Synchronization of stochastic hybrid oscillators driven by a common switching environment

Paul C. Bressloff, and James MacLaurin

Synchronization of stochastic hybrid oscillators driven by a common switching environment

Paul C. Bressloff[1] and James MacLaurin[2]

[1]*Department of Mathematics, University of Utah, Salt Lake City, Utah 84112, USA*
[2]*Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, New Jersey 07102, USA*

Many systems in biology, physics, and chemistry can be modeled through ordinary differential equations (ODEs), which are piecewise smooth, but switch between different states according to a Markov jump process. In the fast switching limit, the dynamics converges to a deterministic ODE. In this paper, we suppose that this limit ODE supports a stable limit cycle. We demonstrate that a set of such oscillators can synchronize when they are uncoupled, but they share the same switching Markov jump process. The latter is taken to represent the effect of a common randomly switching environment. We determine the leading order of the Lyapunov coefficient governing the rate of decay of the phase difference in the fast switching limit. The analysis bears some similarities to the classical analysis of synchronization of stochastic oscillators subject to common white noise. However, the discrete nature of the Markov jump process raises some difficulties: in fact, we find that the Lyapunov coefficient from the quasi-steady-state approximation differs from the Lyapunov coefficient one obtains from a second order perturbation expansion in the waiting time between jumps. Finally, we demonstrate synchronization numerically in the radial isochron clock model and show that the latter Lyapunov exponent is more accurate. *Published by AIP Publishing.* https://doi.org/10.1063/1.5054795

There are a growing number of systems in physics and biology where a population of oscillators can be synchronized by a randomly fluctuating external input applied globally to all of the oscillators, even if there are no interactions between the oscillators (noise-induced phase synchronization). Experimental evidence for such an effect has been found in neural oscillations of the olfactory bulb, synthetic genetic oscillators, laser dynamics, and variations in geographically separated animal populations. Most previous studies of noise-induced phase synchronization have taken the oscillators to be driven by common Gaussian noise. Typically, phase synchronization is established by constructing the Lyapunov exponent for the dynamics of the phase difference between a pair of oscillators and averaging with respect to the noise. If the averaged Lyapunov exponent is negative definite, then the phase difference is expected to decay to zero in the large time limit, establishing phase synchronization. In this paper, we extend the theory of noise-induced synchronization to the case of a common randomly switching environment. Each oscillator then evolves according to a piecewise deterministic Markov process, which involves the coupling between a piecewise continuous differential equation and a time-homogeneous Markov chain. In the fast switching limit, the dynamics converges to a deterministic ordinary differential equation (ODE), which is assumed to support a stable limit cycle. We demonstrate that an uncoupled population of such oscillators can synchronize when they share the same switching Markov jump process. We determine the leading order of the Lyapunov coefficient governing the rate of decay of the phase differences in the weak noise regime (fast but finite switching rates) and show that it differs from the standard expression obtained using a Gaussian approximation of the noise.

I. INTRODUCTION

Self-sustained oscillations in biological, physical, and chemical systems are often described in terms of limit cycle oscillators where the timing along each limit cycle is specified in terms of a single phase variable. Phase reduction methods can then be used to analyze synchronization of an ensemble of weakly-coupled oscillators by approximating the high-dimensional limit cycle dynamics as a closed system of equations for the corresponding phase variables.[1–9] More recently, there has been considerable interest in applying phase reduction methods to the analysis of noise-induced phase synchronization.[10–16] This concerns the observation that a population of oscillators can be synchronized by a randomly fluctuating external input applied globally to all of the oscillators, even if there are no interactions between the oscillators. Evidence for such an effect has been found in experimental studies of neural oscillations in the olfactory bulb[17] and the synchronization of synthetic genetic oscillators.[18,19] A related phenomenon is the reproducibility of a dynamical system's response when repetitively driven by the same fluctuating input, even though initial conditions vary across trials. Examples include the spike-time reliability of single neurons,[20,21] improvements in the reproducibility of laser dynamics,[22] and synchronized variations in wild animal populations located in distinct, well-separated areas caused by common environmental fluctuations.[23]
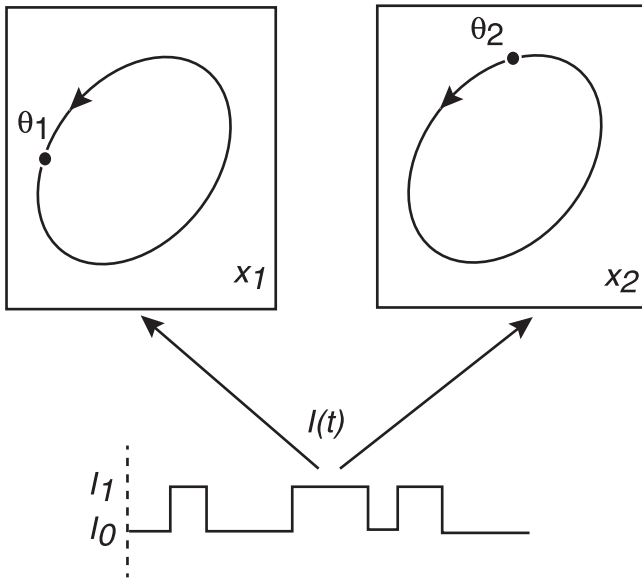
FIG. 1. Pair of non-interacting limit cycle oscillators with phases $\theta_j(t)$, $j = 1, 2$, driven by a common switching external input $I(t)$.

Most studies of noise-induced synchronization take the oscillators to be driven by common Gaussian noise. Typically, phase synchronization is established by constructing the Lyapunov exponent for the dynamics of the phase difference between a pair of oscillators and averaging with respect to the noise. If the averaged Lyapunov exponent is negative definite, then the phase difference is expected to decay to zero in the large time limit, establishing phase synchronization. However, it has also been shown that common Poisson-distributed random impulses, dichotomous or telegrapher noise, and other types of noise generally induce synchronization of limit-cycle oscillators.[24–26] Consider, in particular, the case of an additive dichotomous noise signal $I(t)$ driving a population of M identical non-interacting oscillators according to the system of equations $\dot{x}_j = F(x_j) + I(t)$, where $x_j \in \mathbb{R}^d$ is the state of the jth oscillator, $j = 1, \ldots, M$[25] (see Fig. 1). Here, $I(t)$ switches between two values I_0 and I_1 at random times generated by a two-state Markov chain.[27] That is, $I(t) = I_0[1 - N(t)] + I_1 N(t)$ for $N(t) \in \{0, 1\}$, with the time T between switching events taken to be exponentially distributed with mean switching time τ. Suppose that each oscillator supports a stable limit cycle for each of the two input values I_0 and I_1. It follows that the internal state of each oscillator randomly jumps between the two limit cycles. Nagai *et al.*[25] show that in the slow switching limit (large τ), the dynamics can be described by random phase maps. Moreover, if the phase maps are monotonic, then the associated Lyapunov exponent is generally negative and phase synchronization is stable.

The dichotomous noise-driven system is just one example of a much more general class of randomly switching processes known as *piecewise deterministic Markov processes* (PDMPs).[28,29] More explicitly, let $N(t) \in \Gamma \equiv \{0, \ldots, \Lambda_0 - 1\}$ denote the state of the randomly switching environment. When the environmental state is $N(t) = n$, each oscillator $x_i(t)$ evolves according to the piecewise deterministic ordinary differential equation (ODE) $\dot{x}_i = F_n(x_i)$, $i = 1, \ldots, M$, where the vector field $F_n : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth function. The

discrete stochastic variable $N(t)$ evolves according to a stationary, continuous-time Markov chain with transition matrix $\mathbf{W}$. The additive dichotomous noise case is recovered by taking $\Lambda_0 = 2$ and $F_n(x) = F(x) + I_n$. One major application of PDMPs is stochastic gene regulatory networks, where the continuous variables x_j are the concentrations of protein products (and possibly mRNAs) and the discrete variables represent the various activation/inactivation states of the genes.[30–40] The common randomly switching environment could represent the state of a promoter site that is common to a pair of genes within the same cell or the state of the extracellular environment that drives gene expression in a population of cells. It is thought that synchronous oscillations may have an important functional purpose in systems biology.[34]

In this paper, we develop the theory of noise-induced synchronization for a population of non-interacting PDMPs evolving under a common randomly switching environment. (The population model is presented in Sec. II.) In the slow switching limit, one could generalize the approach of Nagai *et al.*[25] by assuming that each of the vector fields $F_n(x_i)$, $n \in \Gamma$, supports a stable limit cycle and constructing the associated random phase maps. Here, instead, we consider the fast switching regime in which the transition rates between the discrete states $n \in \Gamma$ are much faster than the relaxation rates of the piecewise deterministic dynamics for $x_i \in \mathbb{R}^d$. Thus, there is a separation of time scales between the discrete and continuous processes, so that if r is the characteristic relaxation rate of the continuous dynamics, then r/ϵ is the characteristic transition rate of the Markov chain for some small positive dimensionless parameter ϵ. If the Markov chain is ergodic, then, in the fast switching limit $\epsilon \to 0$, one obtains a deterministic dynamical system in which one averages the piecewise dynamics with respect to the corresponding unique stationary distribution. Suppose that in the deterministic limit, we have a population of independent limit cycle oscillators. Since there is no coupling or remaining external drive to the oscillators, their phases are uncorrelated. The basic issue we wish to address is whether or not phase synchronization occurs when $\epsilon > 0$; we will refer to the resulting oscillators as *stochastic hybrid limit cycle oscillators*. We will proceed by constructing the Lyapunov exponent for a pair of such oscillators driven by a common randomly switching environment.

In Sec. III, we obtain an approximate expression for the Lyapunov exponent by considering a quasi-steady-state (QSS) diffusion approximation of the underlying PDMPs[41] (see also Appendix A), in which each oscillator is approximated by a stochastic differential equation (SDE) with a common Gaussian input. This allows us to adapt previous work on the phase reduction of stochastic limit cycle oscillators[10–12,14] and thus establish that phase synchronization occurs under the diffusion approximation. However, the QSS approximation is only intended to be accurate over timescales that are longer than $O(\epsilon)$. Hence, it is unclear whether or not the associated Lyapunov exponent is accurate, since it is obtained from averaging the fluctuations in the noise over infinitesimally small timescales. Therefore, in Sec. IV, we derive a more accurate expression for the Lyapunov exponent by working directly with an exact implicit equation for the phase dynamics. We exploit the fact that multiple switching events (jumps) occur

during small excursions around the limit cycle for small ϵ, which allows us to express the Lyapunov exponent in terms of discrete sums over these events. Taking expectations then yields an expression for the Lyapunov exponent that differs significantly from the one obtained using the diffusion approximation. Note, however, that both are negative definite, so they both imply phase synchronization but at different rates. Our derivation of the Lyapunov exponent from the exact phase equation also allows us to obtain greater insights into the nature of the QSS approximation and the meaning of the associated Brownian motion (see also Appendix B). Finally, we illustrate the theory by considering the particular example of radial isochron clocks (Sec. V).

