

Augmented Phase Reduction of (Not So) Weakly Perturbed Coupled Oscillators*

Dan Wilson[†]
Bard Ermentrout[‡]

Abstract. While phase reduction is a tremendously useful tool for understanding the dynamics of weakly perturbed limit cycle oscillators, its assumptions break down as perturbations become larger, limiting its practical utility in many applications. This fundamental limitation is often apparent when studying coupled populations of oscillators when the collective behavior approaches a limit cycle with transient behavior that decays slowly. Using the notion of isostables of periodic orbits, which define a coordinate system transverse to the limit cycle, we develop a strategy to augment the standard phase reduction in large populations of coupled oscillators to allow for the application of larger perturbations. The resulting augmented phase reduction yields a tremendous decrease in model complexity with a resulting dimensionality that does not depend on the number of oscillators in the population. Additionally, the augmented phase reduction replicates model behavior that standard phase reduction cannot. Finally, we propose and implement a direct method for the application of augmented phase reduction that would be applicable in systems of biological oscillators. Applying the augmented phase reduction to a model of coupled circadian oscillators yields novel insight about the possible mechanism behind the difference in reported jet-lag severity between eastward and westward travel.

Key words. phase reduction, isostable, isochron, jet-lag, circadian rhythms, coupled oscillators

AMS subject classifications. 70G60, 70K05, 92B25, 34E05

DOI. 10.1137/18M1170558

Contents

I Introduction	278
2 Background: Isochrons and Isostables of Coupled Theta Model Oscillators	280
2.1 Isochrons and Phase Reduction	280
2.2 Isostables and Augmented Phase Reduction	282
2.3 Augmented Phase Reduction for the Coupled Theta Model	285
3 Augmented Phase Reduction of a Large Population of Oscillators	286
3.1 A Minimal Reduction of a Large Population of Oscillators	287

*Received by the editors February 13, 2018; accepted for publication (in revised form) October 16, 2018; published electronically May 8, 2019.

<http://www.siam.org/journals/sirev/61-2/M117055.html>

Funding: This work was supported by National Science Foundation grants DMS-1602841 and DMS-1712922.

[†]Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996 (dwilso81@utk.edu).

[‡]Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15213 (bard@math.pitt.edu).

3.2	A Computationally Efficient Reduction Strategy	291
4	Example Reduction of a Coupled Population of Circadian Oscillators	294
5	Direct Method for Augmented Phase Reduction	298
5.1	A List of Steps to Infer the Functions and Parameters in the Augmented Phase Reduction	300
5.2	Example of the Direct Method Applied to a Circadian System	301
6	Discussion and Conclusions	305
Appendix A. Computation of the iPRC, iIRC, and All Required Functions in the Augmented Phase Reduction		306
Appendix B. General Structure of the Eigenvectors of the Linearization to the Poincaré Map for a Population of Identical Oscillators		309
Appendix C. An Approximation to the Augmented Phase Reduction in the Limit of Small Forcing		311
References		313

1. Introduction. The biological oscillations necessary for sustaining life are usually composed of a large number of periodic oscillators, each communicating on a microscale to give rise to an organism's macroscale behavior. For instance, periodically firing neurons in different brain regions give rise to the brain rhythms that are associated with perception, cognition, and sensation [38], [1], [8], [24]; oscillators in the suprachiasmatic nucleus (SCN) entrain to an external 24-hour light-dark cycle in order to maintain circadian time [70], [16]; and pancreatic islets must coordinate their periodic release of insulin in order to control blood sugar levels [54], [45]. Due to their often parallelizable nature, and with the advent of distributed computing hardware, it is now possible to simulate detailed numerical models of these biological systems. While such numerical simulations are useful for hypothesis testing and developing intuition about these systems, their sheer size and complexity can make it difficult to understand the underlying mechanisms driving a system's behavior.

One strategy that can be used to reduce the complexity and dimensionality of these large and complicated models is phase reduction, whereby a system of the form

$$(1) \quad \dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + \mathbf{P}(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^n,$$

where $\mathbf{x}$ is the state vector, $\mathbf{F}$ gives the unperturbed dynamics, and $\mathbf{P}$ is an external perturbation, can be reduced to a scalar system,

$$(2) \quad \dot{\theta} = \omega + \mathbf{Z}^T(\theta) \mathbf{P}(\mathbf{x}(\theta), t).$$

Here, $\theta \in [0, 2\pi)$ is the oscillator's phase, ω is the unperturbed natural frequency, and $\mathbf{Z}(\theta) \in \mathbb{R}^n$ is the infinitesimal phase response curve (iPRC) to external perturbations. Phase reduction, a well-established and particularly useful strategy for studying weakly perturbed oscillatory dynamical systems, has had many applications in the physical and biological sciences [70], [20], [36], [29], [15].

In principle, phase reduction can be applied to any system that admits a stable periodic orbit, including those comprised of large populations of coupled oscillators. Previous authors have studied the phase response properties when (1) is itself a population of phase oscillators [33], [31], [40], [35], where in many cases it can be shown that the iPRC of the population oscillation is a function of the iPRCs of the constituent oscillators. Furthermore, the population need not be finite, and in the limit as the number of oscillators in a population approaches infinity, the system can be well approximated by studying the dynamics of its probability distribution function [31], [40], from which iPRCs can be obtained. Practically, when studying a large population of oscillators, phase reduction works best when perturbations are small, or close to infinitesimal. Larger perturbations can drive the population away from its limit cycle over time and invalidate the phase reduction (2). Indeed, in large populations of oscillators, when the phase response curve is measured using the direct method results can vary depending on the distribution of oscillators during the application of a perturbation [68]. This phenomenon, along with noise in the system, can lead to significant variation in the measurement of iPRCs experimentally and in numerical simulations [4], [26], [27].

