $\epsilon/\omega < 2$, this line is $\omega_0 = \epsilon$, along which $F' = 0$; for $\epsilon = \epsilon_- > \omega_0$, $F' = 0$ for all values of ω_0 and ϵ; the line of maximal stability then corresponds to the line where the system spends most of its time in ϕ_2, which is the line $\omega_0 = \epsilon$ when $\epsilon < 2\omega$ and $\omega_0 = 2\omega$ when $\epsilon \geqslant \omega$; this is further illustrated in panel (b). The dashed black line shows a parametric plot of the optimal system, i.e., the combination $(\epsilon_{\rm opt}, \omega_0^{\rm opt})$ that maximizes the mutual information as a function of D [values of D along this solid line are indicated by the colored circles; see also Fig. 4(b)]. It is seen that for low diffusion constant, the optimal system that maximizes the mutual information (black line) follows the dashed blue line where the input-output curve is most linear, while for high noise the optimal system moves towards the dashed red line, where the system is most stable. How this tradeoff between linearity and stability maximizes information transmission is further illustrated in panels (c) and (d). Panel (b) shows the average input-output curves for the three points labeled (i), (ii), and (iii) in panel (a). It is seen that as the system moves towards the line of maximal stability, the time the system spends in ϕ_2 increases; for $\epsilon/\omega > 2$, at $\omega_0 = 2\omega$, the system starts the day at ϕ_2. Panel (c) shows the two average input-output curves corresponding to the two points (1) and (2) in panel (a), together with the output noise, for a high value of the diffusion constant, $D = 0.1/T$. Panel (d) shows the same, but then for a low value of the diffusion constant, $D = 10^{-3}/T$. It is seen that when the noise is small (panel d), the output noise of the more stable system (red line) is hardly smaller than that of the more linear system (blue line); consequently, the optimal input-output curve can be linear to maximize information transmission. In contrast, when the noise is large [panel (c)], the system with a more linear input-output curve (the blue line) has significantly more output noise than the more stable but more nonlinear system (red line); in this regime, stability becomes important for taming the output noise, making the optimal system more nonlinear (red line). Other parameters: $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$.

that, for $\epsilon_- < \omega_0$, $F'(\phi^*(t=0)) = (\omega_0 - \epsilon_-)/\omega_0$. Here, we consider the case that $\epsilon_- = \epsilon_+ = \epsilon$. Clearly, in both scenarios the stability is maximized when ϵ approaches ω_0 and $F'(\phi^*)$ becomes zero. This defines the line $\epsilon = \omega_0$, along which $F'(\phi^*) = 0$; it is the part of the red dashed line of maximal stability in Fig. 5(a) that corresponds to $\epsilon < 2\omega$.

For $\epsilon = \epsilon_- > \omega_0$, $F'(\phi^*) = 0$ for both scenarios (1) and (2), because during the day the phase evolution of the system comes to a standstill at ϕ_2; any perturbation in ϕ will fully relax during one period. Can we nonetheless differentiate in the stability strength, even though the *linear* stability $F'(\phi^*) = 0$ for all points (ϵ,ω) above the line $\epsilon = \omega_0$? To answer this question, we turn to a global stability measure, which is defined by the amount of time the deterministic system spends at ϕ_2, which is the bottom of the potential well when $\epsilon = \epsilon_- \geqslant \omega_0$ (see Fig. 1). The value of ω_0 that maximizes the stability for a

given ϵ according to this measure is $\omega_0 = \epsilon$ when $\epsilon \leqslant 2\omega$ and $\omega_0 = 2\omega$ when $\epsilon \geqslant 2\omega$. This fully specifies the line of maximum stability shown in Fig. 5(a). The reason why the stability is maximized along this line is illustrated in Fig. 5(b). During the night, the trajectories evolve freely, and because of noise they will arrive at the beginning of the day with a distribution of phases. Along the line of maximum stability, the stochastic trajectories are most likely to reach the bottom of the potential well at ϕ_2 during the day (see Fig. 1), where they will be confined before they are released again during the night. Indeed, along this stability line the variance in the phase, $\langle \delta\phi^2 \rangle$, will be lowest which tends to increase information transmission. However, the input-output relation $\overline{\phi}(t)$ is then highly nonlinear. In fact, the globally most stable solution, for all possible values of ϵ and ω_0, is

$$\phi^{\text{stab}}(t) \equiv \phi_2 \theta(T/2 - t) + \omega_0 t \theta(t - T/2), \text{ with } \omega_0 = 2\omega, \tag{47}$$

which is the most stable solution for any $\epsilon \geqslant 2\omega$. It is shown in Fig. 5(b)—it is the solution at the high-frequency boundary of the Arnold tongue of scenario (2). This solution maximizes the probability that trajectories that start at the limit cycle at the beginning of the day will return to the limit cycle ϕ_2 before the end of the day. While this solution is maximally stable, no time points t can be inferred from $\phi(t)$ during the day, because $\overline{\phi}(t)$ is completely flat. This dramatically reduces information transmission.

The optimal values of ω_0 and ϵ that maximize the mutual information as a function of the noise in the system can now be understood as a tradeoff between linearity and stability. This tradeoff is illustrated in the bottom panels of Fig. 5, which show the average input-output curves, together with their output noise, for the two points 1 and 2 in the map of Fig. 5(a), both for a high diffusion constant [Fig. 5(c)] and a low diffusion constant [Fig. 5(d)]. When the diffusion constant is low [Fig. 5(d)], the noise in the more stable but more nonlinear system (red line, corresponding to point 2) is hardly lower than that in the more linear but less stable system (blue line, corresponding to point 1), which means that the benefit of linearity dominates and the mutual information is maximized in the more linear system. In contrast, when the noise is larger [Fig. 5(c)], the output noise in the more stable but more nonlinear system (red line) is so much smaller than that in the less stable but more linear system (blue line) that it outweighs the cost of higher nonlinearity, thus maximizing mutual information.

Finally, Fig. 5(a) also shows a parametric plot of the optimal (ϵ,ω_0) that maximizes the mutual information, with the noise D the parameter that is being varied (dashed black line; the colors of circles denote values of the diffusion constant). It is seen that for low D the optimal system traces the dashed blue line of maximal linearity, but then at a higher D makes a transition towards the dashed red line line of maximal stability.

D. Optimal shape of the phase-response curve

In the previous section, we showed how the optimal values of the coupling strength ϵ and the intrinsic frequency ω_0 depend on the noise D in the system, while keeping the shape of the coupling function $Z(\phi)$ constant. In this section, we will relax this restriction.

We first checked the effect of changing the magnitude of the positive and negative lobe of the coupling function $Z(\phi)$ as characterized by ϵ_+ and ϵ_-, respectively (see Fig. 1), keeping $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$ constant. We varied ϵ_+ and ϵ_- via a parameter α, defined as $\epsilon_+ = (1-\alpha)\epsilon$ and $\epsilon_- = \alpha\epsilon$; changing α thus keeps the total absolute coupling strength (the integrated modulus) constant. We found, however, that the results are not very sensitive to the precise values of ϵ_+ and ϵ_- (see Appendix D).

We then decided to compute the mutual information $I(\phi,t)$ as a function of $\Delta\phi_{12}$ and $\Delta\phi_{23}$ for different values of ϵ, ω_0, and D, keeping $\epsilon_+ = \epsilon_- = \epsilon$. We found that the mutual information is essentially independent of $\Delta\phi_{23}$. This can be understood as follows: The deterministic Arnold tongue, and, to a good approximation, the dynamics of the stochastic system, does not depend on the absolute values of ϕ_1,ϕ_2,ϕ_3, but only on $\Delta\phi_{12}$ and $\Delta\phi_{23}$ (see Sec. III A). Moreover, as long as ϕ_3 is crossed during the night (see Fig. 1), we can change ϕ_3 at will, because during the night, when $L(t) = 0$, the clock is not coupled to light [see Eq. (25)], meaning that the clock runs with its intrinsic frequency ω_0. Changing $\Delta\phi_{23}$ by changing ϕ_3 will thus have no effect. Changing $\Delta\phi_{23}$ by changing ϕ_2 will also have no effect when ϕ_1 is simultaneously changed such that $\Delta\phi_{12}$ remains constant: while changing ϕ_2 and ϕ_1 keeping $\Delta\phi_{12}$ and ϕ_3 constant will alter $\Delta\phi_{23}$, we can always change ϕ_3 such that $\Delta\phi_{23}$ remains unchanged. In short, as long as ϕ_3 is crossed during the night (which it will be for most values of ϕ_1 and ϕ_2), changing ϕ_1 and ϕ_2 keeping $\Delta\phi_{12}$ constant does not change the dynamics; the times t_1 and t_2 at which ϕ_1 and ϕ_2 are crossed, respectively, do not change.

Because ϕ_{23} is not critical, we kept $\Delta\phi_{23} = \pi/2$, and then performed very extensive simulations to determine the optimal coupling strength ϵ^*, speed ω_0^*, and optimal dead zone $\Delta\phi_{12}^*$ that maximize the mutual information, as a function of D. Figure 6 shows a parametric plot of $\epsilon^*(D)$, $\omega_0^*(D)$, and $\Delta\phi_{12}^*(D)$, with D being the parameter that is varied. It is seen that for very low D, the optimal coupling strength ϵ^* is small, the optimal intrinsic frequency ω_0^* is close to ω, and the optimal value of $\Delta\phi_{12}^*$ is small. As the diffusion constant is increased, ϵ^* rises but ω_0^* initially remains close to ω and then increases too. The optimal value of $\Delta\phi_{12}$, however, first rises and then falls again.

The behavior of $\Delta\phi_{12}^*$ can again be understood as a tradeoff between linearity and stability. This is illustrated in Fig. 7. The figure shows for different values of ϵ the linearity and the stability of the input-output relation $\overline{\phi}(t)$ as a function of $\Delta\phi_{12}$ and ω_0, computed within the deterministic Arnold tongue of scenario (1) (where the mutual information is highest). The linearity of $\overline{\phi}(t)$ is quantified via $\int_0^T dt[\overline{\phi}(t) - \phi^{\text{lin}}(t)]^2$, which is the average deviation of $\overline{\phi}(t)$ away from the most linear input-output relation, $\phi^{\text{lin}}(t) = \omega t$. The stability of $\overline{\phi}(t)$ is quantified via $\int_0^T dt[\overline{\phi}(t) - \phi^{\text{stab}}(t)]^2$, which is the average deviation of $\overline{\phi}(t)$ away from the most stable input-output relation $\phi^{\text{stab}}(t)$, given by Eq. (47).

The following observations can be made. First, the width of the Arnold tongue (the range of ω_0 that permits a deterministic

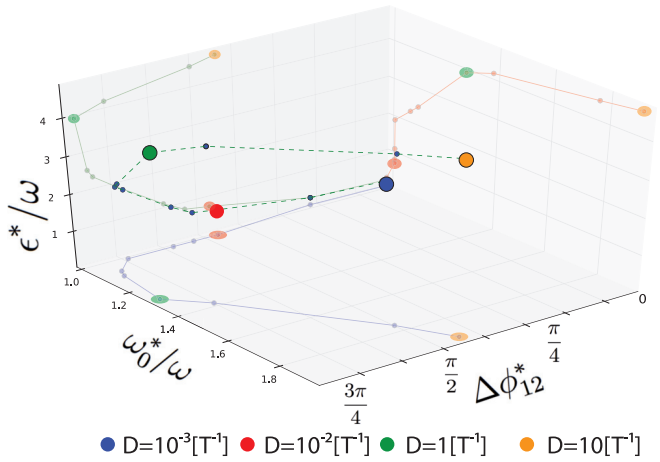

FIG. 6. A parametric plot of the optimal coupling strength $\epsilon^*(D)$, the optimal intrinsic frequency $\omega_0^*(D)$, and the optimal width of the dead zone $\Delta\phi_{12}^*(D)$ that maximize the mutual information, with the noise D being the parameter that is varied. The value of $\Delta\phi_{23} = \pi/2$ was kept constant. It is seen that ϵ^* rises with D, while ω_0^* remains initially close to ω, but then rises too. In contrast, $\Delta\phi_{12}^*$ first increases and then decreases. Colored dots give the diffusion constants for which $(\epsilon^*, \omega_0^*, \Delta\phi_{12}^*)$ are optimal.

solution) decreases as $\Delta\phi_{12}$ increases. Second, the linearity is maximal in the range $1 < \omega_0/\omega < 1.5$, and tends to increase with $\Delta\phi_{12}$: in the dead zone $\Delta\phi_{12}$ the system evolves freely with speed ω_0, which makes $\overline{\phi}(t)$ more linear, especially when $\omega_0 \sim \omega$. In contrast, the stability is highest when ω_0/ω is large and $\Delta\phi_{12}$ is small, particularly for higher values of ϵ. The large magnitudes of ω_0 and ϵ mean that at the beginning of the day the system is strongly driven, $\langle\dot{\phi}\rangle \approx \epsilon + \omega_0$, and the small dead zone $\Delta\phi_{12}$ means that after the system has crossed ϕ_1, it quickly reaches ϕ_2, where, with $\epsilon = \epsilon_- > \omega_0$, the system is then confined (see Fig. 1).