II. POPULATION OF STOCHASTIC HYBRID LIMIT CYCLE OSCILLATORS

Consider a population of identical, noninteracting dynamical systems labeled $i = 1, \ldots, M$, whose states are described by the pair $[x_i(t), N(t)] \in \Sigma \times \Gamma$, where $x_i(t)$ is a continuous variable in a connected bounded domain $\Sigma \subset \mathbb{R}^d$ and $N(t)$ is an i-independent discrete stochastic variable taking values in the finite set $\Gamma \equiv \{0, \ldots, \Lambda_0 - 1\}$. The latter represents the state of an environment that is common to all members of the population. When the environmental state is $N(t) = n$, $x_i(t)$ evolves according to the piecewise deterministic ODE

$$\dot{x}_i = F_n(x_i), \quad i = 1, \ldots, M, \tag{2.1}$$

where the vector field $F_n : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth function. We assume that the dynamics of x_i is confined to the domain Σ. The discrete stochastic variable is taken to evolve according to a homogeneous, continuous-time Markov chain with x-independent generator $\mathbf{A}$, where

$$A_{nm} = W_{nm} - \delta_{n,m} \sum_{k \in \Gamma} W_{kn}, \tag{}$$

and $\mathbf{W}$ is the transition matrix. We make the further assumption that the chain is irreducible, that is, there is a non-zero probability of transitioning, possibly in more than one step, from any state to any other state of the Markov chain. This implies the existence of a unique invariant probability distribution on Γ, denoted by ρ, such that

$$\sum_{m \in \Gamma} A_{nm} \rho_m = 0, \quad \forall n \in \Gamma. \tag{2.2}$$

As a simple example, suppose that $N(t)$ evolves according to a two-state Markov chain. That is, $N(t) \in \Gamma \equiv \{0, 1\}$, and the generator of the Markov chain is given by the matrix

$$\mathbf{A} = \begin{pmatrix} -k_- & k_+ \\ k_- & -k_+ \end{pmatrix}. \tag{2.3}$$

The corresponding stationary distribution of the Markov chain then has components

$$\rho_0 = \frac{k_+}{k_+ + k_-}, \quad \rho_1 = \frac{k_-}{k_+ + k_-}. \tag{2.4}$$

Equation (2.1) defines a PDMP[28] on $\mathbb{R}^d$ for each $i = 1, \ldots, M$, also known as a stochastic hybrid system (SHS). A useful way to implement the PDMP is to decompose

the transition matrix of the Markov chain as $W_{nm} = P_{nm} \lambda_m$, with $\sum_{n \neq m} P_{nm} = 1$. Here, λ_m is the rate of the exponential waiting time density for transitions from state m, whereas P_{nm} is the probability of the transition $m \to n$, $n \neq m$. Suppose that $N(t) = n_0$ and let t_1 be an exponentially distributed random variable with rate λ_{n_0}. That is,

$$\mathbb{P}(t_1 < t) = 1 - \exp\left(-\lambda_{n_0} t\right).$$

Then, in the random time interval $s \in [0, t_1)$, the state of the ith system is $[x^{(0)}(s), n_0]$ with $x^{(0)}$ evolving according to Eq. (2.1) for $n = n_0$. (For the moment, we drop the population label i.) At the random time t_1, we choose an internal state $n_1 \in \Gamma$ with probability $P_{n_1 n_0}$ and call $x^{(1)}(t)$ the solution of the following Cauchy problem on $[t_1, \infty)$:

$$\begin{cases} \dot{x}^{(1)}(t) = F_{n_1}[x^{(1)}(t)], & t \geq t_1, \\ x^{(1)}(t_1) = x^{(0)}(t_1). \end{cases}$$

Iterating this procedure, we construct a sequence of increasing jumping times $(t_k)_{k \geq 0}$ (setting $t_0 = 0$) and a corresponding sequence of internal states $(n_k)_{k \geq 0}$. The evolution $[x(t), N(t)]$ is then defined as (see Fig. 2)

$$[x(t), N(t)] = [x^{(k)}(t), n_k] \quad \text{if } t_k \leq t < t_{k+1}. \tag{2.5}$$

We introduce the population vector $\mathbf{x}(t) = [x_1(t), \ldots, x_M(t)]$ and define the probability density $p_n(\mathbf{x}, t)$, given the initial conditions $x_i(0) = x_{i,0}, N(0) = n_0$, according to

$$p_n(\mathbf{x}, t) d\mathbf{x} = \mathbb{P}\{\mathbf{x}(t) \in (\mathbf{x}, \mathbf{x} + d\mathbf{x}), N(t) = n | \mathbf{x}_0, n_0\}.$$

It can be shown that p evolves according to the forward differential Chapman-Kolmogorov (CK) equation[29,42]

$$\frac{\partial p_n}{\partial t} = \mathbb{L} p_n, \tag{2.6}$$

with the generator $\mathbb{L}$ defined according to

$$\mathbb{L} p_n(\mathbf{x}, t) = -\sum_{i=1}^{M} \nabla_i \cdot [F_n(x_i) p_n(\mathbf{x}, t)] + \frac{1}{\epsilon} \sum_{m \in \Gamma} A_{nm} p_m(\mathbf{x}, t). \tag{2.7}$$

Here, ∇_i denotes the d-dimensional gradient operator with respect to x_i. The first term on the right-hand side represents the probability flow associated with the piecewise deterministic dynamics for a given n, whereas the second term represents jumps in the discrete state n. Note that we have rescaled
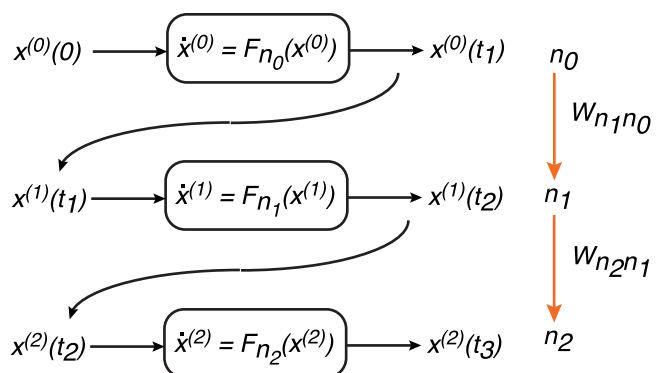
FIG. 2. Schematic diagram of a PDMP for a sequence of jump times $\{t_1, \ldots\}$ and a corresponding set of discrete states $\{n_0, n_1, \ldots\}$. See text for details.

the matrix $\mathbf{A}$ (and hence the rates λ_n) by introducing the dimensionless parameter ϵ, $\epsilon > 0$. This is motivated by the observation that many applications of PDMPs involve a separation of time-scales between the relaxation time for the dynamics of the continuous variables x and the rate of switching between the different discrete states n of the environment. The fast switching limit then corresponds to the case $\epsilon \to 0$. Now, we introduce the averaged vector field $\overline{F} : \mathbb{R}^d \to \mathbb{R}^d$ by

$$\overline{F}(x) = \sum_{n \in \Gamma} \rho_n F_n(x) \tag{2.8}$$

and define the averaged system

$$\begin{cases} \dot{x}_i(t) = \overline{F}[x_i(t)], & i = 1, \dots, M, \\ x_i(0) = x_0. \end{cases} \tag{2.9}$$

Intuitively speaking, one expects the PDMP (2.1) to reduce to the deterministic dynamical system (2.9) in the fast switching limit $\epsilon \to 0$. That is, for sufficiently small ϵ, the Markov chain undergoes many jumps over a small time interval Δt during which $\Delta x \approx 0$ and, thus, the relative frequency of each discrete state n is approximately ρ_n. This can be made precise in terms of a law of large numbers for PDMPs proven in Ref. 43.

In the fast switching (deterministic) limit, each member of the population becomes independent, since the dependence on the current state of the environment disappears. In this paper, we will assume that for each $i = 1, \dots, M$, the averaged dynamical system (2.9) supports a set of stable periodic solutions with the same natural frequency $\overline{\omega} = 2\pi/\overline{\Delta}$, where $\overline{\Delta}$ is the period. That is, we have a population of identical, independent oscillators in the fast switching limit. In state space, each periodic solution is an isolated attractive trajectory or limit cycle. The dynamics on the limit cycle can be described by a uniformly rotating phase such that

$$\frac{d\theta_i}{dt} = \overline{\omega}, \tag{2.10}$$

and $x_i = \Phi[\theta_i(t)] = \Phi(\overline{\omega}t + \psi_i)$ with Φ being a 2π-periodic function and ψ_i the initial phase. Note that Φ satisfies the equation

$$\overline{\omega} \frac{d\Phi}{d\theta} = \overline{F}[\Phi(\theta)]. \tag{2.11}$$

Differentiating both sides with respect to θ gives

$$\frac{d}{d\theta}\left(\frac{d\Phi}{d\theta}\right) = \overline{\omega}^{-1} \overline{J}(\theta) \cdot \frac{d\Phi}{d\theta}, \tag{2.12}$$

where $\overline{J}$ is the 2π-periodic Jacobian matrix

$$\overline{J}_{ab}(\theta) \equiv \left.\frac{\partial \overline{F}_a}{\partial x_b}\right|_{x=\Phi(\theta)} \tag{2.13}$$

for $a, b = 1, \dots, d$.

In the deterministic limit, there is no mechanism for phase synchronizing the population of oscillators, since $\theta_j(t) - \theta_i(t) = \psi_j - \psi_i$ for all t. The main issue we wish to address in this paper is whether or not the presence of a common switching environment can synchronize the population of stochastic hybrid oscillators when $\epsilon > 0$, analogous to the noise-driven synchronization of SDEs.[10–14]

III. DIFFUSION APPROXIMATION AND PHASE SDE

One approach to analyzing synchronization in the fast switching regime ($0 < \epsilon \ll 1$) is to use a QSS diffusion approximation, in which the CK equation (2.6) is approximated by a Fokker-Planck (FP) equation for the total density $C(\mathbf{x}, t) = \sum_n p_n(\mathbf{x}, t)$.[41] The latter determines the probability distribution of solutions of a corresponding SDE for a population of oscillators driven by a common $O(\sqrt{\epsilon})$ multiplicative noise term, which can then be reduced to an effective SDE for the phases along the lines of Ref. 10. The resulting FP equation in the Stratonovich representation takes the form (see Appendix A)

$$\frac{\partial C}{\partial t} = -\sum_{i=1}^{M} \nabla_i \cdot \left[\overline{F}(x_i)C\right] \tag{3.1a}$$

$$- \epsilon \sum_{i,j=1}^{M} \sum_{m,n \in \Gamma} A_{mn}^{\dagger} \rho_n \nabla_i \cdot \left\{ G_m(x_i) \nabla_j \cdot [G_n(x_j)C] \right\},$$

with

$$G_m(x) = F_m(x) - \overline{F}(x). \tag{3.1b}$$

Since Eq. (3.1a) is symmetric with respect to the exchange $(ni, mj) \leftrightarrow (mj, ni)$, we can replace $\mathbf{A}\rho$ by its symmetric part

$$\tilde{A}_{mn} = \frac{1}{2}(A_{mn}^{\dagger} \rho_n + A_{nm}^{\dagger} \rho_m). \tag{3.2}$$

Note that the matrix $\tilde{\mathbf{A}}$ is negative semi-definite, which can be inferred from (B.7) in Appendix B. For example, in the case of a two-state Markov chain, $\tilde{\mathbf{A}} = \text{diag}[-\rho_0/(k_+ + k_-), -\rho_1/(k_+ + k_-)]$.

It follows that under the diffusion approximation, the PDMP (2.1) can be approximated by the Stratonovich SDE

$$dX_i(t) = \overline{F}(X_i)dt + \sqrt{2\epsilon} \sum_{m,n \in \Gamma} G_m(X_i)B_{mn}dW_n(t) \tag{3.3}$$

for $i = 1, \dots, M$, where $\mathbf{B}\mathbf{B}^{\top} = -\tilde{\mathbf{A}}$, and $W(t)$ is a vector of uncorrelated Brownian motions in $\mathbb{R}^M$,

$$\mathbb{E}\left[\mathbf{W}(t)\mathbf{W}(t)^{\top}\right] = t\mathbf{I},$$

and $\mathbf{I}$ is the identity matrix. [Note that it does not matter which Hermitian square root of $-\tilde{\mathbf{A}}$ we take for $\mathbf{B}$, since they all yield the same statistical behavior of $X_i(t)$.] Equation (3.3) represents a population of independent, non-interacting limit cycle oscillators, driven by common external white noise. One can now use phase reduction methods developed for SDEs.

A. Phase reduction

First, suppose that the noise amplitude ϵ is sufficiently small relative to the rate of attraction to the limit cycle, so that deviations transverse to the limit cycle are also small (up to some exponentially large stopping time). This suggests that the definition of a phase variable persists in the stochastic setting, and one can derive a stochastic phase equation by decomposing the solution to the SDE (3.3) according to

$$X_i(t) = \Phi[\theta_i(t)] + \sqrt{\epsilon}v_i(t), \tag{3.4}$$

with $\theta_i(t)$ and $v_i(t)$ corresponding to the phase and amplitude components, respectively, of the ith oscillator. However, there
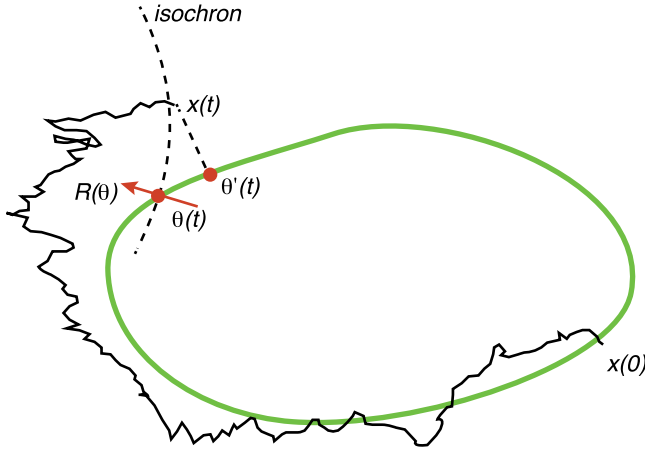
FIG. 3. Different choices of amplitude-phase decomposition. Two possibilities are orthogonal projection with phase $\theta'(t)$ and isochronal projection with phase $\theta(t)$. In the latter case, the response to perturbations depends on the phase response curve $R(\theta)$, which is normal to the isochron at the point of intersection with the limit cycle.

is not a unique way to define the phase θ_i, which reflects the fact that there are different ways of projecting the exact solution onto the limit cycle[44–47] (see Fig. 3). One well-known approach is to use the method of isochrons,[10,12–14] which we briefly outline here.