Other experimental evidence suggests that while phase reduction can be a useful tool to help gain theoretical insight about an oscillatory population of coupled oscillators, standard phase reduction (2) is usually not sufficient to fully characterize the behavior of these complicated systems in response to external perturbations. For instance, circadian misalignment due to jet-lag is known to last longer after eastward travel, as compared to westward travel [51], [62]. This pronounced asynchrony is difficult to explain using a phase reduced model since for small enough perturbations, reentrainment should occur on similar time scales regardless of the direction of the shift. Also, the phase change of the circadian pacemaker in response to light perturbations has been shown to scale nonlinearly with the applied light intensity in humans [71] and fruit flies [12], a feature that cannot be predicted from the standard phase reduced equations.

A number of authors have proposed to extend the applicability of phase reduction when larger perturbations are applied. For instance, [37] and [53] propose a phase reduction method that is valid for strongly perturbed oscillators with slowly varying external input. The authors of [55] consider a related problem of an oscillator subject to large amplitude, high-frequency perturbations. Others have taken an alternative approach of computing phase coordinates throughout the basin of attraction of the limit cycle [52], [11]; this strategy does not result in a reduction of dimensionality, but is valid for arbitrary perturbations and has been shown to be useful in control applications in low-dimensional settings [48]. Alternatively, others have explored the possibility of augmenting the phase coordinate with amplitude coordinates that represent perturbations in directions transverse to the limit cycle [5], [34], [69], [58]. Most of these methods are based on Floquet theory [30], which provides an elegant characterization of the structure of solutions of linear, time-varying periodic systems. Floquet theory stipulates that for small perturbations to a limit cycle, the decay in transverse directions can be described by a set of exponentially decaying, time-varying vector-valued functions. With this in mind, the authors of [23] and [9] postulated using a globally exponentially decaying coordinate system to represent the approach toward a limit cycle and developed methods to compute such a coordinate system both in the basin of attraction of the limit cycle and along specific trajectories for low-dimensional systems. Additionally, [69] suggested using a coordinate system based on the notion

of isostables (cf., [43]), which represent level sets of initial conditions that approach the periodic orbit together. Recently, it was shown that isostable coordinates could be used to improve the accuracy of standard phase reduction [66].

In this work, we will focus on the augmented phase reduction developed in [66], which provides a useful balance between accuracy and computational tractability. Generally, periodic systems composed of a large number of coupled periodic oscillators have a large number of nonnegligible isostable coordinates, limiting the utility of this augmented phase reduction. However, as we will show, when the individual oscillators are identical and coupled through a mean-field, only a small subset of these isostable coordinates are necessary. With this augmented reduction, we find that many behaviors can be replicated when standard phase reduction alone is insufficient.

This paper is organized as follows: In section 2 we give necessary background information on isostables and isochrons and illustrate shortcomings of standard phase reduction when applied to coupled oscillator models; this section represents a review of the theory presented in both [69] and [66] with new examples chosen to provide intuition about the utility of augmented phase reduction. The remainder of the paper presents new results which use augmented phase reduction as a starting point. Section 3 gives a strategy for augmenting the standard phase reduction of a large population of oscillators by a small number of isostable coordinates to understand behavior as perturbations become larger. We illustrate this strategy for a model of coupled circadian oscillators in section 4 and find that the augmented reduction gives a close approximation of the full system while adding only a small number of isostable coordinates. For the same model, standard phase reduction does not provide a good approximation of the system dynamics. Section 5 investigates a strategy for calculating an augmented phase reduction in experimental systems, akin to the direct method [29], [50] for measuring phase response curves, and section 6 gives concluding remarks.

2. Background: Isochrons and Isostables of Coupled Theta Model Oscillators. In order to illustrate many of the concepts that will be important in the remainder of this work, consider the following model comprised of two coupled theta oscillators [13]:

$$(3) \quad \begin{aligned} \dot{T}_1 &= 1 - \cos(T_1) + (1 + \cos(T_1))I_1(t) + p_1(t), \\ \dot{T}_2 &= 1 - \cos(T_2) + (1 + \cos(T_2))I_2(t) + p_2(t). \end{aligned}$$

Here, $T_1, T_2 \in [0, 2\pi)$ so that when either T_1 or T_2 reaches 2π , it resets to zero, $I_1(t) = I_2(t) = 2 + \sin(T_1(t)) + \sin(T_2(t))$ represent identical feedback coupling to each oscillator, and $p_1(t)$ and $p_2(t)$ represent external perturbations. The coupled theta model is often used as a minimal model to represent periodically spiking neural behavior [46], [13]. In the absence of external perturbations, (3) admits a stable periodic orbit, which we will refer to as γ (shown in the top panel of Figure 1), for which both oscillators are synchronized, i.e., $T_1(t) = T_2(t)$.

2.1. Isochrons and Phase Reduction. In many cases, it is useful to analyze a system such as (3) in reference to its periodic orbit. To this end, one can define a scalar phase $\theta \in [0, 2\pi)$ such that along the limit cycle, $\theta = \omega t$, where $\omega = 2\pi/T$ is the natural frequency and T is the unperturbed natural period of oscillation. We will refer to the state of the system along the periodic orbit $\mathbf{x}^\gamma(\theta) \equiv [T_1(\theta) \ T_2(\theta)]^T$ and let $\theta = 0$ correspond to the moment at which $T_1, T_2 = 0$ on the periodic orbit. One can also extend the notion of phase to the basin of attraction of the limit cycle using the concept of isochrons [70], [20], which are defined as follows: let $\mathbf{a}(0)$ correspond

to an initial state of (3) in γ , then the isochron associated with $\mathbf{a}(0)$ is the set of all initial conditions $\mathbf{b}(0)$ for which

$$(4) \quad \lim_{t \rightarrow \infty} \|\mathbf{a}(t) - \mathbf{b}(t)\| = 0,$$

where $\|\cdot\|$ can be any norm. Using this definition, the isochron of an initial condition gives a sense of its asymptotic convergence to the periodic orbit. The middle-left panel of Figure 1 shows the isochrons calculated for the nonlinear two-dimensional system (3).