Figure 7 also shows superimposed a parametric plot of the optimal $\Delta\phi_{12}^*(D)$ against the optimal $\omega_0^*(D)$. The colored dots denote the diffusion constants for which $(\omega_0^*, \Delta\phi_{12}^*)$ are optimal; the diffusion constant for which the ϵ of a panel is the optimal coupling strength ϵ^* is shown near the top of the Arnold tongue. It is seen that for very small D, the optimal system parameters $(\omega_0^*, \Delta\phi_{12}^*, \epsilon^*)$ put the system in the regime where $\phi(t)$ is linear [top left panel (a)]; increasing $\Delta\phi_{12}^*$ would not make the system significantly more linear, since ϵ^* is still very small. Increasing D raises ϵ^*, while ω^* remains close to ω. The optimal width of the dead zone $\Delta\phi_{12}^*$ now increases, because for the higher value of ϵ^* the system becomes significantly more linear when $\Delta\phi_{12}^*$ is increased. Beyond $D = 1/T$, however, linearity is sacrificed for stability. The optimal coupling strength ϵ^* and intrinsic frequency ω^* increase, while the optimal size of the dead zone decreases, to maximize stability. Indeed, when the noise is even larger still, the width of the dead zone reduces to zero and the coupling strength and intrinsic frequency become even larger: during the day the system is rapidly driven to ϕ_2, where it then remains

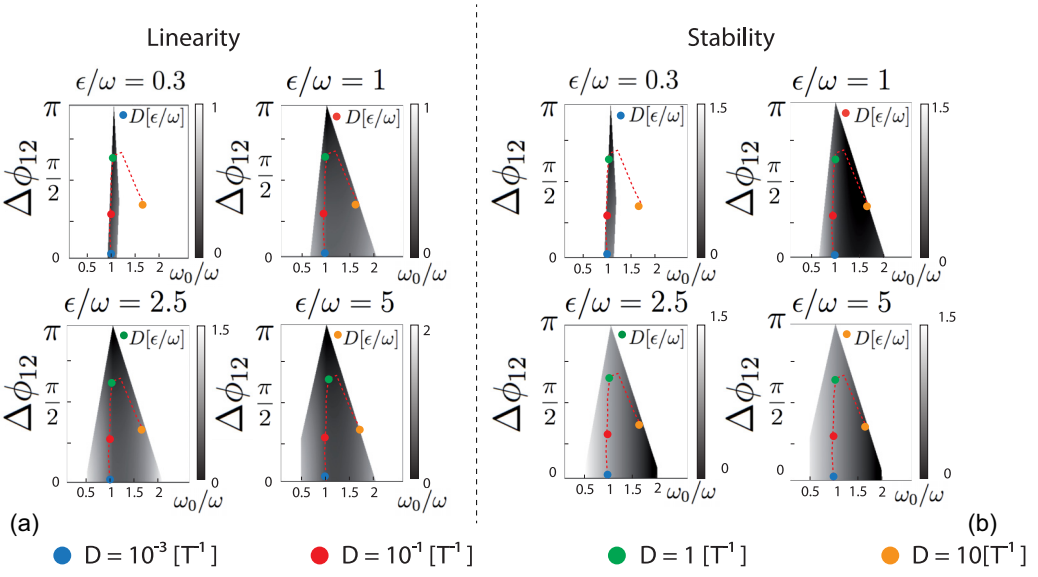

FIG. 7. The optimal shape of the instantaneous phase-response curve $Z(\phi)$ arises as a tradeoff between linearity and stability. The linearity (a) is quantified via $\int_0^T dt[\overline{\phi}(t) - \phi^{\text{lin}}(t)]^2$, which is the average deviation of the mean input-output relation $\overline{\phi}(t)$ away from the most linear solution $\phi^{\text{lin}}(t) = \omega t$. The stability (b) is quantified via $\int_0^T dt[\overline{\phi}(t) - \phi^{\text{stab}}(t)]^2$, which is the average deviation of $\overline{\phi}(t)$ away from the most stable solution $\phi^{\text{stab}}(t)$, given by Eq. (47). These measures are computed as a function of the intrinsic frequency ω_0 and the width of the dead zone $\Delta\phi_{12}$, for different values of ϵ, inside the Arnold tongue of scenario (1); note that smaller values correspond to higher linearity and stability, respectively. Superimposed is a parametric plot of the optimal intrinsic frequency $\omega_0^*(D)$ and optimal width of the dead zone $\Delta\phi_{12}^*(D)$ that maximize the mutual information for a given D. The dots denote the values of D to which $\omega_0^*(D)$ and $\Delta\phi_{12}^*(D)$ correspond; the value of D for which the ϵ of a panel is the optimal coupling strength ϵ^* is given near the top of the Arnold tongue. It is seen that for small D, the optimal parameters $(\omega_0^*(D), \Delta\phi_{12}^*(D), \epsilon^*(D))$ that maximize the mutual information are those that make the input-output relation $\overline{\phi}(t)$ most linear [top left panel (a)], while for large D, the optimal parameters are those that make the system very stable [bottom right panel (b)]. Other parameters: $\Delta\phi_{23} = \pi/2$.

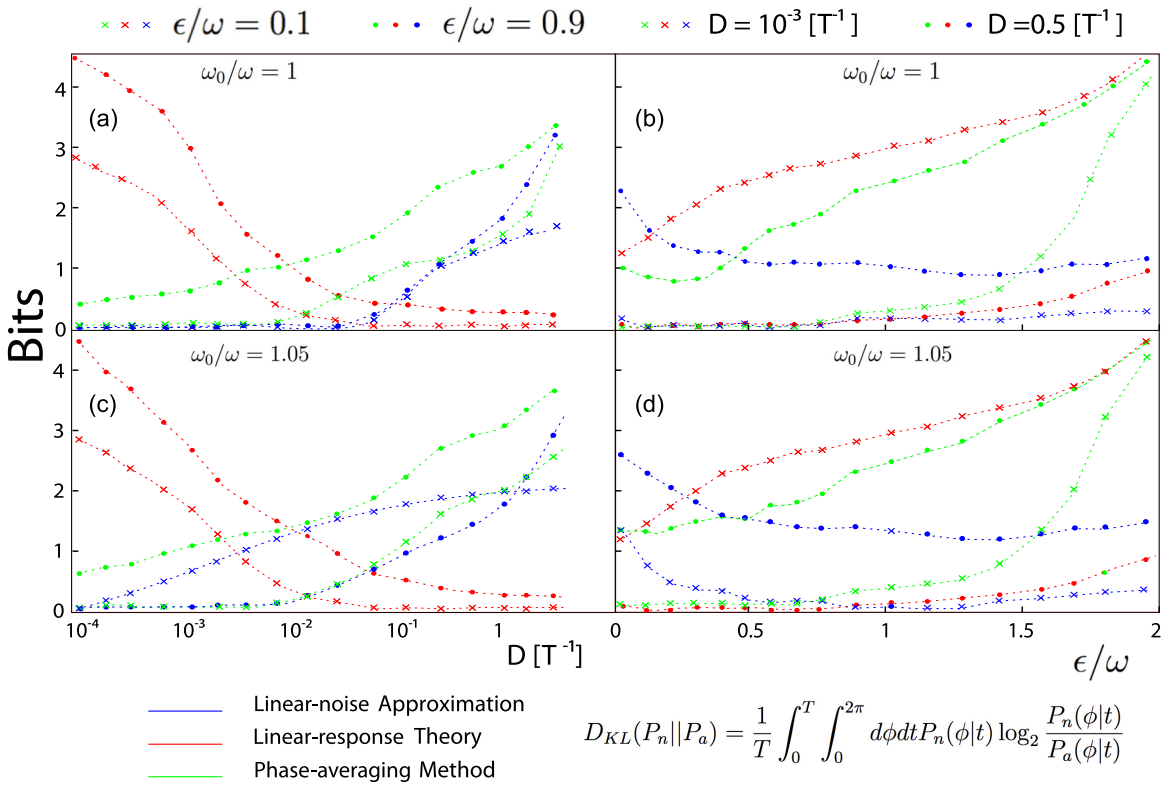

FIG. 8. Comparison between simulation results and three different theories: linear-noise approximation (LNA), phase-average method (PAM), and linear-response theory (LRT). The comparison is performed by computing the Kullback-Leibler divergence $D_{KL}(P_n||P_a)$ between $P_n(\phi|t)$ as obtained in the simulations and $P_a(\phi|t)$ as predicted by the theory. For two values of ω_0, namely $\omega_0/\omega = 1$ [panels (a) and (b)] and $\omega_0/\omega = 1.05$ [panels (c) and (d)], we show $D_{KL}(P_n||P_a)$ as a function of D for two values of ϵ [panels (a) and (c)] and $D_{KL}(P_n||P_a)$ as a function of ϵ for two values of D [panels (b) and (d)]. It is seen that the LNA accurately predicts the regime of strong coupling and low noise, PAM the regime of weak coupling and weak noise, and LRT the regime of high noise and weak coupling. Other parameters: $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$ for all data points.

strongly confined till the beginning of the night [see Fig. 1 and also Fig. 5(c)]. In this limit, the clock transmits one bit of information, and the system can only distinguish between day and night.

Figure 6 thus generalizes the finding of Fig. 5 that corresponds to a fixed dead zone and shows that the optimal shape of the instantaneous phase-response curve can be understood as a tradeoff between linearity and stability.

IV. THEORY

The simulation results can be described quantitatively via three different theories, which each accurately describes a particular regime of parameters: The linear-noise approximation (LNA) describes the regime of strong coupling and low diffusion; the phase-averaging method (PAM) holds in the low diffusion, weak coupling regime; and the linear-response theory (LRT) applies in the regime of high noise and weak coupling. Here, we have borrowed the terminology LNA from the name of the theory to describe biochemical networks that is based on the same underlying principles: indeed, rather than linearizing the chemical Langevin equation around the fixed point given by the mean-field chemical rate equations and taking the noise at that fixed point, we here linearize the return map $F(\phi)$ around its fixed point, and compute the noise at that fixed point. The results of the respective theories in their regime of validity are shown in Fig. 4. A more detailed comparison between the simulation results and the theoretical predictions, discussed below, is shown in Fig. 8, where ϵ and D are varied for two different values of ω_0.

A. Linear-noise approximation

The linear-noise approximation (LNA) is expected to be accurate when the driving is strong compared to the diffusion constant, so that the system closely follows the deterministic solution $\phi^*(t)$, which is given by the return map of Eq. (44): $\phi^*(t) = \phi^*(t+T) = F(\phi^*(t))$. Because in this regime the deviations from the deterministic solution are small, we can expand $F(\phi)$ up to linear order in $\delta\phi = \phi - \phi^*$ to obtain $F(\phi^* + \delta\phi)$; see Eq. (45). This makes it possible to derive how a deviation from the deterministic solution at time t will relax to the limit cycle at time $t+T$: $\delta\phi(t+T) = F'(\phi^*)\delta\phi(t)$ [see Eq. (46)]. The quantity $F'(\phi^*)$ thus determines the stability of the system near the deterministic fixed point. It can be readily obtained from the deterministic solutions.