Consider the unperturbed deterministic system $\dot{x}_i = \overline{F}(x_i)$. Stroboscopically observing the system at time intervals of length $\overline{\Delta}$ leads to a Poincare mapping

$$x_i(t) \rightarrow x_i(t + \overline{\Delta}) \equiv \mathcal{P}[x_i(t)],$$

for which all points on the limit cycle are fixed points. We choose a point x_i^* on the limit cycle and consider all points in the vicinity of x_i^* that are attracted to it under the action of $\mathcal{P}$. They form a $(d - 1)$-dimensional hypersurface $\mathcal{I}$ called an isochron,[1,2,4,6,7] crossing the limit cycle at x_i^*. A unique isochron can be drawn through each point on the limit cycle (at least locally) so that the isochrons can be parameterized by the phase, $\mathcal{I} = \mathcal{I}(\theta_i)$. Finally, the definition of phase is extended by taking all points $x_i \in \mathcal{I}(\theta_i)$ to have the same phase, $\Theta(x_i) = \theta_i$, which then rotates at the natural frequency $\overline{\omega}$ (in the unperturbed case). Hence, for an unperturbed oscillator in the vicinity of the limit cycle, we have

$$\overline{\omega} = \frac{d\Theta(x_i)}{dt} = \nabla\Theta(x_i) \cdot \frac{dx_i}{dt} = \nabla\Theta(x_i) \cdot F(x_i).$$

Now, we consider the Stratonovich SDE (3.3). For the moment, we replace the $O(\epsilon^{1/2})$ term by a bounded deterministic function $H \in \mathbb{R}^d$ so that we have the perturbed deterministic equation

$$\frac{dx_i}{dt} = \overline{F}(x_i) + \sqrt{\epsilon} H(x_i, t).$$

The additional complications arising from a stochastic perturbation will be addressed below. Differentiating the isochronal phase using the chain rule gives

$$\frac{d\Theta}{dt} = \nabla\Theta(x_i) \cdot [\overline{F}(x_i) + \sqrt{\epsilon} H(x_i, t)]$$

$$= \overline{\omega} + \sqrt{\epsilon}\nabla\Theta(x_i) \cdot H(x_i, t).$$

We now make the approximation that deviations of x_i from the limit cycle are ignored on the right-hand side by setting $x_i(t) = \Phi[\theta_i(t)]$ with Φ being the 2π-periodic solution on the limit cycle. This then yields the closed phase equation

$$\frac{d\theta_i}{dt} = \overline{\omega} + \sqrt{\epsilon}\sum_{a=1}^{d} R_a(\theta_i)H_a[\Phi(\theta_i), t], \qquad (3.5)$$

where

$$R_a(\theta) = \left.\frac{\partial\Theta}{\partial x_a}\right|_{x_a=\Phi(\theta)} \qquad (3.6)$$

is a 2π-periodic function of θ known as the ath component of the *phase resetting curve* (PRC).[1,2,4,6,7] One way to evaluate the PRC is to exploit the fact that it is the 2π-periodic solution of the linear equation

$$\overline{\omega}\frac{dR(\theta)}{d\theta} = -\overline{J}(\theta)^{\top} \cdot R(\theta), \qquad (3.7)$$

under the normalization condition

$$R(\theta) \cdot \frac{d\Phi(\theta)}{d\theta} = 1. \qquad (3.8)$$

$\overline{J}(\theta)^{\top}$ is the transpose of the Jacobian matrix $\overline{J}(\theta)$.

Returning to the Stratonovich SDE (3.3), treating the stochastic perturbation along identical lines to the deterministic case (i.e., substituting $H(t)dt = \sqrt{2}\sum_{m,n\in\Gamma}\{F_m[\Phi(\theta_i)] - \overline{F}[\Phi(\theta_i)]\}B_{mn}dW_n(t)$, and exploiting the fact that the usual rules of calculus hold (in contrast to Ito SDEs) would then lead to the following SDE for the phase $\theta_i(t)$:

$$d\theta_i = \overline{\omega}dt + \sqrt{2\epsilon}\sum_{m,n\in\Gamma}\mathcal{F}_m(\theta_i)B_{mn}dW_n(t), \qquad (3.9)$$

where

$$\mathcal{F}_n(\theta_i) = \sum_{a=1}^{d} R_a(\theta_i)\left\{F_{n,a}[\Phi(\theta_i)] - \overline{F}_a[\Phi(\theta_i)]\right\}. \qquad (3.10)$$

Introducing the population phase vector $\boldsymbol{\theta}(t) = [\theta_1(t), \ldots, \theta_M(t)]$, the corresponding phase FP equation for the population probability density $C(\boldsymbol{\theta}, t)$ is

$$\frac{\partial C}{\partial t} = -\overline{\omega}\sum_{i=1}^{M}\frac{\partial C}{\partial \theta_i} \qquad (3.11)$$

$$+ \epsilon\sum_{i,j=1}^{M}\sum_{m,n\in\Gamma}\tilde{A}_{mn}\frac{\partial}{\partial\theta_i}\left\{\mathcal{F}_m(\theta_i)\frac{\partial}{\partial\theta_j}[\mathcal{F}_n(\theta_j)C]\right\}.$$

However, there are a number of major differences from the deterministic case. Probably, the most significant is that a Wiener process is not bounded, which means that over sufficiently long time intervals, there is a small but non-zero probability that the stochastic term induces large deviations from the limit cycle, resulting in a breakdown of the perturbation analysis. This issue can be addressed using variational methods and large deviation theory,[47,48] which show that for sufficiently small ϵ, the system remains in a neighborhood of the limit cycle up to exponentially long times. The second issue is that there are no $O(\sqrt{\epsilon})$ corrections to the deterministic part of the phase equation so that one has to go to $O(\epsilon)$ in order to determine the leading order corrections to the

drift term. There are two sources of $O(\epsilon)$ terms: one arises from the coupling between the phase and amplitude fluctuations transverse to the limit cycle, and the second arises from changing between Stratonovich and Ito versions of the SDE based on Ito's formula.[14,46–48] The precise form of these terms will also depend on the particular choice of the phase reduction method. However, if the limit cycle is sufficiently attracting, then they tend to have a small effect.[14] Moreover, such drift terms do not contribute to the leading order expression for the Lyapunov exponent describing the evolution of phase differences [see Eq. (3.16)]. Therefore, we shall drop such contributions in our subsequent analysis and reinterpret Eq. (3.9) as an Ito SDE.

B. Lyapunov exponent

We now use the phase SDE (3.9) interpreted in the Ito sense to investigate the effects of a common switching environment in the small ϵ regime, following previous studies.[10,12–15] The first step is to consider the SDE for the phase difference $\psi = \theta_j - \theta_1$ for any fixed j such that $j \neq 1$. Assuming ψ is infinitesimally small, we have

$$\frac{d\psi}{dt} = \sqrt{2\epsilon}\,\psi(t)\left[\sum_{m,n\in\Gamma}\mathcal{F}'_m(\phi)B_{mn}dW_n(t)\right], \qquad (3.12)$$

where $'$ denotes differentiation with respect to ϕ, and we have set $\theta_1(t) = \phi(t)$ so that $\phi(t)$ evolves according to Eq. (3.9) for $i = 1$. Introducing a new variable $y = \log(\psi)$ and using Ito's formula yields the SDE

$$dy = -\epsilon\sum_{m,n\in\Gamma}\mathcal{F}'_m(\phi)\tilde{A}_{mn}\mathcal{F}'_n(\phi)dt$$
$$+ \sqrt{2\epsilon}\left[\sum_{m,n\in\Gamma}\mathcal{F}'_m(\phi)B_{mn}dW_n(t)\right]. \qquad (3.13)$$

We define the Lyapunov exponent according to

$$\lambda_{\text{QSS}} = \lim_{T\to\infty}\frac{1}{T}[y(T) - y(0)]. \qquad (3.14)$$

(More precisely, we only take the limit in T up until the time that the system leaves a neighborhood of the limit cycle.) In the previous work,[49] we have demonstrated that such times are typically of exponential length, so there is plenty of time for the Lyapunov exponent to converge to the expected value. It follows that λ corresponds to the long-time average of the right-hand of Eq. (3.13). Assuming that the system is ergodic, we can replace the time average by an ensemble average with respect to the Wiener processes. Given

$$\mathbb{E}\left[\sum_{m,n\in\Gamma}\mathcal{F}'_m(\phi)B_{mn}dW_n(t)\right] = 0$$

for the Ito stochastic process, it follows that

$$\lambda_{\text{QSS}} = \epsilon\,\mathbb{E}\left[\sum_{m,n\in\Gamma}\mathcal{F}'_m(\phi)\tilde{A}_{mn}\mathcal{F}'_n(\phi)\right] < 0, \qquad (3.15)$$

provided $\mathcal{F}_n(\phi)$ is not a constant (since the matrix $\tilde{A}$ is negative definite). Since the Lyapunov exponent is then negative definite, we infer that the population of phases evolving

according to Eq. (3.9) synchronize, in the sense that

$$\lim_{t\to\infty}[\theta_j(t) - \theta_1(t)] = 0, \quad \text{for all } j = 1,\dots,M.$$

Moreover, assuming that a stationary density $P_s(\phi)$ exists, with $P_s(\phi) \approx 1/(2\pi)$ in the weak noise regime, then we can approximate the expectation by an integral around the limit cycle:

$$\lambda_{\text{QSS}} = \epsilon\int_0^{2\pi}\left[\sum_{m,n\in\Gamma}\mathcal{F}'_m(\phi)\tilde{A}_{mn}\mathcal{F}'_n(\phi)\right]\frac{d\phi}{2\pi} < 0. \qquad (3.16)$$

This then implies that if we had included $O(\epsilon)$ contributions to the drift, then these would yield a total derivative in the phase difference equation, which would vanish when averaged around the limit cycle.

In conclusion, we have established that under the diffusion approximation, a population of identical, stochastic hybrid limit cycle oscillators will phase synchronize when driven by a common switching environment in the fast switching limit. This then raises the issue as to whether or not this ensures synchronization of the corresponding population of PDMPs evolving according to the exact dynamics of Eq. (2.1). In particular, the QSS approximation is only intended to be accurate over timescales that are longer than $O(\epsilon)$. Hence, it is not clear to what extent the above Lyapunov exponent λ_{QSS} is accurate, because it is obtained from averaging the noise fluctuations over infinitesimally small timescales. Indeed, it follows from the smoothness of the functions $\{F_m\}$ that the infinitesimal of the exact isochronal phase $\theta_i = \Theta(x_i)$ satisfies $d\theta_i \sim O(dt)$, whereas under the diffusion approximation, $d\theta_i \sim \sqrt{\epsilon}O(dt^{1/2})$. This implies that no matter how small we take ϵ, the isochronal phase will never exhibit the relatively large fluctuations over very small timescales that is characteristic of SDEs.

In fact, the above issue also raises some questions about the conventional approach to phase synchronization in stochastic differential equations. We are not aware of any application of SDEs that is intended to be accurate over infinitely short timescales (i.e., for infinitely high frequencies): in practice, there is always a very short timescale over which the noise is not white, but highly correlated. For example, in stochastic models of stock price fluctuations, this timescale must be at least as long as the time it takes for the central computer to process a single trade. However, the Lyapunov coefficient that one obtains from the conventional stochastic phase synchronization analysis derives from averaging over these infinitesimally fine fluctuations. In fact, one finds that phase synchronization still occurs in the case of more realistic forms of environmental noise.[26]

IV. STOCHASTIC HYBRID PHASE EQUATION

In this section, we consider an alternative approach to deriving the Lyapunov exponent, which avoids the need for the QSS diffusion approximation. The method involves analyzing the PDMP for the exact isochronal phase defined according to

$$\theta_i(t) = \Theta[x_i(t)], \qquad (4.1)$$

where $x_i(t)$ now evolves according to the exact PDMP (2.1), rather than the approximate SDE (3.3).