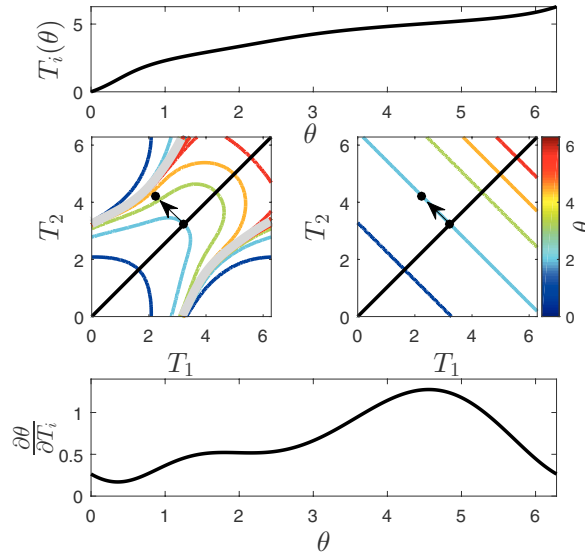


Fig. 1 The top panel shows the state of T_1 and T_2 as a function of θ for the system (3) on the stable periodic orbit. The middle-left and middle-right panels show the fully nonlinear isochrons and linearized approximation, respectively, for (3). The gray line in the left panel represents an unstable periodic orbit which is not in the basin of attraction of the limit cycle. Using the nonlinear isochrons, one can see that perturbations orthogonal to the periodic orbit (black arrow and dots) can significantly advance the phase, a characteristic that is not observed when using only a linear approximation to the isochrons. The bottom panel shows the phase response curve for perturbations to either T_1 or T_2 as a function of phase.

While the global isochrons of a nonlinear dynamical system can display complicated and even fractal patterns [52], it is often useful to reduce the dimensionality of a system by assuming that it stays near its stable periodic orbit. In this case, one can approximate (3) as a phase reduced equation of the form (2),

$$(5) \quad \dot{\theta} = \omega + \mathbf{Z}^T(\theta) \mathbf{P}(t),$$

where $\mathbf{P}(t) \equiv [p_1(t) \ p_2(t)]^T$ and $\mathbf{Z}(\theta) \equiv \frac{\partial \theta}{\partial \mathbf{x}}|_{\mathbf{x} \in \gamma(\theta)}$. In essence, a phase reduction is a linearization of the isochrons for a system close to its periodic orbit: the middle-right panel of Figure 1 shows the result of this linearization for the system (3), and the bottom panel shows the approximated phase response curve for perturbations to either T_1 or T_2 . Notice that for the linearized approximation, $\frac{\partial \theta}{\partial T_1} = \frac{\partial \theta}{\partial T_2}$, so that perturbations of the form $p_1(t) = -p_2(t)$ do not change the phase of the system. However, from the fully nonlinear isochron calculations, perturbations of this form can indeed alter the phase of the system.

The coupled theta oscillator model (3) illustrates a significant limitation of phase reduction when applied to small populations of oscillators. Specifically, it predicts that the overall phase shift will be the sum of the phase shifts resulting from perturbations to each constituent oscillator. As illustrated in Figure 1, such a reduction strategy cannot predict the phase shift caused by modifying the relative position of these oscillators, for instance, with a desynchronizing pulse. While these points are illustrated for two coupled oscillators, these limitations are also present for larger populations of coupled oscillators. As we will show in the section to follow, augmenting this reduction with isostable coordinates can overcome this limitation.

2.2. Isostables and Augmented Phase Reduction. In order to make the phase reduction more accurate, one can also track the isostable coordinates, which represent directions transverse to the periodic orbit. This section reviews some of the main findings from [69], [66] which will be important in the results to follow. Of course, when considering the example system (3), a transformed system with one phase and one isostable coordinate is not really reduced, as the number of states in the original model and transformed models are identical. Nevertheless, this example is helpful for developing intuition for more complicated models.

To begin, consider a general system of the form (1) which admits a stable periodic orbit, γ . Isostable coordinates of this periodic orbit can be defined by considering the transient behavior of (1) near γ . If we consider any initial condition $\mathbf{x}(0) \in \Gamma_0$, where Γ_0 is defined as the $\theta = 0$ isochron, then by the definition of isochrons (4), this initial condition first returns to Γ_0 at time T . By defining Γ_0 as our Poincaré surface, this allows for the construction of a Poincaré map [64],

$$(6) \quad \Upsilon : \Gamma_0 \rightarrow \Gamma_0; \quad \mathbf{x} \mapsto \phi(T, \mathbf{x}),$$

where $\phi(t, \mathbf{x})$ is the unperturbed flow of (1), and $\Upsilon(\mathbf{x}_0) = \mathbf{x}_0$ for $\mathbf{x}_0 \in \gamma$. At all locations in an n -dimensional neighborhood of $\mathbf{x}_0$ (not just on the Poincaré surface), $\phi(T, \mathbf{x})$ is well approximated by

$$(7) \quad \phi(T, \mathbf{x}) = \mathbf{x}_0 + J_\phi(\mathbf{x} - \mathbf{x}_0) + \mathcal{O}(\|\mathbf{x} - \mathbf{x}_0\|^2),$$

where J_ϕ is the Jacobian of $\phi(T, \mathbf{x})$ evaluated at $\mathbf{x}_0$ (not to be confused with J , the Jacobian of the vector field of (1) as defined in Appendix A). If we suppose J_ϕ is diagonalizable with real eigenvalues, let $V \in \mathbb{R}^{n \times n}$ be a matrix whose columns form a basis of unit length right eigenvectors $\mathbf{v}_j$ with corresponding left eigenvectors $\mathbf{u}_j$ and eigenvalues λ_j , for $j = 1, \dots, n$. These eigenvalues are also known as Floquet multipliers, or characteristic multipliers of the system [21]. Because the linearization (7) is defined for a stable periodic orbit, there is exactly one eigenvalue equal to one, and the remaining eigenvalues have magnitude strictly less than 1. We will assume that all eigenvalues are real. Without loss of generality, assume that all eigenvalues are positive (if they are not, one can take the period to be $2T$ so that eigenvalues are strictly positive), then for any $0 < \lambda_j < 1$, we can define an associated isostable coordinate

$$(8) \quad \psi_j(\mathbf{x}) = \lim_{k \rightarrow \infty} [\mathbf{u}_j^T (\phi(t_\Gamma^k, \mathbf{x}) - \mathbf{x}_0) \cdot \exp(-\log(\lambda_j)t_\Gamma^k/T)],$$

where t_Γ^k is the k th return time to Γ_0 under the flow and $\mathbf{u}_j$ is used to select for the component of the decay toward $\mathbf{x}_0$ in the $\mathbf{v}_j$ direction. The definition of isostable coordinates considers the position of an oscillator in the basis of the eigenvectors of

J_ϕ as it exponentially converges to the periodic orbit and compares it with another function which grows exponentially at the same rate. The limit of these two values gives a scalar isostable coordinate; cf., [43], [69], [66]. We emphasize that (8) is valid in the basin of attraction of the limit cycle, not just on Γ_0 .

