

Phase Reduction of Stochastic Biochemical Oscillators*

Paul C. Bressloff[†] and James N. MacLaurin[‡]

Abstract. A common method for analyzing the effects of molecular noise in chemical reaction networks is to approximate the underlying chemical master equation by a Fokker–Planck equation and to study the statistics of the associated chemical Langevin equation. This so-called system-size expansion involves performing a perturbation expansion with respect to a small dimensionless parameter $\epsilon = \Omega^{-1}$, where Ω characterizes the system size. For example, Ω could be the mean number of proteins produced by a gene regulatory network. In the deterministic limit $\Omega \rightarrow \infty$, the chemical reaction network evolves according to a system of ordinary differential equations based on classical mass action kinetics. In this paper we develop a phase reduction method for chemical reaction networks that support a stable limit cycle in the deterministic limit. We present a variational principle for the phase reduction, yielding an exact analytic expression for the resulting phase dynamics. We demonstrate that this decomposition is accurate over timescales that are exponential in the system size Ω . This contrasts with the phase equation obtained under the system-size expansion, which is only accurate up to times $O(\Omega)$. In particular, we show that for a constant C , the probability that the system leaves an $O(\zeta)$ neighborhood of the limit cycle before time T scales as $T \exp(-C\Omega b\zeta^2)$, where b is the rate of attraction to the limit cycle. We illustrate our analysis using the example of a chemical Brusselator.

Key words. chemical reaction networks, stochastic oscillators, phase reduction, Markov processes, chemical oscillators

AMS subject classifications. 92C40, 92C45, 34C15

DOI. 10.1137/18M1221205

1. Introduction. Genetically identical cells exposed to the same environmental conditions can show significant variation in molecular content and marked differences in phenotypic characteristics. This intrinsic variability is linked to the fact that many cellular events at the genetic level involve small numbers of molecules (low copy numbers) and has led to a large number of studies focused on investigating the origins and consequences of intrinsic or molecular noise in gene expression (see the reviews [25, 38, 40, 44, 12, 26]). In the presence of intrinsic noise, standard deterministic models of chemical reaction networks (CRNs), which are based on mass action kinetics, have to be replaced by discrete stochastic models. The stochastic dynamics of the CRN can then be formulated in terms of a chemical master equation, which determines the evolution of the probability distribution $P_{\mathbf{n}}(t)$ for $\mathbf{n} = (n_1, \dots, n_K)$, where n_i is the number of each molecular species labeled $i = 1, \dots, K$. In general, it is not possible to obtain exact solutions of the master equation, even in the case of a stationary solution. (Note, however, that recent progress has been made by generalizing the theory of determin-

*Received by the editors October 17, 2018; accepted for publication (in revised form) by Y. Bakhtin October 1, 2019; published electronically January 8, 2020.

<https://doi.org/10.1137/18M1221205>

Funding: This work was supported by the National Science Foundation (DMS-1613048).

[†]Department of Mathematics, University of Utah, Salt Lake City, UT 84112 (bressloff@math.utah.edu).

[‡]Department of Mathematical Sciences, NJIT, Newark, NJ 07102 (james.n.maclaurin@njit.edu).

istic chemical reaction networks to stochastic models, using notions of weak reversibility and deficiency zero [2, 4].) Therefore, one often resorts to some form of approximation scheme. The most common approach is to carry out a Taylor expansion of the master equation with respect to a dimensionless variable Ω that characterizes the size of the system [45, 14]. This so-called system-size expansion yields a Fokker–Planck (FP) equation, whose solution determines the probability density $p(\mathbf{x}, t)$ for the concentrations $\mathbf{x} = \mathbf{n}/\Omega$, which evolve according to a chemical Langevin equation.

A general issue is to what extent the system-size expansion of a chemical master equation accounts for the statistics of the underlying CRN. In cases where the deterministic mass action kinetics have a unique stable fixed point, the corresponding stationary distribution of the master equation can be well approximated by Gaussian-like fluctuations around the fixed point. One can thus obtain good estimates for the mean and variance of molecular concentrations using the system-size expansion [42, 14, 38]. However, such an approximation can break down when there are multiple fixed points (metastability). Although one can still account for the effects of fluctuations well within the basin of attraction of each metastable fixed point, there is now a small probability that there is a noise-induced transition to the basin of attraction of another fixed point. Since the probability of such a transition is usually of order $e^{-\tau\Omega}$ with $\tau = O(1)$, except close to the boundary of the basin of attraction, such a contribution cannot be analyzed accurately using standard Fokker–Planck methods. That is, for large Ω the transitions between different metastable states typically involve rare transitions (large fluctuations) that lie in the tails of the associated probability distribution, where the Gaussian approximation no longer necessarily holds. These exponentially small transitions play a crucial role in allowing the system to approach the unique stationary state (if it exists) in the asymptotic limit $t \rightarrow \infty$. In other words, for bistable (or multistable) systems, the limits $t \rightarrow \infty$ and $N \rightarrow \infty$ do not necessarily commute [24, 46]. In order to ensure accurate estimates of transition rates, one has to use alternative approximation schemes based on large deviation theory and perturbation methods, which are applied directly to the underlying master equation or CK equation [12]. We also note that several recent papers bound the probability of large deviations over fixed timescales [29, 1].

Similar issues hold for other types of attractors such as stable limit cycles. In the absence of noise, one can represent the dynamics of a limit cycle in terms of motion around a closed curve in state space. Each point on the limit cycle can be assigned a phase $\theta \in [0, 2\pi]$ such that the dynamics is given by $\dot{\theta} = \omega_0$, where $\omega_0 = 2\pi/\Delta_0$ and Δ_0 is the period of the oscillation. When noise is included, one tends to observe an irregular trajectory around the limit cycle. The resulting stochastic phase of the oscillator (assuming that it can be defined) undergoes diffusion around the limit cycle, so all phase information is eventually lost. The effective diffusion coefficient of this process is one measure of robustness to phase decoherence [23]. (For a general discussion of robustness in biological processes, see [27, 37, 34].) A second effect of noise is that it can extend the parameter regime over which oscillations can occur. That is, even though the deterministic system converges to a fixed point, the system exhibits oscillatory behavior when noise is included. This can be established by looking at the power spectra of the concentrations [6, 7]. As far as we are aware, almost all previous studies of stochastic CRN oscillators are based on the system-size expansion of chemical master equations or modifications of such approximations, which generate more accurate estimations

of the statistics that remain uniformly accurate for long times [35]. One exception is a hybrid computational study, which combines simulations of the discrete Markov process via Gillespie and solutions of the approximate SDE [41].

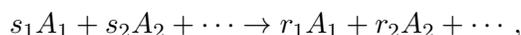
Recently, we have developed a variational method for carrying out a phase reduction of stochastic limit cycle oscillators evolving according to a Langevin equation, which yields an exact analytical expression for the phase dynamics [9]. Within the variational framework, the minimization scheme determines the phase by projecting the full stochastic solution on to the limit cycle using Floquet vectors. Hence, in a neighborhood of the limit cycle the phase variable coincides with the well-known isochronal phase. This has the advantage that the phase dynamics decouples from transverse fluctuations to leading order. Another major feature of the variational approach is that it allows one to obtain an exponential bound on the growth of transverse fluctuations. This issue, which is typically ignored in studies of stochastic phase oscillators, is important since any phase reduction scheme ultimately breaks down over sufficiently long timescales, since there is a nonzero probability of leaving a bounded neighborhood of the limit cycle, and the notion of phase no longer makes sense. The mean time to escape a neighborhood of the limit cycle is thus a measure of robustness to transverse fluctuations. In a subsequent paper, we extended the variational method to piecewise-deterministic Markov processes (PDMPs) that support a limit cycle in the adiabatic or fast switching limit [10]. (A PDMP evolves deterministically except at a sequence of random times where the deterministic dynamics switches to a different mode according to a discrete Markov process [11, 12].)

In this paper, we further extend our variational method in order to explore the robustness of stochastic biochemical oscillators to transverse fluctuations, which we analyze by working directly with the underlying discrete Markov process rather than resorting to a system-size expansion. The limitation of the latter within this context is that it is only accurate up to times $O(\Omega)$, whereas we are interested in obtaining exponential bounds on the typical time that the system spends in a neighborhood of the limit cycle. We first derive an approximate stochastic phase equation for the exact discrete Markov process using Kurtz's random time change representation [31, 32, 33, 3] and the method of isochrons. We then show how this equation can be derived more rigorously from a variational principle, which yields an exact implicit equation for the phase dynamics. We also use the variational principle to prove that, with very high probability, the time to escape a neighborhood of the limit cycle scales as $\exp(C\Omega b)$, where C is a constant, and b is the rate of attraction toward the limit cycle. The fact that the timescale of the bound is so large demonstrates the utility of using an exact representation of the phase.

The organization of the paper is as follows. In section 2 we define a CRN and review the method of isochrons for analyzing deterministic biochemical oscillators. We then use the method of isochrons to derive a stochastic phase equation under the system-size expansion. In section 3 we show how to derive a corresponding phase equation for the exact discrete Markov process using Kurtz's random time change representation [31, 32, 33, 3]. The variational method for CRNs is presented in section 4 and we illustrate our theory by numerically simulating the stochastic Brusselator. Rigorous exponential bounds on the probability of escape from a neighborhood of the limit cycle are obtained in section 5. Finally, in section 6, we summarize the paper and draw some conclusions.

2. Biochemical oscillators and the system-size expansion.

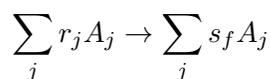
2.1. Chemical reaction network. Consider a well-mixed compartment containing a set of chemical species A_j , $j = 1, \dots, K$. Let n_j be the number of molecules of A_j and set $\mathbf{n} = (n_1, \dots, n_K)$. A typical single-step chemical reaction takes the form



where s_j, r_j are known as *stoichiometric coefficients*. When one such reaction occurs the state \mathbf{n} is changed according to

$$n_i \rightarrow n_i + r_i - s_i.$$

Introducing the vector \mathbf{S} with $S_i = r_i - s_i$, we have $\mathbf{n} \rightarrow \mathbf{n} + \mathbf{S}$. The reverse reaction



would then have $\mathbf{n} \rightarrow \mathbf{n} - \mathbf{S}$. More complicated multistep reactions can always be decomposed into these fundamental single-step reactions with appropriate stoichiometric coefficients.

In the case of a large number of molecules, one can describe the dynamics of a single step chemical reaction in terms of a kinetic or rate equation involving the concentrations $x_j = n_j/\Omega$ —the *law of mass action*. Here Ω is a dimensionless quantity representing the system size, which is usually taken to be the mean number of molecules or some volume scale factor. In a well-mixed container there is a spatially uniform distribution of each type of molecule, and the probability of a collision depends on the probability that each of the reactants is in the same local region of space. Ignoring any statistical correlations, the latter is given by the product of the individual concentrations. It then follows that the kinetic equations take the form

$$(2.1) \quad \frac{dx_i}{dt} = \kappa(r_i - s_i) \prod_{j=1}^K x_j^{s_j} \equiv S_i \lambda(\mathbf{x}),$$

where $\mathbf{x} = (x_1, \dots, x_K)$ and κ is a rate constant. Now suppose that there are $a = 1, \dots, M$ separate single-step reactions. Then

$$(2.2) \quad \frac{dx_i}{dt} := F_i(\mathbf{x}) = \sum_{a=1}^M S_{ia} \lambda_a(\mathbf{x}), \quad i = 1, \dots, K,$$

where a labels a single-step reaction and \mathbf{S} is the so-called $K \times M$ stoichiometric matrix for K molecular species and M reactions. Thus S_{ia} specifies the change in the number of molecules of species i in a given reaction a . The functions λ_a are known as transition intensities or *propensities*.

Mathematically speaking, the kinetic equations hold in the thermodynamic limit $\Omega \rightarrow \infty$. For finite Ω , it is necessary to take into account fluctuations in the number $N_i(t)$ of each chemical species. Let $P_{\mathbf{n}}(t) = \mathbb{P}[N_1(t) = n_1, \dots, N_K(t) = n_K]$ with $\mathbf{n} = (n_1, \dots, n_K)$. Given

the kinetic equations (2.2), the probability distribution $P_{\mathbf{n}}(t)$ evolves according to the chemical master equation

$$(2.3) \quad \frac{dP_{\mathbf{n}}(t)}{dt} = \Omega \sum_{a=1}^R \left(\prod_{i=1}^K \mathbb{L}^{-S_{ia}} - 1 \right) \lambda_a(\mathbf{n}/\Omega) P_{\mathbf{n}}(t).$$

Here $\mathbb{L}^{-S_{ia}}$ is a step or ladder operator such that for any function $g(\mathbf{n})$,

$$(2.4) \quad \mathbb{L}^{-S_{ia}} g(n_1, \dots, n_i, \dots, n_K) = g(n_1, \dots, n_i - S_{ia}, \dots, n_K).$$

In general, it is not possible to obtain exact solutions of the master equation (2.3) even in the case of a stationary solution. (Note, however, that progress has been made by generalizing the theory of deterministic chemical reaction networks to stochastic models, using notions of weak reversibility and deficiency zero [2].) Therefore, one often resorts to some form of approximation scheme.