A. Exact PDMP for the isochronal phase

Suppose that there is a finite sequence of jump times $\{t_1, \ldots t_r\}$ within the time interval $(0, T)$, and let n_k be the corresponding discrete state in the interval (t_k, t_{k+1}) with $t_0 = 0$ (see Sec. II). Introducing the set

$$\mathcal{T} = [0, T] \setminus \cup_{k=1}^{r} \{t_k\},$$

it follows that Eq. (2.1) holds for all $t \in \mathcal{T}$. Hence, using the chain rule for $t \in \mathcal{T}$,

$$\frac{d\theta_i}{dt} = \nabla\Theta(x_i) \cdot F_n(x_i)$$
$$= \overline{\omega} + \nabla\Theta(x_i) \cdot [F_n(x_i) - \overline{F}(x_i)]$$
$$= \overline{\omega} + \mathcal{H}_n(x_i), \tag{4.2}$$

where

$$\mathcal{H}_n(x) := \nabla\Theta(x) \cdot [F_n(x) - \overline{F}(x)]. \tag{4.3}$$

We will use Eq. (4.2) to derive a more accurate expression for the Lyapunov exponent that has the same sign as Eq. (3.16) but differs in explicit form. The direct method has a number of other advantages. First, it is more intuitive to preserve the piecewise nature of the stochastic dynamics, rather than replacing it by a continuous Markov process. Indeed, it is not clear which aspect of the PDMP the Brownian motion $W_n(t)$ corresponds to. (In Appendix B, we explicitly identify a random variable that corresponds to the Brownian motion in the QSS SDE. This then allows us to identify over what timescale the QSS approximation is accurate.) Second, since Eq. (4.2) is exact, it is possible to numerically solve for $\theta_i(t)$ outside the fast switching regime and thus determine how phase synchronization varies as ϵ is increased.

For sufficiently small ϵ, there is a high probability that the environmental state switches multiple times during one period Δ_0. Hence, although $\mathcal{H}_n(x_i)$ is not necessarily $O(\epsilon)$, it only applies for a small time interval before switching, and the accumulative effect of the perturbation over one cycle remains small. This suggests that we can set $x_i = \Phi(\theta_i)$ on the right-hand side of Eq. (4.2), which yields the closed PDMP for the phase

$$\frac{d\theta_i}{dt} = \overline{\omega} + \mathcal{F}_n(\theta_i), \quad t \in \mathcal{T}, \quad N(t) = n, \tag{4.4}$$

with $\mathcal{F}_n(\theta) = \mathcal{H}_n[\Phi(\theta)]$. The corresponding probability density $p_n(\boldsymbol{\theta}, t)$ evolves according to the CK equation

$$\frac{\partial p_n}{\partial t} = -\sum_{i=1}^{M} \frac{\partial}{\partial \theta_i} \{[\overline{\omega} + \mathcal{F}_n(\theta_i)]p_n(\boldsymbol{\theta}, t)\} + \frac{1}{\epsilon} \sum_{m \in \Gamma} A_{nm} p_m(\boldsymbol{\theta}, t).$$
$$\tag{4.5}$$

One way to proceed, by analogy with the analysis of SDEs,[12] would be to consider a pair of oscillators, introduce slow phase variables $\varphi_j = \theta_i - \overline{\omega}t$, and average over a single period of the limit cycle. One could then attempt to find the steady-state probability density for the resulting phase difference and establish that the phase difference is localized around zero.

However, it is difficult to make this approach rigorous, and finding the stationary solution of the CK equation for a PDMP is non-trivial.

B. Lyapunov exponent

Therefore, we will proceed by considering a pair of isochronal phases $\theta_1(t)$ and $\theta_2(t)$ evolving according to Eq. (4.2). We set $\theta_1 = \phi$, $\theta_2 = \phi + \psi$ and define $y = \log\psi$ as before. [Without loss of generality, we take $\theta_2(0) > \theta_1(0)$.] Exploiting the fact that the inter-switching times $\Delta t_k = t_{k+1} - t_k$ are exponentially distributed with an $O(1/\epsilon)$ rate, we Taylor expand $\Delta y_k := y(t_{k+1}) - y(t_k)$ to second order in Δt_k,

$$\Delta y_k = \psi(t_k)^{-1}\Delta\psi_k - \frac{1}{2}\psi(t_k)^{-2}\Delta\psi_k^2 + O(\Delta\psi_k^3),$$

where $\Delta\psi_k = \psi(t_{k+1}) - \psi(t_k)$, and

$$\Delta\psi_k = \Delta t_k[\mathcal{H}_{n_k}(x_2) - \mathcal{H}_{n_k}(x_1)]$$
$$+ \frac{1}{2}\Delta t_k^2[\mathcal{Q}_{n_k}(x_2) - \mathcal{Q}_{n_k}(x_1)] + O(\Delta t_k^3),$$

with x_j evaluated at time $t = t_k$,

$$\mathcal{Q}_n(x) = \nabla\Theta(x) \cdot \{J_n(x)F_n(x)\} + \{J_\Theta(x)F_n(x)\} \cdot F_n(x),$$
$$\tag{4.6}$$

and J_Θ is the Hessian of the isochronal phase map, and J_n is the Jacobian of F_n. Hence, we have

$$\Delta y_k = \psi(t_k)^{-1}[\mathcal{H}_{n_k}(x_2) - \mathcal{H}_{n_k}(x_1)]\Delta t_k \tag{4.7}$$
$$- \frac{1}{2}\psi(t_k)^{-2}[\mathcal{H}_{n_k}(x_2) - \mathcal{H}_{n_k}(x_1)]^2\Delta t_k^2$$
$$+ \frac{1}{2}\psi(t_k)^{-1}[\mathcal{Q}_{n_k}(x_2) - \mathcal{Q}_{n_k}(x_1)]\Delta t_k^2 + O(\Delta t_k^3).$$

Our goal is to understand the asymptotics of the rate of increase of y with respect to time, i.e., the typical value of

$$\frac{\sum_{k=0}^{K} \Delta y_k}{\sum_{k=0}^{K} \Delta t_k},$$

for large K. We will show that the second term in the decomposition (4.7) dominates the numerator of the above fraction. It is not immediately obvious why this should be the case, particularly since the second term is itself asymptotically small in ϵ, since Δt_k^2 scales as ϵ^2. In fact, over short time scales, the fluctuations due to $\psi(t_k)^{-1}[\mathcal{H}_{n_k}(x_1) - \mathcal{H}_{n_k}(x_2)]\Delta t_k$ are dominant. However, we will see that the reason that they are not dominant in the long time average is that their mean is zero, and the fluctuations decorrelate exponentially quickly in time.

The first step is to note that for sufficiently small ϵ, the system will switch many times during one period $\overline{\Delta}$, while ψ will hardly change. To this end, we introduce the cycle times $\tau_p = p\overline{\Delta}$ and set $\hat{\Delta} y_p = y(\tau_{p+1}) - y(\tau_p)$. If there are an average of $\mathcal{N}$ jumps during one cycle, we then have

$$\hat{\Delta} y_p \simeq \sum_{l=k}^{l=k+\mathcal{N}} \Delta y_l. \tag{4.8}$$

We recall that Δt_k is exponentially distributed with rate $\lambda_n \epsilon^{-1}$ when the current discrete state is n. This means that

$$\mathbb{E}[\Delta t_k | n_k = n] = \frac{\epsilon}{\lambda_n}, \quad \mathbb{E}[\Delta t_k^2 | n_k = n] = \frac{2\epsilon^2}{\lambda_n^2}.$$

Noting that $\mathbb{E}[\Delta t_k | n_k = n]$ is the mean waiting time in state n, over any particular time interval δt, we can estimate that the number of jumps from state n to some other state per cycle to be $\delta t \rho_n \lambda_n / \epsilon$ (obtained by dividing the expected time spent in state n by the average time it takes to leave state n). Summing over n, the total number of jumps is approximately

$$\delta t \epsilon^{-1} \sum_{n \in \Gamma} \rho_n \lambda_n. \tag{4.9}$$

In particular, over the course of one cycle, we expect that the total number of jumps is approximately

$$\mathcal{N} \approx \frac{\overline{\Delta}}{\epsilon} \sum_{n \in \Gamma} \rho_n \lambda_n. \tag{4.10}$$

Since x_1, x_2 will not be static over the limit cycle, we must further partition the set of $\mathcal{N}$ jumps into blocks of $\mathcal{M}$ jumps, for $1 \ll \mathcal{M} \ll \epsilon^{-1}$, such that over these $\mathcal{M}$ jumps, x_1, and x_2 (as well as ψ) are approximately constant. (The motivation for this choice of scaling is as follows: $\mathcal{M}$ must be small enough that x_1 and x_2 do not change substantially over all $\mathcal{M}$ jumps, but large enough that the number of jumps that the system makes to each state is determined by ρ.) It follows from Eq. (4.9) that the total elapsed time over $\mathcal{M}$ jumps is approximately

$$\tau_{\mathcal{M}} := \sum_{l=k}^{l=k+\mathcal{M}} \Delta t_l \simeq \frac{\epsilon \mathcal{M}}{\sum_{n \in \Gamma} \rho_n \lambda_n}. \tag{4.11}$$

Furthermore, the number of these $\mathcal{M}$ jumps that enter state n is approximately $\mathcal{M} \rho_n \lambda_n / \sum_{m \in \Gamma} \rho_m \lambda_m$.

The next step is to evaluate the sum of each of the terms on the right-hand side of Eq. (4.7) over $\mathcal{M}$ jumps. Because there are many jumps, the Law of Large Numbers implies that the sum can be approximated by its average. (We discuss this in more detail in Sec. IV C.) For the sake of illustration, let us focus on the second term and consider the following summation:

$$\sum_{l=k}^{l=k+\mathcal{M}} \psi(t_l)^{-2} [\mathcal{H}_{n_l}(x_2) - \mathcal{H}_{n_l}(x_1)]^2 \Delta t_l^2$$

$$\simeq \sum_{l=k}^{l=k+\mathcal{M}} \mathbb{E}\left[\psi(t_l)^{-2} [\mathcal{H}_{n_l}(x_2) - \mathcal{H}_{n_l}(x_1)]^2 \Delta t_l^2 \right]$$

$$\simeq \frac{\mathcal{M}}{\sum_{m \in \Gamma} \lambda_m \rho_m} \psi(t_k)^{-2} \sum_{m \in \Gamma} \rho_m \lambda_m \{\mathcal{H}_m(x_2) - \mathcal{H}_m(x_1)\}^2$$

$$\times \mathbb{E}\left[\Delta t_l^2 | n_l = m \right]$$

$$= 2\epsilon^2 \frac{\mathcal{M}}{\sum_{m \in \Gamma} \lambda_m \rho_m} \psi(t_k)^{-2} \sum_{m \in \Gamma} \frac{\rho_m}{\lambda_m} \{\mathcal{H}_m(x_2) - \mathcal{H}_m(x_1)\}^2$$

$$\simeq 2\epsilon \tau_{\mathcal{M}} \psi(t_k)^{-2} \sum_{m \in \Gamma} \frac{\rho_m}{\lambda_m} [\mathcal{H}_m(x_2) - \mathcal{H}_m(x_1)]^2.$$

Ignoring transverse amplitude fluctuations, which can be justified using methods from Ref. 49, we have

$$x_1 \simeq \Phi(\phi), \quad x_2 \simeq \Phi(\phi + \psi) \simeq x_1 + \Phi'(\phi)\psi.$$

Therefore,

$$\mathcal{H}_m(x_2) - \mathcal{H}_m(x_1) \simeq \mathcal{H}_m'[\Phi(\phi)]\Phi'(\phi)\psi,$$

and

$$\psi(t_k)^{-1}\{\mathcal{H}_m(x_2) - \mathcal{H}_m(x_1)\} \simeq \mathcal{F}_m'(\phi_k),$$

where $\mathcal{F}_m(\phi) = \mathcal{H}_m[\Phi(\phi)]$. Hence,

$$\tau_{\mathcal{M}}^{-1} \sum_{l=k}^{l=k+\mathcal{M}} \psi(t_l)^{-2} [\mathcal{H}_{n_l}(x_1) - \mathcal{H}_{n_l}(x_2)]^2 \Delta t_l^2$$

$$\simeq 2\epsilon \sum_{m \in \Gamma} \frac{\rho_m}{\lambda_m} \mathcal{F}_m'(\phi_k)^2. \tag{4.12}$$

We similarly find that

$$\tau_{\mathcal{M}}^{-1} \sum_{l=k}^{l=k+\mathcal{M}} \psi(t_l)^{-1} [\mathcal{Q}_{n_l}(x_1) - \mathcal{Q}_{n_l}(x_2)] \Delta t_l^2$$

$$\simeq 2\epsilon \sum_{m \in \Gamma} \frac{\rho_m}{\lambda_m} \frac{d}{d\phi_k} \mathcal{Q}_m[\Phi(\phi_k)]. \tag{4.13}$$

If we apply the same analysis to the summation of the linear term in Δt_k in Eq. (4.7), we notice that its expectation is approximately zero. That is,

$$\mathbb{E}\left\{ \sum_{l=k}^{l=k+\mathcal{M}} \psi(t_l)^{-1} [\mathcal{H}_{n_l}(x_1) - \mathcal{H}_{n_l}(x_2)] \Delta t_l \right\}$$

$$\simeq \frac{\mathcal{M}}{\sum_{m \in \Gamma} \lambda_m \rho_m} \psi(t_k)^{-1} \sum_{m \in \Gamma} \rho_m \lambda_m \{\mathcal{H}_m(x_1) - \mathcal{H}_m(x_2)\}$$

$$\times \mathbb{E}[\Delta t_l | n_l = m]$$

$$= \epsilon \frac{\mathcal{M}}{\sum_{m \in \Gamma} \lambda_m \rho_m} \psi(t_k)^{-1} \sum_{m \in \Gamma} \rho_m \{\mathcal{H}_m(x_1) - \mathcal{H}_m(x_2)\}$$

$$= 0,$$

since $\sum_{m \in \Gamma} \rho_m \mathcal{H}_m(x_i) = 0$. This suggests that its contribution to the Lyapunov exponent is negligible over long times. A more rigorous analysis indeed establishes that (see Appendix B)

$$\frac{1}{q\overline{\Delta}} \sum_{l=0}^{q\mathcal{N}} \psi(t_l)^{-1} [\mathcal{H}_{n_l}(x_1) - \mathcal{H}_{n_l}(x_2)] \Delta t_l \to 0 \tag{4.14}$$

as $q \to \infty$, as long as both oscillators stay close to the limit cycle.