Given a variance at time t, $\langle\delta\phi(t)^2\rangle$, the variance at time $t+T$, $\langle\delta\phi(t+T)^2\rangle$, is given by two contributions:

$$\langle\delta\phi(t+T)^2\rangle = F'^2(\phi^*)\langle\delta\phi(t)^2\rangle + V[\phi(t+T)|\phi^*(t)]. \quad (48)$$

The first contribution is a deterministic contribution, which is determined by how a deviation $\delta\phi(t) = \phi(t) - \phi^*(t)$ at

time t regresses deterministically to the mean at time $t+T$: $\delta\phi(t+T) = F'(\phi^*)\delta\phi(t)$. The second contribution describes the variance of the distribution $P(\phi(t+T)|\phi^*(t))$ of $\phi(t+T)$ at time $t+T$, given that at time t the system was at the deterministic solution $\phi^*(t)$; in general, we should instead compute the variance at $t+T$ for an arbitrary initial $\phi(t) = \delta\phi(t) + \phi^*(t)$, but to leading order in small $\delta\phi$ it is sufficient to evaluate the noise at the deterministic solution ϕ^*. It is important to note that the variance $V[\phi(t+T)|\phi^*(t)]$ depends not only on the diffusion constant, but also on the deterministic force, as in a canonical LNA description: For example, in the simplest possible noisy dynamics, $\dot{\delta x} = -k\delta x(t) + \eta(t)$, with $\langle\eta(t)\eta(t')\rangle = 2D\delta(t-t')$, the deterministic contribution to the variance $\langle\delta x(t+T)^2\rangle$ at time $t+T$, given the variance $\langle\delta x(t)^2\rangle$ at time t, is $\langle\delta x(t)^2\rangle e^{-2kT}$, while the stochastic contribution to the variance at time $t+T$ is $V[\delta x(t+T)|x^*(t)] = (D/k)(1-e^{-2kT})$, which indeed depends on the force constant k. However, in the limit that the force is weak, the stochastic contribution is given by the variance of free diffusion: $V[\delta x(t+T)|x^*(t)] = 2DT$. We assume, and subsequently verify numerically, that a similar simplification applies for our phase oscillator model. Indeed, except at the boundaries ϕ_1, ϕ_2, and ϕ_3, our phase dynamics reduces to diffusion with a constant drift, for which it is rigorously true that $V[\phi(t+T)|\phi^*(t)] = 2DT$; our assumption hence amounts to neglecting any corrections to the integrated noise due to the brief "kicks" at these boundaries. Equation (48) then reduces to

$$\langle\delta\phi(t+T)^2\rangle = F'^2(\phi^*)\langle\delta\phi(t)^2\rangle + 2DT. \tag{49}$$

This expression constitutes the fluctuation-dissipation relation for this system. In steady state, $\langle\delta\phi(t+T)^2\rangle = \langle\delta\phi(t)^2\rangle$, from which it follows that

$$\langle\delta\phi(t)^2\rangle = \frac{2DT}{1-F'^2(\phi^*)}. \tag{50}$$

Clearly, the variance depends not only on the diffusion constant, but also on the stability, which increases with the coupling strength; as derived below Eq. (46), for scenario (1), $F'(\phi^*) = (\omega_0-\epsilon_-)/(\omega_0+\epsilon_+)$ decreases (meaning the system becomes more stable) as ϵ_- and ϵ_+ increase.

In this linear-noise approximation, the distribution of the phase at time t is a simple Gaussian with a mean $\overline{\phi}(t)$ that is given by the deterministic solution, $\overline{\phi}(t) = \phi^*(t)$, and a variance that is given by Eq. (50):

$$P(\phi|t) = \frac{1}{\sqrt{2\pi}\sigma_\phi}\exp-\frac{[\phi-\overline{\phi}(t)]^2}{2\sigma_\phi^2}, \tag{51}$$

where $\sigma_\phi \equiv \sqrt{\langle\delta\phi^2\rangle}$. This variance is, in this approximation, independent of the phase.

To derive the mutual information, it is convenient to invert the problem and look for the distribution of possible times t, given ϕ. This can be obtained from Bayes' rule:

$$P(t|\phi) = P(t)\frac{P(\phi|t)}{P(\phi)}, \tag{52}$$

where $P(t) = 1/T$ is the uniform prior probability of having a certain time and $P(\phi)$ is the steady-state distribution of ϕ, which in the small noise limit can be computed via $P(t)dt = P(\phi)d\phi$. If the noise ξ is small compared to the mean, then $P(t|\phi)$ will be a Gaussian distribution that is peaked around $t^*(\phi)$, which is the best estimate of the time given the phase [17,24,35]:

$$P(t|\phi) \simeq \frac{1}{\sqrt{2\pi\sigma_t^2}}\exp\left[-\frac{[t-t^*(\phi)]^2}{2\sigma_t^2}\right]. \tag{53}$$

Here $\sigma_t^2 = \sigma_t^2(t^*)$ is the variance in the estimate of the time, and it is given by [17]

$$\sigma_t^2 = \sigma_\phi^2\left(\frac{dt}{d\overline{\phi}}\right)^2. \tag{54}$$

We note that σ_t^2 does depend on t because the slope $d\overline{\phi}/dt$ depends on t. Indeed, while the LNA assumes that σ_ϕ^2 is independent of ϕ, it does capture the fact that changing ϵ and ω_0 can affect the mutual information not only by changing the noise σ_ϕ^2 but also via the slope $d\overline{\phi}/dt$ of the input-output relation $\overline{\phi}(t)$.

The mutual information can now be obtained from

$$I(\phi;t) = H(t) - \langle H(t|\phi)\rangle_\phi \tag{55}$$

$$= \log_2 T - \left\langle\frac{1}{2}\log_2\left(2\pi e\sigma_\phi^2\left(\frac{dt}{d\overline{\phi}}\right)^2\right)\right\rangle_\phi \tag{56}$$

$$= \log_2\left(\frac{T}{\sqrt{2\pi e\sigma_\phi^2}}\right) + \frac{1}{T}\int_0^T dt\log_2\frac{d\overline{\phi}}{dt}, \tag{57}$$

where $\langle\ldots\rangle_\phi$ denotes an average over $P(\phi)$, and we have exploited that in the LNA the variance σ_ϕ^2 is independent of ϕ. For the model presented here, $\overline{\phi}(t) = \phi^*(t)$ is piecewise linear, and the second integral can be obtained analytically, for each of the scenarios; for scenario (1), for example, the second term is $1/T[t_1\log_2(\omega_0+\epsilon_+) + (t_2-t_1)\log_2\omega_0 + (T/2-t_2)\log_2(\omega_0-\epsilon_-) + T/2\log_2\omega_0]$.

Figure 4 shows that the LNA accurately predicts the mutual information $I_{\omega_0^{\rm opt}}(\phi;t)$ in the regime that the coupling strength ϵ is large and the diffusion constant D is small. A more detailed comparison is shown in Fig. 8, which shows the Kullback-Leibler divergence $D_{KL}(P_n||P_a)$ between the distribution $P_n = P_n(\phi|t)$ obtained in the simulations and $P_a = P_a(\phi|t)$ as predicted by LNA. Figures 8(a) and 8(b) show the result for $\omega_0/\omega = 1$, while Figs. 8(c) and 8(d) show the results for $\omega_0/\omega = 1.05$. Moreover, Figs. 8(a) and 8(c) show the results as a function of D for two values of ϵ, while Figs. 8(b) and 8(d) show the results as a function of ϵ for two values of D.

Figures 8(a) and 8(c) show that as D is decreased at fixed ϵ, the LNA becomes accurate for small D, as expected. Figures 8(b) and 8(d) show that for large D, the LNA never becomes accurate, even for large ϵ. However, for large values of ϵ, the assumption that the stochastic contribution to the variance is given by that of free diffusion, $V[\delta\phi(t+T)|\phi^*(t)] \simeq 2DT$, breaks down. This is also the reason why for the smaller value of D [crosses in Figs. 8(b) and 8(d)], the LNA works very well for low values of ϵ, but then becomes slightly less accurate for higher values of ϵ. Indeed, for $\epsilon = \epsilon_- > \omega_0$, $F' = 0$, and the key assumption of LNA—namely that the dynamics can

be expanded to linear order around the deterministic fixed point—breaks down.

Comparing Fig. 8(c) against Fig. 8(a) and Fig. 8(d) against Fig. 8(b) shows that LNA is less accurate in the small D (ϵ) regime when $\omega_0/\omega = 1.05$ [Fig. 8(c)]([Fig. 8(d)]) than when $\omega_0/\omega = 1.0$ [Fig. 8(a)]([Fig. 8(b)]). More specifically, while LNA is very accurate for $D < 10^{-2}/T$ for both values of ϵ when $\omega_0/\omega = 1.0$ [Fig. 8(a)], LNA becomes less accurate for $D < 10^{-2}/T$ when $\omega_0/\omega = 1.05$ and ϵ is small, i.e., $\epsilon/\omega = 0.1$ [Fig. 8(c)]; only for $\epsilon/\omega = 0.9$ is LNA still accurate in this regime. Similarly, while LNA is very accurate for $\epsilon/\omega < 1$ when $D = 10^{-3}/T$ and $\omega_0/\omega = 1.0$ [Fig. 8(b)], LNA becomes less accurate for $\epsilon/\omega < 0.5$ when $D = 10^{-3}/T$ yet $\omega_0/\omega = 1.05$ [Fig. 8(d)]. This observation can be understood by noting that when ω_0 is increased, the system moves to the boundary of the Arnold tongue of scenario (1), especially when ϵ is small (see Fig. 2). The system then switches under the influence of noise between the solution of scenario (1) and that of scenario (2), meaning that the response becomes nonlinear and LNA breaks down. Interestingly, however, another method, described in the next section, accurately describes this regime.

B. Phase-averaging method

In the limit that the coupling ϵ is weak, the diffusion constant D is small, and the intrinsic frequency ω_0 is close to the driving frequency ω, we expect that the evolution of ϕ is close to that of the free-running oscillator, $\phi_0(t) = \omega_0 t + \phi_0$. In this regime the phase will exhibit fluctuations that are slow, occurring on time scales much larger than the intrinsic period T_0. The detailed coupling within a clock cycle becomes irrelevant, and only the average coupling over a clock period matters. This leads to the notion of phase averaging, in which $P(\phi(t) - \omega t|t)$ no longer depends on t: $P(\phi(t) - \omega t|t) = P(\phi(t) - \omega t) \equiv P(\psi)$, with $\psi \equiv \phi(t) - \omega t$.

Following Pikovsky [1], we now make this intuitive notion concrete by rewriting the coupling term as

$$Q(\phi,t) = Z(\phi)L(t) \tag{58}$$

$$= \sum_k \sum_l a_k b_l e^{i(k\phi + l\omega t)}. \tag{59}$$

If the coupling and the noise are weak, $\epsilon \to 0, D \to 0$, we may expect that $\phi \simeq \omega_0 t + \phi_0$ for all times t. If we substitute this into Eq. (59), we find

$$Q(\phi,t) = \sum_k \sum_l a_k b_l e^{ik\phi_0} e^{i(k\omega_0 + l\omega)t}. \tag{60}$$

When $\omega \approx \omega_0$, the terms $k = -l$ contribute most strongly to the integral. These terms correspond to variations in the force on long time scales. We thus expect that in the regime that $\epsilon, D \to 0$ and $\omega \approx \omega_0$, where the phase is expected to follow $\phi \approx \omega_0 t + \phi_0$, the terms $k = -l$ yield the strongest contributions to the force:

$$Q(\phi,t) = \sum_k a_k b_{-k} e^{ik(\phi - \omega t)} \tag{61}$$

$$= \int_0^T dt' Z(\psi + \omega t') L(t') \tag{62}$$

$$= Q(\psi), \tag{63}$$

where in Eq. (62) we have introduced the new phase variable $\psi \equiv \phi - \omega t$. The force $Q(\psi)$ is commonly referred to as the phase-response curve; it is thus a convolution of the instantaneous phase-response curve $Z(\phi)$ and the light-signal $L(t)$.

The temporal evolution of ψ, $\dot{\psi} = \dot{\phi} - \omega$, is, using Eq. (25),

$$\frac{d\psi}{dt} = \omega_0 - \omega + \epsilon Q(\psi) + \xi(t) \tag{64}$$

$$= -\nu + \epsilon Q(\psi) + \xi(t), \tag{65}$$

with $\nu = \omega - \omega_0$. The first two terms on the right-hand side are the deterministic force, which can be written as the derivative of a potential $V(\psi)$,

$$-\nu + \epsilon Q(\psi) = -\frac{dV(\psi)}{d\psi}, \tag{66}$$

with the potential given by

$$V(\psi) = \nu\psi - \epsilon \int_{-\pi}^{\psi} Q(x)dx. \tag{67}$$

Indeed, the evolution of ψ can be described as that of a particle in a potential $V(\psi)$, which is a 2π-periodic potential with a slope given by $\nu = \omega - \omega_0$.