2.2. Biochemical oscillators and isochrons. The next step is to assume that the deterministic dynamical system (2.2) supports a stable limit cycle. That is, there exists a stable periodic solution $\mathbf{x} = \mathbf{x}^*(t)$ with $\mathbf{x}^*(t) = \mathbf{x}^*(t + \Delta_0)$, where $\omega_0 = 2\pi/\Delta_0$ is the natural frequency of the oscillator. The dynamics on the limit cycle can be described by a uniformly rotating phase such that

$$(2.5) \quad \frac{d\theta}{dt} = \omega_0,$$

and $x^*(t) = \Phi(\theta(t))$ with Φ a 2π -periodic function and $\theta(t) = \theta_0 + \omega_0 t$. The dynamical equations for Φ can thus be written as

$$(2.6) \quad \omega_0 \frac{d\Phi}{d\theta} = F(\Phi(\theta)).$$

Differentiating both sides with respect to θ gives

$$(2.7) \quad \frac{d}{d\theta} \left(\frac{d\Phi}{d\theta} \right) = \omega_0^{-1} J(\theta) \cdot \frac{d\Phi}{d\theta},$$

where J is the 2π -periodic Jacobian matrix

$$(2.8) \quad J_{jk}(\theta) \equiv \left. \frac{\partial F_j}{\partial x_k} \right|_{x=\Phi(\theta)}.$$

Since the dynamical system is autonomous, it follows that it is neutrally stable with respect to perturbations along the limit cycle.

Our main concern in this paper is to characterize how intrinsic noise associated with fluctuations in the number of each chemical species for finite Ω affects the phase of the oscillator. This will require understanding how the deterministic system responds to small perturbations. Therefore, consider the perturbed deterministic dynamical system

$$(2.9) \quad \frac{dx_i}{dt} := F_i(\mathbf{x}) + \sqrt{\epsilon} G_i(\mathbf{x}), \quad i = 1, \dots, K,$$

where $0 < \epsilon \ll 1$ and $G_i(\mathbf{x})$ is bounded. A classical method for deriving the corresponding phase equation is based on the construction of isochrons [48, 30, 19, 13, 5]. Suppose that the unperturbed system (2.2) is observed stroboscopically at time intervals of length Δ_0 . This can be represented by the Poincaré mapping

$$\mathbf{x}(t) \rightarrow \mathbf{x}(t + \Delta_0) \equiv \mathcal{P}(\mathbf{x}(t)),$$

which has all points on the limit cycle as fixed points. The isochron through a point \mathbf{x}^* on the limit cycle is the $(K - 1)$ -dimensional hypersurface \mathcal{I} , consisting of all points in the vicinity of \mathbf{x}^* that are attracted to \mathbf{x}^* under the action of \mathcal{P} . Thus each isochron can be uniquely labeled by the phase of its point of intersection with the limit cycle, that is, $\mathcal{I} = \mathcal{I}(\theta)$ for $\mathbf{x}^* = \Phi(\theta)$. This then allows us to extend the definition of phase to a neighborhood of the limit cycle by taking all points $\mathbf{x} \in \mathcal{I}(\theta)$ to have the same phase, $\Theta(\mathbf{x}) = \theta$, which then rotates at the natural frequency ω_0 (in the unperturbed case). Hence, for an unperturbed oscillator in the vicinity of the limit cycle we have

$$\omega_0 = \frac{d\Theta}{dt} = \sum_{j=1}^K \frac{\partial \Theta}{\partial x_j} \frac{dx_j}{dt} = \sum_{j=1}^K \frac{\partial \Theta}{\partial x_j} F_j(\mathbf{x}).$$

Extending the definition of isochrons to the perturbed system (2.9), we have to leading order

$$\frac{d\Theta}{dt} = \sum_{j=1}^K \frac{\partial \Theta}{\partial x_j} (F_j(\mathbf{x}) + \sqrt{\epsilon} G_j(\mathbf{x}, t)) = \omega_0 + \sqrt{\epsilon} \sum_{j=1}^K \frac{\partial \Theta}{\partial x_j} G_j(\mathbf{x}, t).$$

As a further leading order approximation, deviations of \mathbf{x} from the limit cycle are ignored on the right-hand side. Hence, setting $\mathbf{x}(t) = \Phi(\theta(t))$ with Φ the 2π -periodic solution on the limit cycle,

$$\frac{d\Theta}{dt} = \omega_0 + \sqrt{\epsilon} \sum_{j=1}^K \left. \frac{\partial \Theta}{\partial x_k} \right|_{\mathbf{x}=\Phi} G_k(\Phi, t).$$

We can now obtain a closed equation for the phase by noting that $\Theta(\Phi(\theta)) = \theta$:

$$(2.10) \quad \frac{d\theta}{dt} = \omega_0 + \sqrt{\epsilon} \sum_{j=1}^K R_j(\theta) G_j(\Phi(\theta), t),$$

where

$$(2.11) \quad R_k(\theta) = \left. \frac{\partial \Theta}{\partial x_k} \right|_{\mathbf{x}=\Phi(\theta)}$$

is a 2π -periodic function of θ known as the k th component of the *phase response curve* (PRC). One way to calculate the PRC $R(\theta)$ is to note that it is a 2π -periodic solution of the linear equation [16, 17, 36]

$$(2.12) \quad \omega_0 \frac{dR(\theta)}{d\theta} = -J(\theta)^\top \cdot R(\theta),$$

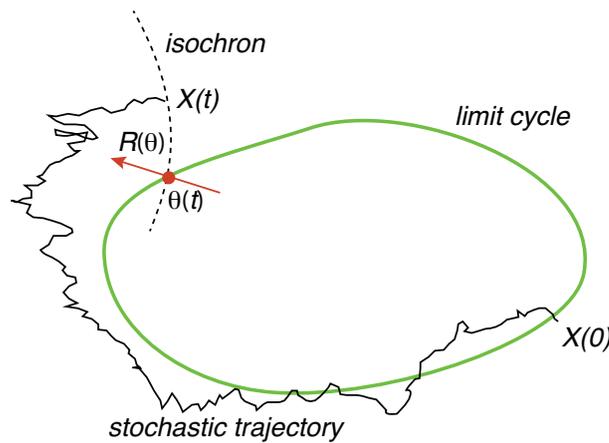


Figure 1. Stochastic trajectory of chemical reaction network for finite Ω , assuming that the corresponding mass action kinetics supports a stable limit cycle. The method of isochrons determines the phase $\theta(t)$ by tracing where the isochron through $X(t)$ intersects the limit cycle. The response 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.

with the normalization condition

$$(2.13) \quad R(\theta) \cdot \frac{d\Phi(\theta)}{d\theta} = 1.$$

Here $J(\theta)^\top$ is the transpose of the Jacobian matrix $J(\theta)$ given by (2.8).

2.3. System-size expansion and phase reduction. Now suppose that we include the effects of intrinsic noise by taking Ω to be large but finite, with the corresponding deterministic system operating in an oscillatory regime. Intuitively, one might expect that simulating the chemical reaction network using Gillespie’s stochastic simulation algorithm [20, 21, 22], for example, will generate a stochastic trajectory that remains in a neighborhood of the limit cycle (up to some stopping time), as illustrated schematically in Figure 1. While it remains within such a neighborhood, one can effectively treat the noise as a small perturbation of the deterministic limit cycle and apply the method of isochrons.

As we noted in the introduction, most studies of stochastic CRN oscillators [23, 6, 7, 35] are based on the system-size expansion of the chemical master equation (2.3) [18, 45, 14]. This involves setting $\lambda_a(\mathbf{n}/\Omega)P_{\mathbf{n}}(t) \rightarrow \lambda_a(\mathbf{x})p(\mathbf{x}, t)$ with $\mathbf{x} = \mathbf{n}/\Omega$ and treating \mathbf{x} as a continuous vector.¹ Then

$$\begin{aligned} \prod_{i=1}^K \mathbb{L}^{-S_{ia}} h(\mathbf{x}) &= h(\mathbf{x} - \mathbf{S}_a/\Omega) \\ &= h(\mathbf{x}) - \Omega^{-1} \sum_{i=1}^K S_{ia} \frac{\partial h}{\partial x_i} + \frac{1}{2\Omega^2} \sum_{i,j=1}^K S_{ia} S_{ja} \frac{\partial^2 h(\mathbf{x})}{\partial x_i \partial x_j} + O(\Omega^{-3}). \end{aligned}$$

¹Note that $N_j(t)$, $j = 1, \dots, K$, are discrete random variables so that $X_j(t) = N_j(t)/\Omega$ are also discrete. However, if Ω is sufficiently large, then the trajectory consists of small jumps so that it can be approximated by a continuous trajectory, assuming that the number of jumps in a small time interval is bounded appropriately.

Carrying out a Taylor expansion of the master equation to second order thus yields the multivariate Fokker–Planck (FP) equation

$$(2.14) \quad \frac{\partial p}{\partial t} = - \sum_{i=1}^K \frac{\partial F_i(\mathbf{x})p(\mathbf{x}, t)}{\partial x_i} + \frac{1}{2\Omega} \sum_{i,j=1}^K \frac{\partial^2 D_{ij}(\mathbf{x})p(\mathbf{x}, t)}{\partial x_i \partial x_j},$$

where

$$(2.15) \quad F_i(\mathbf{x}) = \sum_{a=1}^R S_{ia} \lambda_a(\mathbf{x}), \quad D_{ij}(\mathbf{x}) = \sum_{a=1}^R S_{ia} S_{ja} \lambda_a(\mathbf{x}).$$

The FP equation (2.14) corresponds to the multivariate Ito stochastic differential equation (SDE)

$$(2.16) \quad dX_i = F_i(\mathbf{X})dt + \frac{1}{\sqrt{\Omega}} \sum_{a=1}^R B_{ia}(\mathbf{X})dW_a(t),$$

where $W_a(t)$ are independent Wiener processes [18],

$$(2.17) \quad \langle dW_a(t) \rangle = 0, \quad \langle dW_a(t)dW_b(t') \rangle = \delta_{a,b}\delta(t-t')dt dt',$$

and $\mathbf{D} = \mathbf{B}\mathbf{B}^T$, that is,

$$(2.18) \quad B_{ia} = S_{ia} \sqrt{\lambda_a(\mathbf{x})}.$$

(Note that the SDE (2.16) can also be derived more rigorously using Kurtz's random time change representation [31, 32, 33, 3].)

In terms of the isochronal phase reduction method highlighted in section 2.2, suppose that we treat the sum over Wiener processes in (2.16) as a perturbation of the deterministic equation, by setting $\epsilon = 1/\Omega$ and

$$(2.19) \quad G_j(\mathbf{x})dt = \sum_{a=1}^R B_{ja}(\mathbf{X})dW_a(t).$$

This leads to a stochastic version of the phase equation (2.10). However, one needs to be careful in applying the phase reduction procedure to the Ito SDE. First, in order to apply the usual rules of calculus it is necessary to convert the Ito SDE to Stratonovich form using Ito's lemma. This yields an $O(1/\Omega)$ correction to the drift term [43, 17, 36]. Since we are interested in the leading order contributions to the phase equation, we can ignore such a term and simply take the phase to evolve according to the Ito SDE

$$(2.20) \quad d\theta = \omega_0 dt + \sqrt{\frac{1}{\Omega}} \sum_{j=1}^K R_j(\theta) \sum_{a=1}^R B_{ja}(\Phi(\theta(t)))dW_a(t).$$

Equation (2.20) follows from substituting (2.19) into the right-hand side of (2.10) and setting $\epsilon = \Omega^{-1}$. The second, and more significant issue, is that even for large Ω there is a small

but nonzero probability that a stochastic trajectory generated by the SDE (2.16) leaves any neighborhood of the limit cycle, in which case the notion of phase breaks down. In previous work, we have developed a variational principle for SDEs that provides strong exponential bounds on the escape probability [9]. (We have also extended the variational formulation to the case of piecewise-deterministic Markov processes [10].) One could directly apply our analysis to the SDE (2.16). However, the latter is already an approximation of the exact discrete Markov process. This motivates the analysis of this paper, namely, to derive an explicit phase equation to $O(1/\Omega)$ and to obtain bounds on the probability of escape by working directly with the discrete Markov process, rather than its system-size expansion. The mathematical framework underlying our approach is the random time change representation of chemical reaction networks, which has been developed by Kurtz [31, 32, 33, 3]. As we explained in the introduction, a fundamental aim in systems biology is to understand the robustness of the oscillation: that is, for a given system size, how long should one typically expect the oscillation to persist? The system-size expansion is limited because it only aims to reproduce the statistics of the oscillation over timescales of $O(\Omega)$. However, as we show in the following sections, one should expect the oscillation to persist for exponentially long periods of time, a metastable timescale over which the system-size expansion breaks down. Moreover, our variational definition of the phase will be accurate for the entire time that the system is in a neighborhood of the limit cycle.

3. Random time change representation and phase reduction. Let the random variable $N_i(t)$ denote the number of molecules in species i . An explicit representation of $N_i(t)$ is given by

$$(3.1) \quad N_i(t) = N_i(0) + \sum_{a=1}^M S_{ia} \mathcal{N}_a(t),$$

where $\mathcal{N}_a(t)$ denotes the number of times reaction a has occurred by time t and is a jump process for which the jump rate is locally given by $\Omega \lambda_a(\mathbf{N}(t)/\Omega)$. It can be proved that if $Y(t)$ is a unit rate Poisson process, that is,

$$\mathbb{P}[Y(t) = m | Y(0) = 0] = \frac{t^m e^{-t}}{m!},$$

then a representation of $\mathcal{N}_a(t)$ is given by the *time change representation* [31, 3],

$$(3.2) \quad \mathcal{N}_a(t) = Y_a \left(\Omega \int_0^t \lambda_a(\mathbf{N}(s)/\Omega) ds \right).$$

It follows that the expected number of reactions will grow linearly with Ω . Moreover, using the law of large numbers for Poisson processes,

$$\lim_{\Omega \rightarrow \infty} \frac{Y_a(\Omega T)}{\Omega} = T,$$

one has

$$\lim_{\Omega \rightarrow \infty} \frac{\mathcal{N}_a(t)}{\Omega} = \int_0^t \lambda_a(\mathbf{x}(s)) ds$$

with $\mathbf{x}(s) = \lim_{\Omega \rightarrow \infty} \mathbf{N}(s)/\Omega$. These identities hold with probability 1 [31]. This fact can be used to derive the deterministic equations (2.2) in the thermodynamic limit [4].