Finally, we can take the rate of change of the phase to be approximately that of the deterministic system, over one course of the limit cycle. That is, we assume that for all $t \in [\tau_p, \tau_{p+1}]$, $\phi(t) - \phi(\tau_p) = \overline{\omega}(t - \tau_p) + O(\epsilon)$. Since, by assumption, $\mathcal{M} \ll \epsilon^{-1}$, the expressions in (4.12) and (4.13) are excellent approximations to the derivative with respect to time. Combining the above results and choosing k to be such

that $\tau_p \simeq t_k$,

$$y(\tau_{p+1}) - y(\tau_p) = \int_{\tau_p}^{\tau_{p+1}} y'(t) dt$$

$$\simeq \epsilon \int_{\tau_p}^{\tau_{p+1}} \left(\sum_{m \in \Gamma} \frac{\rho_m}{\lambda_m} \left\{ \frac{d}{d\phi} \mathcal{Q}_m[\Phi(\phi_t)] - \mathcal{F}'_m(\phi_t)^2 \right\} \right) dt$$

$$\simeq \frac{\epsilon \bar{\Delta}}{2\pi} \int_0^{2\pi} \left(\sum_{m \in \Gamma} \frac{\rho_m}{\lambda_m} \left\{ \frac{d}{d\alpha} \mathcal{Q}_m[\Phi(\alpha)] - \mathcal{F}'_m(\alpha)^2 \right\} \right) d\alpha,$$

upon a change of variable. Now

$$\int_0^{2\pi} \frac{d}{d\alpha} \mathcal{Q}_m[\Phi(\alpha)] d\alpha = \mathcal{Q}_m[\Phi(2\pi)] - \mathcal{Q}_m[\Phi(0)] = 0$$

due to the periodicity of Φ. We thus find that taking q revolutions around the limit cycle (which take a time $q\bar{\Delta}$)

$$\frac{1}{q\bar{\Delta}} \left[y(\tau_q) - y(\tau_1) \right] \simeq -\frac{\epsilon}{2\pi} \int_0^{2\pi} \sum_{m \in \Gamma} \frac{\rho_m}{\lambda_m} \mathcal{F}'_m(\alpha)^2 d\alpha.$$

We thus obtain the Lyapunov exponent

$$\lambda = -\epsilon \int_0^{2\pi} \sum_{n \in \Gamma} \rho_n \lambda_n^{-1} \mathcal{F}'_n(\phi)^2 \frac{d\phi}{2\pi} < 0. \qquad (4.15)$$

C. Remarks

The above analysis establishes that the Lyapunov exponent λ obtained from the exact isochronal phase equation differs significantly from the Lyapunov exponent λ_{QSS} obtained under the diffusion approximation, i.e., $\lambda \neq \lambda_{\mathrm{QSS}}$, with λ_{QSS} given by Eq. (3.16). Since they are both negative definite [assuming $\mathcal{F}'_n(\phi) \neq 0$], they both predict that phase synchronization will occur, but at different rates. The origin of the discrepancy is that λ_{QSS} is obtained by averaging with respect to noise fluctuations over infinitesimally small timescales that do not occur in the exact PDMP. As we show in Appendix B, one can write

$$\lambda_{\mathrm{QSS}} \simeq -\lim_{K \to \infty} \frac{1}{2t_K} \mathbb{E} \left[\left(\sum_{l=0}^K \mathcal{F}'_l \Delta t_l \right)^2 \right], \qquad (4.16)$$

whereas

$$\lambda \simeq -\lim_{K \to \infty} \frac{1}{2t_K} \mathbb{E} \left[\sum_{l=0}^K \left(\mathcal{F}'_l \Delta t_l \right)^2 \right]. \qquad (4.17)$$

(Again, the limit should only really be taken up until the time that the system leaves a neighborhood of the limit cycle.) Our derivation of λ is also useful in helping us understand how the QSS approximation works. It is not immediately obvious where the Brownian motion of Eq. (3.3) comes from. In Appendix B, we demonstrate that in Eq. (4.7), the term linear in Δt_l corresponds to the stochastic integral of the QSS reduction in (3.13). More precisely, the probability law of $\sqrt{2\epsilon} \int_0^{t_K} \sum_{m,n \in \Gamma} \mathcal{F}'_m(\phi) B_{mn} dW_n(t)$ is very close to the law of $\sum_{l=0}^K \mathcal{F}'_{n_l} \Delta t_l$. Indeed, it can be shown that their first two moments are equal to leading order in ϵ. (One could straightforwardly extend the analysis of Appendix B to demonstrate that their higher order moments must also converge as $\epsilon \to 0$.) However, the probability laws of the above two random variables are only convergent over timescales

much larger than $O(\epsilon)$, and this essentially accounts for the discrepancy in the Lyapunov exponents.

V. EXAMPLE: RADIAL ISOCHRON CLOCK

In order to illustrate the above general theory, we will consider a particularly simple model of an oscillator based on the complex amplitude equation that arises for a limit cycle oscillator close to a Hopf bifurcation:

$$\frac{dA}{dt} = (\mu + i\eta)A - (1 + i\alpha)|A|^2 A, \quad A \in \mathbb{C}. \qquad (5.1)$$

In polar coordinates, $A = re^{i\phi}$,

$$\frac{dr}{dt} = r(\mu - r^2), \quad \frac{d\phi}{dt} = \eta - \alpha r^2. \qquad (5.2)$$

This system is also known as a modified radial isochron clock model. The solution for arbitrary initial data $r(0) = r_0, \theta(0) = \theta_0$ is

$$r(t) = \sqrt{\mu} \left[1 + \frac{\mu - r_0^2}{r_0^2} e^{-2\mu t} \right]^{-1/2}, \qquad (5.3a)$$

$$\phi(t) = \phi_0 + \omega t - \frac{\alpha}{2} \log[r_0^2 + (\mu - r_0^2)e^{-2\mu t}], \qquad (5.3b)$$

where $\omega := \eta - \alpha\mu$ is the natural frequency of the stable limit cycle at $r_0^2 = \mu$. In Cartesian coordinates,

$$\frac{dx}{dt} := F_1(x, y) = \mu x - \eta y - (x^2 + y^2)(x - \alpha y), \qquad (5.4a)$$

$$\frac{dy}{dt} := F_2(x, y) = \mu y + \eta x - (x^2 + y^2)(y + \alpha x). \qquad (5.4b)$$

One of the useful features of the radial isochron clock model is that the isochronal phase can be calculated explicitly. Strobing the explicit solution (5.3) at times $t = n\Delta_0$, we see that

$$\lim_{n \to \infty} \phi(n\Delta_0) = \phi_0 - \alpha \ln r_0 \mod 2\pi.$$

Hence, we can define an isochronal phase on the whole plane according to

$$\Theta(r, \phi) = \phi - \alpha \ln r. \qquad (5.5)$$

It follows that the isochrones are logarithmic spirals with $\phi - \alpha \ln r = \text{constant}$. Now, we rewrite the phase (5.5) in Cartesian coordinates,

$$\Theta(\mathbf{x}) = \tan^{-1} \frac{y}{x} - \frac{\alpha}{2} \log(x^2 + y^2),$$

so that

$$\frac{\partial \Theta}{\partial x} = -\frac{y}{x^2 + y^2} - \alpha \frac{x}{x^2 + y^2}, \quad \frac{\partial \Theta}{\partial y} = \frac{x}{x^2 + y^2} - \alpha \frac{y}{x^2 + y^2}.$$

On the limit cycle $\Phi(\theta) = \sqrt{\mu}(\cos\theta, \sin\theta)$, so that the components of the PRC are

$$R_x(\theta) = \frac{\partial \Theta[\Phi(\mathbf{x})]}{\partial x} = \frac{1}{\sqrt{\mu}} [-\sin\theta - \alpha\cos\theta], \qquad (5.6a)$$

$$R_y(\theta) = \frac{\partial \Theta[\Phi(\mathbf{x})]}{\partial x} = \frac{1}{\sqrt{\mu}} [\cos\theta - \alpha\sin\theta]. \qquad (5.6b)$$

Given the deterministic model, the next step is to specify the corresponding PDMP for a single oscillator. One possibility is to assume that one or more of the coefficients switch. For example, in polar coordinates, we could take

$$\frac{dr}{dt} = r(\mu_n - r_i^2), \quad \frac{d\phi}{dt} = \eta_n - \alpha r^2. \qquad (5.7)$$

That is, both the amplitude and angular frequency of each oscillator switch. In Cartesian coordinates, we have

$$\frac{dx}{dt} := F_{n,1}(\mathbf{x}) = \mu_n x - \eta_n y - (x^2 + y^2)(x - \alpha y), \quad (5.8a)$$

$$\frac{dy}{dt} := F_{n,2}(\mathbf{x}) = \mu_n y + \eta_n x - (x^2 + y^2)(y + \alpha x). \quad (5.8b)$$

It immediately follows that the averaged system is given by

$$\frac{dx}{dt} = \overline{F}_1(\mathbf{x}) := \overline{\mu} x - \overline{\eta} y - (x^2 + y^2)(x - \alpha y), \qquad (5.9a)$$

$$\frac{dy}{dt} = \overline{F}_2(\mathbf{x}) := \overline{\mu} y + \overline{\eta} x - (x^2 + y^2)(y + \alpha x), \qquad (5.9b)$$

where

$$\overline{\mu} = \sum_{n \in \Gamma} \mu_n \rho_n, \quad \overline{\eta} = \sum_{n \in \Gamma} \eta_n \rho_n. \qquad (5.10)$$

The corresponding natural frequency is $\overline{\omega} = \overline{\eta} - \alpha \overline{\mu}$. Also, note from Eq. (3.1b) that

$$G_{n,1}(\mathbf{x}) = (\mu_n - \overline{\mu})x - (\eta_n - \overline{\eta})y, \qquad (5.11a)$$

$$G_{n,2}(\mathbf{x}) = (\mu_n - \overline{\mu})y + (\eta_n - \overline{\eta})x. \qquad (5.11b)$$

The phase PDMP (4.4) for the radial isochron clock takes the form of a simple velocity jump process

$$\frac{d\theta}{dt} = \omega_n, \qquad (5.12)$$

since

$$\mathcal{F}_n(\theta) = \sum_{a=1,2} R_a(\theta) \left\{ F_{n,a}[\Phi(\theta)] - \overline{F}_a[\Phi(\theta)] \right\}$$

$$= [-\sin\theta - \alpha\cos\theta][(\mu_n - \overline{\mu})\cos\theta - (\eta_n - \overline{\eta})\sin\theta]$$

$$+ [\cos\theta - \alpha\sin\theta][(\mu_n - \overline{\mu})\sin\theta + (\eta_n - \overline{\eta})\cos\theta]$$

$$= (\eta_n - \overline{\eta}) - \alpha(\mu_n - \overline{\mu}) = \omega_n - \overline{\omega}. \qquad (5.13)$$

In this case, $\mathcal{F}_n' = 0$, so Eq. (4.15) implies that the Lyapunov exponents is zero (to leading order in ϵ). This means that one cannot establish that synchronization occurs using the analysis of this paper. It is possible that synchronization might still occur, but the Lyapunov coefficient would be $o(\epsilon)$, and it is not easy to ascertain the synchronization in the numerical results.