The evolution of the probability density $P(\psi,t)$ is given by the Fokker-Planck equation corresponding to Eq. (65):

$$\partial_t P(\psi,t) = -\partial_\psi\{[-\nu + \epsilon Q(\psi)]P(\psi,t)\} + D\partial_\psi^2 P(\psi,t) \tag{68}$$

$$= -\frac{\partial J(\psi,t)}{\partial \psi}, \tag{69}$$

where we have defined the probability current

$$J(\psi,t) = -P(\psi,t)\frac{dV(\psi)}{dx} - D\frac{\partial P(\psi,t)}{\partial t}. \tag{70}$$

In steady state, $\partial P(\psi,t)/\partial t = 0$, which yields the following stationary solution that is 2π-periodic in ψ:

$$\overline{P}(\psi) = \frac{1}{C}\int_{\psi}^{\psi+2\pi} e^{[V(\psi') - V(\psi)]/D} d\psi'. \tag{71}$$

Here, C is the normalization constant.

Figure 4 shows that the phase-averaging method (PAM) accurately predicts the mutual information $I(\phi;t)$ in the regime that both the coupling strength ϵ and the diffusion constant D are small. The more detailed comparison based on the Kullback-Leibler divergence $D_{KL}(P_n||P_a)$ between the distribution $P_n = P_n(\phi|t)$ obtained in the simulations and $P_a = P_a(\phi|t)$ as predicted by PAM confirms this interpretation: as shown in Fig. 8(b), when $\omega_0/\omega = 1.05$, PAM is accurate for $D < 10^{-2}/T$ when $\epsilon/\omega = 0.1$ (green crosses), while LNA breaks down in this regime (blue crosses). Similarly, as illustrated in Fig. 8(d), when $\omega_0/\omega = 1.05$, PAM is accurate for $\epsilon/\omega < 0.7$ when $D = 10^{-3}/T$ (green crosses), whereas LNA again breaks down in this regime (blue crosses).

While the LNA breaks down when the distribution $P(\phi|t)$ becomes non-Gaussian as the coupling becomes too weak, the PAM accurately describes $P(\phi|t)$ in the low-coupling, low-noise regime, as it allows for non-Gaussian distributions. However, the PAM does assume that $\phi(t)$ follows ωt. As a result it breaks down when the coupling becomes large, causing the average input-output relation $\bar{\phi}(t)$ to deviate markedly from ωt, an effect that can be captured by the LNA. PAM also breaks

down when ϵ is small and $\omega \approx \omega_0$, yet D is large: now the large diffusion constant causes the instantaneous $\phi(t)$ to deviate markedly from ωt. This regime can, however, be described by linear-response theory.

C. Linear-response theory

When the coupling strength is weak yet the diffusion constant is large, $\phi(t)$ at any moment in time will tend to deviate strongly from $\omega_0 t$, but the steady-state distribution will be close to that of a noisy, free-running oscillator, $P_0(\phi) = 1/(2\pi)$. The full distribution can then be obtained as a perturbation to this distribution. This is the central idea of linear-response theory (LRT).

We start with the Fokker-Planck equation for the evolution of $P(\phi,t)$:

$$\partial_t P(\phi,t) = D\partial_\phi^2 P(\phi,t) + \omega_0 \partial_\phi P(\phi,t) + L(t)\partial_\phi[Z(\phi)P(\phi,t)]. \tag{72}$$

We now consider the external signal $L(t)Z(\phi)$ to be a weak perturbation of the free-running system. To this end, we rewrite the above equation as

$$\partial_t P(\phi,t) = [\mathcal{F}_0 + \epsilon \mathcal{F}_1(t)]P(\phi,t), \tag{73}$$

where $\mathcal{F}_0$ is the operator that defines the time evolution of the unperturbed system and $\mathcal{F}_1$ is that due to the perturbation:

$$\mathcal{F}_0 = +D\partial_\phi^2 + \omega_0\partial_\phi, \tag{74}$$

$$\mathcal{F}_1(t) = +L(t)\partial_\phi Z(\phi) + L(t)Z(\phi)\partial_\phi. \tag{75}$$

Furthermore, we expand $P(\phi,t)$ as

$$P(\phi,t) \simeq p_0(\phi,t) + \epsilon p_1(\phi,t) + \epsilon^2 p_2(\phi,t) + \mathcal{O}(\epsilon^3). \tag{76}$$

Substituting this expression into Eq. (73), and keeping only terms up to order ϵ, we find

$$\mathcal{O}(0): \quad \mathcal{F}_0 p_0(\phi,t) = \partial_t p_0(\phi,t), \tag{77}$$

$$\mathcal{O}(\epsilon): \quad \partial_t p_1(\phi,t) - \mathcal{F}_0 p_1(\phi,t) = \mathcal{F}_1 p_0(\phi,t). \tag{78}$$

We are interested in the solutions that satisfy the periodic boundary conditions,

$$p_i(\phi,t) = p_i(\phi + 2\pi,t), \tag{79}$$

$$\partial_\phi p_i(\phi,t) = \partial_\phi p_i(\phi + 2\pi,t), \tag{80}$$

for both $i = 0,1$. Moreover, in steady state, for $t \to \infty$, it must hold that

$$p_i(\phi,t) = p_i(\phi,t+T). \tag{81}$$

Equation (77) describes the diffusion of a particle with drift. The steady-state solution, which obeys Eqs. (79)–(81), is

$$\lim_{t\to\infty} p_0(\phi,t) = \frac{1}{2\pi}. \tag{82}$$

Clearly, $p_0(\phi,t)$ in steady state is flat, which means that any deviation in the steady-state solution for $P(\phi,t)$ from the flat distribution must be contained in $p_1(\phi,t)$. Since $p_1(\phi,t)$ is, by construction, a small perturbation, this approach will be accurate only when the full distribution is sufficiently flat, which means that the diffusion constant cannot be too small.

To obtain $p_1(\phi,t)$, we proceed by substituting the solution for $p_0(\phi,t)$, Eq. (82), into Eq. (78), yielding

$$\partial_t p_1(\phi,t) - D\partial_\phi^2 p_1(\phi,t) - \omega_0\partial_\phi p_1(\phi,t) = L(t)p_0(\phi,t)\partial_\phi Z(\phi). \tag{83}$$

The solution to this nonhomogeneous heat equation is given by

$$p_1(\phi,t) = \int_0^{2\pi} d\xi\, G(\phi - \omega_0 t,\xi,t) f(\xi) + \int_0^{2\pi}\int_0^t d\tau d\xi\, G(\phi - \omega_0 t,\xi,t-\tau)A(\xi,\tau), \tag{84}$$

where $f(\phi)$ is the initial condition, $G(\phi - \omega_0 t,\phi_0,t,t_0)$ is the Green's function of the unperturbed diffusion operator with drift, and $A(\phi,t) \equiv L(t)p_0(\phi,t)\partial_\phi Z(\phi)$. This expression holds for any t, not only for the steady-state solution.

To obtain the steady-state solution, we aim to find the initial condition $P(\phi,t) = f(\phi)$ that folds back onto itself after a time T: $P(\phi,t+T) = P(\phi,t) = f(\phi)$. To this end, we evaluate Eq. (84) for $t = T$, to arrive at the Fredholm equation of the second kind:

$$f(\phi) = \int_0^{2\pi} d\xi f(\xi) G(\phi,\xi,t=T) + Q(\phi), \tag{85}$$

where $Q(\phi)$ is given by Eq. (C17). The above equation can be solved analytically; see Appendix C.

Figures 8 and 4 show, respectively, that the LRT accurately describes $P(\phi,t)$ and hence the mutual information in the regime that the coupling is weak and the diffusion constant is large. In contrast to the phase-averaging method, the LRT breaks down for smaller diffusion constant. The reason is that then $P(\phi,t)$ deviates increasingly from the uniform distribution, $p_0(\phi,t) = 1/(2\pi)$, and the full solution $P(\phi,t)$ can no longer be treated as a weak perturbation to p_0.

V. DISCUSSION

The phase-response curves that have been measured experimentally often have a positive lobe and a negative one, separated by a dead zone where the coupling strength is zero [2]. However, the width of the dead zone varies considerably from organism to organism. Here, we asked how the optimal phase-response curve depends on the intrinsic noise in the system, using the mutual information as a performance measure.

Information theory predicts that the number of signals that can be transmitted reliably through a communication channel depends on the shape of the input distribution, the input-output relation, and the noise in the system. These arguments apply to any signaling system and the circadian clock is no exception.

When the input distribution is flat and the noise is low, then, in general, the optimal input-output relation is linear. The phase-oscillator model of the clock obeys this rule: the input distribution $p(t) = 1/T$ is flat, and the optimal input-output relation $\phi(t)$ is indeed linear in the low-noise regime [Figs. 5(b) and 5(c)]. Such a linear input-output relation is obtained for an

intrinsic period that is close to 24 h and for a dead zone that is relatively large (Figs. 6 and 7). Our analysis thus predicts that less-noisy circadian clocks exhibit a relatively large dead zone. Interestingly, the rule also explains why for a constant dead zone, in the low-noise limit, the optimal intrinsic frequency decreases as the coupling strength increases [see Fig. 5(a)].

In the large-noise regime, containment of noise becomes paramount. This inevitably requires a large coupling strength. While a strong coupling distorts the input-output relation, which tends to reduce information transmission, it also reduces the noise, enhancing information transmission [Figs. 5(b) and 5(c)]. The stability is further enhanced by increasing the intrinsic frequency and reducing the width of the dead zone (Fig. 7). Indeed, our results predict that noisy circadian systems feature a smaller dead zone and a higher intrinsic frequency.

These results have been obtained by reducing the circadian clock to a phase-oscillator model. It is useful to briefly review the generality and limitations of this approach. The mutual information obeys $I(\mathbf{n};t) \geqslant I(R,\phi;t) \geqslant I(\phi;t)$. Hence, any mapping of $\mathbf{n}$ to ϕ makes it possible to put a lower bound on the mutual information. The bound will be tight when the phase, according to this mapping, contains most of the information on time.

Another question is whether the model that we use to describe the evolution of the phase is accurate. Phase-oscillator models have commonly been employed to describe oscillatory systems, yet they are typically described as being valid in the limits of weak driving and low noise: this ensures that the coupled system stays close to the limit cycle of the unperturbed, deterministic system, so that the coupling function and the diffusion constant can be approximated by their values on that limit cycle [1]. Here, having derived the phase oscillator description in the weak coupling limit, we then proceed to study it for arbitrary values of ϵ and D. This might at first glance seem self-contradictory. It should be realized, however, that biochemical noise and coupling can have two distinct effects: they can affect the dynamics *along* the limit cycle, i.e., of ϕ, and/or they can cause the system to move *away* from the limit cycle. Only perturbations in the latter direction, orthogonal to the limit cycle, need to be small for the phase oscillator description to apply. Moreover, ϵ and D are dimensionful parameters that can only be meaningfully said to be large or small in comparison to another parameter, and the appropriate parameter for comparison is different for perturbations along and orthogonal to the limit cycle. Thus, it is entirely possible for ϵ and D to be small compared to the rate of relaxation to the limit cycle, implying that neither the external driving nor the noise can force the system far from the limit cycle and that the phase oscillator model is a good approximation, but simultaneously for one or both of ϵ and D to be large compared to ω_0, so that perturbations to the phase dynamics are not weak. We imagine that just such a situation holds here: D and ϵ can become bigger than ω_0—meaning that the noise and the coupling can induce large changes in ϕ—but, even for large D/ω_0 and ϵ/ω_0, the system in our model does not significantly move off the limit cycle. It remains an open question how, for a given, particular clock biochemical noise and strong coupling to an entrainment signal affect the dynamics: How far does the system move away from its limit cycle, and how much do the diffusion constant and the coupling function then change? The detailed and minimal biochemical network models that have been developed for the cyanobacterium *Synechococcus elongatus* would make it possible to investigate these questions in detail [36–43].