As before, we define $X_i(t) = N_i(t)/\Omega$ to be the concentration of species i at time t , which has the deterministic dynamics (2.2) in the large Ω limit. Let $\mathbf{S}_a = \{S_{ia}\}_{i=1,\dots,S}$ be the vector representing the change in species number by reaction a , and $\dot{Y}_a(z) = Y_a(z) - z$ be the a^{th} Poisson process minus its mean. We then find that

$$(3.3) \quad \mathbf{X}(t) = \mathbf{X}(0) + \int_0^t \mathbf{F}(\mathbf{X}(s)) ds + \Omega^{-1} \sum_{a=1}^M \dot{Y}_a \left(\Omega \int_0^t \lambda_a(\mathbf{X}(s)) ds \right) \mathbf{S}_a.$$

Hence, we have represented the evolution of the concentration as the sum of a deterministic dynamics and a stochastic noise of zero mean. (Taking the stochastic term to have zero mean will be useful later, when we use the theory of martingales; see section 5.) It has been proved that in the large Ω limit, the dynamics converges to the ODE [31]

$$(3.4) \quad \mathbf{X}(t) = \mathbf{X}(0) + \int_0^t \mathbf{F}(\mathbf{X}(s)) ds.$$

We now sketch a heuristic method for deriving a phase equation from (3.3). In section 4 we will put our derivation on a more rigorous footing using a variational approach. One way to specify the Poisson process $\mathcal{N}_a(t)$ is in terms of the event or jump times t_k^a , $k = 1, 2, \dots$. That is,

$$(3.5) \quad \mathcal{N}_a(t) = \sum_{k \geq 1} H(t - t_k^a),$$

where $H(t)$ is the Heaviside function, and the interevent times $t_{k+1}^a - t_k^a$ are generated from an exponential distribution:

$$\mathbb{P}(t_{k+1}^a \in [\tau, \tau + \Delta\tau] \text{ and } t_k^a \leq \tau | \mathbf{X}(\tau)) \simeq \Delta\tau \lambda_a(\mathbf{X}(\tau)).$$

Formally speaking, $dH(t - t_k^a) = \delta(t - t_k^a) dt$, where $\delta(t)$ is the Dirac delta function, so that

$$\frac{d\mathcal{N}_a}{dt} = \sum_{k \geq 1} \delta(t - t_k^a).$$

Differentiating (3.3) using $\dot{Y}_a(z) = Y_a(z) - z$ and (3.2) thus gives

$$(3.6) \quad \frac{dX_i}{dt} = F_i(\mathbf{X}) + \Omega^{-1} \sum_{a=1}^M S_{ia} \left(\sum_k \delta(t - t_k^a) - \Omega \lambda_a(\mathbf{X}(t)) \right).$$

If we treat the sum on the right-hand side of the above equation as a perturbation of the deterministic equation, then we can formally use the isochronal phase reduction method to derive the following stochastic phase equation to $O(1/\Omega)$:

$$(3.7) \quad \frac{d\theta}{dt} = \omega_0 + \frac{1}{\Omega} \sum_{j=1}^K R_j(\theta(t)) \sum_{a=1}^M S_{ja} \left(\sum_k \delta(t - \bar{t}_k^a) - \Omega \lambda_a(\Phi(\theta(t))) \right).$$

Here \bar{t}_k^a are the sequence of jump times for the inhomogeneous Poisson process

$$(3.8) \quad \bar{Y}_a(t) = Y_a \left(\Omega \int_0^t \lambda_a(\Phi(\theta(s))) ds \right).$$

(In general, the times \bar{t}_k^a will differ from the exact times t_k^a , since in the former case X is evaluated on the limit cycle.) In terms of differentials (see [39] or section A.3 of [4] for a definition of the stochastic integral with respect to \bar{Y}),

$$(3.9) \quad d\theta(t) = \omega_0 t + \frac{1}{\Omega} \sum_{j=1}^K R_j(\theta(t)) \sum_{a=1}^M S_{ja} \{ d\bar{Y}_a(t) - \Omega \lambda_a(\Phi(\theta(t))) dt \}.$$

Finally, integrating (3.7) with respect to t , we obtain the stochastic phase equation

$$(3.10) \quad \theta(t) = \theta(0) + \omega_0 t + \frac{1}{\Omega} \sum_{j=1}^K \sum_{a=1}^M S_{ja} \int_0^t R_j(\theta(s)) (d\bar{Y}_a(s) - \Omega \lambda_a(\Phi(\theta(s)))) ds,$$

or, equivalently

$$(3.11) \quad \theta(t) = \theta(0) + \omega_0 t + \frac{1}{\Omega} \sum_{j=1}^K \sum_{a=1}^M S_{ja} \left[\sum_{k \geq 1} R_j(\theta(\bar{t}_a^k)) H(t - \bar{t}_a^k) - \Omega \int_0^t R_j(\theta(s)) \lambda_a(\Phi(\theta(s))) ds \right].$$

We expect the above phase equation to be accurate up to times $O(\Omega)$ due to the central limit theorem for jump Markov processes [31]. In the subsequent sections we will define a variational phase that is accurate over much longer timescales, and we will use this phase to bound the probability of the system leaving the basin of attraction of the limit cycle. This in turn requires bounds on the reaction rates λ_a in some neighborhood \mathcal{U} of the limit cycle, which could be identified with the basin of attraction. (Once the system leaves the neighborhood of the limit cycle, one cannot sensibly define a phase dynamics in the first place. The analysis in subsequent sections will demonstrate that the system stays in a neighborhood of the limit cycle for exponentially long times, with very high probability.) Therefore, we assume the following uniform bounds on the reaction rates λ_a ,

$$(3.12) \quad \sup_{1 \leq a \leq M} \sup_{\theta \in [0, 2\pi]} \sup_{z: \|z - \Phi(\theta)\|_\theta \leq \eta} |\lambda_a(z)| \leq \bar{\lambda}.$$

Here $\|\cdot\|_\theta$ is a weighted norm that we define in the next section and $\eta > 0$ is a constant denoting the boundary of the attracting manifold of the limit cycle. Let $\mathcal{M}_{u,t}^a$ be the number of times that reaction a occurs over a time interval $[u, t]$; an estimate of the error in (3.11) is $O(\mathcal{M}_{0,t}/\Omega^2)$, where $\mathcal{M}_{0,t} = \sum_{a=1}^M \mathcal{M}_{0,t}^a$. In fact, for the error bounds in the coming sections, we need to bound the number of reactions that occur over arbitrary intervals of a certain length, i.e., we need a bound for arbitrary $u \in [0, t]$. To this end, for a constant integer

$c \geq m\bar{\lambda}(t-u)\Omega$, where $m > 0$ is such that $2m = m \log(m) + 1$,

$$\begin{aligned} \mathbb{P}\left(\mathcal{M}_{u,t}^a \geq c\right) &= \mathbb{P}\left(Y_a\left(\Omega \int_0^t \lambda_a(\mathbf{X}(s))ds\right) - Y_a\left(\Omega \int_0^u \lambda_a(\mathbf{X}(s))ds\right) \geq c\right) \\ &\leq \mathbb{P}\left(Y_a\left(\Omega \int_0^u \lambda_a(\mathbf{X}(s))ds + \Omega\bar{\lambda}(t-u)\right) - Y_a\left(\Omega \int_0^u \lambda_a(\mathbf{X}(s))ds\right) \geq c\right) \\ &= \frac{1}{c!} \left\{ \Omega\bar{\lambda}(t-u) \right\}^c \exp\left(-\Omega\bar{\lambda}(t-u)\right), \end{aligned}$$

using the definition of the Poisson distribution. Using the discrete version of Chebyshev's exponential inequality, for some $s > 0$ to be specified more precisely below,

$$\begin{aligned} \mathbb{P}\left(\mathcal{M}_{u,t}^a \geq c\right) &\leq \sum_{a=0}^{\infty} \left\{ \frac{1}{a!} \left\{ \Omega\bar{\lambda}(t-u) \right\}^a \exp\left(-\Omega\bar{\lambda}(t-u) + sa - sc\right) \right\} \\ &= \exp\left(\Omega\bar{\lambda}(t-u)\{e^s - 1\} - sc\right) \\ &\leq \exp\left(\Omega\bar{\lambda}(t-u)\{e^s - 1 - sm\}\right), \end{aligned}$$

using the formula for the moment generating function of the Poisson distribution in the penultimate line, and using the fact that $c \geq m\Omega\bar{\lambda}(t-u)$ in the last line. We substitute $s = \log(m)$ to minimize the above bound and obtain that

$$\exp\left(\Omega\bar{\lambda}(t-u)\{e^s - 1 - sm\}\right) = \exp\left(\Omega\bar{\lambda}(t-u)\{m - 1 - m \log(m)\}\right).$$

Now since $2m = m \log(m) + 1$, it must be that $m - 1 - m \log(m) = -m$. We thus obtain that

$$\exp\left(\Omega\bar{\lambda}(t-u)\{e^s - 1\} - sc\right) \leq \exp\left(-m\Omega\bar{\lambda}(t-u)\right).$$

In conclusion, we have proved that for the above value of m ,

$$(3.13) \quad \mathbb{P}\left(\mathcal{M}_{u,t}^a \geq m\Omega\bar{\lambda}(t-u)\right) \leq \exp\left(-m\Omega\bar{\lambda}(t-u)\right).$$

Since

$$\mathbb{P}\left(\mathcal{M}_{u,t} \geq mM\Omega\bar{\lambda}(t-u)\right) \leq \sum_{a=1}^M \mathbb{P}\left(\mathcal{M}_{u,t}^a \geq m\Omega\bar{\lambda}(t-u)\right),$$

it must be that

$$(3.14) \quad \mathbb{P}\left(\mathcal{M}_{u,t} \geq mM\Omega\bar{\lambda}(t-u)\right) \leq M \exp\left(-m\Omega\bar{\lambda}(t-u)\right) \leq \exp\left(-m\Omega\bar{\lambda}(t-u)/2\right),$$

where the last inequality follows for sufficiently large $(t-u)\Omega$.

4. Amplitude-phase decomposition. One of the limitations of carrying out a phase reduction using the method of isochrons is that it does not take proper account of fluctuations transverse to the limit cycle. This is important when trying to estimate the time for a stochastic trajectory to leave a neighborhood of the limit cycle. Following previous studies of SDEs, suppose that we decompose the random time change solution (3.3) into longitudinal (phase) and transverse (amplitude) fluctuations of the limit cycle according to

$$(4.1) \quad \mathbf{X}(t) = \Phi(\theta(t)) + \frac{1}{\sqrt{\Omega}} \mathbf{v}(t),$$

where $X_i(t) = N_i(t)/\Omega$. (One could carry out an analogous decomposition of the solution to the SDE (2.16), obtained under the system-size expansion [23, 7, 28, 8, 9].) Here the scalar random variable $\theta(t)$ represents the undamped random phase shift along the limit cycle, and $\mathbf{v}(t)$ is a transversal perturbation (see Figure 1). As it stands, the decomposition (4.1) is nonunique unless an additional mathematical constraint is imposed. Recently, we introduced a variational principle for SDEs that uniquely specifies the amplitude-phase decomposition by minimizing the weighted norm $\|\mathbf{X}(t) - \Phi(\theta(t))\|_\rho$ for a specific choice of ρ , which has two significant features: (i) sufficiently close to the limit cycle, the phase corresponds to the isochronal phase; (ii) one can obtain rigorous exponential bounds on the probability of escape [9, 10]. We will apply the variational approach to the stochastic CRN model and show that both of these results carry over; (ii) will be established in section 5. In order to specify the weight ρ , we first need to introduce a little Floquet theory.

4.1. Floquet decomposition. Suppose that the ODE (2.2) supports a limit cycle $\mathbf{x} = \Phi(\omega_0 t)$. In order to determine the linear stability of the limit cycle, we set $\mathbf{x}(t) = \Phi(\omega_0 t) + \mathbf{y}(t)$ and Taylor expand equation (2.2) to first order in \mathbf{y} . This yields the nonautonomous linear ODE

$$(4.2) \quad \frac{d\mathbf{y}}{dt} = \mathbf{J}(\omega_0 t)\mathbf{y}(t),$$

where the Jacobian $\mathbf{J}(\omega_0 t)$ is a Δ_0 -periodic function of time t . Determining stability thus reduces to the problem of analyzing the asymptotic behavior of solutions to a time-periodic linear ODE, which involves the use of Floquet theory. (The solutions are not necessarily periodic.)