A second possibility is to assume that the environments drive the x coordinate with a switching input I_n such that $\overline{I} = \sum_n \rho_n I_n = 0$. As a concrete example, we take the evolution in

Cartesian coordinates to be given by

$$\frac{dx}{dt} = \mu x - \eta y - (x^2 + y^2)(x - \alpha y) + \frac{x}{\sqrt{x^2 + y^2}} v_n^1, \quad (5.14a)$$

$$\frac{dy}{dt} = \mu y + \eta x - (x^2 + y^2)(y + \alpha x) + \frac{y}{\sqrt{x^2 + y^2}} v_n^2. \quad (5.14b)$$

Here, $v_n = (v_k^1, v_k^2)$ is a jump Markov process, assuming the following 4 states

$$v_1 = (2, -1), \quad v_2 = (-4, -4),$$

$$v_3 = (-3, 2), \quad v_4 = (8.8, 7.2),$$

with transition matrix

$$\begin{pmatrix} 0 & 2 & 2.5 & 0.1 \\ 1 & 0 & 0.5 & 4 \\ 0.5 & 0.7 & 0 & 2 \\ 3 & 0.4 & 0.25 & 0 \end{pmatrix}. \qquad (5.15)$$

The state vector and transition matrix were chosen arbitrarily, except for the normalization

$$\sum_{k=1}^{4} \rho_k v_k = 0,$$

to ensure that the averaged system supports a deterministic limit cycle. Hence, the averaged system is given by Eq. (5.4), whereas the phase PDMP takes the form

$$\frac{d\theta}{dt} = \omega - (\sin\theta + \alpha\cos\theta)v_n^1 + [\cos\theta - \alpha\sin\theta]v_n^2. \quad (5.16)$$

In this case, $\mathcal{F}_n' \neq 0$, and we expect a pair of hybrid oscillators to synchronize.

Numerical simulations of a pair of radial isochron oscillators evolving according to Eq. (5.14) with a common environmental drive confirm that synchronization does occur. An example set of results are shown in Fig. 4. Note, in particular, that the quantity

$$-\frac{1}{t} \log \left| \phi^1(t) - \phi^2(t) \right|$$

converges to the Lyapunov exponent λ given by Eq. (4.15), which was calculated directly from the underlying PDMP, rather than the quasi-steady-state Lyapunov exponent λ_{QSS} of Eq. (3.16).

VI. DISCUSSION

In this paper, we proved that stable oscillators subject to a common rapidly-switching noise will synchronize, in the vast majority of cases. We identified the leading order contribution to the Lyapunov exponent and explained why this is different from the Lyapunov exponent predicted by the QSS diffusion approximation. These results were shown to be consistent with a simulation of the radial isochron oscillator subject to a common environmental noise.

In more detail, we considered a population of M identical stochastic hybrid systems in $\mathbb{R}^d$, evolving according to the piecewise deterministic ODE (2.1) and subject to a common switching environment. In the fast switching limit, the resulting dynamical system (2.9) was assumed to support a stable
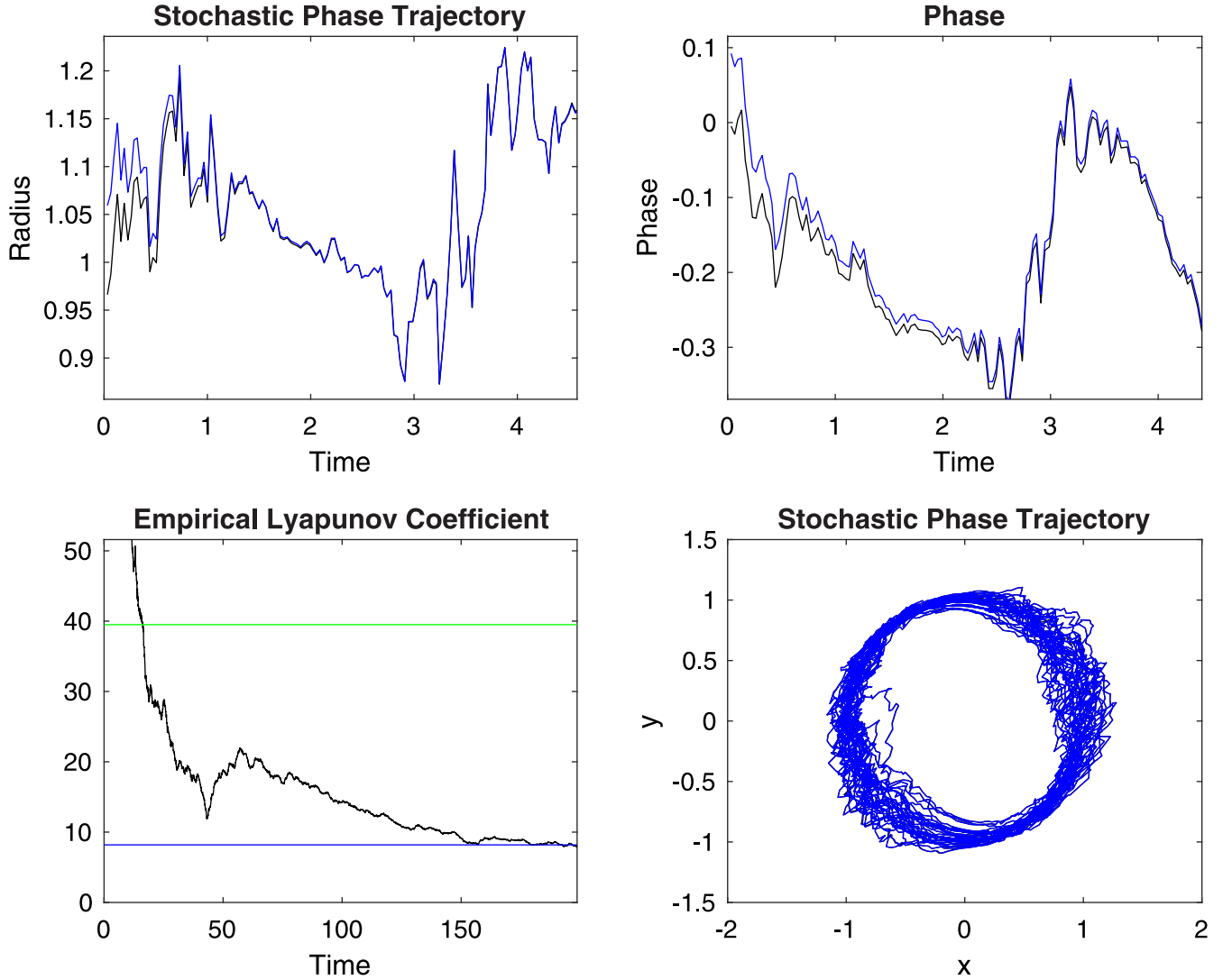
FIG. 4. Synchronization of two radial isochron oscillators with common switching noise. We set $\epsilon = 0.01$. The radii of the initial conditions differ by 0.1, and the phases of the initial condition differ by 0.1. In the top left, we plot the radii of the two oscillators, and in the top right, we plot the angle (one oscillator is plotted in black and one in blue). In the bottom right, we plot the $x - y$ trajectories of the oscillators: one in black and one in blue. In the bottom left, we plot $-\epsilon^{-1}t^{-1}\log|\phi^1(t) - \phi^2(t)|$ in black. We plot $\epsilon^{-1}\lambda_{QSS}$ in green and $\epsilon^{-1}\lambda$ in blue. It can be seen that λ is the correct Lyapunov exponent.

limit cycle. We showed that the phases of rapidly-switching oscillators converge at a rate $e^{-\epsilon\kappa t}$, where ϵ is the timescale of the switching, and the Lyapunov exponent $\lambda = -\epsilon\kappa$ is given by Eq. (4.15):

$$\kappa = \int_0^{2\pi} \sum_{n \in \Gamma} \rho_n \lambda_n^{-1} \mathcal{F}_n'(\phi)^2 \frac{d\phi}{2\pi} > 0.$$

Here, $\{\rho_n\}_{n\in\Gamma}$ is the stationary distribution of the Markovian switching process, λ_n is the rate of the exponential waiting time density for transitions from state n, and $\mathcal{F}_n$ is given by (3.10),

$$\mathcal{F}_n(\phi) = \sum_{a=1}^{d} R_a(\phi) \left\{ F_{n,a}[\Phi(\phi)] - \overline{F}_a[\Phi(\phi)] \right\},$$

where $\mathbf{x} = \Phi(\phi)$ specifies the deterministic limit cycle in $\mathbb{R}^d$, $\{R_a(\phi), a = 1, \ldots, d\}$ is the phase resetting curve, and

$$\overline{F}_a(\mathbf{x}) = \sum_{n\in\Gamma} \rho_n F_n(\mathbf{x}).$$

We also established that errors occur in estimating the rate of phase convergence under the QSS diffusion approximation. That is, the latter yields a rate $e^{-\epsilon\kappa_{QSS}t}$, where $\lambda_{QSS} = -\epsilon\kappa_{QSS}$ is given by Eq. (3.16):

$$\kappa_{QSS} = -\int_0^{2\pi} \left[\sum_{m,n\in\Gamma} \mathcal{F}_m'(\phi) \tilde{A}_{mn} \mathcal{F}_n'(\phi) \right] \frac{d\phi}{2\pi} > 0,$$

where

$$\tilde{A}_{mn} = \frac{1}{2}(A_{mn}^\dagger \rho_n + A_{nm}^\dagger \rho_m),$$

and $\mathbf{A}^\dagger$ is the pseudo-inverse of the generator $\mathbf{A}$ of the switching process.

In the limit as the switching gets faster and faster (i.e., $\epsilon \to 0$), the rate of synchronization gets slower and slower. This is not surprising, since we know that in the deterministic $\epsilon = 0$ limit, the oscillators will in general never synchronize if their starting conditions are different. These results are contingent on the two oscillators staying in the attracting neighborhood of the limit cycle. Indeed, we have shown in

Ref. [49] that the timescale over which the oscillators remain close to the limit cycle scales as $\exp\left(c\epsilon^{-1}\right)$, for a constant c. In fact, if one were to continue the analysis of this paper, and develop precise error bounds for the probability of the two oscillators synchronizing, then one would find that the smaller ϵ is, then the more likely it is that the oscillators, once they are almost synchronized, stay almost synchronized. In summary, as $\epsilon \to 0$, the oscillators synchronize at a slower and slower rate but stay synchronized with a higher and higher probability.

ACKNOWLEDGMENTS

APPENDIX A: QUASI-STEADY-STATE (QSS) REDUCTION

The basic steps of the QSS reduction of the population equations (3.3) are as follows:

(a) Decompose the probability density as

$$p_n(\mathbf{x}, t) = C(\mathbf{x}, t)\rho_n + \epsilon w_n(\mathbf{x}, t),$$

where $\sum_n p_n(\mathbf{x}, t) = C(\mathbf{x}, t)$ and $\sum_n w_n(\mathbf{x}, t) = 0$. Substituting into Eq. (2.6) yields

$$\rho_n \frac{\partial C}{\partial t} + \epsilon \frac{\partial w_n}{\partial t} = -\sum_{i=1}^{M} \nabla_i \cdot (F_n(x_i)[\rho_n C + \epsilon w_n]) \\ + \frac{1}{\epsilon}\sum_{m\in\Gamma} A_{nm}[\rho_m C + \epsilon w_m].$$

Summing both sides with respect to n then gives

$$\frac{\partial C}{\partial t} = -\sum_{i=1}^{M}\left\{\nabla_i \cdot \left[\overline{F}(x_i)C\right] + \epsilon\sum_{n\in\Gamma}\nabla_i \cdot [F_n(x_i)w_n]\right\}. \tag{A1}$$

(b) Using the equation for C and the fact that $\mathbf{A}\rho = 0$, we have

$$\epsilon\frac{\partial w_n}{\partial t} = \sum_{m\in\Gamma} A_{nm}w_m \\ - \rho_n\sum_{i=1}^{M}\left\{\nabla_i \cdot [F_n(x_i)C] - \nabla_i \cdot [\overline{F}(x_i)C]\right\} \\ - \epsilon\sum_{i=1}^{M}\left\{\left[\nabla_i \cdot [F_n(x_i)\omega_n] - \rho_n\sum_{m\in\Gamma}\nabla_i \cdot [F_m(x_i)w_m]\right]\right\}.$$

(c) Introduce the asymptotic expansion

$$w_n \sim w_n^{(0)} + \epsilon w_n^{(1)} + \epsilon^2 w_n^{(2)} + \cdots$$

and collect $O(1)$ terms:

$$\sum_{m\in\Gamma} A_{nm}w_m^{(0)} = \rho_n\sum_{i=1}^{M}\nabla_i \cdot \left[F_n(x_i) - \overline{F}(x_i)]C\right].$$

The Fredholm alternative theorem show that this has a solution, which is unique on imposing the condition

$\sum_n w_n^{(0)}(\mathbf{x}, t) = 0$:

$$w_n^{(0)}(\mathbf{x}) = \sum_{m\in\Gamma} A_{nm}^\dagger\rho_m\sum_{j=1}^{N}\left\{\nabla_j \cdot \left[F_m(x_j) - \overline{F}(x_j)\right]C\right\},$$

where $\mathbf{A}^\dagger$ is the pseudo-inverse of the generator $\mathbf{A}$.