As an example, consider the coupled theta model (3). Near the periodic orbit, Γ_0 can be approximated as the surface orthogonal to $\mathbf{Z}(0)$. Here, for states close to the periodic orbit, the section $\text{mod}(T_1 + T_2, 2\pi) = 0$ corresponds to Γ_0 . A Poincaré map of the form (6) can be defined with respect to this surface, whose linearization is found numerically to be

$$(9) \quad \begin{bmatrix} T_1(t+T) \\ T_2(t+T) \end{bmatrix} = \mathbf{x}_0 + \begin{bmatrix} 0.55 & 0.45 \\ 0.45 & 0.55 \end{bmatrix} \left(\begin{bmatrix} T_1(t) \\ T_2(t) \end{bmatrix} - \mathbf{x}_0 \right) \quad \text{for } \begin{bmatrix} T_1(t) \\ T_2(t) \end{bmatrix} \in \Gamma_0.$$

The eigenvectors of this linearized map are $\mathbf{v}_2 = [1 \ 1]^T$, $\mathbf{v}_1 = [1 \ -1]^T$ with associated eigenvalues $\lambda_2 = 1$ and $\lambda_1 = 0.1$. Here $\mathbf{v}_2$ corresponds to the periodic orbit so that $\mathbf{v}_1$ can be used to define the isostable coordinate ψ_1 . Figure 2 shows the isostable coordinates of (3) calculated according to (8). Compare this coordinate system to the one defined by isochrons shown in Figure 1. While isochrons give a sense of the progression through the periodic orbit, isostable coordinates give a sense of how long it will take to approach the periodic orbit, with values of larger magnitude corresponding to longer times. For instance, near the unstable periodic orbit, $|\psi_1|$ approaches infinity while values that are already close to the stable periodic orbit have isostable coordinates with relatively small magnitude.

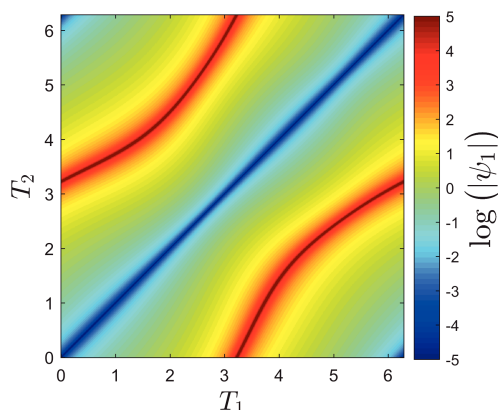


Fig. 2 For the coupled theta model (3), ψ_1 is evaluated directly by approximating the limit in (8). Isostable coordinates give a sense of how long it will take the unperturbed dynamics to approach the stable periodic orbit.

For a general system (1), isostable coordinates can be particularly useful, as in the absence of external perturbation, one can verify by direct differentiation of (8) that $\dot{\psi}_j = \kappa_j \psi_j$, where $\kappa_j = \log(\lambda_j)/T$. Switching to isostable coordinates via the chain rule and assuming that the state is close to the periodic orbit we have

$$(10) \quad \frac{d\psi_j}{dt} = \mathbf{I}_j^T(\theta) (\mathbf{F}(\mathbf{x}^\gamma(\theta)) + \mathbf{P}(t)),$$

where $\mathbf{x}^\gamma(\theta)$ is the state on the periodic orbit as a function of θ and the infinitesimal isostable response curve (iIRC) $\mathbf{I}_j(\theta) \equiv \frac{\partial \psi_j}{\partial \mathbf{x}} \big|_{\mathbf{x}^\gamma(\theta)}$ and can be calculated either directly

or using strategies discussed later. Because $\dot{\psi}_j = \kappa_j \psi_j$ when $\mathbf{P}(t) = 0$, this implies that $\mathbf{I}_j^T \mathbf{F}(\mathbf{x}^\gamma(\theta)) = \kappa_j \psi_j$, so that the above equation can be simplified to

$$(11) \quad \dot{\psi}_j = \kappa_j \psi_j + \mathbf{I}_j^T(\theta) \mathbf{P}(t),$$

which yields the isostable reduction from [69] and is analogous to the phase reduction (5).

Equations (5) and (11) are first order accurate approximations to the phase and isostable dynamics, respectively. In order to extend the accuracy of these equations, it is necessary to relate the isostable coordinates to perturbations from the limit cycle. As illustrated in [66], this can be accomplished using Floquet theory [19], [30], which states that near the periodic orbit, the general structure of solutions can be written as

$$(12) \quad \mathbf{x}(t) - \mathbf{x}^\gamma(t) = \sum_{j=1}^n c_j e^{\kappa_j t} \mathbf{q}_j^t(t),$$

where $\mathbf{q}_j^t(t) \in \mathbb{R}^n$ are T -periodic functions, κ_j are the Floquet exponents defined earlier, and c_j are constants which are chosen to satisfy the initial conditions. Using the definition of isostable coordinates (8), it was shown in [66] that for small enough perturbations from the periodic orbit, the position of an oscillator can be written as

$$(13) \quad \mathbf{x}(\theta, \psi_1, \dots, \psi_{n-1}) = \mathbf{x}^\gamma(\theta) + \sum_{j=1}^{n-1} \mathbf{q}_j(\theta) \psi_j,$$