Our work shows that the behavior of the coupled phase oscillator can be accurately described by three different theories, which each work best in a different parameter regime. In the regime of weak coupling, low noise, and intrinsic frequency close to the driving frequency, the phase-averaging method is very accurate. In the regime that the driving is strong compared to the diffusion constant, the linear-noise approximation is most accurate. These are the two most relevant regimes for understanding the design of circadian clocks. There is also another regime, however, namely that of weak coupling and high noise, and in this regime linear-response theory is very accurate. That linear-response theory can describe any regime at all is perhaps surprising, since it has been argued that this theory should be applied to phase oscillators only with the greatest care [1]. The argument is that small but resonant forcing can have effects on ϕ that build up over time, meaning that the effect of perturbations that are nominally of order ϵ, and thus small, will eventually become large with time. However, noise can pre-empt this accumulation of resonant perturbations by effectively randomizing the phase and erasing the memory of earlier perturbations before they are able to accumulate over time. As a result, the full distribution of the phase can be written as a small perturbation around the uniform distribution, and this does make it possible to apply linear-response theory. While this regime is probably less relevant for understanding biological clocks, this approach may be useful in other contexts.

Finally, we have focused on the optimal design of the clock as a function of the intrinsic noise in the system. As Pfeuty *et al.* have shown, fluctuations in the input signal are an important consideration for understanding the design of circadian clocks [2]. It will be interesting to see whether maximizing the mutual information will reveal new design principles for clocks driven by fluctuating signals.

ACKNOWLEDGMENTS

We thank J. van Zon for a critical reading of the manuscript. This work is part of the research program of the Netherlands Organisation for Scientific Research (NWO), and by NSF Grant No. DMR-1056456 (D.K.L.).

APPENDIX A: ARNOLD TONGUE OF THE DETERMINISTIC MODEL

For completeness, we give here the inequalities for all scenarios.

Scenario (1). As discussed in the main text: $\phi_3 - 2\pi < \phi_s < \phi_1; t_2 < T/2 < t_3$. If $\epsilon_- \leqslant \omega_0$, then

$$T \leqslant \frac{2\pi - \epsilon_- \Delta\phi_{12}/\omega_0}{\omega_0 - \epsilon_-/2}, \tag{A1}$$

$$T > \frac{2\pi + \epsilon_+ \Delta\phi_{12}/\omega_0}{\epsilon_+/2 + \omega_0}, \tag{A2}$$

$$T < \frac{2\pi + \epsilon_+ \Delta\phi_{12}/\omega_0 + \Delta\phi_{23}(\epsilon_+ + \epsilon_-)/(\omega_0 - \epsilon_-)}{\epsilon_+/2 + \omega_0}, \tag{A3}$$

$$T > \frac{(\Delta\phi_{13} - 2\pi)(\epsilon_+ + \epsilon_-)/(\omega_0 + \epsilon_+) + 2\pi - \epsilon_-\Delta\phi_{12}/\omega_0}{\omega_0 - \epsilon_-/2}. \tag{A4}$$

If $\epsilon_- > \omega_0$ then

$$T \leqslant \frac{2(2\pi - \Delta\phi_{12})}{\omega_0}, \tag{A5}$$

$$T > \frac{2\pi - \Delta\phi_{12} + (\Delta\phi_{12}/\omega_0)(\epsilon_+ + \omega_0)}{\epsilon_+/2 + \omega_0}, \tag{A6}$$

$$T > \frac{2\Delta\phi_{23}}{\omega_0}. \tag{A7}$$

Scenario (2). $\phi_1 < \phi_s < \phi_2; 0 < t_2 < T/2 < t_3 < t_1 < T$. For $\epsilon_- < \omega_0$, the evolution of $\phi(t)$ is given by

$$\phi_s + \omega_0 t_2 + (-\epsilon_- + \omega_0)(T/2 - t_2) + \omega_0 T/2 = \phi_s + 2\pi. \tag{A8}$$

This yields

$$t_2 = \frac{2\pi - T(\omega_0 - \epsilon_-/2)}{\epsilon_-} < T/2 \quad \text{and} \quad > 0, \tag{A9}$$

$$t_3 = \frac{\Delta\phi_{23}}{\omega_0 - \epsilon_-} + t_2 > T/2, \tag{A10}$$

$$t_1 = t_2 - \Delta\phi_{12}/\omega_0 + T < T, \tag{A11}$$

$$\phi_s = \phi_2 - \omega_0 t_2 > \phi_1. \tag{A12}$$

This yields the following inequalities:

$$T > \frac{2\pi}{\omega_0}, \tag{A13}$$

$$T < \frac{2\pi}{\omega_0 - \epsilon_-/2}, \tag{A14}$$

$$T < \frac{2\pi + \Delta\phi_{13}\epsilon_-}{\omega_0 - \epsilon_-/2}, \tag{A15}$$

$$T > \frac{2\pi - \epsilon_-\Delta\phi_{12}/\omega_0}{\omega_0 - \epsilon_-/2}. \tag{A16}$$

If $\epsilon_- > \omega_0$, the equation to solve is

$$\phi_s + \omega_0 t_2 + \omega_0 T/2 = \phi_s + 2\pi. \tag{A17}$$

The solution is

$$t_2 = \frac{2\pi}{\omega_0} - T/2 < T/2, \tag{A18}$$

$$t_3 = \infty > T/2, \tag{A19}$$

$$t_1 = t_2 - \Delta\phi_{12}/\omega_0 + T < T, \tag{A20}$$

$$\phi_s = \phi_2 - \omega_0 t_2 > \phi_1 \quad \text{and} \quad < \phi_2. \tag{A21}$$

This yields the following inequalities:

$$T > 2\pi/\omega_0, \tag{A22}$$

$$T > \frac{2(2\pi - \Delta\phi_{12})}{\omega_0}, \tag{A23}$$

$$T < 4\pi/\omega_0. \tag{A24}$$

This scenario is stable, because $\phi(t)$ between $t = 0$ and $t = t_2$ is steeper than $\phi(t)$ between t_2 and $T/2$.

Scenario (3). $\phi_2 < \phi_s < \phi_3; 0 < t_2 < T/2$. If $\epsilon_- < \omega_0$ then

$$\phi_s + (-\epsilon_- + \omega_0)T/2 + \omega_0 T/2 = \phi_s + 2\pi. \tag{A25}$$

This equation does not depend on t_i. There is only one period that fits the solution:

$$T = \frac{2\pi}{\omega_0 - \epsilon_-/2}. \tag{A26}$$

This period is on the boundary of the Arnold tongue of scenario (2). This solution seems degenerate, being neither stable nor unstable.

If $\epsilon_- > \omega_0$, the equation that solves $\phi(t)$ is

$$\phi_s + (-\epsilon_- + \omega_0)t_2 + \omega_0 T/2 = \phi_s + 2\pi. \tag{A27}$$

The solution is

$$t_2 = \frac{2\pi - \omega_0 T/2}{-\epsilon_- + \omega_0}, \tag{A28}$$

$$\phi_s = \phi_2 + \omega_0 T/2 - 2\pi. \tag{A29}$$

The requirement that $t_2 > 0$ yields the inequality

$$T > \frac{4\pi}{\omega_0}, \tag{A30}$$

because the denominator of Eq. (A28) is negative. The requirement that $t_2 < T/2$ yields

$$\frac{2\pi - T(\omega_0 - \epsilon_-/2)}{\omega_0 - \epsilon_-} < 0. \tag{A31}$$

Since the denominator is negative for $\epsilon_- > \omega_0$, this means that $[2\pi - T(\omega_0 - \epsilon_-/2)] > 0$. When $\epsilon_- > 2\omega_0$, this is true for any T. When $\epsilon_- < 2\omega_0$ (but still larger than ω_0 because otherwise there is no solution at all; see above), then

$$T < \frac{2\pi}{\omega_0 - \epsilon_-/2}. \tag{A32}$$

The constraints $\phi_2 < \phi_s < \phi_3$ yield

$$T > \frac{4\pi}{\omega_0}, \tag{A33}$$

$$T < \frac{2(\Delta\phi_{23} + 2\pi)}{\omega_0}. \tag{A34}$$

This solution is rather strange. When the light comes up, the clock is being driven backwards. The solution seems stable, though. In fact, it seems extremely stable: after one period, the system is back on its limit cycle.

Scenario (4). $\phi_3 - 2\pi < \phi_s < \phi_1; 0 < t_1 < T/2 < t_2$. The equation that determines the steady state is

$$\phi_s + (\omega_0 + \epsilon_+)t_1 + \omega_0(T/2 - t_1) + \omega_0 T/2 = \phi_s + 2\pi. \tag{A35}$$

The solution is

$$t_1 = \frac{2\pi - \omega_0 T}{\epsilon_+} < T/2 \quad \text{and} \quad > 0, \tag{A36}$$

$$t_2 = t_1 + \frac{\Delta\phi_{12}}{\omega_0} > T/2, \tag{A37}$$

$$\phi_s = \phi_1 - (\epsilon_+ + \omega_0)t_1 = \phi_1 - (\epsilon_+ + \omega_0)(2\pi - \omega_0 T)/\epsilon_+. \tag{A38}$$

The conditions for T are

$$T \leqslant \frac{2\pi}{\omega_0}, \tag{A39}$$

$$T > \frac{2\pi}{\omega_0 + \epsilon_+/2}, \tag{A40}$$

$$T < \frac{2\pi + \Delta\phi_{12}\epsilon_+/\omega_0}{\epsilon_+/2 + \omega_0}, \tag{A41}$$

$$T > \frac{\epsilon_+\Delta\phi_{13} + 2\pi\omega_0}{\omega_0(\omega_0 + \epsilon_+)}. \tag{A42}$$

Scenario (5). $\phi_3 - 2\pi < \phi_s < \phi_1$; $t_1 > T/2$. The governing equation is

$$\phi_s + (\epsilon_+ + \omega_0)T/2 + \omega_0 T/2 = \phi_s + 2\pi. \tag{A43}$$

This means that

$$T = \frac{2\pi}{\omega_0 + \epsilon_+/2}. \tag{A44}$$

Clearly, for each ϵ_+ there is only one period, not a range of periods. Since $\phi(T/2) = \phi_s + (\epsilon_+ + \omega)T/2$, which must be smaller than ϕ_1, and $\phi_s > \phi_3 - 2\pi$, we find that there exists only a solution if $\Delta\phi_{13} < 2\pi\omega_0/(\epsilon_+ + 2\omega_0)$. Hence, for given ϕ_1 and ϕ_3, this puts an upper bound on ϵ_+. If a solution exists, the starting phase ϕ_s must lie in the range $\phi_3 - 2\pi < \phi_s < \phi_1 - \pi(\epsilon_+ + \omega_0)/(\epsilon_+/2 + \omega_0)$. Moreover, the solution is neutral; it does not relax back to a unique ϕ_s. In fact, this is a very general observation: if the solution is neutral, it means that there can only be locking for one value of the period. Being able to lock over a range of periods of the driving signal means that the clock should be able to adjust its period by changing the phase, but a neutral solution means that changing the phase does not lead to a change in its period.

Scenario (6). $\phi_3 - 2\pi < \phi_s < \phi_1$; $0 < t_1 < t_2 < t_3 < T/2$. This scenario can only arise when $\epsilon_- < \omega_0$, because otherwise the system never makes it to ϕ_3 before the sun sets. The equation to be solved is then

$$\phi_s + (\epsilon_+ + \omega_0)t_1 + \Delta\phi_{13} + (\omega_0 + \epsilon_+)(T/2 - t_3) + \omega_0 T/2 = \phi_s + 2\pi. \tag{A45}$$

This equation can be solved by noting that $\Delta\phi_{12} = \omega_0(t_2 - t_1)$ and $\Delta\phi_{23} = (-\epsilon_- + \omega_0)(t_3 - t_2)$. It follows that there is only one period that satisfies the above equation:

$$T = \frac{2\pi - \Delta\phi_{23} + \epsilon_+\Delta\phi_{12}/\omega_0 + (\epsilon_+ + \omega_0)\Delta\phi_{23}/(-\epsilon_- + \omega_0)}{\omega_0 + \epsilon_+/2}. \tag{A46}$$

Clearly, for a given ϵ_- and ϵ_+ there is only one period, not a range of periods, to which the system can entrain. This means that the solution is neutral, which can indeed be understood by noting that the initial slope at $t = 0, \omega_0 + \epsilon_+$, is the same as $t = T/2$. The condition for the solution to exist is that $\phi(T/2) = 2\pi + \phi_s - \omega_0 T/2 > \phi_3$. This yields for ϕ_s

$$\phi_3 - 2\pi + \omega_0 T/2 < \phi_s < \phi_1. \tag{A47}$$

There is thus only a solution when

$$T < \frac{2(2\pi - \Delta\phi_{13})}{\omega_0}. \tag{A48}$$

One could use this condition to determine the range of $\epsilon_{+/-}$ over which there is a solution, given ϕ_1,ϕ_2,ϕ_3. But since this scenario only yields one line in the phase diagram, we do not pursue this further.