We begin by noting that there are d linearly independent solutions of (4.2). Suppose that the corresponding vector space is spanned by a particular set of K -dimensional basis vectors $\mathbf{v}_j(t)$, $j = 1, \dots, K$. These are used to construct a fundamental matrix $\Phi(t)$, whose columns are given by the solutions $\mathbf{v}_j(t)$ so that $d\Phi(t)/dt = \mathbf{M}(t)\Phi(t)$. Since $\mathbf{M}(t + \Delta_0) = \mathbf{M}(t)$, it follows that $\hat{\Phi}(t) = \Phi(t + \Delta_0)$ is also a fundamental matrix:

$$\frac{d\hat{\Phi}}{dt} = \mathbf{M}(t + \Delta_0)\hat{\Phi}(t) = \mathbf{M}(t)\hat{\Phi}(t).$$

From linear independence we can write

$$\Phi_{ij}(t + \Delta_0) = \sum_{k=1}^K \Phi_{ik}(t)E_{kj}$$

for some constants E_{kj} . The matrix \mathbf{E} is nonsingular, since $\det \Phi(t + \Delta_0) = \det \Phi(t) \det \mathbf{E}$ and thus $\det \mathbf{E} \neq 0$. Let (μ, \mathbf{u}) be an eigenvalue pair of \mathbf{E} : $(\mathbf{E} - \mu \mathbf{I})\mathbf{u} = 0$, and consider the solution $\chi(t) = \Phi(t)\mathbf{u}$. Then

$$\chi(t + \Delta_0) = \Phi(t + \Delta_0)\mathbf{u} = \Phi(t)\mathbf{E}\mathbf{u} = \Phi(t)\mu\mathbf{u} = \mu\chi(t).$$

The eigenvalues of \mathbf{E} are called the Floquet multipliers or characteristic numbers of (4.2) and are independent of the choice of $\Phi(t)$. In order to establish the latter, suppose that $\Phi(t)$ and $\Phi^*(t)$ are two fundamental matrices. There then exists a nonsingular constant matrix \mathbf{C} for which $\Phi^*(t) = \Phi(t)\mathbf{C}$. Hence,

$$\Phi^*(t + \Delta_0) = \Phi(t + \Delta_0)\mathbf{C} = \Phi(t)\mathbf{E}\mathbf{C} = \Phi^*(t)\mathbf{C}^{-1}\mathbf{E}\mathbf{C} = \Phi^*(t)\mathbf{E}^*.$$

As \mathbf{E} and \mathbf{E}^* are related by a similarity transformation, they both have the same eigenvalues. Typically, $\Phi(t)$ is chosen to be real so that the characteristic equation for the eigenvalues μ has real coefficients. Hence, if there exists a complex multiplier μ , then its complex conjugate $\bar{\mu}$ is also a multiplier (complex conjugate pair). It is also common to express each Floquet multiplier μ in terms of a Floquet exponent ρ according to $\mu = e^{\rho\Delta_0}$. Note that ρ is defined up to an additive multiple of $2\pi i/\Delta_0$ so it can be uniquely specified by requiring $-\pi < \text{Im}(\rho\Delta_0) \leq \pi$. In terms of the Floquet exponent, we have $\chi(t + \Delta_0) = e^{\rho\Delta_0}\chi(t)$ and thus

$$\chi(t + \Delta_0)e^{-\rho(t+\Delta_0)} = \chi(t)e^{-\rho t}.$$

Hence we can decompose the solution as $\chi(t) = \mathbf{p}(t)e^{\rho t}$, where $\mathbf{p}(t)$ is Δ_0 -periodic. In order to use this Floquet decomposition to study the asymptotic behavior of solutions, we will assume that \mathbf{E} has K distinct eigenvalues μ_i so that there exist d independent solutions $\chi_i(t) = \mathbf{p}_i(t)e^{\rho_i t}$ with $\mathbf{p}_i(t + \Delta_0) = \mathbf{p}_i(t)$. The general solution thus takes the form

$$(4.3) \quad \mathbf{y}(t) = \sum_{i=1}^K c_i \mathbf{p}_i(t) e^{\rho_i t}$$

for constant coefficients c_i . Hence, the asymptotic behavior of the perturbation $\mathbf{y}(t)$ will depend on the sign of $\text{Re}(\rho_i)$ for $i = 1, \dots, K$.

It is important to note that there always exists at least one solution with a unit multiplier $\mu = 1$ ($\rho = 0$), namely, $\chi(t) = \dot{\mathbf{x}}(t)|_{\mathbf{x}(t)=\Phi(\omega_0 t)}$. This follows from differentiating equation (2.2) with respect to t and reflects the fact that the limit cycle is marginally stable with respect to perturbations along the limit cycle (phase shift invariance). If we identify $\mathbf{p}_1(t) = \chi(t)$, then we can express the unique fundamental matrix for which $\Phi(0) = 1$ according to

$$(4.4) \quad \Pi(t) = P(\omega_0 t) \exp(t\mathcal{S})P^{-1}(0),$$

where $\mathcal{S} = \text{diag}(\rho_1, \dots, \rho_K)$, and $P(\omega_0 t)$ is the 2π -periodic matrix whose columns are given by the vectors $\mathbf{p}_j(t)$, such that $P(\theta)^{-1}\Phi'(\theta) = \mathbf{e}$ with $\mathbf{e}_j = \delta_{1,j}$. In order to simplify the following notation, we will assume throughout this paper that the Floquet multipliers are real and hence $P(\theta)$ is a real matrix. One could readily generalize these results to the case that \mathcal{S}

is complex. The limit cycle is taken to be asymptotically stable, meaning that for a constant $b > 0$, for all $2 \leq i \leq K$,

$$(4.5) \quad \nu_i \leq -b.$$

Since $F \in \mathcal{C}^2$, it must be that $P \in \mathcal{C}^2$. Furthermore $P^{-1}(\theta)$ exists for all θ , since $\Pi^{-1}(t)$ exists for all t .

The above Floquet decomposition motivates the following weighted inner product: For any $\theta \in \mathbb{R}$, denoting the standard Euclidean dot product on \mathbb{R}^d by $\langle \cdot, \cdot \rangle$,

$$\langle u, v \rangle_\theta = \langle P^{-1}(\theta)u, P^{-1}(\theta)v \rangle,$$

and $\|u\|_\theta = \sqrt{\langle u, u \rangle_\theta}$. This weighting is useful for two reasons: it leads to a leading order separation of the phase from the amplitude, and it facilitates the strong bounds derived below, because the weighted amplitude always decays, no matter what the phase. The former is a consequence of the fact that the matrix $P^{-1}(\theta)$ generates a coordinate transformation, in which the phase in a neighborhood of the limit cycle coincides with the asymptotic phase defined using isochrons (see also [8]). This is reflected by the following relationship between the tangent vector to the limit cycle, $\Phi'(\theta)$, and the PRC $R(\theta)$ of (2.11) [9]:

$$(4.6) \quad \mathfrak{B}(\theta)P^\top(\theta)R(\theta) = P^{-1}(\theta)\Phi'(\theta),$$

where $\mathfrak{B}(\theta) := \|P^{-1}(\theta)\Phi'(\theta)\|^2 = 1$ from the definition of the Floquet matrix P .

4.2. Variational principle. Let $\beta(t)$ be the phase at time t . Since $\mathbf{X}(t)$ is constant between jumps, $\beta(t)$ is also taken to be constant between jumps. We can now state the variational principle for the stochastic phase [9]: $\beta(t)$ for $t \in \mathcal{T}$ is determined by requiring $\beta(t) = b_t(\theta_t)$, where $b_t(\theta_t)$ is a local minimum of the variational problem

$$(4.7) \quad \inf_{b \in \mathcal{N}(b_t(\theta_t))} \|\mathbf{X}(t) - \Phi(b)\|_{\theta_t} = \|\mathbf{X}(t) - \Phi(b_t(\theta_t))\|_{\theta_t}, \quad t \in \mathcal{T},$$

where θ_t is a prescribed time-dependent weight, and $\mathcal{N}(b_t(\theta_t))$ denotes a sufficiently small neighborhood of $b_t(\theta_t)$. The variational problem is equivalent to performing an orthogonal projection of the solution on to the limit cycle with respect to the weighted Euclidean norm at some θ_t . Write the partial derivative of (4.7) as

$$(4.8) \quad \mathcal{G}_0(\mathbf{x}, b, \theta) := \frac{\partial}{\partial b} \|\mathbf{x} - \Phi(b)\|_\theta^2 = -2 \langle \mathbf{x} - \Phi(b), \Phi'(b) \rangle_\theta.$$

At the minimum,

$$(4.9) \quad \mathcal{G}_0(\mathbf{X}(t), \beta(t), \theta_t) = 0.$$

We then impose the constraint that the location of the weight must coincide with the location of the minimum, i.e., $\beta(t) = \theta_t$. It follows that $\beta(t)$ satisfies the implicit equation

$$(4.10) \quad \mathcal{G}(\mathbf{X}(t), \beta(t)) := \mathcal{G}_0(\mathbf{X}(t), \beta(t), \beta(t)) = 0.$$

Let τ_k denote the time of the k th reaction in a given realization of the stochastic CRN. (For the moment, we don't specify which reaction occurs.) Since $\mathbf{X}(t)$ is constant for $t \in [\tau_k, \tau_{k+1})$, $\beta(t)$ is also constant over this interval. It is clear that there could be more than one minimum of (4.7), and therefore this equation may be insufficient to determine $\beta(\tau_{k+1})$. We therefore further stipulate that $\beta(\tau_{k+1})$ is the local minimum of (4.8) that also minimizes $|\beta(\tau_{k+1}) - \beta(\tau_{k+1}^-)|$. In other words, we emphasize continuity over finding the global minimum. This is consistent with our amplitude/phase decomposition, as well as the classical isochron phase decomposition of stochastic oscillators. There are several reasons for this choice. We wish to identify a linear phase equation, which requires the increments in $\beta(t)$ to be small. Furthermore, situations where a global minimum switches to being a local minimum, and the corresponding jump in phase is much greater than $O(\Omega^{-1})$, are relatively rare. Trying to understand this highly nonlinear phenomenon would take us beyond the scope of this paper.

4.3. Linear phase approximation. As it stands, the variational principle generates an exact, implicit equation for the stochastic phase. Here we carry out a perturbation expansion to show that the leading order dynamics of the phase over small time intervals recovers the explicit isochronal phase equation (3.11); we will refer to this as the *linear phase approximation*. The latter is a direct consequence of the fact that our variational principle defines the weighted inner product in terms of the Floquet matrix introduced in section 4.1. Let $\delta\mathbf{X}_k = \mathbf{X}(\tau_{k+1}) - \mathbf{X}(\tau_k)$. It is immediate from (3.1) that $|\delta\mathbf{X}_k| \leq \Omega^{-1} \sup_{1 \leq a \leq M} \|S_a\|$. We use (4.10) to estimate $\delta\beta_k = \beta(\tau_{k+1}) - \beta(\tau_k)$ to leading order. It must be that

$$(4.11) \quad \mathcal{G}_0(\mathbf{X}(\tau_{k+1}), \beta(\tau_{k+1}), \beta(\tau_{k+1})) - \mathcal{G}_0(\mathbf{X}(\tau_k), \beta(\tau_k), \beta(\tau_k)) = 0.$$

This means that, to leading order,

$$(4.12) \quad \frac{\partial}{\partial \mathbf{X}} \mathcal{G}_0(\mathbf{X}(\tau_k), \beta(\tau_k), \beta(\tau_k)) \delta\mathbf{X}_k + \frac{\partial}{\partial \beta} \mathcal{G}_0(\mathbf{X}(\tau_k), \beta(\tau_k), \beta(\tau_k)) \delta\beta_k = 0.$$

Rearranging, we find that

$$(4.13) \quad \delta\beta_k \simeq \frac{1}{\mathfrak{M}(\mathbf{X}(\tau_k), \beta(\tau_k))} \left\langle \Phi'(\beta(\tau_k)), \delta\mathbf{X}_k \right\rangle_{\beta(\tau_k)},$$

where

$$(4.14) \quad \mathfrak{M}(\mathbf{X}, \beta) := \frac{1}{2} \frac{\partial \mathcal{G}(\mathbf{X}, \beta)}{\partial \beta} = - \left\langle \Phi''(\beta), \mathbf{X} - \Phi(\beta) \right\rangle_{\beta} + \left\langle \Phi'(\beta), \Phi'(\beta) \right\rangle_{\beta} - \left\langle \Phi'(\beta), \frac{d}{d\beta} (P(\beta)P(\beta)^T)^{-1} (\mathbf{X} - \Phi(\beta)) \right\rangle.$$

Recall from the definition of the Floquet matrix P below (4.4) that $P(\theta)^{-1}\Phi'(\theta) = \mathbf{e}$ and, hence,

$$\langle \Phi'(\theta), \Phi'(\theta) \rangle_{\theta=\beta(\tau_k)} = \|P^{-1}(\theta)\Phi'(\theta)\|_{\theta=\beta(\tau_k)}^2 = 1.$$

Now as long as $\|\mathbf{X} - \Phi(\beta)\| \ll 1$ (we determine the probability of this in section 5), $\mathfrak{M}(\mathbf{X}, \beta) \geq \frac{1}{2}$ (say). We can then use Taylor's theorem to infer that, since $\delta\mathbf{X}_k = O(\Omega^{-1})$, the error in

(4.13) is $O(\Omega^{-2})$. Note that, since by assumption β is a local minimum, $\mathfrak{M}(\mathbf{X}, \beta)$ must be positive, because it is the curvature at the local minimum.