where

$$(14) \quad \begin{aligned} \mathbf{q}_j(\theta) &= \mathbf{q}_j^t(t\omega), \\ \mathbf{q}_j^t(t) &\equiv \frac{[\phi(t, \epsilon \mathbf{v}_j + \mathbf{x}_0) - \phi(t, \mathbf{x}_0)] \exp(-\kappa_j t)}{\epsilon} \end{aligned}$$

for $0 < \epsilon \ll 1$. The representation (13) above arises as a consequence of Floquet theory [30]. Here, each function $\mathbf{q}_j(\theta)$ is related to the decay of the eigenvalues of the linearized map (7). Furthermore, by calculating the Hessian matrix of partial derivatives of isochrons and isostables along the periodic orbit, i.e., $H_{\theta, \mathbf{x}^\gamma(\theta)} \equiv \nabla(\nabla\theta)$ and $H_{\psi_j, \mathbf{x}^\gamma(\theta)} \equiv \nabla(\nabla\psi_j)$, both evaluated at $\mathbf{x}^\gamma(\theta)$, we can derive the second order accurate augmented phase reduction [66]

$$(15) \quad \begin{aligned} \dot{\theta} &= \omega + (\mathbf{Z}(\theta) + H_{\theta, \mathbf{x}^\gamma(\theta)} \Delta \mathbf{x})^T \mathbf{P}(t), \\ \dot{\psi}_j &= \kappa_j \psi_j + (\mathbf{I}_j + H_{\psi_j, \mathbf{x}^\gamma(\theta)} \Delta \mathbf{x})^T \mathbf{P}(t), \end{aligned}$$

where $\Delta \mathbf{x} = \mathbf{x}(\theta, \psi_1, \dots, \psi_{n-1}) - \mathbf{x}^\gamma(\theta)$ and $j = 1, \dots, n-1$. Substituting (13) into (15) yields the second order reduced dynamics in terms of the phase and isostable coordinates,

$$(16) \quad \begin{aligned} \dot{\theta} &= \omega + \mathbf{Z}^T(\theta) \mathbf{P}(t) + \sum_{k=1}^{n-1} [\mathbf{B}^{kT}(\theta) \psi_k] \mathbf{P}(t), \\ \dot{\psi}_j &= \kappa_j \psi_j + \mathbf{I}_j^T(\theta) \mathbf{P}(t) + \sum_{k=1}^{n-1} [\mathbf{C}_j^{kT}(\theta) \psi_k] \mathbf{P}(t) \end{aligned}$$

for $j = 1, \dots, n-1$ with $\mathbf{B}^k(\theta) \equiv \left(\frac{\partial}{\partial \psi_k} \frac{\partial \theta}{\partial \mathbf{x}} \right) \big|_{\mathbf{x}^\gamma(\theta)} = H_{\theta, \mathbf{x}^\gamma(\theta)} \mathbf{q}_k(\theta)$ and $\mathbf{C}_j^k(\theta) \equiv \left(\frac{\partial}{\partial \psi_k} \frac{\partial \psi_j}{\partial \mathbf{x}} \right) \big|_{\mathbf{x}^\gamma(\theta)} = H_{\psi_j, \mathbf{x}^\gamma(\theta)} \mathbf{q}_k(\theta)$. Finally, in high-dimensional models, it is often the case that some λ_j will be close to zero so that the resulting κ_j is negative and large in magnitude. In this case, the corresponding ψ_j can generally be assumed to be identical to zero and ignored from (16), resulting in a reduction in the dimension compared to the original equation (1).

It should be noted that the amplitude coordinates used in (16) are not unique, and other authors have been successful in using related strategies to understand the behavior of stochastic limit cycle oscillators. For instance, [5] investigates both an orthogonal basis and a basis chosen according to Floquet theory (similar to the one used here) to investigate the behavior of stochastic differential equations; when using the Floquet basis, it is possible to exploit the effective decoupling between the phase and amplitude coordinates for small magnitude perturbations. Furthermore, [34] proposes a drift-free stochastic differential equation to study the resulting phase noise due to the inherent randomness of chemical kinetics. In [60], a generalized set of phase and amplitude coordinates is used to show how correlated noise can influence the behavior of limit cycle oscillators. Additionally, [6] suggests defining a stochastic phase using a variational approach in order to increase the accuracy of the resulting phase equation over long time scales. The authors of [18] take an entirely different approach in the investigation of oscillatory biochemical reactions under the influence of molecular noise by directly analyzing a Fokker–Planck equation derived from the chemical master equation in the weak-noise limit. Clearly, phase and amplitude coordinate systems have been immensely useful for characterizing behavior of perturbed and noisy limit cycle oscillators.

2.3. Augmented Phase Reduction for the Coupled Theta Model. As shown above, the coupled theta model (3) can be transformed to a system with one phase variable θ and one isostable coordinate ψ_1 which gives information about its state with respect to its stable periodic orbit. Therefore, the augmented phase reduction of the form (16) will require the calculation of four functions: $\mathbf{Z}(\theta)$, $\mathbf{B}^1(\theta)$, $\mathbf{I}_1(\theta)$, and $\mathbf{C}_1^1(\theta)$. These can be calculated using one of two different strategies, either directly from the globally calculated isochron and isostable coordinates as shown in Figures 1 and 2, respectively, or as the solution to the adjoint equation [14], [7] and its second order analogue detailed in [66]. Appendix A details strategies for calculation of these functions. Having performed the calculation, the transformation (16) can be used on (3) to investigate its utility.