Scenario (7). $\phi_1 < \phi_s < \phi_2$; $0 < T/2 < t_2 < t_3 < t_1$. The governing equation is

$$\phi_s + \omega_0 T/2 + \omega_0 T/2 = \phi_s + 2\pi. \tag{A49}$$

This indeed yields only one solution,

$$T = \frac{2\pi}{\omega_0}. \tag{A50}$$

Indeed, there only exists a solution when the driving frequency equals the intrinsic frequency, which is to be expected, since with this solution the system does not see the driving. The solution exists only if $\Delta\phi_{12} > \pi$. This solution is neutral, in that all solutions $\phi_1 < \phi_s < \phi_2$ are valid, for all values of $\epsilon_{-/+}$. One may wonder what that implies for the dynamics. If one would perform a simulation for $\epsilon_{-/+} > 0$ and $\omega = \omega_0$, and if one would then start with $\phi_1 < \phi_s < \phi_2$, then due to the noise the simulation would initially perform a random walk where initially, at the beginning of each day, the phase of the clock would fluctuate between ϕ_1 and ϕ_2. However, once the oscillator due to noise would cross the boundary ϕ_1, then the system will be driven to a solution that is described under scenario (1).

Scenario (8). $\phi_1 < \phi_s < \phi_2$; $0 < t_2 < t_3 < T_2 < t_1$. There can only be a solution, if it exists, when $\epsilon_- < \omega_0$. For $\epsilon_- > \omega_0$ the system never makes it to ϕ_3 before $T/2$. The governing equation is

$$\phi_s + \omega_0 t_2 + \Delta\phi_{23} + (\epsilon_+ + \omega_0)(T/2 - t_3) + \omega_0 T/2 = \phi_s + 2\pi. \tag{A51}$$

To solve this, we note that

$$t_3 = t_2 + \Delta\phi_{23}/(-\epsilon_- + \omega_0). \tag{A52}$$

This yields

$$t_2 = \frac{T(\omega_0 + \epsilon_+/2) - 2\pi - \Delta\phi_{23}(\epsilon_+ + \epsilon_-)/(\omega_0 - \epsilon_-)}{\epsilon_+}. \tag{A53}$$

We further have

$$\phi_s = \phi_2 - \omega_0 t_2. \tag{A54}$$

The condition $t_2 > 0$ yields

$$T > \frac{2\pi + \Delta\phi_{23}(\epsilon_+ + \epsilon_-)/(\omega_0 - \epsilon_-)}{\omega_0 + \epsilon_+/2}. \tag{A55}$$

The condition $t_3 < T/2$ yields

$$T < \frac{2\pi + \Delta\phi_{23}\epsilon_-/(\omega_0 - \epsilon_-)}{\omega_0}. \tag{A56}$$

The condition $\phi_1 < \phi_s = \phi_2 - \omega_0 t_2$ yields

$$T < \frac{2\pi + \epsilon_+ \Delta\phi_{12}/\omega_0 + \Delta\phi_{23}(\epsilon_+ + \epsilon_-)/(\omega_0 - \epsilon_-)}{\omega_0 + \epsilon_+/2}. \tag{A57}$$

The Arnold tongue of this scenario is embedded in those of scenarios (1) and (2). The solution corresponding to this scenario is indeed unstable: the system either converges to the solution of scenario (1) or (2). This can be easily proven by noting that the time it takes to cross $\Delta\phi_{23}$ is constant, as is the time to cross the night. The change in the phase a period later is then the change in the phase at $\phi(T/2)$. This is given by $\delta\phi(T/2) = \partial\phi(T/2)/\partial t_3 \delta t_1 = \partial\phi(T/2)/\partial t_3 \delta\phi_s/\omega_0 = (\epsilon_+ + \omega_0)/\omega_0 \delta\phi_s$, where we have noted that $\delta t_1 = -\delta\phi_s/\omega_0$ and $\partial\phi(T/2)/\partial t_3 = -(\epsilon_+ + \omega_0)$. Because $(\epsilon_+ + \omega_0)/\omega_0 > 1$, the change in the phase after a full period is larger than the initial change in the phase: $\delta\phi(T) = \delta\phi(T/2) > \delta\phi_s$. The solution is unstable.

Scenario (9). $\phi_1 < \phi_s < \phi_2$; $t_2 < t_3 < t_1 < T_2$. There can only be a solution if $\epsilon_- < \omega_0$. The equation to be solved is

$$\phi_s + \omega_0 t_2 + 2\pi - \Delta\phi_{12} + \omega_0(T - t_1) = \phi_s + 2\pi, \tag{A58}$$

which gives

$$T = \Delta\phi_{12}/\omega_0 + t_1 - t_2. \tag{A59}$$

We further have

$$t_1 - t_2 = \frac{2\pi - \Delta\phi_{13}}{\omega_0 + \epsilon_+} + \frac{\Delta\phi_{23}}{\omega_0 - \epsilon_-}. \tag{A60}$$

Hence,

$$T = \frac{\Delta\phi_{12}}{\omega_0} + \frac{2\pi - \Delta\phi_{13}}{\omega_0 + \epsilon_+} + \frac{\Delta\phi_{23}}{\omega_0 - \epsilon_-}, \tag{A61}$$

which we could have written down right away upon somewhat more careful thinking. We can obtain a bound on the parameters that allow a solution by noting that $0 < t_1 - t_2 < T/2$. Combining with Eq. (A59) yields $\Delta\phi_{12}/\omega_0 < T < 2\Delta\phi_{12}/\omega_0$. Combing this with Eq. (A61) yields

$$\frac{\Delta\phi_{12}}{\omega_0} < \frac{2\pi - \Delta\phi_{13}}{\omega_0 + \epsilon_+} + \frac{\Delta\phi_{23}}{\omega_0 - \epsilon_-}. \tag{A62}$$

A visual inspection illustrates this constraint very clearly. The parameter ϵ_- should be small, that is not close to unity. A large ϵ_+ also helps.

Scenario (10). $\phi_2 < \phi_s < \phi_3$; $0 < T_2 < t_3, t_1, t_2$. Both for $\epsilon_- < \omega_0$ and $\epsilon_- > \omega_0$, the scenario corresponds to that of scenario (3), but with $\epsilon_- < \omega_0$ in that scenario. There is only a solution for

$$T = 2\pi/(\omega_0 - \epsilon_-/2). \tag{A63}$$

Scenario (11). $\phi_2 < \phi_s \phi_3$; $0 < t_3 < T/2 < t_1, t_2$. Only if $\epsilon_- < \omega_0$ may a solution exist: if $\epsilon_- > \omega_0$, we are back to scenario (3) or (10). The governing equation is

$$\phi_s + (-\epsilon_- + \omega_0)t_3 + (\epsilon_+ + \omega_0)(T/2 - t_3) + \omega_0 T/2 = \phi_s + 2\pi. \tag{A64}$$

The solution is

$$t_3 = \frac{T(\omega_0 + \epsilon_+/2) - 2\pi}{\epsilon_+ + \epsilon_-}, \tag{A65}$$

$$\phi_s = \phi_3 - (\omega_0 - \epsilon_-)t_3. \tag{A66}$$

The condition $t_3 > 0$ yields

$$T > \frac{2\pi}{\omega_0 + \epsilon_+/2}. \tag{A67}$$

The condition $t_3 < T/2$ yields the inequality

$$T < \frac{2\pi}{\omega_0 - \epsilon_-/2}. \tag{A68}$$

The condition $\phi_s > \phi_2$ yields

$$T < \frac{\Delta\phi_{23}(\epsilon_+ + \epsilon_-)/(\omega_0 - \epsilon_-) + 2\pi}{\omega_0 + \epsilon_+/2}. \tag{A69}$$

The condition $t_1 > T/2$ yields the inequality

$$T > \frac{2\pi - (2\pi - \Delta\phi_{13})(\epsilon_+ + \epsilon_-)/(\epsilon_+ + \omega_0)}{\omega_0 - \epsilon_-/2}. \tag{A70}$$

The solution space overlaps with those of scenarios (1)–(3). Interestingly, we find again that this solution is unstable: $\delta\phi(T) = \delta\phi(T/2) = \partial\phi(T/2)/\partial t_3 \delta t_3 = -(\omega_0 + \epsilon_+)\delta t_3 = -(\omega_0 + \epsilon_+)\partial t_3/\partial\phi_s \delta\phi_s = (\omega_0 + \epsilon_+)/(\omega_0 - \epsilon_-)\delta\phi_s > \delta\phi_s$. We thus can see that when $\phi(t)$ is convex for $0 < t < T/2$, the solution tends to be unstable.

Scenario (12). $\phi_2 < \phi_s < \phi_3$; $t_3, t_1 < T/2 < t_2$. Only if $\epsilon_- < \omega_0$ may a solution exist. The governing equation is

$$\phi_s + (-\epsilon_- + \omega_0)t_3 + (2\pi - \Delta\phi_{13}) + \omega_0(T/2 - t_1) + \omega_0 T/2 = \phi_s + 2\pi. \tag{A71}$$

Exploiting that $t_1 = t_3 + (2\pi - \Delta\phi_{13})/(\epsilon_+ + \omega_0)$, the solution is

$$t_3 = \frac{\omega_0 T - \Delta\phi_{13} - \omega_0(2\pi - \Delta\phi_{13})/(\epsilon_+ + \omega_0)}{\epsilon_-}, \tag{A72}$$

$$\phi_s = \phi_3 - (\omega_0 - \epsilon_-)t_3. \tag{A73}$$

The condition $t_3 > 0$ yields the inequality

$$T > \frac{\Delta\phi_{13}}{\omega_0} + \frac{2\pi - \Delta\phi_{13}}{\epsilon_+ + \omega_0}. \tag{A74}$$

The condition $t_1 < T/2$ gives

$$T < \frac{\Delta\phi_{13} + (\omega_0 - \epsilon_-)(2\pi - \Delta\phi_{13})/(\epsilon_+ + \omega_0)}{\omega_0 - \epsilon_-/2}. \tag{A75}$$

The condition $t_2 = t_1 + \Delta\phi_{12}/\omega_0 > T/2$ yields

$$T > \frac{\Delta\phi_{13} + (\omega_0 - \epsilon_-)(2\pi - \Delta\phi_{13})/(\epsilon_+ + \omega_0) - \epsilon_- \Delta\phi_{12}/\omega_0}{\omega_0 - \epsilon_-/2}. \tag{A76}$$

The condition $\phi_s > \phi_2$ yields the inequality

$$T < \frac{\Delta\phi_{13} + \omega_0(2\pi - \Delta\phi_{13})/(\epsilon_+ + \omega_0) + \epsilon_-\Delta\phi_{23}/(\omega_0 - \epsilon_-)}{\omega_0}. \tag{A77}$$

This curve is convex, that is, the part of $\phi(t)$ that really matters is convex: the initial slope near $t = 0$, $\omega_0 - \epsilon_-$, is smaller than the slope near $t = T/2$, which is ω. This gives an unstable solution.

Scenario (13). $\phi_2 < \phi_s < \phi_3$; $t_3, t_1, t_2 < T/2$. Again, a solution may only exist if $\epsilon_- < \omega_0$. The central equation is

$$\phi_s + (-\epsilon_- + \omega_0)t_3 + (2\pi - \Delta\phi_{23}) + (-\epsilon_- + \omega_0)(T/2 - t_2) + \omega_0 T/2 = \phi_s + 2\pi. \tag{A78}$$

The solution is

$$T = \frac{\Delta\phi_{23}}{\omega_0} + \frac{(\omega_0 - \epsilon_-)(t_2 - t_3)}{\omega_0}. \tag{A79}$$

The time difference is

$$t_2 - t_3 = \frac{\Delta\phi_{12}}{\omega_0} + \frac{2\pi - \Delta\phi_{13}}{\omega_0 + \epsilon_+}, \tag{A80}$$

which gives for the period

$$T = \frac{\Delta\phi_{23}}{\omega_0} + \frac{\omega_0 - \epsilon_-}{\omega_0}\left(\frac{\Delta\phi_{12}}{\omega_0} + \frac{2\pi - \Delta\phi_{13}}{\omega_0 + \epsilon_+}\right). \tag{A81}$$

APPENDIX B: HEAT MAPS MUTUAL INFORMATION AS A FUNCTION OF COUPLING STRENGTH AND INTRINSIC FREQUENCY

Figure 3(a) shows the mutual information as a function of the coupling strength $\epsilon = \epsilon_+ = \epsilon_-$ and intrinsic frequency ω_0, for one value of the diffusion constant, $D = 0.1/T$. Figure 9 shows the same plot, but then also for $D = 1/T$ and $D = 10^{-4}/T$. For $D = 10^{-4}/T$, the mutual information shows very rich behavior, corresponding to intricate locking behavior.