(d) Combining Eqs. (A2) and (A1) shows that C evolves according to the Fokker-Planck (FP) equation

$$\frac{\partial C}{\partial t} = -\sum_{i=1}^{M}\nabla_i \cdot \left[\overline{F}(x_i)C\right] - \epsilon\sum_{i,j=1}^{M}\sum_{n,m\in\Gamma} A_{nm}^\dagger\rho_m \\ \times \nabla_i \cdot \left(F_n(x_i)\nabla_j \cdot \left[F_m(x_j) - \overline{F}(x_j)C\right]\right).$$

Using the fact that $\sum_m w_n = 0$, this can be rewritten in the Stratonovich form (3.1a). One typically has to determine the pseudo-inverse of $\mathbf{A}$ numerically.

APPENDIX B: ANALYSIS OF $O(\Delta t_k)$ TERM IN LYAPUNOV EXPONENT

In this appendix, we show that the $O(\Delta t_k)$ term in Eq. (4.7) does not contribute to the Lyapunov exponent in the long time limit. First, recall that

$$\psi(t_l)^{-1}[\mathcal{H}_{n_l}(x_1) - \mathcal{H}_{n_l}(x_2)] \simeq \mathcal{F}'_{n_l}(\phi).$$

From now on, we drop the dependence of $\mathcal{F}$ on ϕ to simplify notation. We start by understanding the leading order second moment of $\mathcal{M}$ terms, i.e.,

$$\mathcal{V}_\mathcal{M} := \mathbb{E}\left[\left(\sum_{l=k}^{l=k+\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l\right)^2\right].$$

(An important reason that we do this is that, as noted in Sec. IV C, the rate of change of this term with respect to time yields the Lyapunov exponent of the quasi-steady-state reduction.) $\mathcal{M}$ will be taken to be large enough that the system has switched sufficiently many times for the quasi-steady-state approximation to be accurate, but $\mathcal{M}$ is also taken to be sufficiently small that $x_1, x_2,$ and ϕ are approximately constant. In other words, $1 \ll \mathcal{M} \ll \epsilon^{-1}$. To this end, we expand out the square to obtain

$$\mathbb{E}\left[\left(\sum_{l=k}^{l=k+\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l\right)^2\right] \\ = \sum_{l=k}^{l=k+\mathcal{M}}\sum_{j=k}^{j=k+\mathcal{M}}\mathbb{E}\left[\Delta t_l\Delta t_j\mathcal{F}'_{n_l}\mathcal{F}'_{n_j}\right] \\ \simeq 2\sum_{l=k}^{l=k+\mathcal{M}}\sum_{j=0}^{\mathcal{M}}\mathbb{E}\left[\Delta t_l\Delta t_{l+j}\mathcal{F}'_{n_l}\mathcal{F}'_{n_{j+l}}\right].$$

The reason why the above approximation is very accurate is that the number of extra terms obtained through the reindexing is negligible compared to the total number of terms.

Now, we can approximate the above equation by using the fact that the correlations between $\Delta t_l\mathcal{F}'_{n_l}$ and $\Delta t_{l+j}\mathcal{F}'_{n_{l+j}}$ decay exponentially fast in j, thanks to the Perron-Frobenius theorem. The Perron-Frobenius theorem applies because the Markov chain is irreducible. Since, as demonstrated in

Sec. IV, the mean of $\Delta t_{l+j}\Delta F_{n_{l+j}}$ is 0, the contribution of the terms with asymptotically large j will be negligible. We now explain these statements in more detail.

Let $P = (P_{km})$ be the matrix with elements $P_{km} = \lambda_m^{-1}W_{km}$. If one knows that the system was in state m, and that it has jumped, then it jumps to state k with probability P_{km}. The Perron-Frobenius theorem implies that

$$\lim_{q\to\infty} P^q = B, \tag{B1}$$

where B is the rank-1 matrix with the ith element of each column equal to $\frac{1}{\sum_{a\in\Gamma}\rho_a\lambda_a}\lambda_i\rho_i$. Furthermore, the convergence is exponentially fast, i.e.,

$$||P^q - B|| = O\big[\exp(-q)\big], \tag{B2}$$

for any matrix norm.[50]

Now,

$$2\sum_{l=k}^{l=k+\mathcal{M}}\sum_{j=0}^{\mathcal{M}}\mathbb{E}\bigg[\Delta t_l\Delta t_{l+j}\mathcal{F}'_{n_l}\mathcal{F}'_{n_{j+l}}\bigg]$$

$$= 2\sum_{l=k}^{l=k+\mathcal{M}}\sum_{p,m\in\Gamma}\mathbb{P}(n_l=m)\mathbb{E}\big[\Delta t_l\mathcal{F}'_{n_l}|n_l=m\big]$$

$$\times \sum_{j=0}^{\mathcal{M}}(P^j)_{pm}\mathbb{E}\big[\Delta t_{l+j}\mathcal{F}'_{n_{j+l}}|n_{j+l}=p\big]$$

$$\simeq 2\epsilon^2\sum_{l=k}^{l=k+\mathcal{M}}\sum_{p,m\in\Gamma}\frac{\mathcal{F}'_p\mathcal{F}'_m}{\lambda_m\lambda_p}\mathbb{P}(n_l=m)\sum_{j=0}^{\mathcal{M}}(P^j)_{pm}.$$

Because x_1, x_2 change by a negligible amount over this timescale [as $\mathcal{M} = o(\epsilon^{-1})$], if $n_i = n_j$, then $\mathcal{F}'_{n_i} \simeq \mathcal{F}'_{n_j}$.

If we substitute P^j for its limit B in the above, we obtain

$$\frac{1}{\sum_{a\in\Gamma}\rho_a\lambda_a}\sum_{j=0}^{\mathcal{M}}\sum_{l=k}^{l=k+\mathcal{M}}\sum_{p,m\in\Gamma}\frac{\mathcal{F}'_m\mathcal{F}'_p}{\lambda_m\lambda_p}\mathbb{P}(n_l=m)\rho_p\lambda_p = 0,$$

since $\sum_{p\in\Gamma}\rho_p\Delta\mathcal{F}'_p = 0$. This is what we expect in light of the above discussion, because the system decorrelates after infinitely many jumps, and the mean is zero.

In light of (B2), the above discussion means that for large $\mathcal{M}$, we can extend the summation to ∞, without much of a loss of accuracy

$$2\sum_{l=k}^{l=k+\mathcal{M}}\sum_{j=0}^{\mathcal{M}}\mathbb{E}\bigg[\Delta t_l\Delta t_{l+j}\mathcal{F}'_{n_l}\mathcal{F}'_{n_{j+l}}\bigg]$$

$$\simeq 2\epsilon^2\sum_{l=k}^{l=k+\mathcal{M}}\sum_{p,m\in\Gamma}\frac{\mathcal{F}'_p\mathcal{F}'_m}{\lambda_m\lambda_p}\mathbb{P}(n_l=m)\sum_{j=0}^{\infty}(P^j)_{pm}$$

$$\simeq \frac{2\epsilon^2\mathcal{M}}{\sum_{a\in\Gamma}\rho_a\lambda_a}\sum_{p,m\in\Gamma}\frac{\rho_m\mathcal{F}'_p\mathcal{F}'_m}{\lambda_p}\sum_{j=0}^{\infty}(P^j)_{pm}, \tag{B3}$$

since, as explained above, the number of jumps to state m is approximately $\mathcal{M}\rho_m\lambda_m/(\sum_{a\in\Gamma}\rho_a\lambda_a)$.

Now, we wish to understand how the above term relates to the pseudo-inverse of A (which occurred in the QSS SDE

of Appendix A). In fact, we claim that

$$-\sum_{j=0}^{\infty}\sum_{p,q\in\Gamma}A_{mp}\lambda_p^{-1}(P^j)_{pq}\rho_q\mathcal{F}'_q = \rho_m\mathcal{F}'_m, \tag{B4}$$

which will allow us to establish the identity in (B5).

Note that the summation on the left converges because $\sum_{p\in\Gamma}A_{mp}\lambda_p^{-1}B_{pq} = 0$ for every $q, m \in \Gamma$, and, therefore, the convergence is exponential in j thanks to (B2). We expand out the left hand side, substituting A and using a truncated summation, i.e., for some $R \in \mathbb{Z}^+$,

$$-\sum_{j=0}^{R}\sum_{p,q\in\Gamma}A_{mp}\lambda_p^{-1}(P^j)_{pq}\mathcal{F}'_q\rho_q$$

$$= -\sum_{j=0}^{R}\sum_{p,q\in\Gamma}\big[W_{mp} - \lambda_p\delta(m,p)\big]\lambda_p^{-1}(P^j)_{pq}\mathcal{F}'_q\rho_q$$

$$= \sum_{q\in\Gamma}\bigg\{\sum_{j=0}^{R}(P^j)_{mq}\mathcal{F}'_q\rho_q - \sum_{j=1}^{R+1}(P^j)_{mq}\mathcal{F}'_q\rho_q\bigg\}$$

$$= \mathcal{F}'_m\rho_m - \sum_{q\in\Gamma}(P^{R+1})_{mq}\mathcal{F}'_q\rho_q.$$

Now, $\sum_{q\in\Gamma}B_{mq}\mathcal{F}'_q\rho_q = 0$, because, as noted just below (B1), the columns of B are all the same, and $\sum_{q\in\Gamma}\mathcal{H}_q\rho_q = 0$. Since $P^{R+1} \to B$ as $R \to \infty$, when we take $R \to \infty$, we obtain (B4).

It follows from (B4) that

$$-\sum_{j=0}^{\infty}\sum_{q\in\Gamma}\lambda_p^{-1}(P^j)_{pq}\rho_q\mathcal{F}'_q = \sum_{m\in\Gamma}A^\dagger_{pm}\rho_m\mathcal{F}'_m + C\rho_p, \tag{B5}$$

for some C, where we recall the pseudo-inverse $A^\dagger$, defined in Appendix A on the QSS approximation.

Substituting this identity into (B3), we find that

$$\sum_{l=k}^{l=k+\mathcal{M}}\sum_{j=0}^{\mathcal{M}}\mathbb{E}\bigg[\Delta t_l\Delta t_{l+j}\mathcal{F}'_{n_l}\mathcal{F}'_{j+l}\bigg]$$

$$\simeq -\frac{\epsilon^2\mathcal{M}}{\sum_{a\in\Gamma}\rho_a\lambda_a}\sum_{p,q\in\Gamma}A^\dagger_{pq}\mathcal{F}'_p\mathcal{F}'_q\rho_q. \tag{B6}$$

This implies that

$$\frac{1}{2(t_{k+\mathcal{M}} - t_k)}\mathcal{V}_\mathcal{M} \simeq -\epsilon\sum_{p,q\in\Gamma}A^\dagger_{pq}\mathcal{F}'_p\mathcal{F}'_q\rho_q, \tag{B7}$$

using the expression for $\tau_\mathcal{M} = t_{k+\mathcal{M}} - t_k$ in (4.11). Just like the end of Sec. IV B, we find that

$$\frac{1}{2t_K}\mathbb{E}\bigg[\bigg(\sum_{l=0}^{K}\mathcal{F}'_l\Delta t_l\bigg)^2\bigg]$$

$$\simeq -\frac{\epsilon}{2\pi}\int_0^{2\pi}\sum_{p,q\in\Gamma}A^\dagger_{pq}\mathcal{F}'_p(\theta)\mathcal{F}'_q(\theta)\rho_q d\theta$$

$$= \lambda_{QSS}. \tag{B8}$$

We have thus established the claim in (4.16).

The final step in the proof is to show that

$$\mathbb{P}\left(t_{q\mathcal{M}}^{-1}\sum_{l=0}^{q\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l \geq O(\epsilon)\right) \ll 1. \tag{B9}$$

(Technically, this probability is conditional on both systems staying in a neighborhood of the limit cycle until the time $t_{q\mathcal{M}}$. We have demonstrated that this occurs for very long times, with very high probability, elsewhere.[49]) To this end, we use Chebyshev's inequality, noting that

$$\mathbb{P}\left(t_{q\mathcal{M}}^{-1}\sum_{l=0}^{q\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l \geq a\right)$$
$$\leq \frac{1}{(at_{q\mathcal{M}})^2}\mathbb{E}\left[\left(\sum_{l=0}^{q\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l\right)^2\right]. \tag{B10}$$

We now bound the rate of growth in time of the expectation on the right. The immediate use of this bound will be to show that the probability is negligible. A secondary use is that it will help us understand how this discrete approximation compares to the QSS SDE.