Equation (16) can be thought of as a second order correction to the standard phase reduction (5). To illustrate the advantage of this correction, consider a second order approximation of the isochrons of (3) shown in the left panel of Figure 3 calculated according to a second order accurate expansion

$$(17) \quad \theta^{(2)}(\mathbf{y}) = \theta(\bar{\mathbf{y}}) + \frac{\partial \theta}{\partial \psi_1} \Delta \psi_1 + \frac{1}{2} \frac{\partial^2 \theta}{\partial \psi_1^2} \Delta \psi_1^2,$$

where $\mathbf{y} \equiv [T_1 \ T_2]^T$, $\bar{\mathbf{y}} = \frac{1}{2} [(T_1 + T_2) \ (T_1 + T_2)]^T$, $\Delta \psi_1 = \frac{\partial \psi_1}{\partial \mathbf{y}}^T (\mathbf{y} - \bar{\mathbf{y}})$, and all derivatives are evaluated at $\bar{\mathbf{y}}$. Compared to the isochrons of the full system in the left panel of Figure 1, the second order approximation is significantly closer than the linear approximation. For this reason, we would expect the transformation of the form (16) to be significantly more accurate than the standard phase reduction (5) in predicting the behavior of the unreduced equations (3), particularly for perturbations for which

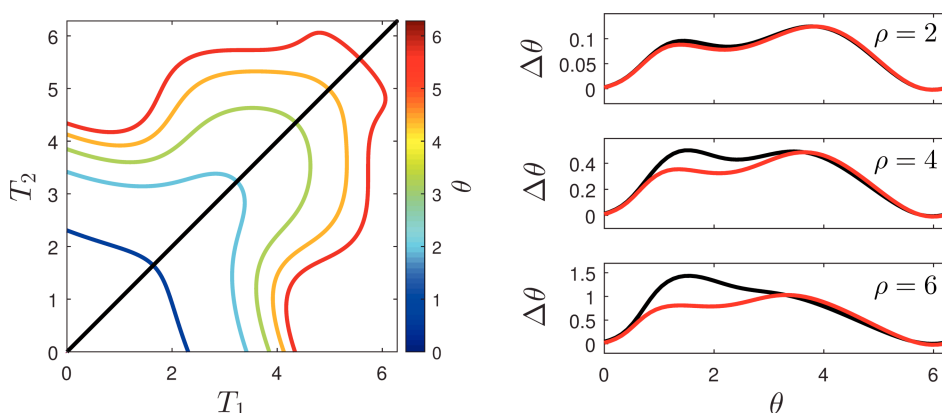


Fig. 3 The left panel shows a second order approximation to the isochrons of (3) calculated according to (17). The right panel compares the effect of pulsatile perturbations of the form $[p_1(t) \ p_2(t)] = [\rho \ -\rho]$ applied for $T/20$ time units to the full equations (3) (black lines) and the transformed equations (16) (red lines). For comparison, standard phase reduction predicts $\Delta\theta = 0$ at all phases.

$p_1(t) \neq p_2(t)$. Indeed, the right panels of Figure 3 compare the results of applying perturbations of the form $[p_1(t) \ p_2(t)] = [\rho \ -\rho]$ for different initial conditions which start on the limit cycle and last for $T/20$ units of time. The resulting $\Delta\theta$ values show the change in phase caused by these perturbations. For relatively small perturbations, the phase differences from the full equations (3) (black lines) and the transformed equations (16) (red lines) are nearly identical. As the magnitude of the perturbation grows, however, discrepancies do emerge between the two models (recall that the standard phase reduction (5) would predict $\Delta\theta = 0$ for all perturbations). Also note that the change in phase between these perturbations grows nonlinearly with the magnitude, an effect that could not be predicted using the standard phase reduction (5).

As mentioned earlier, the transformation of (3) to (16) neither reduces the dimensionality nor simplifies the model equations. Instead, this example provides an intuitive illustration of the advantages of using (16) over the standard phase reduced equations (5) when studying coupled oscillators. In the examples to follow we will show that when we have more than two oscillators, under certain conditions we can use an equation of the form (16) to significantly reduce the dimensionality of the full equations.

3. Augmented Phase Reduction of a Large Population of Oscillators. Consider a population of N identical oscillators with mean-field coupling:

$$(18) \quad \frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} F(x_1) + G(\bar{x}, x_1) \\ F(x_2) + G(\bar{x}, x_2) \\ \vdots \\ F(x_N) + G(\bar{x}, x_N) \end{bmatrix} + \begin{bmatrix} \alpha_1 \delta \\ \alpha_2 \delta \\ \vdots \\ \alpha_n \delta \end{bmatrix} u(t),$$

where $x_i \in \mathbb{R}^M$ so that n (the overall dimension) equals NM , $F(x)$ gives the nominal dynamics of each oscillator, $G(\bar{x}, x)$ gives the coupling as a function of $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$, $\delta \in \mathbb{R}^M$ is a perturbation with α_i giving the weighted contribution felt by each oscillator, and $u(t) = \mathcal{O}(\epsilon)$ uniformly in time with $0 < \epsilon \ll 1$ is a

time-dependent scalar function. Equation (18) might, for instance, describe a population of synaptically coupled neurons under the influence of deep brain stimulation [44], where the effective voltage perturbations are dependent on each neuron's spatial location relative to the voltage probe.

To proceed, suppose that in the absence of perturbation, the system (18) admits an asymptotically stable periodic orbit, γ_N , with period T . Furthermore, suppose that γ_N is a synchronous solution, i.e., a solution for which $\mathbf{x}_1 = \mathbf{x}_2 = \cdots = \mathbf{x}_N$. We will denote the full state of the system $\mathbf{x} = [\mathbf{x}_1^T \ \cdots \ \mathbf{x}_N^T]^T$ so that we can define the Poincaré map of the form (6). For this map, Υ , recall that $\mathbf{x}_0$ is defined as the location for which $\Upsilon(\mathbf{x}_0) = \mathbf{x}_0$ so that it can be approximated by the linearization $\phi(T, \mathbf{x}) = \mathbf{x}_0 + J_\phi(\mathbf{x} - \mathbf{x}_0) + \mathcal{O}(\|\mathbf{x} - \mathbf{x}_0\|^2)$. The $N \times M$ eigenvalues of J_ϕ correspond to the Floquet multipliers of the periodic orbit, and provided that J_ϕ is diagonalizable, as defined in (8), the $NM - 1$ nonunity eigenvalues of J_ϕ and their associated eigenvectors can be used to define isostable coordinates for this system. The system (18) could be transformed using a single phase variable and $NM - 1$ isostable coordinates using methods described in [66]. However, this transformation would generally not be very useful as it would not reduce the number of state variables and would be difficult to calculate for such a large system. As we will show here, by carefully choosing isostable coordinates, only a maximum of $2M - 1$ isostable coordinates are necessary in the augmented reduction of (18).