If we now take note of the reaction label a , it follows by analogy with (3.3) that

$$(4.15) \quad \beta(t) = \beta(0) + \frac{1}{\Omega} \sum_{a=1}^M \int_0^t \frac{1}{\mathfrak{M}(\mathbf{X}(s), \beta(s))} \left\langle \Phi'(\beta(s)), \mathbf{S}_a \right\rangle_{\beta(s)} dY_a(s) + O\left(\Omega^{-2} \mathcal{M}_{0,t}\right).$$

Here

$$Y_a(s) = Y_a\left(\Omega \int_0^s \lambda_a(\mathbf{X}(u)) du\right),$$

which reduces to (3.8) on taking $\mathbf{X}(u) = \Phi(\beta(u))$. We note that the above Poisson integral is, by definition, just the summation over the set of jumps (see section 3). That is, by definition,

$$\int_0^t \frac{1}{\mathfrak{M}(\mathbf{X}(s), \beta(s))} \left\langle \Phi'(\beta(s)), \mathbf{S}_a \right\rangle_{\beta(s)} d\bar{Y}^a(s) = \Omega^{-1} \sum_{k>0} \frac{H(t - t_k^a)}{\mathfrak{M}(\mathbf{X}(t_k^a), \beta(t_k^a))} \left\langle \Phi'(\beta(t_k^a)), \mathbf{S}_a \right\rangle_{\beta(t_k^a)},$$

where $\{t_i^a\}_{i \in \mathbb{Z}^+}$ are the jump times of $Y_a(\Omega \int_0^s \lambda_a(\mathbf{X}(u)) du)$.

We wish to represent the above equation as a time-integral plus a *martingale*. (A martingale is a stochastic process whose increments over small time-intervals are independent of the past history and of mean zero [39, 4]. For example, any stochastic integral with respect to a Wiener process is a martingale. The martingales in the following equations are often referred to as *compensated Poisson processes*.) In this way, the form of the equation will resemble that of a standard stochastic differential equation. Another important reason that we wish to put the equation in this form is that it will allow us to use a martingale inequality to obtain bounds on the system leaving a neighborhood of the limit cycle (see section 5). To this end, (4.15) can be rewritten as

$$(4.16) \quad \beta(t) = \beta(0) + \sum_{a=1}^M \int_0^t \frac{1}{\mathfrak{M}(\mathbf{X}(s), \beta(s))} \left\langle \Phi'(\beta(s)), \mathbf{S}_a \right\rangle_{\beta(s)} \lambda_a(\mathbf{X}(s)) ds + \Omega^{-1} \tilde{Y}(t) + O\left(\Omega^{-2} \mathcal{M}_{0,t}\right),$$

where $\tilde{Y}(t)$ is the martingale

$$(4.17) \quad \tilde{Y}(t) = \sum_{a=1}^M \int_0^t \frac{1}{\mathfrak{M}(\mathbf{X}(s), \beta(s))} \left\langle \Phi'(\beta(s)), \mathbf{S}_a \right\rangle_{\beta(s)} \{dY^a(s) - \Omega \lambda_a(\mathbf{X}(s)) ds\}.$$

(Note that $\tilde{Y}(t)$ is discontinuous, unlike the stochastic integral.)

Equation (4.16) has the same basic structure as the isochronal phase equation (3.11). Indeed, it reduces to (3.11) if we set $\mathbf{X}(s) = \Phi(\beta(s))$ on the right-hand side. This follows from the observations

$$\mathfrak{M}(\Phi(\beta), \beta) = \left\langle \Phi'(\beta), \Phi'(\beta) \right\rangle_{\beta} = \left\langle P^{-1}(\beta) \Phi'(\beta), P^{-1}(\beta) \Phi'(\beta) \right\rangle = 1$$

and

$$\langle \Phi'(\beta), \mathbf{S}_a \rangle_{\beta(s)} = \left\langle [P(\theta)P^\top(\theta)]^{-1} \Phi'(\theta), \mathbf{S}_a \right\rangle = \sum_{j=1}^K S_{ja} R_j(\theta),$$

using (4.6). In particular [9],

$$\sum_{a=1}^M \frac{1}{\mathfrak{M}(\mathbf{X}(s), \beta(s))} \left\langle \Phi'(\beta(s)), \mathbf{S}_a \right\rangle_{\beta(s)} \lambda_a(\mathbf{X}(s)) = \omega_0 + O(\|\mathbf{X}(s) - \Phi(\beta(s))\|^2).$$

We thus obtain the leading order phase equation

$$(4.18) \quad \beta(t) \simeq \beta(0) + \omega_0 t + \Omega^{-1} \tilde{Y}(t) + O\left(\mathcal{M}_{0,t} \Omega^{-2} + t \sup_{s \in [0,t]} \|\mathbf{X}(s) - \Phi(\beta(s))\|^2\right).$$

Furthermore, this linearized phase is typically accurate over timescales of $O(\Omega)$, as long as

$$t \sup_{s \in [0,t]} \|\mathbf{X}(s) - \Phi(\beta_s)\|^2$$

remains small. (We are going to bound this in the next section.) This is because, as we saw in section 3, the total number of reactions over the time interval $[0, t]$ is of order $\Omega \bar{\lambda} t$, with very high probability. For large Ω , the probability law of the martingale $\tilde{Y}(t)$ may be approximated by the stochastic integral

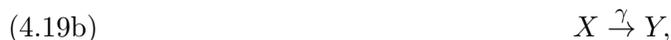
$$\tilde{Y}(t) \simeq \sqrt{\Omega} \sum_{a=1}^M \int_0^t \frac{1}{\mathfrak{M}(\mathbf{X}(s), \beta(s))} \left\langle \Phi'(\beta(s)), \mathbf{S}_a \sqrt{\lambda_a(\mathbf{X}(s))} \right\rangle_{\beta(s)} dW^a(s),$$

where $\{W^a(s)\}$ are independent Wiener processes [31, 32]. In this way, we see that the probability law of the phase equation (4.18) may be approximated to be the law of the SDE

$$\beta(t) \simeq \beta(0) + t\omega_0 + \frac{1}{\sqrt{\Omega}} \sum_{a=1}^M \int_0^t \frac{1}{\mathfrak{M}(\Phi(\beta(s)), \beta(s))} \left\langle \Phi'(\beta(s)), \mathbf{S}_a \sqrt{\lambda_a(\Phi(\beta(s)))} \right\rangle_{\beta(s)} dW^a(s).$$

Note that this agrees with the system-size expansion for the phase in (2.20), after using (4.6).

4.4. Example: The Brusselator. One benchmark model of a biochemical oscillator is the Brusselator, which is an idealized model of an autocatalytic reaction, in which at least one of the reactants is also a product of the reaction [15]. The model consists of two chemical species X and Y interacting through the following reaction scheme:



These reactions describe the production and degradation of an X molecule, an X molecule spontaneously transforming into a Y molecule, and two molecules of X reacting with a single molecule of Y to produce three molecules of X . The corresponding mass-action kinetic equations for $x_1 = [X], x_2 = [Y]$ are (after rescaling so that $c = d = 1$)

$$(4.20a) \quad \frac{dx_1}{dt} = \kappa - (\gamma + 1)x_1 + x_1^2x_2 = F_1(x_1, x_2),$$

$$(4.20b) \quad \frac{dx_2}{dt} = \gamma x_1 - x_1^2x_2 = F_2(x_1, x_2).$$

The system has a fixed point at $x_1^* = \kappa, x_2^* = \gamma/\kappa$, which is stable when $\gamma < \kappa^2 + 1$ and unstable when $\gamma > \kappa^2 + 1$. Moreover, the fixed point undergoes a Hopf bifurcation at the critical value $\gamma = \kappa^2 + 1$ for fixed κ , leading to the formation of a stable limit cycle.

It is straightforward to write down the propensities and stoichiometric coefficients, and thus construct a stochastic version of the Brusselator that keeps track of the number of molecules $n_1(t)$ and $n_2(t)$. There are two chemical species $A_j, j = 1, 2$, with $A_1 = X$ and $A_2 = Y$, and four single step-reactions $a = 1, \dots, 4$ corresponding to (4.19a)–(4.19d), respectively. The propensities and stoichiometric coefficients are as follows:

$$\begin{aligned} \lambda_1 &= \kappa, & S_{11} &= 1, & S_{21} &= 0, \\ \lambda_2 &= \gamma n_1/\Omega, & S_{12} &= -1, & S_{22} &= 1, \\ \lambda_3 &= n_1(n_1 - 1)n_2/\Omega^3, & S_{13} &= 1, & S_{23} &= -1, \\ \lambda_4 &= n_1/\Omega, & S_{14} &= -1, & S_{24} &= 0. \end{aligned}$$

Substituting into the master equation (2.3) for $K = 2, R = 4$, we have

$$(4.21) \quad \begin{aligned} \frac{dP_{n_1n_2}}{dt} &= \kappa\Omega P_{n_1-1,n_2} + \gamma(n_1 + 1)P_{n_1+1,n_2-1} + (n_1 - 1)(n_1 - 2)(n_2 + 1)P_{n_1-1,n_2+1}, \\ &(n_1 + 1)P_{n_1+1,n_2} - (\kappa\Omega + \gamma n_1 + n_1(n_1 - 1)n_2 + n_1)P_{n_1n_2}. \end{aligned}$$

Given the propensities and stoichiometric coefficients, we can construct both the SDE (2.16) under the system-size expansion and the equal time representation (3.3). For example, suppose that the system size Ω is sufficiently large so that we can carry out a system-size expansion of the master equation with $x_j = n_j/\Omega$. This yields the multivariate FP equation (2.14) and corresponding SDE (2.16), with $F_j(\mathbf{x})$ given by (4.20) and

$$(4.22) \quad \mathbf{D} = \begin{pmatrix} \kappa + (\gamma + 1)x_1 + x_1^2x_2 & -(\gamma x_1 + x_1^2x_2) \\ -(\gamma x_1 + x_1^2x_2) & \gamma x_1 + x_1^2x_2 \end{pmatrix}.$$

We simulated the stochastic Brusselator for $\kappa = c = d = 1, \gamma = 2.5$, and two different values of Ω , a large system size $\Omega = 3000$ and a small system-size $\Omega = 30$. The results are plotted in Figures 2 and 3. It can be seen that the linear phase evolving according to (2.20) is initially in close agreement with the exact variational phase evolving according to (4.10). However, after one orbit of the limit cycle, there is already a significant difference between the two. Despite this nonlinear diffusion of the phase, the system still stays in close proximity

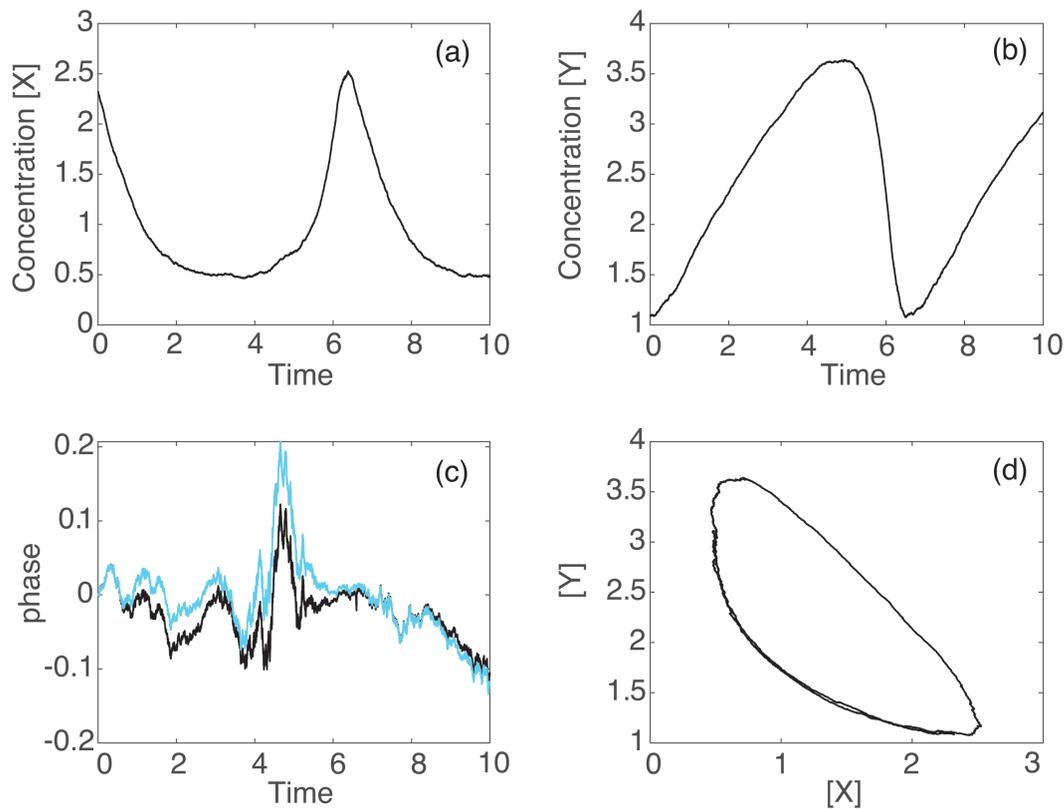


Figure 2. Numerical simulation of the stochastic Brusselator, with $\kappa = c = d = 1$, $\gamma = 2.5$, and the total number of particles $\Omega = 3000$. Initially, $[X]$ and $[Y]$ are taken to lie exactly on the limit cycle. (a) Plot of concentration $[X]$ as a function of time t . (b) Corresponding plot for $[Y]$. (c) Plot of the linear phase (light trajectory) and the exact variational phase (dark trajectory). (We have subtracted $\omega_0 t$ off both solutions.) (d) Phase portrait in $[X] - [Y]$ plane. It is notable that even with such a large Ω , the linear phase differs from the variational phase.

to the deterministic limit cycle for the entire length of the simulation. This can be explained using the analysis of section 5, where we demonstrate that with very high probability the system stays close to the limit cycle for times of $O(\exp(b\Omega C))$, where b is the rate of decay toward the limit cycle, and C is a constant.