APPENDIX C: LINEAR-RESPONSE THEORY

As shown in the main text, the evolution of $p_1(\phi,t)$ is given by

$$\partial_t p_1(\phi,t) - D\partial_\phi^2 p_1(\phi,t) - \omega_0\partial_\phi p_1(\phi,t) = L(t)p_0(\phi,t)\partial_\phi Z(\phi). \tag{C1}$$

The solution to this nonhomogeneous heat equation is

$$p_1(\phi,t) = \int_0^{2\pi} d\xi G(\phi - \omega_0 t,\xi,t) f(\xi) + \int_0^{2\pi}\int_0^t d\tau d\xi G(\phi - \omega_0 t,\xi,t-\tau)A(\xi,\tau), \tag{C2}$$

where $f(\phi)$ is the initial condition, $G(\phi - \omega_0 t,\phi_0,t,t_0)$ is the Green's function of the unperturbed diffusion operator, and $A(\phi,t) \equiv L(t)p_0(\phi,t)\partial_\phi Z(\phi) = L(t)/(2\pi)[-\delta(\phi - \phi_1) - \delta(\phi - \phi_2) + 2\delta(\phi - \phi_3)]$.

The Green's function is given by

$$G(\phi - \omega_0 t,\phi_0,t) = \sum_{j=0}^{\infty} e^{-j^2 Dt}\{A_j(\phi_0)\cos[j(\phi - \omega_0 t)] + B_j(\phi_0)\sin[j(\phi - \omega_0 t)]\}, \tag{C3}$$

with

$$A_j(\phi_0) = \frac{1}{\pi}\int d\phi\delta(\phi - \omega_0 t - \phi_0)\cos[j(\phi - \omega_0 t)] = \frac{1}{\pi}\cos j\phi_0, \tag{C4}$$

$$B_j(\phi_0) = \frac{1}{\pi}\int d\phi\delta(\phi - \omega_0 t - \phi_0)\sin(j\phi') = \frac{1}{\pi}\sin j\phi_0, \tag{C5}$$

$$A_0 = \frac{1}{2\pi}, \tag{C6}$$

$$B_0 = 0. \tag{C7}$$

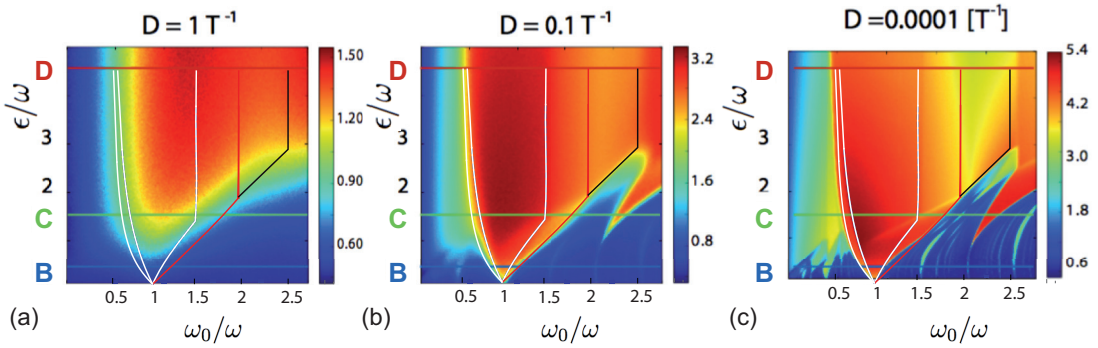

FIG. 9. The mutual information as a function of the coupling strength ϵ and the intrinsic frequency ω_0, for three different values of the diffusion constant D. In all panels, $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$. Superimposed in black is the deterministic Arnold tongue for scenarios (1) and (4). (a) $D = 1/T$. (b) $D = 0.1/T$ [the same panel as Fig. 3(a)]. (c) $D = 10^{-4}/T$. Note the rich behavior of the mutual information, corresponding to higher-order locking scenarios.

This yields

$$G(\phi,\phi_0,t) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{j=1}^{\infty} e^{-i^2Dt}\{\cos(j\phi_0)\cos[j(\phi-\omega_0 t)] + \sin(j\phi_0)\sin[j(\phi-\omega_0 t)]\}. \quad \text{(C8)}$$

Substituting this expression into Eqs. (C1) and (C2) gives

$$p_1(\phi,t) = \int_0^{2\pi} d\xi\, G(\phi,\xi,t) f(\xi) + \int_0^{2\pi}\int_0^t d\tau d\xi\, G(\phi,\xi,t-\tau)\frac{L(\tau)}{2\pi} \times [-\delta(\xi-\phi_1) - \delta(\xi-\phi_2) + 2\delta(\xi-\phi_3)] \quad \text{(C9)}$$

$$= G_0(\phi,t) + \frac{1}{2\pi}\int_0^t d\tau L(\tau)\Delta G(\phi,t-\tau), \quad \text{(C10)}$$

where

$$G_0(\phi,t) = \int_0^{2\pi} d\xi\, G(\phi,\xi,t) f(\xi),$$
$$\Delta G(\phi,t-\tau) = -G(\phi,\phi_1,t-\tau) - G(\phi,\phi_2,t-\tau) + 2G(\phi,\phi_3,t-\tau). \quad \text{(C11)}$$

We can integrate the second term of Eq. (C10) by parts. Calling the primitive of ΔG

$$C(\phi,\tau;t) = \int d\tau \Delta G(\phi,t-\tau), \quad \text{(C12)}$$

we find

$$p_1(\phi,t) = G_0(\phi,t) + [L(\tau)C(\phi,\tau;t)]_{\tau=0}^{\tau=t} - \int_0^t d\tau \frac{dL(\tau)}{d\tau} C(\phi,\tau;t). \quad \text{(C13)}$$

Since $L(\tau)$ is a sequence of step functions,

$$\frac{dL(\tau)}{d\tau} = \sum_{n=0}^{\infty}\delta(\tau-nT) - \delta[\tau-(nT+T/2)], \quad \text{(C14)}$$

which yields

$$p_1(\phi,t) = G_0(\phi,t) + [L(\tau)C(\phi,\tau;t)]_0^t - \sum_{n=0}^{nT<t}[C(\phi,nT;t) - C(\phi,nT+T/2;t)]. \quad \text{(C15)}$$

Equation (C1) was derived assuming that the system is in steady state, and $p(\phi,t) = p(\phi,t+T)$. This means that we only have to consider times $0 < t < T$, in which case only the first two terms in the last sum on the right-hand side remain. More specifically, in steady state, the initial condition $f(\phi)$ equals the steady-state distribution, and $f(\phi) = p(\phi,t=0) = p(\phi,t=T)$, meaning that the above expression reduces to

$$f(\phi) = G_0(\phi,T) + Q(\phi) = \int_0^{2\pi} f(\xi)G(\phi,\xi,t=T) + Q(\phi), \quad \text{(C16)}$$

where $Q(\phi)$ is defined as

$$Q(\phi) \equiv -2C(\phi,\tau=0;T) + C(\phi,\tau=T;T) + C(\phi,\tau=T/2;T). \quad \text{(C17)}$$

Equation (C16) is an integral equation, more specifically a Fredholm equation of the second type. The integration kernel $G(\phi,\xi,T)$ has the form

$$G(\phi,\xi,T) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{j=1}^{\infty} e^{-j^2DT}\{\cos[j(\phi-\omega_0 T)]\cos(j\xi) + \sin[j(\phi-\omega_0 T)]\sin(j\xi)\}. \quad \text{(C18)}$$

We define $G^*(\phi,\xi) = G(\phi,\xi) - 1/(2\pi)$, and rewrite Eq. (C16) as

$$f(\phi) = \int_0^{2\pi} d\xi f(\xi) G^*(\phi,\xi,t=T) + \frac{1}{2\pi}\int_0^{2\pi} d\xi f(\xi) + Q(\phi) \quad \text{(C19)}$$

$$= \int_0^{2\pi} d\xi f(\xi) G^*(\phi,\xi,t=T) + \frac{1}{2\pi} + Q(\phi) \quad \text{(C20)}$$

$$= \int_0^{2\pi} d\xi f(\xi) G^*(\phi,\xi,t=T) + Q^*(\phi), \quad \text{(C21)}$$

where in going from the first to the second line we have exploited that $f(\phi)$ is normalized, and in the last line we have defined $Q^*(\phi) \equiv Q(\phi) - 1/(2\pi)$. The kernel $G^*(\phi,\xi,T)$ is separable, and we can rewrite Eq. (C16) as

$$f(\phi) = \sum_{j=1}^{\infty} e^{-j^2DT}\int_0^{2\pi} d\xi f(\xi)\{\cos[j(\phi-\omega_0 T)]\cos(j\xi) + \sin[j(\phi-\omega_0 T)]\sin(j\xi)\} + Q_j^*(\phi) \quad \text{(C22)}$$

with $Q^*(\phi) = \sum_j Q_j^*(\phi)$.

To solve this integral equation, we define

$$c1_j \equiv \int_0^{2\pi} d\xi e^{-j^2DT} f(\xi)\cos(j\xi), \quad \text{(C23)}$$

$$c2_j \equiv \int_0^{2\pi} d\xi e^{-j^2DT} f(\xi)\sin(j\xi), \quad \text{(C24)}$$

so that

$$f(\phi) = \sum_j \{\cos[j(\phi-\omega_0 T)]c1_j + \sin[j(\phi-\omega_0 T)]c2_j + Q_j^*(\phi)\}. \quad \text{(C25)}$$

We now multiply both sides, once with $e^{-j^2DT}\cos(j\phi)$ and once with $e^{-j^2DT}\sin(j\phi)$, and integrate from 0 to 2π. On the left-hand side, this gives c_{1j} and c_{2j}, respectively. We then arrive at the following set of linear equations:

$$c_{1j} = \sum_k A_{jk}c_{1k} + B_{jk}c_{2k} + Q_{1k}^*, \quad \text{(C26)}$$

$$c_{2j} = \sum_k C_{jk}c_{1k} + D_{jk}c_{2k} + Q_{2k}^*, \quad \text{(C27)}$$

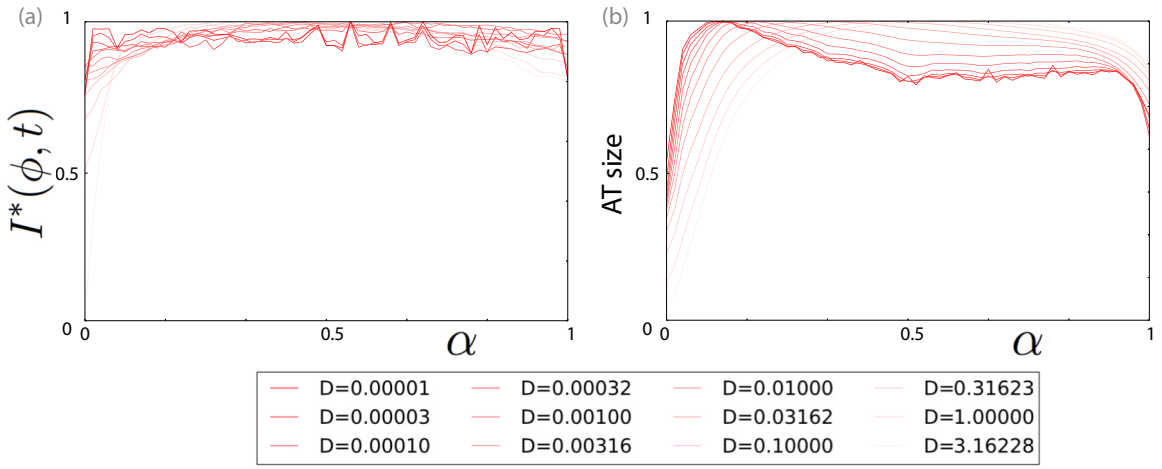

FIG. 10. Information transmission is not much affected by the relative magnitudes of ϵ_+ and ϵ_- in the coupling function $Z(\phi)$ (see Fig. 1). We vary ϵ_+ and ϵ_- via a parameter α, defined as $\epsilon_+ = (1-\alpha)\epsilon$ and $\epsilon_- = \alpha\epsilon$; varying α thus keeps the total absolute coupling strength constant. We vary α and the diffusion constant D, and optimize over ϵ and ω_0, keeping $\Delta\phi_{12} = \Delta\phi_{23} = \pi/2$ constant in all simulations. (a) The maximal mutual information $I^*(\phi,t)$, obtained by optimizing $I(\phi,t)$ over ϵ and ω_0, as a function of α, for different values of D. It is seen that for most values of D, $I^*(\phi;t)$ is quite independent of α. (b) The size of the Arnold tongue of the stochastic system as a function of α, for different values of D. The size is defined as $\int_{\omega_0^{\min}}^{\omega_0^{\max}} d\omega_0 \int_{\epsilon^{\min}}^{\epsilon^{\max}} d\epsilon I(\phi;t)/I^*(\phi;t)$, with $\omega_0^{\min}/\omega = 0.4$, $\omega_0^{\max} = 2.7$, $\epsilon_{\min}/\omega = 0$, $\epsilon_{\max}/\omega = 5$. It is seen that, except for the low and high values of α, the size of the Arnold tongue of the stochastic system is fairly independent of α.