3.1. A Minimal Reduction of a Large Population of Oscillators. We can determine a set of necessary isostable coordinates by considering the eigenvectors of J_ϕ . As shown in Appendix B, all eigenvectors of J_ϕ are either of the form

$$(19) \quad \mathbf{v}_i^S = \begin{bmatrix} \mu_i^S \\ \mu_i^S \\ \vdots \\ \mu_i^S \end{bmatrix} \in \mathbb{R}^{MN},$$

where $1 \leq i \leq M$, or are in the span of the linearly independent columns of a matrix of the form

$$(20) \quad E(\mu_i^A) = \begin{bmatrix} \mu_i^A & -\mu_i^A/(N-1) & & -\mu_i^A/(N-1) \\ -\mu_i^A/(N-1) & \mu_i^A & & -\mu_i^A/(N-1) \\ -\mu_i^A/(N-1) & -\mu_i^A/(N-1) & & -\mu_i^A/(N-1) \\ \vdots & \vdots & \dots & \vdots \\ -\mu_i^A/(N-1) & -\mu_i^A/(N-1) & & \mu_i^A \\ -\mu_i^A/(N-1) & -\mu_i^A/(N-1) & & -\mu_i^A/(N-1) \end{bmatrix},$$

where μ_i^S and $\mu_i^A \in \mathbb{R}^M$, $E(\mu_i^A) \in \mathbb{R}^{MN \times (N-1)}$, and where $1 \leq i \leq M$. All eigenvectors in the span of a given $E(\mu_i^A)$ correspond to the same eigenvalue. As shown in Appendix B there exist M eigenvectors of the form $\mathbf{v}_i^S$ and M sets of $N - 1$ eigenvectors of the form (20), all of which are linearly independent and comprise the NM eigenvectors of J_ϕ . We will refer to the eigenvectors which make up the columns of $E(\mu_i^A)$ as a set of asynchronous eigenvectors, and the eigenvectors of the form $\mathbf{v}_i^S$ will be referred to as synchronous eigenvectors. We emphasize that there are M total matrices of the form $E(\mu_i^A)$ with each matrix corresponding to the M asynchronous eigenvectors.

Synchronous eigenvectors of the matrix J_ϕ with the form $\mathbf{v}_i^S$ will yield one phase variable and $M - 1$ isostable coordinates. All of these coordinates with Floquet

multipliers that are not near zero (and therefore do not decay rapidly) must be used in the augmented phase reduction. In general, each of the $N - 1$ columns of each $E(\mu_i^A)$ would result in $N - 1$ isostable coordinates in the reduction (16). However, we will show with the analysis starting from (21) and ending just before (27) that for any $E(\mu_i^A)$ only one isostable coordinate is necessary to capture the behavior. These arguments can be applied for each $E(\mu_1^A), \dots, E(\mu_M^A)$, resulting in M corresponding isostable coordinates. Ultimately, considering isostable coordinates resulting from all synchronous and asynchronous eigenvectors, we will show that we require $2M - 1$ isostable coordinates for the overall reduction. In order to show this result, for the moment consider a single set of asynchronous eigenvectors $E(\mu_i^A)$ with associated eigenvalue λ_i . In order to reduce (18) using a minimal number of coordinates, it will be useful to represent these eigenvectors in a different basis:

$$(21) \quad V_B = E(\mu_i^A) [\mathbf{y}_1 \quad \mathbf{y}_2 \quad \dots \quad \mathbf{y}_{N-1}],$$

where $\mathbf{y}_k \in \mathbb{R}^{N-1}$ for all k so that $E(\mu_i^A)\mathbf{y}_k$ gives the k^{th} element of our new basis. The procedure below details how we will choose each $\mathbf{y}_k$ in this change of basis.

Consider the multiplication of one of these basis elements by our perturbation from (18),

$$(22) \quad \mathbf{y}_k^T E^T(\mu_i^A) \begin{bmatrix} \alpha_1 \boldsymbol{\delta} \\ \alpha_2 \boldsymbol{\delta} \\ \vdots \\ \alpha_N \boldsymbol{\delta} \end{bmatrix} = \sum_{j=1}^M \left[\delta_j \mathbf{y}_k^T E_{\text{sub}}^{(j)T}(\mu_i^A) \boldsymbol{\alpha} \right] \\ = \sum_{j=1}^M \left[\delta_j \boldsymbol{\alpha}^T E_{\text{sub}}^{(j)}(\mu_i^A) \mathbf{y}_k \right],$$

where δ_j is the j th element of the vector $\boldsymbol{\delta}$, $\boldsymbol{\alpha} \equiv [\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_N]^T$, and $E_{\text{sub}}^{(j)}$ is a subset of the rows of $E(\mu_i^A)$, specifically the j th, $M + j$ th, $2M + j$ th, $\dots$, and $(N - 1)M + j$ th rows. Equivalence in the first line of (22) can be shown by appropriately rearranging the rows of both $E(\mu_i^A)$ and $[\alpha_1 \boldsymbol{\delta}^T \quad \alpha_2 \boldsymbol{\delta}^T \quad \dots \quad \alpha_N \boldsymbol{\delta}^T]^T$ and then rewriting the multiplication. The second line is obtained by taking the transpose of the first. The structure of $E_{\text{sub}}^{(j)}$ is given in (B7) of Appendix B. Because of this highly regular structure, we can conclude that if $\mathbf{y}_k$ is in the null space of $\boldsymbol{\alpha}^T E_{\text{sub}}^{(i)}(\mu_i^A)$ for any i , then it will also be in the null space of $\boldsymbol{\alpha}^T E_{\text{sub}}^{(j)}(\mu_i^A)$ for any j . If this is the case, the perturbation $[\alpha_1 \boldsymbol{\delta}^T \quad \alpha_2 \boldsymbol{\delta}^T \quad \dots \quad \alpha_N \boldsymbol{\delta}^T]^T u(t)$ will be orthogonal to the k th element of our new basis. We will therefore choose our new basis so that the resulting elements are orthogonal to each other and so that $N - 2$ elements of this basis are orthogonal to the perturbation $\mathbf{P}_\delta \equiv [\alpha_1 \boldsymbol{\delta}^T \quad \alpha_2 \boldsymbol{\delta}^T \quad \dots \quad \alpha_N \boldsymbol{\delta}^T]^T$. See Figure 4 for a visual representation of this choice.