5. Bounding the probability of the CRN leaving a neighborhood of the limit cycle.

The probability bounds of this section are accurate in the limit that $\Omega b \gg 1$. (Recall that b is the rate of decay toward the limit cycle in the deterministic dynamics.) Recalling the definition of the matrix P in (4.4), we will assume throughout this analysis that

$$(5.1) \quad \sup_{\theta \in [0, 2\pi]} \max \left\{ \|P(\theta)\|, \|P(\theta)^{-1}\| \right\} = O(1),$$

$$(5.2) \quad \sup_{\mathbf{Z} \in \mathbb{R}^{\mathbf{K}}, 1 \leq i, j, k \leq \mathbf{K}} \left| \frac{\partial^2}{\partial Z^j \partial Z^k} F_i(\mathbf{Z}) \right| = O(1),$$

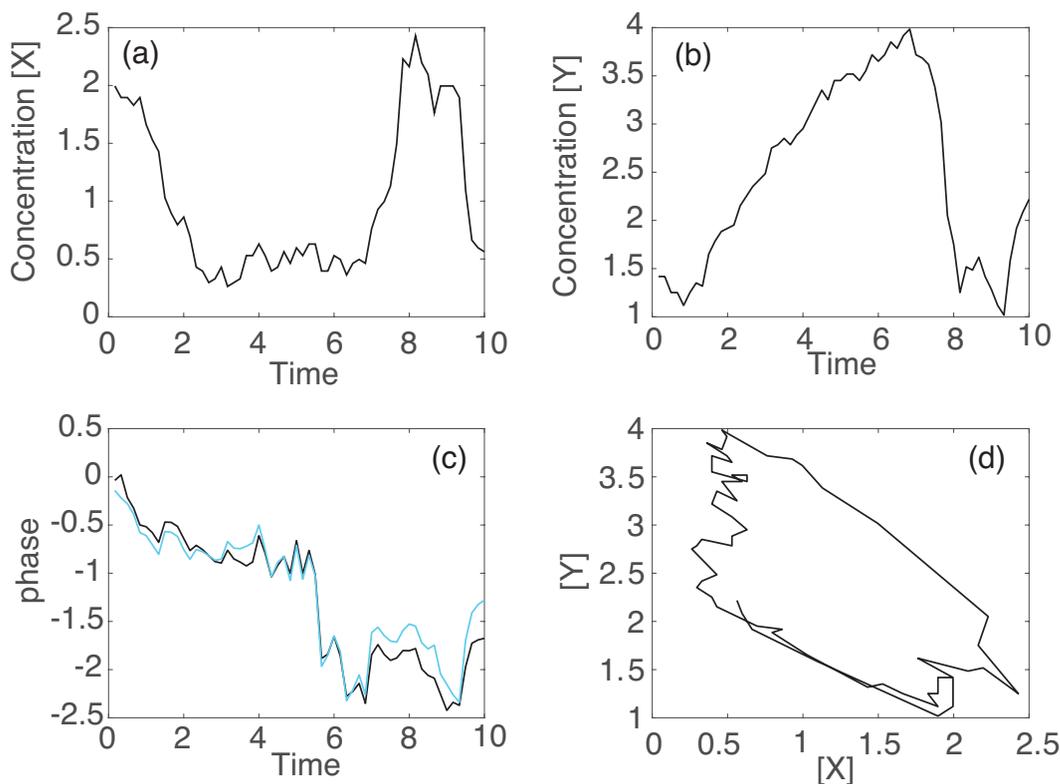


Figure 3. Numerical simulation of the stochastic Brusselator. The simulation is the same as Figure 2, except that we use $\Omega = 30$, instead of $\Omega = 3000$. The stochasticity is clearly much greater than when $\Omega = 3000$. Furthermore the difference between the linear phase and the variational phase is more substantial after noting the different scale for the phase.

where we are using the Euclidean norm for the matrices. It should also be noted that ω_0 can be arbitrarily large and the bounds would still be accurate (This is precisely what was observed in the numerical simulation of the chemical Brusselator—the diffusion of the phase was already substantial after one orbit of the limit cycle, but the system still stayed close to the limit cycle.) Define the amplitude \mathbf{v} and weighted amplitude \mathbf{w} according to

$$(5.3) \quad \mathbf{v}(t) = \mathbf{X}(t) - \Phi(\beta(t)), \quad \mathbf{w}(t) = P(\beta(t))^{-1} \mathbf{v}(t).$$

As is explained in more detail in [9], the reason that we make this change of variable is that, to leading order, the Euclidean vector norm $\|\mathbf{w}(t)\|$ decays uniformly in time. The main result that we prove in this section is that there exists a constant $C > 0$ such that

$$(5.4) \quad \mathbb{P} \left(\sup_{t \in [0, T]} \|\mathbf{w}(t)\| \geq \zeta \right) \leq bT \left\{ \exp(-\Omega b \zeta^2 C) \right\}$$

for all sufficiently small ζ . (We require in particular that $\zeta \leq b/4$.) We assume throughout this section that

$$(5.5) \quad \|\mathbf{w}(0)\| \leq \frac{\zeta}{8}.$$

As in section 4.3, suppose that τ_k is the time of the k th reaction. To leading order in $\delta\beta_k, \delta X_k$,

$$(5.6) \quad \mathbf{w}(\tau_{k+1}) - \mathbf{w}(\tau_k) \simeq \frac{\partial}{\partial \beta} \left\{ P(\beta(\tau_k))^{-1} \left(\mathbf{X}(\tau_k) - \Phi(\beta(\tau_k)) \right) \right\} \delta\beta_k + P(\beta(\tau_k))^{-1} \delta \mathbf{X}_k,$$

and the error in the above approximation is $O(\|\delta \mathbf{X}_k\|^2, \delta\beta_k^2) = O(\Omega^{-2})$.

Note that

$$(5.7) \quad \frac{\partial}{\partial \beta} \left\{ P(\beta)^{-1} \left(\mathbf{X}(t) - \Phi(\beta(t)) \right) \right\} = -\omega_0^{-1} P(\beta)^{-1} J(\beta) \left(\mathbf{X}(t) - \Phi(\beta) \right) + \omega_0^{-1} \mathcal{S} \mathbf{w}(t) - P(\beta)^{-1} \Phi'(\beta).$$

This follows from the identity

$$\omega_0 \frac{d}{d\theta} P(\theta) = J(\theta) P(\theta) - P(\theta) \mathcal{S},$$

which can be obtained by differentiating both sides of (4.4) with respect to t and using (4.2). Substituting (3.3) and (4.18), we thus find that for any $u \in [0, T]$ and $t > u$,

(5.8)

$$\mathbf{w}(t) = \mathbf{w}(u) + O\left(\Omega^{-2} \mathcal{M}_{u,t} + (t-u) \|\mathbf{v}(\mathbf{t})\|^2\right) + \Omega^{-1} [\check{Y}(t) - \check{Y}(u)]$$

(5.9)

$$+ \int_u^t \left\{ \mathcal{S} \mathbf{w}(s) + P(\beta(s))^{-1} \left(F(\mathbf{X}(s)) - F(\Phi(\beta(s))) - J(\beta(s)) \{ \mathbf{X}(s) - \Phi(\beta(s)) \} \right) \right\} ds,$$

where $\mathcal{M}_{u,t}$ is, as before, the total number of reactions over the time interval $[u, t]$. We have used the fact that $\omega_0 \Phi'(\beta) = F(\Phi(\beta))$ and the increments in $\{\mathbf{X}, \beta, \mathbf{w}\}$ are all $O(\Omega^{-1})$ (which we noted in the previous section). Here $\check{Y}(t)$ is the martingale

$$(5.10) \quad \check{Y}(t) = \sum_{a=1}^M \int_0^t P(\beta(s))^{-1} \mathbf{S}_a d\check{Y}_a(s)$$

$$(5.11) \quad + \int_0^t \left\{ -\omega_0^{-1} P(\beta(s))^{-1} J(\beta(s)) \left(\mathbf{X}(s) - \Phi(\beta(s)) \right) + \omega_0^{-1} \mathcal{S} \mathbf{w}(s) - P(\beta(s))^{-1} \Phi'(\beta(s)) \right\} d\check{Y}(s).$$

By Taylor's theorem,

$$(5.12) \quad F(\mathbf{X}(s)) - J(\beta(s)) \{ \mathbf{X}(s) - \Phi(\beta(s)) \} - F(\Phi(\beta(s))) = O\left(\|\mathbf{X}(s) - \Phi(\beta(s))\|^2\right).$$

Since $\|\mathbf{v}(t)\| = O(\|\mathbf{w}(t)\|)$, noting the assumption in (5.1), this means that

$$\mathbf{w}(t) = \mathbf{w}(u) + \int_u^t \mathcal{S}\mathbf{w}(s)ds + O\left(\Omega^{-2}\mathcal{M}_{u,t} + (t-u) \sup_{r \in [u,t]} \|\mathbf{w}(r)\|^2\right) + \Omega^{-1}[\check{Y}(t) - \check{Y}(u)].$$

Applying the change of variable $\mathbf{w}(t) \rightarrow \langle \mathbf{w}(t), \mathbf{w}(t) \rangle = \|\mathbf{w}(t)\|^2$, we find that

$$(5.13) \quad \|\mathbf{w}(t)\|^2 = \|\mathbf{w}(u)\|^2 + 2 \int_u^t \langle \mathbf{w}(s), \mathcal{S}\mathbf{w}(s) \rangle ds + 2\Omega^{-1} \int_u^t \langle \mathbf{w}(s), d\check{Y}(s) \rangle + O\left(\sup_{r \in [u,t]} \|\mathbf{w}(r)\| \left\{ \Omega^{-2}\mathcal{M}_{u,t} + (t-u) \sup_{r \in [u,t]} \|\mathbf{w}(r)\|^2 \right\}\right).$$

Performing the second change of variable $\|\mathbf{w}(t)\|^2 \rightarrow \|\mathbf{w}(t)\|$ then yields

$$(5.14) \quad \|\mathbf{w}(t)\| = \|\mathbf{w}(u)\| + \int_u^t \frac{1}{\|\mathbf{w}(s)\|} \langle \mathbf{w}(s), \mathcal{S}\mathbf{w}(s) \rangle ds + \frac{1}{\Omega} \int_u^t \frac{1}{\|\mathbf{w}(s)\|} \langle \mathbf{w}(s), d\check{Y}(s) \rangle + O\left((t-u) \sup_{r \in [u,t]} \|\mathbf{w}(r)\|^2 + \sup_{r \in [u,t]} \frac{\mathcal{M}_{u,t}}{\Omega^2 \|\mathbf{w}(r)\|}\right).$$

Furthermore our stability assumption on the limit cycle (see equation (4.5)) implies that

$$\langle \mathbf{w}(s), \mathcal{S}\mathbf{w}(s) \rangle \leq -b \|\mathbf{w}(s)\|^2.$$

Note that the first element of \mathbf{w} is identically zero, as a direct consequence of its definition, which is what leads to the strict decay in the above equation.

Hence, as long as $\|\mathbf{w}(s)\| \leq \frac{b}{4}$,

$$(5.15) \quad \|\mathbf{w}(t)\| \leq \|\mathbf{w}(u)\| - \frac{3b}{4} \int_u^t \|\mathbf{w}(s)\| ds + O\left(\sup_{r \in [u,t]} \frac{\mathcal{M}_{u,t}}{\Omega^2 \|\mathbf{w}(r)\|}\right) + \mathcal{Y}(t) - \mathcal{Y}(u),$$

where

$$(5.16) \quad \mathcal{Y}(t) = \frac{1}{\Omega} \int_0^t \frac{1}{\|\mathbf{w}(s)\|} \langle \mathbf{w}(s), d\check{Y}(s) \rangle.$$

Define the stopping time

$$(5.17) \quad \bar{\tau} = \inf \{t : \|\mathbf{w}(t)\| \geq \zeta\}$$

and the stopped process

$$\tilde{\mathcal{Y}}(t) = \mathcal{Y}(t \wedge \bar{\tau}).$$

We now outline a set of events that, taken together, ensure that $\sup_{t \leq T} \|\mathbf{w}(t)\| \leq \zeta$. This means that, to bound the probability that $\sup_{t \leq T} \|\mathbf{w}(t)\| > \zeta$, it suffices to bound each of the

following events. These events are

$$(5.18a) \quad \|\mathbf{w}(0)\| \leq \frac{\zeta}{8},$$

$$(5.18b) \quad \sup_{t \in [0, T]} \sup_{s \in [0, 8b^{-1}]} \left\{ \tilde{\mathcal{Y}}(t+s) - \tilde{\mathcal{Y}}(t) \right\} \leq \frac{\zeta}{2},$$

$$(5.18c) \quad \sup_{s \in [0, T]} \sup_{t \in [0, 8b^{-1}]} \Omega^{-2} \mathcal{M}_{s, s+t} \leq \frac{\zeta^2}{64}.$$

The timescale of $O(b^{-1})$ has been chosen for a reason: it is the timescale of the decay toward the limit cycle. To see why the above three events are sufficient to ensure that $\sup_{t \leq T} \|\mathbf{w}(t)\| \leq \zeta$, let us suppose for a contradiction that (5.18a)–(5.18c) all hold, but that $\|\mathbf{w}(t)\| = \zeta$ for some $t \in [0, T]$, and that $\|\mathbf{w}(s)\| < \zeta$ for all $s \in [0, t)$ (i.e., we are taking t to be the first time that $\|\mathbf{w}\|$ attains ζ). By definition, $\mathcal{Y}(s) = \tilde{\mathcal{Y}}(s)$ for all $s \in [0, t]$. Let $\tau = \sup \{u \leq t : \|\mathbf{w}(u)\| = \frac{\zeta}{8}\}$, noting that τ exists since, by assumption, $\|\mathbf{w}(0)\| \leq \zeta/8$.