where

$$A_{jk} = \int_0^{2\pi} d\phi e^{-j^2 DT} \cos(j\phi)\cos[k(\phi-\omega_0 T)], \quad \text{(C28)}$$

$$B_{jk} = \int_0^{2\pi} d\phi e^{-j^2 DT} \cos(j\phi)\sin[k(\phi-\omega_0 T)], \quad \text{(C29)}$$

$$C_{jk} = \int_0^{2\pi} d\phi e^{-j^2 DT} \sin(j\phi)\cos[k(\phi-\omega_0 T)], \quad \text{(C30)}$$

$$D_{jk} = \int_0^{2\pi} d\phi e^{-j^2 DT} \sin(j\phi)\sin[k(\phi-\omega_0 T)], \quad \text{(C31)}$$

$$Q^*_{1k} = \int_0^{2\pi} d\phi e^{-j^2 DT} \cos(j\phi) Q^*_k(\phi), \quad \text{(C32)}$$

$$Q^*_{2k} = \int_0^{2\pi} d\phi e^{-j^2 DT} \sin(j\phi) Q^*_k(\phi). \quad \text{(C33)}$$

We can define the vectors $\mathbf{c}_1$ and $\mathbf{c}_2$ with elements c_{1j} and c_{2j}, respectively, as well as the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, with elements $A_{jk}, B_{jk}, C_{jk}, D_{jk}$, respectively, and the vectors $\mathbf{q}_1$ and $\mathbf{q}_2$ with elements Q^*_{1j} and Q^*_{2j}, respectively. This allows us to define the vectors $\mathbf{c}^T \equiv (\mathbf{c}_1^T : \mathbf{c}_2^T)$ and $\mathbf{q}^T \equiv (\mathbf{q}_1^T : \mathbf{q}_2^T)$, where T denotes the transpose, and the matrix

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}. \quad \text{(C34)}$$

We can then rewrite Eqs. (C26) and (C27) as

$$\mathbf{c} = \mathbf{M}\mathbf{c} + \mathbf{q}, \quad \text{(C35)}$$

which has as its solution

$$\mathbf{c} = (\mathbf{I} - \mathbf{M})^{-1}\mathbf{q}, \quad \text{(C36)}$$

with $\mathbf{I}$ the identity matrix. With the coefficients c_{1j} and c_{2j} thus found, $f(\phi)$ can be obtained from Eq. (C25), yielding, finally, the steady-state solution $p_{ss}(\phi) = 1/(2\pi) + f(\phi)$.

APPENDIX D: MUTUAL INFORMATION AS A FUNCTION OF ϵ_+ AND ϵ_-

Figure 10 addresses how the mutual information depends on ϵ_+ and ϵ_-. To this end, the parameters are varied as $\epsilon_+ = (1-\alpha)\epsilon$ and $\epsilon_- = \alpha\epsilon$; varying α thus keeps the total absolute coupling strength ϵ constant. The figure shows that the mutual information is rather insensitive to the relative values of ϵ_+ and ϵ_-.

[1] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronisation: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, England, 2003).
[2] B. Pfeuty, Q. Thommen, and M. Lefranc, Robust entrainment of circadian oscillators requires specific phase response curves, Biophys. J. **100**, 2557 (2011).
[3] Y. Hasegawa and M. Arita, Circadian clocks optimally adapt to sunlight for reliable synchronization, J. R. Soc. Interface **11**, 20131018 (2013).
[4] Y. Hasegawa and M. Arita, Optimal Implementations for Reliable Circadian Clocks, Phys. Rev. Lett. **113**, 108101 (2014).
[5] C. E. Shannon, The mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).
[6] E. Ziv, I. Nemenman, and C. H. WIggins, Optimal signal processing in small stochastic biochemical networks, PLoS One **2**, e1077 (2007).

[7] F. Tostevin and P. R. ten Wolde, Mutual Information between Input and Output Trajectories of Biochemical Networks, Phys. Rev. Lett. **102**, 218101 (2009).

[8] P. Mehta, S. Goyal, T. Long, Bonnie L. Bassler, and Ned S. Wingreen, Information processing and signal integration in bacterial quorum sensing, Mol. Syst. Biol. **5**, 325 (2009).

[9] G. Tkačik, A. M. Walczak, and W. Bialek, Optimizing information flow in small genetic networks, Phys. Rev. E **80**, 031920 (2009).

[10] F. Tostevin and P. R. ten Wolde, Mutual information in time-varying biochemical systems, Phys. Rev. E **81**, 061917 (2010).

[11] W. H. de Ronde, F. Tostevin, and P. Rein ten Wolde, Effect of feedback on the fidelity of information transmission of time-varying signals, Phys. Rev. E **82**, 031914 (2010).

[12] G. Tkacik, J. S. Prentice, V. Balasubramanian, and E. Schneidman, Optimal population coding by noisy spiking neurons, Proc. Natl. Acad. Sci. U.S.A. **107**, 14419 (2010).

[13] A. M. Walczak, G. Tkačik, and W. Bialek, Optimizing information flow in small genetic networks. II. Feed-forward interactions, Phys. Rev. E **81**, 041905 (2010).

[14] W. de Ronde, F. Tostevin, and P. R. ten Wolde, Multiplexing Biochemical Signals, Phys. Rev. Lett. **107**, 048101 (2011).

[15] R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko, Information transduction capacity of noisy biochemical signaling networks, Science **334**, 354 (2011).

[16] W. de Ronde, F. Tostevin, and P. ten Wolde, Feed-forward loops and diamond motifs lead to tunable transmission of information in the frequency domain, Phys. Rev. E **86**, 021913 (2012).

[17] J. O. Dubuis, G. Tkacik, E. F. Wieschaus, T. Gregor, and W. Bialek, Positional information, in bits, Proc. Natl. Acad. Sci. U.S.A. **110**, 16301 (2013).

[18] C. G. Bowsher, M. Voliotis, and P. S. Swain, The fidelity of dynamic signaling by noisy biomolecular networks, PLoS Comput. Biol. **9**, e1002965 (2013).

[19] J. Selimkhanov, B. Taylor, J. Yao, A. Pilko, J. Albeck, A. Hoffmann, L. Tsimring, and R. Wollman, Systems biology. Accurate information transmission through dynamic biochemical signaling networks, Science **346**, 1370 (2014).

[20] W. de Ronde and P. R. ten Wolde, Multiplexing oscillatory biochemical signals, Phys. Biol. **11**, 026004 (2014).

[21] C. C. Govern and P. R. ten Wolde, Energy Dissipation and Noise Correlations in Biochemical Sensing, Phys. Rev. Lett. **113**, 258102 (2014).

[22] T. R. Sokolowski and G. Tkačik, Optimizing information flow in small genetic networks. IV. Spatial coupling, Phys. Rev. E **91**, 062710 (2015).

[23] N. B. Becker, A. Mugler, and P. R. ten Wolde, Optimal Prediction by Cellular Signaling Networks, Phys. Rev. Lett. **115**, 258103 (2015).

[24] M. Monti and P. R. ten Wolde, The accuracy of telling time via oscillatory signals, Phys. Biol. **13**, 1 (2016).

[25] D. T. Gillespie, The chemical Langevin equation, J. Chem. Phys. **113**, 297 (2000).

[26] I. Mihalcescu, W. H. Hsing, and S. Leibler, Resilient circadian oscillator revealed in individual cyanobacteria, Nature (London) **430**, 81 (2004).

[27] P. Westermark, D. K. Welsh, H. Okamura, and H. Herzel, Quantification of circadian rhythms in single cells, PLoS Comput. Biol. **5**, e1000580 (2009).

[28] A. B. Webb, N. Angelo, J. E. Huettner, and E. D. Herzog, Intrinsic, nondeterministic circadian rhythm generation in identified mammalian neurons, Proc. Natl. Acad. Sci. U.S.A. **106**, 16493 (2009).

[29] G. Süel, R. P. Kulkarni, J. Dworkin, J. Garcia-Ojalvo, and M. B. Elowitz, Tunability and noise dependence in differentiation dynamics, Science **315**, 1716 (2007).

[30] S.-ya Miyagishima, C. P. Wolk, and K. W. Osteryoung, Identification of cyanobacterial cell division genes by comparative and mutational analyses, Mol. Microbiol. **56**, 126 (2005).

[31] J. Aschoff and H. Pohl, Phase relations between a circadian rhythm and its zeitgeber within the range of entrainment, Naturwissenschaften **65**, 80 (1978).

[32] A. E. Granada, G. Bordyugov, A. Kramer, and H. Herzel, Human chronotypes from a theoretical perspective, PLoS One **8**, e59464 (2013).

[33] G. Bordyugov, U. Abraham, A. Granada, P. Rose, K. Imkeller, A. Kramer, and H. Herzel, Tuning the phase of circadian entrainment, J. R. Soc. Interface **12**, 20150282 (2015).

[34] C. Schmal, J. Myung, H. Herzel, and G. Bordyugov, A theoretical study on seasonality, Front. Neurol. **6**, 94 (2015).

[35] G. Tkačik and A. M. Walczak, Information transmission in genetic regulatory networks: A review, J. Phys.: Condens. Matter **23**, 153102 (2011).

[36] J. S. van Zon, D. K. Lubensky, P. R. H. Altena, and P. R. ten Wolde, An allosteric model of circadian KaiC phosphorylation, Proc. Natl. Acad. Sci. U.S.A. **104**, 7420 (2007).

[37] M. J. Rust, J. S. Markson, W. S. Lane, D. S. Fisher, and E. K. O'Shea, Ordered phosphorylation governs oscillation of a three-protein circadian clock, Science **318**, 809 (2007).

[38] D. Zwicker, D. K. Lubensky, and P. R. ten Wolde, Robust circadian clocks from coupled protein-modification and transcription-translation cycles. Supporting info, Proc. Natl. Acad. Sci. U.S.A. **107**, 22540 (2010).

[39] C. Phong, J. S. Markson, C. M. Wilhoite, and M. J. Rust, Robust and tunable circadian rhythms from differentially sensitive catalytic domains, Proc. Natl. Acad. Sci. U.S.A. **110**, 1124 (2013).

[40] J. Paijmans, M. Bosman, P. R. ten Wolde, and D. K. Lubensky, Discrete gene replication events drive coupling between the cell cycle and circadian clocks, Proc. Natl. Acad. Sci. U.S.A. **113**, 4063 (2016).

[41] J. Paijmans, D. K. Lubensky, and P. R. ten Wolde, A thermodynamically consistent model of the post-translational Kai circadian clock, PLoS Comput. Biol. **13**, e1005415 (2017).

[42] J. Paijmans, D. K Lubensky, and P. R. ten Wolde, Robustness of synthetic oscillators in growing and dividing cells, Phys. Rev. E **95**, 052403 (2017).

[43] J. Paijmans, D. K. Lubensky, and P. R. ten Wolde, Period robustness and entrainability of the Kai system to changing nucleotide concentrations, Biophys. J. **113**, 157 (2017).