We must thus understand the second order moment of $q\mathcal{M}$ terms, i.e.,

$$\mathbb{E}\left[\left(\sum_{l=0}^{q\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l\right)^2\right] = \sum_{l=0}^{q\mathcal{M}}\sum_{j=0}^{q\mathcal{M}}\mathbb{E}\left[\Delta t_l \Delta t_j \mathcal{F}'_{n_l}\mathcal{F}'_{n_j}\right]$$
$$\simeq 2\sum_{l=0}^{q\mathcal{M}}\sum_{j=0}^{q\mathcal{M}}\mathbb{E}\left[\Delta t_l \Delta t_{l+j} \mathcal{F}'_{n_l}\mathcal{F}'_{n_{j+l}}\right]$$
$$= 2\sum_{r=1}^{q}\sum_{l=(r-1)\mathcal{M}}^{l=r\mathcal{M}}\sum_{j=0}^{q\mathcal{M}}\mathbb{E}\left[\Delta t_l \Delta t_{l+j} \mathcal{F}'_{n_l}\mathcal{F}'_{n_{j+l}}\right].$$

Using the estimate derived in (B6), we find that

$$\mathbb{E}\left[\left(\sum_{l=0}^{q\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l\right)^2\right] - \mathbb{E}\left[\left(\sum_{l=0}^{(q-1)\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l\right)^2\right]$$
$$\simeq -\frac{2\epsilon^2\mathcal{M}}{\sum_{a\in\Gamma}\rho_a\lambda_a}\sum_{p,u\in\Gamma}\mathcal{A}^\dagger_{pu}\mathcal{F}'_p\mathcal{F}'_u\rho_u.$$

Hence, using the expression for the time for $\mathcal{M}$ jumps in (4.11),

$$\frac{1}{\sum_{l=(q-1)\mathcal{M}+1}^{q\mathcal{M}}\Delta t_l}\left\{\mathbb{E}\left[\left(\sum_{l=0}^{q\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l\right)^2\right]\right.$$
$$\left.-\mathbb{E}\left[\left(\sum_{l=0}^{(q-1)\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l\right)^2\right]\right\}$$
$$\simeq -\frac{\sum_{n\in\Gamma}\rho_n\lambda_n}{\epsilon\mathcal{M}}\frac{2\epsilon^2\mathcal{M}}{\sum_{a\in\Gamma}\rho_a\lambda_a}\sum_{p,u\in\Gamma}\mathcal{A}^\dagger_{pu}\mathcal{F}'_p\mathcal{F}'_u\rho_u$$
$$= -2\epsilon\sum_{p,u\in\Gamma}\mathcal{A}^\dagger_{pu}\mathcal{F}'_p\mathcal{F}'_u\rho_u.$$

Since the above expression approximates the derivative with respect to time, after re-integrating, we find that

$$\mathbb{E}\left[\left(\sum_{l=0}^{q\mathcal{M}}\mathcal{F}'_l\Delta t_l\right)^2\right] \simeq -2\epsilon\int_0^{t_{q\mathcal{M}}}\sum_{p,u\in\Gamma}\mathcal{A}^\dagger_{pu}\mathcal{F}'_p(\theta_s)\mathcal{F}'_u(\theta_s)\rho_u ds. \tag{B11}$$

It then follows from the above expression and (B10) that

$$\mathbb{P}\left(t_{q\mathcal{M}}^{-1}\sum_{l=0}^{q\mathcal{M}}\mathcal{F}'_{n_l}\Delta t_l \geq a\right) = O\left(\frac{\epsilon}{a^2 t_{q\mathcal{M}}}\right). \tag{B12}$$

This clearly goes to zero as $q \to \infty$. We have thus justified why the linear terms in Δt_l in (4.7) are $o(\epsilon)$ over long periods of time, and this implies (4.14). This is why the quadratic terms in Δt_l dominate the linear ones over long periods of time. More powerful bounds could be obtained by taking exponential moments, but we do not go into this in this paper.

[1]A. Winfree, *The Geometry of Biological Time* (Springer-Verlag, New York, 1980).

[2]Y. Kuramoto, *Chemical Oscillations, Waves and Turbulence* (Springer-Verlag, New-York, 1984).

[3]G. B. Ermentrout and N. Kopell, "Frequency plateaus in a chain of weakly coupled oscillators. I," SIAM J. Math. Anal. **15**, 215–237 (1984).

[4]L. Glass and M. C. Mackey, *From Clocks to Chaos* (Princeton University Press, Princeton, NJ, 1988).

[5]G. B. Ermentrout and N. Kopell, "Multiple pulse interactions and averaging in systems of coupled neural oscillators," J. Math. Biol. **29**, 195–217 (1991).

[6]G. B. Ermentrout, "Type I membranes, phase resetting curves, and synchrony," Neural Comput. **8**, 979 (1986).

[7]E. Brown, J. Moehlis, and P. Holmes, "On the phase reduction and response dynamics of neural oscillator populations," Neural Comput. **16**, 673–715 (2004).

[8]P. Ashwin, S. Coombes, and R. Nicks, "Mathematical frameworks for oscillatory network dynamics in neuroscience," J. Math. Neurosci. **6**, 1–92 (2016).

[9]H. Nakao, "Phase reduction approach to synchronization of nonlinear oscillators," Contemp. Phys. **57**, 188–214 (2016).

[10]J. N. Teramae and D. Tanaka, "Robustness of the noise-induced phase synchronization in a general class of limit cycle oscillators," Phys. Rev. Lett. **93**, 204103 (2004).

[11]D. S. Goldobin and A. Pikovsky, "Synchronization and desynchronization of self-sustained oscillators by common noise," Phys. Rev. E **71**, 045201 (2005).

[12]H. Nakao, K. Arai, and Y. Kawamura, "Noise-induced synchronization and clustering in ensembles of uncoupled limit cycle oscillators," Phys. Rev. Lett. **98**, 184101 (2007).

[13]K. Yoshimura and K. Arai, "Phase reduction of stochastic limit cycle oscillators," Phys. Rev. Lett. **101**, 154101 (2008).

[14]J. N. Teramae, H. Nakao, and G. B. Ermentrout, "Stochastic phase reduction for a general class of noisy limit cycle oscillators," Phys. Rev. Lett. **102**, 194102 (2009).

[15]C. Ly and G. B. Ermentrout, "Synchronization of two coupled neural oscillators receiving shared and unshared noisy stimuli," J. Comput. Neurosci. **26**, 425–443 (2009).

[16]G. B. Ermentrout, "Noisy oscillators," in *Stochastic Methods in Neuroscience*, edited by C. R. Laing and G. J. Lord (Oxford University Press, Oxford, 2009).

[17]R. F. Galan, G. B. Ermentrout, and N. N. Urban, "Optimal time scale for spike-time reliability: Theory, simulations and experiments," J. Neurophysiol. **99**, 277–283 (2008).

[18]T. Zhou, J. Zhang, Z. Yuan, and L. Chen, "Synchronization of genetic oscillators," Chaos **18**, 037126 (2008).

[19]L. Potvin-Trottier, N. D Lord, G. Vinnicombe, and J. Paulsson, "Synchronous long-term oscillations in a synthetic gene circuit," Nature **538**, 514–517 (2016).

[20]Z. F. Mainen and T. J. Sejnowski, "Reliability of spike timing in neocortical neurons," Science **268**, 1503 (1995).

[21]R. F. Galan, N. Fourcaud-Trocme, G. B. Ermentrout, and N. N. Urban, "Correlation-induced synchronization of oscillations in olfactory bulb neurons," J. Neurosci. **26**, 3646 (2006).

[22]A. Uchida, R. McAllister, and R. Roy, "Consistency of nonlinear system response to complex drive signals," Phys. Rev. Lett. **93**, 244102 (2004).

[23]B. T. Grenfell, K. Wilson, B. F. Finkenstadt, T. N. Coulson, S. Murray, S. D. Albon, J. M. Pemberton, T. H. Clutton-Brock, and M. J. Crawley, "Noise and determinism in synchronized sheep dynamics," Nature **394**, 674 (1998).

[24]H. Nakao, K. Arai, K. Nagai, Y. Tsubo, and Y. Kuramoto, "Synchrony of limit-cycle oscillators induced by random external impulses," Phys. Rev. E **72**, 026220 (2005).

[25]K. Nagai, H. Nakao, and Y. Tsubo, "Synchrony of neural oscillators induced by random telegraphic currents," Phys. Rev. E **71**, 036217 (2005).

[26]D. S. Goldobin, J. Teramae, H. Nakao, and G. B. Ermentrout, "Dynamics of limit-cycle oscillators subject to general noise," Phys. Rev. Lett. **105**, 154101 (2010).

[27]I. Bena, "Dichotomous Markov noise: Exact results for out-of-equilibrium systems," Int. J. Mod. Phys. B **20**, 2825 (2006).

[28]M. H. A. Davis, "Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models," J. R. Stat. Soc. Series B Stat. Methodol. **46**, 353–388 (1984).

[29]P. C. Bressloff, "Stochastic switching in biology: From genotype to phenotype (topical review)," J. Phys. A **50**, 133001 (2017).

[30]T. B. Kepler and T. C. Elston, "Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations," Biophys. J. **81**, 3116–3136 (2001).

[31]R. Karmakar and I. Bose, "Graded and binary responses in stochastic gene expression," Phys. Biol. **1**, 197–204 (2004).

[32]S. Zeiser, U. Franz, O. Wittich, and V. Liebscher, "Simulation of genetic networks modelled by piecewise deterministic Markov processes," IET Syst. Biol. **2**, 113–135 (2008).

[33]M. W. Smiley and S. R. Proulx, "Gene expression dynamics in randomly varying environments," J. Math. Biol. **61**, 231–251 (2010).

[34]A. Hilfinger and J. Paulsson, "Separating intrinsic from extrinsic fluctuations in dynamic biological systems," Proc. Natl. Acad. Sci. **108** 12167 (2011).

[35]J. M. Newby, "Isolating intrinsic noise sources in a stochastic genetic switch," Phys. Biol. **9**, 026002 (2012).

[36]A. Singh and M. Soltani, "Quantifying intrinsic and extrinsic variability in stochastic gene expression models," PLoS ONE **8**, e84301 (2013).

[37]C. Zechner and H. Koeppl, "Uncoupled analysis of stochastic reaction networks in fluctuating environments," PLoS Comput. Biol. **10**, e100394 (2014).

[38]J. M. Newby, "Bistable switching asymptotics for the self regulating gene," J. Phys. A **48**, 185001 (2015).

[39]P. G. Hufton, Y. T. Lin, T. Galla, and A. J. McKane, "Intrinsic noise in systems with switching environments," Phys. Rev. E **93**, 052119 (2016).

[40]E. Levien and P. C. Bressloff, "Robustness of stochastic chemical reaction networks to extrinsic noise: The role of deficiency," Multiscale Model. Simul. **16**, 1519–1541 (2018).

[41]J. M. Newby and P. C. Bressloff, "Quasi-steady state reduction of molecular-based models of directed intermittent search," Bull Math. Biol. **72**, 1840–1866 (2010).

[42]C. W. Gardiner, *Handbook of Stochastic Methods*, 4th ed. (Springer, Berlin, 2009).

[43]A. Faggionato, D. Gabrielli, and M. R. Crivellari, "Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors," Markov Process. Relat. Fields **16**, 497–548 (2010).

[44]D. Gonze, J. Halloy, and P. Gaspard, "Biochemical clocks and molecular noise: Theoretical study of robustness factors," J. Chem. Phys. **116**, 10997–11010 (2002).

[45]H. Koeppl, M. Hafner, A. Ganguly, and A. Mehrotra, "Deterministic characterization of phase noise in biomolecular oscillators," Phys. Biol. **8**, 055008 (2011).

[46]M. Bonnin, "Amplitude and phase dynamics of noisy oscillators," Int. J. Circ. Theor. Appl. **45**, 636–659 (2017).

[47]P. C. Bressloff and J. N. Maclaurin, "Variational method for analyzing stochastic limit cycle oscillators," SIAM J. Appl. Dyn. Syst. **17**, 2205–2233 (2018).

[48]G. Giacomin, C. Poquet, and A. Shapira, "Small noise and long time phase diffusion in stochastic limit cycle oscillators," J. Differ. Equ. **264**, 1019–1049 (2018).

[49]P. C. Bressloff and J. N. Maclaurin, "Variational method for analyzing stochastic hybrid oscillators," Chaos **28**, 063105 (2018).

[50]L. Saloff-Coste, "Lectures on finite Markov chains," in *Lectures on Probability Theory and Statistics* (Springer, 1997), pp. 301–413.