Suppose for a contradiction that $t - \tau \geq 8b^{-1}$. Then, since $\|\mathbf{w}(s)\| \geq \zeta/8$ for all $s \in [\tau, t]$, it follows from (5.15) that

$$\|\mathbf{w}(t)\| \leq \|\mathbf{w}(t - b^{-1})\| - \frac{24b}{4b} \frac{\zeta}{8} + \frac{\zeta}{2} + \frac{\zeta}{8} \leq \|\mathbf{w}(t - b^{-1})\|.$$

This contradicts our assumption that $\|\mathbf{w}(s)\| < \zeta$ for all $s \in [0, t)$, since $\|\mathbf{w}(t)\| = \zeta$. We can thus conclude that $t - \tau < 8b^{-1}$. But in this case, it is immediate that

$$\|\mathbf{w}(t)\| \leq \|\mathbf{w}(\tau)\| + \frac{\zeta}{2} + \frac{\zeta}{8} \leq \frac{3\zeta}{4},$$

which contradicts our assumption that $\|\mathbf{w}(t)\| = \zeta$.

We can therefore conclude that the events in (5.18a)–(5.18c) imply that $\sup_{s \in [0, T]} \|\mathbf{w}(s)\| < \zeta$. Given that, by stipulation, $\|\mathbf{w}(0)\| \leq \zeta/8$, this means that

$$(5.19) \quad \mathbb{P} \left(\sup_{s \in [0, T]} \|\mathbf{w}(s)\| \geq \zeta \right) \\ \leq \mathbb{P} \left(\sup_{s \in [0, T]} \sup_{t \in [0, 8b^{-1}]} \Omega^{-2} \mathcal{M}_{s, s+t} > \frac{\zeta^2}{64} \right) + \mathbb{P} \left(\sup_{t \in [0, T]} \sup_{s \in [0, 8b^{-1}]} \left\{ \tilde{\mathcal{Y}}(t+s) - \tilde{\mathcal{Y}}(t) \right\} > \frac{\zeta}{2} \right).$$

We now bound the above two probabilities, showing that they are each $O(\exp(-b\Omega C))$ for some constant $C > 0$.

Let $t_i = 8i/b$. Using a union of events bound,

$$(5.20) \quad \mathbb{P} \left(\sup_{s \in [0, T]} \sup_{t \in [0, 8b^{-1}]} \Omega^{-2} \mathcal{M}_{s, s+t} > \frac{\zeta^2}{64} \right) \leq \sum_{i=0}^{\lfloor bT/8 \rfloor} \mathbb{P} \left(\sup_{t \in [t_i, t_{i+1}]} \Omega^{-2} \mathcal{M}_{t_i, t_i+t} > \frac{\zeta^2}{128} \right).$$

Furthermore, it follows from (3.14) that, since $t_{i+1} - t_i = \frac{8}{b}$,

$$\mathbb{P} \left(\sup_{t \in [t_i, t_{i+1}]} \Omega^{-2} \mathcal{M}_{t_i, t_i+t} > \frac{\zeta^2}{128} \right) \leq \exp(-\Omega b C)$$

for some constant $C > 0$, as long as $b\Omega$ is sufficiently large. We thus find that

$$\mathbb{P}\left(\sup_{s \in [0, T]} \sup_{t \in [0, 8b^{-1}]} \Omega^{-2} \mathcal{M}_{s, s+t} > \frac{\zeta^2}{64}\right) \leq Tb \exp(-Cb\Omega)$$

for large enough $b\Omega$, as required.

We turn to the second inequality in (5.19). We must show that for some constant C ,

$$(5.21) \quad \mathbb{P}\left(\sup_{t \in [0, T]} \sup_{s \in [0, 8b^{-1}]} \{\tilde{\mathcal{Y}}(t+s) - \tilde{\mathcal{Y}}(t)\} > \frac{\zeta}{2}\right) \leq Tb \exp(-Cb\Omega).$$

We first claim that

$$(5.22) \quad \mathbb{P}\left(\sup_{t \in [0, T]} \sup_{s \in [0, 8b^{-1}]} \{\tilde{\mathcal{Y}}(t+s) - \tilde{\mathcal{Y}}(t)\} > \frac{\zeta}{2}\right) \leq \sum_{i=0}^{\lfloor bT/8 \rfloor} \mathbb{P}\left(\sup_{s \in [t_i, t_{i+1}]} |\tilde{\mathcal{Y}}(s) - \tilde{\mathcal{Y}}(t_i)| > \frac{\zeta}{6}\right).$$

To see why this is true, suppose that $t+s \in [t_i, t_{i+1}]$ and $t \in [t_{i-1}, t_i]$. Then

$$|\tilde{\mathcal{Y}}(t+s) - \tilde{\mathcal{Y}}(t)| \leq |\tilde{\mathcal{Y}}(t+s) - \tilde{\mathcal{Y}}(t_i)| + |\tilde{\mathcal{Y}}(t_i) - \tilde{\mathcal{Y}}(t_{i-1})| + |\tilde{\mathcal{Y}}(t) - \tilde{\mathcal{Y}}(t_{i-1})|.$$

Hence if each of the terms on the right is less than $\zeta/6$, their sum must be less than $\zeta/2$. This establishes (5.22).

We now bound the terms on the right-hand side of (5.22), finding that

$$\begin{aligned} & \mathbb{P}\left(\sup_{s \in [t_i, t_{i+1}]} |\tilde{\mathcal{Y}}(s) - \tilde{\mathcal{Y}}(t_i)| > \frac{\zeta}{6}\right) \\ & \leq \mathbb{P}\left(\sup_{s \in [t_i, t_{i+1}]} (\tilde{\mathcal{Y}}(s) - \tilde{\mathcal{Y}}(t_i)) > \frac{\zeta}{6}\right) + \mathbb{P}\left(\sup_{s \in [t_i, t_{i+1}]} (\tilde{\mathcal{Y}}(t_i) - \tilde{\mathcal{Y}}(s)) > \frac{\zeta}{6}\right) \\ & \leq \mathbb{P}\left(\sup_{s \in [t_i, t_{i+1}]} \exp\left(\alpha\{\tilde{\mathcal{Y}}(s) - \tilde{\mathcal{Y}}(t_i)\}\right) > \exp\left\{\frac{\zeta\alpha}{6}\right\}\right) \\ & \quad + \mathbb{P}\left(\sup_{s \in [t_i, t_{i+1}]} \exp\left(-\alpha\{\tilde{\mathcal{Y}}(s) - \tilde{\mathcal{Y}}(t_i)\}\right) > \exp\left\{\frac{\zeta\alpha}{6}\right\}\right), \end{aligned}$$

since the exponential function is positive and increasing. Now for times greater than t_i $\tilde{\mathcal{Y}}$ is a martingale, due to the optional sampling theorem [39]. This means that both $\exp(\alpha\{\tilde{\mathcal{Y}}(t) - \tilde{\mathcal{Y}}(t_i)\})$ and $\exp(-\alpha\{\tilde{\mathcal{Y}}(t) - \tilde{\mathcal{Y}}(t_i)\})$ are submartingales for any positive constant α , since the exponential function is convex (and we have also used Jensen's inequality). We can thus use Doob's submartingale inequality [4, A.2.1] to conclude that

$$(5.23) \quad \mathbb{P}\left(\sup_{s \in [t_i, t_{i+1}]} \exp\left(\alpha\{\tilde{\mathcal{Y}}(s) - \tilde{\mathcal{Y}}(t_i)\}\right) > \exp\left(\frac{\zeta\alpha}{6}\right)\right) \leq \mathbb{E}\left[\exp\left(\alpha\left\{\tilde{\mathcal{Y}}(t_{i+1}) - \tilde{\mathcal{Y}}(t_i) - \frac{\zeta}{6}\right\}\right)\right]$$

and

$$(5.24) \quad \begin{aligned} & \mathbb{P} \left(\sup_{s \in [t_i, t_{i+1}]} \exp \left(-\alpha \{ \tilde{\mathcal{Y}}(s) - \tilde{\mathcal{Y}}(t_i) \} \right) > \exp \left(\frac{\zeta \alpha}{6} \right) \right) \\ & \leq \mathbb{E} \left[\exp \left(-\alpha \left\{ \tilde{\mathcal{Y}}(t_{i+1}) - \tilde{\mathcal{Y}}(t_i) + \frac{\zeta}{6} \right\} \right) \right]. \end{aligned}$$

We now bound the first of the above expectations. In reaching the definition of the martingale $\mathcal{Y}(t)$ in (5.16) (which also appears in (5.14)) we performed a sequence of changes of variable in (4.15), (4.17), (5.9), and (5.13). For each variable, we performed a simple Eulerian interpolation to calculate how it changes over the time interval $[t_i^a, t_{i+1}^a]$, these being successive jump times of the process \mathcal{N}_a . That is, the increment in the variable from time t_i^a to t_{i+1}^a was approximated to be the derivative of the variable (evaluated at time t_i^a) multiplied by the magnitude of the jump. Indeed in principle we could have performed the change-of-variable $\{Y_a(t)\}_{a=1}^M \rightarrow \|w_i\|$ in one go, and the functions g_a in the equation below could be obtained, through the chain-rule, to just be the multiplication of the derivatives of each of the intermediate variables. In other words, the martingale \mathcal{Y}_t is of the form

$$(5.25) \quad \mathcal{Y}_t - \mathcal{Y}_s = \sum_{a=1}^M \left\{ \Omega^{-1} \int_s^t g_a(r, \mathbf{X}(r)) d\mathcal{N}_a(r) - \int_s^t g_a(r, \mathbf{X}(r)) \lambda_a(r) dr \right\}$$

for functions $\{g_a\}_{a=1}^M$ that are uniformly bounded in the tube $\|\mathbf{X}(t) - \Phi(t)\|_{\beta(t)} \leq \zeta$. Let this uniform bound be \bar{g} . Each time we performed this Eulerian interpolation, we also subtracted a compensating time integral to ensure that the resulting variable remained a martingale. This is why the term g_a appears twice: once in the (discrete) stochastic integral and once in the standard Lebesgue integral.

Now the increments in \mathcal{N}_a are Poisson-distributed [4], which means that when we take the derivative of the following expectation, it suffices to use a second order Taylor expansion. Including higher order terms in the Taylor Expansion would change the following expression by a factor of Ω^{-2} , which is negligible for large Ω . Using (5.25),

$$(5.26) \quad \begin{aligned} & \frac{d}{dt} \mathbb{E} \left[\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s) \right] \\ & \simeq \lim_{\delta t \rightarrow 0} (\delta t)^{-1} \mathbb{E} \left[\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s) \left\{ \alpha (\tilde{\mathcal{Y}}(t + \delta t) - \tilde{\mathcal{Y}}(t)) + \frac{\alpha^2}{2} (\tilde{\mathcal{Y}}(t + \delta t) - \tilde{\mathcal{Y}}(t))^2 \right\} \right]. \end{aligned}$$

Now

$$(5.27) \quad \mathbb{E} \left[\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s) \{ \tilde{\mathcal{Y}}(t + \delta t) - \tilde{\mathcal{Y}}(t) \} \right] = 0$$

thanks to the martingale nature of $\tilde{\mathcal{Y}}_t$ (it is a compensated Poisson process [4]). For the other

term, substituting the decomposition in (5.25),

$$\begin{aligned} & \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \mathbb{E} \left[\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s) (\tilde{\mathcal{Y}}(t + \delta t) - \tilde{\mathcal{Y}}(t))^2 \right] \\ &= \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \mathbb{E} \left[\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s) \left(\left\{ \Omega^{-1} \sum_{a=1}^M \int_t^{t+\delta t} g_a(r, \mathbf{X}(r)) d\mathcal{N}_a(r) \right\}^2 \right. \right. \\ & \quad \left. \left. - 2 \left\{ \Omega^{-1} \sum_{a=1}^M \int_t^{t+\delta t} g_a(r, \mathbf{X}(r)) d\mathcal{N}_a(r) \right\} \left\{ \sum_{a=1}^M \int_t^{t+\delta t} g_a(r, \mathbf{X}(r)) \lambda_a(r) dr \right\} \right. \right. \\ & \quad \left. \left. + \left\{ \sum_{a=1}^M \int_t^{t+\delta t} g_a(r, \mathbf{X}(r)) \lambda_a(r) dr \right\}^2 \right) \right]. \end{aligned}$$

Now the expectation of the third term in the above expansion is

$$\lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \mathbb{E} \left[\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s) \left\{ \int_t^{t+\delta t} g_a(r, \mathbf{X}(r)) \lambda_a(r) dr \right\}^2 \right] = \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \mathbb{E} [O(\delta t^2)] = 0.$$

Similarly, the second term is

$$\begin{aligned} & \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \mathbb{E} \left[\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s) \left\{ \Omega^{-1} \sum_{a=1}^M \int_t^{t+\delta t} g_a(r, \mathbf{X}(r)) d\mathcal{N}_a(r) \right\} \left\{ \int_t^{t+\delta t} g_a(r, \mathbf{X}(r)) \lambda_a(r) dr \right\} \right] \\ &= \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \mathbb{E} \left[O(\text{Number of jumps in interval } [t, t + \delta t]) \times O(\delta t) \right] \\ &= \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \mathbb{E} [O(\Omega \delta t^2)] = 0. \end{aligned}$$

The first term can be bounded using the properties of the Poisson process. For $\delta t \ll M^{-1}\Omega^{-1}$, the expectation is dominated by the probability of precisely one of the processes $\mathcal{N}_a(r)$ jumping precisely once over the interval $[t, t + \delta t]$. This probability is upperbounded by $\delta t \Omega M \bar{\lambda}$, since the intensity of each of the processes $\{\mathcal{N}_a\}_{a=1}^M$ is upperbounded by $\Omega \bar{\lambda}$. (Note the upperbound for the intensities in (3.12).) Indeed the probability of two jumps or more over this interval is of order $(\delta t \Omega M)^2$, which is asymptotically negligible in the limit $\delta t \rightarrow 0$. We thus find that, to leading order in Ω^{-1} ,

$$\begin{aligned} & \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} \mathbb{E} \left[\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s) \left\{ \Omega^{-1} \sum_{a=1}^M \int_t^{t+\delta t} g_a(r, \mathbf{X}(r)) d\mathcal{N}_a(r) \right\}^2 \right] \\ & \leq \lim_{\delta t \rightarrow 0} \frac{1}{\delta t} M \Omega^{-2} \bar{g}^2 \bar{\lambda} \Omega \delta t \mathbb{E} [\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s)] \\ & = M \Omega^{-1} \bar{g}^2 \bar{\lambda} \mathbb{E} [\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s)]. \end{aligned}$$

Coming back to (5.26), we have proved that

$$\frac{d}{dt} \mathbb{E} [\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s)] \leq \frac{M \alpha^2}{2\Omega} \bar{g}^2 \bar{\lambda} \mathbb{E} [\exp(\alpha \tilde{\mathcal{Y}}_t - \alpha \tilde{\mathcal{Y}}_s)].$$

Since the above expectation is equal to 1 when $t = s$, an application of Gronwall's inequality yields that

$$(5.28) \quad \mathbb{E}[\exp(\alpha\tilde{\mathcal{Y}}_t - \alpha\tilde{\mathcal{Y}}_s)] \leq \exp\left(\frac{(t-s)\alpha^2\bar{\lambda}\bar{g}^2M}{2\Omega}\right).$$

To optimize the bound in (5.23), and noting that $t - s = 8b^{-1}$, we thus choose

$$\alpha = \frac{6\Omega b\zeta}{8\bar{\lambda}\bar{g}^2M}.$$

We are using the fact that the quadratic $c_1\alpha^2 - c_2\alpha$ (for $c_1, c_2 > 0$) has a minimum of $-c_2^2/4c_1$ when $\alpha = c_2/2c_1$. We thus obtain that

$$(5.29) \quad \mathbb{P}\left(\sup_{s \in [t_i, t_{i+1}]} \left\{ \tilde{\mathcal{Y}}(s) - \tilde{\mathcal{Y}}(t_i) \right\} > \frac{\zeta}{6}\right) \leq \exp(-C\Omega b\zeta^2)$$

for some positive constant C . The bound for (5.24) is analogous. In light of (5.22), we have thus established (5.21).

6. Discussion and conclusion. In this paper we outlined a variational definition of the phase of a stochastic chemical reaction oscillator. This definition is versatile and adaptable to a number of applications. First, we demonstrated that under a linear phase approximation, we recover the phase predicted by the system-size expansion of the underlying discrete Markov process. Second, one can straightforwardly calculate the phase at any particular time through directly performing the minimization in (4.7). Third, we used this definition of the phase to estimate the time that the system stays in close proximity to the limit cycle, demonstrating that this time is $O(\exp(Cb\Omega))$ or larger with probability $O(\exp(-Cb\Omega))$, where b is the rate of decay to the limit cycle, and Ω is the system size. In other words, if $b\Omega \gg 1$, then the system stays close to the limit cycle for extremely long times (much longer than the time over which the system-size expansion is accurate), with extremely high probability. This analysis also provides a means of assessing the substantial diffusive shifts of the phase that can occur over these very long periods of time. Indeed, the typical order of magnitude of the diffusive shift $\tilde{Y}(t)$ of the phase β (defined in (4.17)) is $O(\sqrt{\Omega^{-1}t})$. Hence over timescales of $O(\exp(Cb\Omega))$, the typical diffusive shift of the phase is $O(\Omega^{-1/2}\exp(Cb\Omega/2))$. In most circumstances, one would expect that this shift is much greater than ω_0 , the period of the oscillation. In other words, over this timescale one cannot predict the phase $\beta(t)$ with any accuracy at all: all phases are of almost equal likelihood.

On another note, the probability bound in section 5 of this paper differs from the analogue for a piecewise-deterministic Markov process in [10], because the latter does not depend on b to leading order. This is because, in many circumstances, the most likely means by which the PDMP system can leave a neighborhood of the limit cycle is to not switch over a significantly long period of time. Since the rate of switching does not depend on the rate of decay toward the limit cycle, the leading order probability bound also does not depend on the rate of decay toward the limit cycle. By contrast, in the CRN network of this paper, if the system does not react over a significant timescale, its distance from the limit cycle does not change.

In general, the methods of this paper are extremely helpful for assessing the robustness of oscillations in chemical reaction networks, which are ubiquitous in cell biology [37, 47, 11, 26]. In future work, we plan to extend these methods to chemical reaction networks that support more complicated emergent dynamics.

REFERENCES

- [1] A. AGAZZI, A. DEMBO, AND J. P. ECKMANN, *Large deviations theory for Markov jump models of chemical reaction networks*, Ann. Appl. Probab., 28 (2018).
- [2] D. F. ANDERSON, G. CRACIUN, AND T. G. KURTZ, *Product-form stationary distributions for deficiency zero chemical reaction networks*, Bull. Math. Biol., 72 (2010), pp. 1947–1970.
- [3] D. F. ANDERSON AND T. G. KURTZ, *Continuous time Markov chain models for chemical reaction networks*, Design Anal. Biomol. Circuits, 2011, pp. 3–42.
- [4] D. F. ANDERSON AND T. G. KURTZ, *Stochastic Analysis of Biochemical Systems*, Springer, New York, 2015.
- [5] P. ASHWIN, S. COOMBES, AND R. NICKS, *Mathematical frameworks for oscillatory network dynamics in neuroscience*, J. Math. Neurosci., 6 (2016).
- [6] R. P. BOLAND, T. GALLA, AND A. J. MCKANE, *How limit cycles and quasi-cycles are related in systems with intrinsic noise*, J. Stat. Mech. Theory Exp., 2008, pp. 1–27.
- [7] R. P. BOLAND, T. GALLA, AND A. J. MCKANE, *Limit cycles, complex floquet multipliers, and intrinsic noise*, Phys. Rev. E, 79 (2009), 051131.
- [8] M. BONNIN, *Amplitude and phase dynamics of noisy oscillators*, Int. J. Circuit Theory Appl., 45 (2017), pp. 636–659.
- [9] P. C. BRESSLOFF AND J. N. MACLAURIN, *A variational method for analyzing stochastic limit cycle oscillators*, SIAM J. Appl. Dyn. Syst., 17 (2018), pp. 2205–2233.
- [10] P. C. BRESSLOFF AND J. N. MACLAURIN, *A variational method for analyzing limit cycle oscillations in stochastic hybrid systems*, Chaos, 28 (2018), 063105.
- [11] P. C. BRESSLOFF, *Stochastic Processes in Cell Biology*, Springer, New York, 2014.
- [12] P. C. BRESSLOFF, *Stochastic switching in biology: From genotype to phenotype*, J. Phys. A, 50 (2017), 133001.
- [13] E. BROWN, J. MOEHLIS, AND P. HOLMES, *On the phase reduction and response dynamics of neural oscillator populations*, Neural Comput., 16 (2004), pp. 673–715.
- [14] J. ELF AND M. EHRENBERG, *Fast evaluation of fluctuations in biochemical networks with the linear noise approximation*, Genome Res., 13 (2003) pp. 2475–2484.
- [15] I. R. EPSTEIN AND J. A. POJMAN, *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos*, Oxford University Press, Oxford, 1998.
- [16] G. B. ERMENTROUT, *Type I membranes, phase resetting curves, and synchrony*, Neural Comput., 8 (1996), pp. 979–1001.
- [17] G. B. ERMENTROUT, *Noisy Oscillators*, in Stochastic Methods in Neuroscience, C. R. Laing and G. J. Lord, eds., Oxford University Press, Oxford, 2009.
- [18] C. W. GARDINER, *Handbook of Stochastic Methods*, 4th ed., Springer, New York, 2009.
- [19] L. GLASS AND M. C. MACKEY, *From Clocks to Chaos*, Princeton University Press, Princeton, NJ, 1988.
- [20] D. T. GILLESPIE, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem., 81 (1977), pp. 2340–2361.
- [21] D. T. GILLESPIE, *Approximate accelerated stochastic simulation of chemically reacting systems*, J. Chem. Phys., 115 (2001), pp. 1716–1733.
- [22] D. T. GILLESPIE, A. HELLANDER, AND L. R. PETZOLD, *Perspective: stochastic algorithms for chemical kinetics*, J. Chem. Phys., 138 (2013), 170901.
- [23] D. GONZE, J. HALLOY, AND P. GASPARD, *Biochemical clocks and molecular noise: Theoretical study of robustness factors*, J. Chem. Phys., 116 (2002), pp. 10997–11010.
- [24] P. HANGGI, H. GRABERT, P. TALKNER AND H. THOMAS, *Bistable systems: Master equation versus Fokker–Planck modeling*, Phys. Rev. A, 29 (1984), pp. 371–378.

- [25] M. KAERN, T. C. ELSTON, W. J. BLAKE, AND J. J. COLLINS, *Stochasticity in gene expression: From theories to phenotypes*, Nat. Rev. Genetics, 6 (2005), pp. 451–464.
- [26] S. KESKIN, G. DEVAKANMALAI, S. GNANAPACKIAM, S. KWON, H. VU, Q. HONG, Y. LEE, M. SOLTANI, A. SINGH, A. AY, AND E. OZBUDAK, *Noise in the vertebrate segmentation clock is boosted by time delays but tamed by notch signaling*, Cell Rep., 23 (2018), pp. 2175–2185.
- [27] H. KITANO, *Biological Robustness*, Nat. Rev. Genetics, 5 (2004), pp. 826–837.
- [28] H. KOEPL, M. HAFNER, A. GANGULY, AND A. MEHROTRA, *Deterministic characterization of phase noise in biomolecular oscillators.*, Phys. Biol., 8 (2011), 055008.
- [29] R. KUMAR AND L. POPOVIC, *Large deviations for multi-scale jump-diffusion processes*, Stochastic Process. Appl., 127 (2017), pp. 1297–1321.
- [30] Y. KURAMOTO, *Chemical Oscillations, Waves and Turbulence*, Springer, New York, 1984.
- [31] T. G. KURTZ, *Limit theorems for sequences of jump Markov processes approximating ordinary differential processes*, J. Appl. Probab., 8 (1971) pp. 344–356.
- [32] T. G. KURTZ, *Limit theorems and diffusion approximations for density dependent Markov chains*, Math. Program. Stud., 5 (1976), pp. 67–78.
- [33] T. G. KURTZ, *Representations of Markov processes as multiparameter changes*, Ann. Probab., 8 (1980), pp. 682–715.
- [34] P. K. MAINI, T. E. WOOLLEY, R. E. BAKER, E. A. GAFFNEY, AND S. LEE, *Turing’s model for biological pattern formation and the robustness problem*, Interface Focus, 2 (2012), pp. 487–496.
- [35] G. MINAS AND D. A. RAND, *Long-time analytic approximation of large stochastic oscillators: Simulation, analysis and inference*, PLoS Comput. Biol., 13 (2017), e1005676.
- [36] H. NAKAO, *Phase reduction approach to synchronization of nonlinear oscillators*, Contemp. Phys., 57 (2016), pp. 188–214.
- [37] B. NOVAK AND J. TYSON, *Design principles of biochemical oscillators*, Nat. Rev. Mol. Cell Biol., 9 (2008).
- [38] J. PAULSSON, *Models of stochastic gene expression*, Phys. Life Rev., 2 (2005) pp. 157–175.
- [39] P. PROTTER, *Stochastic Integration and Differential Equations*, 2nd ed., Springer, New York, 2005.
- [40] A. RAJ AND A. VAN OUDENAARDEN, *Nature, nurture, or chance: Stochastic gene expression and its consequences*, Cell, 135 (2008), pp. 216–226.
- [41] O. SUVAK AND A. DEMIR, *Phase computations and phase models for discrete molecular oscillators*, J. Bioinf. Syst. Biol., 6 (2012), pp. 1–28.
- [42] M. B. ELOWITZ, A. J. LEVINE, E. D. SIGGIA, AND P. S. SWAIN, *Stochastic gene expression in a single cell*, Science, 297 (2002), pp. 1183–1186.
- [43] 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 (2009), 194102.
- [44] L. S. TSIMIRING, *Noise in Biology*, Rep. Prog. Phys., 77 (2014), 026601.
- [45] N. G. VAN KAMPEN, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1992.
- [46] M. VELLELA AND H. QIAN, *Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: The Schlögl model revisited*, J. R. Soc. Interface, 6 (2009), pp. 925–940.
- [47] X.-J. WANG, *Neurophysiological and computational principles of cortical rhythms in cognition*, Phys. Rev., 90 (2010), pp. 1195–1268.
- [48] A. WINFREE, *The Geometry of Biological Time*, Springer, New York, 1980.