In our new basis (21), when $\theta = 0$, for each set of asynchronous eigenvectors exactly $N - 2$ will be orthogonal to $\mathbf{P}_\delta$. We will use these eigenvectors to define the associated isostable coordinates from (8). Because of the orthogonality relationships mentioned earlier, for each set of asynchronous eigenvectors there is an associated set of isostable coordinates with isostable response curves for which

$$(23) \quad \mathbf{I}_\zeta^T(0) \mathbf{P}_\delta \neq 0, \\ \mathbf{I}_{\zeta+k}^T(0) \mathbf{P}_\delta = 0 \quad \text{for } k = 1, \dots, N - 2,$$

where $\zeta \in \mathbb{N}$ indicates the index of the isostable response curves within the complete set.

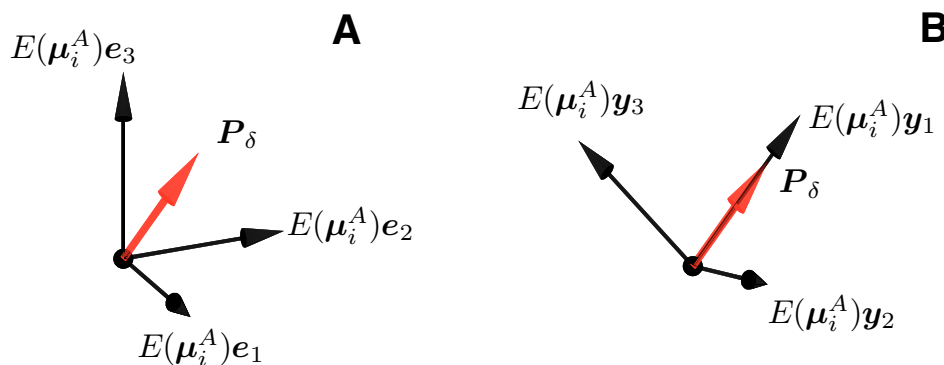


Fig. 4 Panel A shows a projection of $\mathbf{P}_\delta$ into a basis of elements of $E(\mu_i^A)$. Here, $\mathbf{e}_j$ represent elements of the standard basis, e.g., $[1 \ 0 \ 0]^T$. In panel B, after using the transformation from (21) all but one element of the new basis are orthogonal to $\mathbf{P}_\delta$. As explained below, using this basis to define isostable coordinates in (8) allows us to neglect most of the resulting coordinates from the overall reduction in (27) because perturbations are always orthogonal to most of the resulting isostable response curves.

We will now show that for $k = 1, \dots, N-2$, $\mathbf{I}_{\zeta+k}^T(\theta)\mathbf{P}_\delta = 0$ for all θ . To do so, consider that for any k , $\mathbf{I}_k(t)$ is the periodic solution to the adjoint equation (A5) from Appendix A,

$$\begin{aligned} \frac{d\mathbf{I}_k}{dt} &= (\kappa_k \mathbf{I} - \mathbf{J}(\mathbf{x}(t))^T) \mathbf{I}_k(t), \\ (24) \quad &= \begin{bmatrix} \kappa_k \mathbf{I} - \mathbf{A}^T(t) - \mathbf{B}^T(t) & -\mathbf{B}^T(t) & \dots & -\mathbf{B}^T(t) \\ -\mathbf{B}^T(t) & \kappa_k \mathbf{I} - \mathbf{A}^T(t) - \mathbf{B}^T(t) & \dots & -\mathbf{B}^T(t) \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{B}^T(t) & -\mathbf{B}^T(t) & \dots & \kappa_k \mathbf{I} - \mathbf{A}^T(t) - \mathbf{B}^T(t) \end{bmatrix} \mathbf{I}_k(t), \end{aligned}$$

where $\mathbf{I}$ is the identity matrix, $\kappa_k = \log(\lambda_k)/T$, $\mathbf{x}(t) \in \gamma_N$, $\mathbf{J}$ is the Jacobian evaluated at γ_N , $\mathbf{A}(t) \equiv (\frac{\partial F}{\partial \mathbf{x}} + \frac{\partial G}{\partial \mathbf{x}})|_{\mathbf{x}(t)}$, and $\mathbf{B}(t) \equiv \frac{1}{N} \frac{\partial G}{\partial \mathbf{x}}|_{\mathbf{x}(t)}$. Note that (24) and (B1) from Appendix B are similar in structure. Following the same arguments as given in Appendix B, one can show that all periodic solutions of (24) are of the form

$$(25) \quad \mathbf{I}_{\zeta+k}(t) = \begin{bmatrix} \chi_1^{\zeta+k} \mathbf{I}_R(t) \\ \vdots \\ \chi_N^{\zeta+k} \mathbf{I}_R(t) \end{bmatrix} \quad \text{for } k = 0, \dots, N-2,$$

where $\mathbf{I}_R(t) \in \mathbb{R}^M$ is the periodic solution to $\frac{d}{dt} \mathbf{I}_R(t) = (\kappa_k \mathbf{I} - \mathbf{A}^T(t)) \mathbf{I}_R(t)$, and $\chi_1^{\zeta+k}, \dots, \chi_N^{\zeta+k}$ are real coefficients which sum to zero. Next, by direct evaluation,

$$(26) \quad \mathbf{I}_{\zeta+k}^T(t) \mathbf{P}_\delta = \mathbf{I}_R^T(t) \delta \left(\sum_{j=1}^N (\alpha_j \chi_j^{\zeta+k}) \right) \quad \text{for } k = 0, \dots, N-2.$$

From (23), when $\theta(t_0) = 0$, $\mathbf{I}_{\zeta+k}^T(t_0)\mathbf{P}_\delta = 0$ for $k = 1, \dots, N-2$. This can only be true if either $\mathbf{I}_R^T(t_0)\boldsymbol{\delta} = 0$ or $\sum_{j=1}^N(\alpha_j\chi_j^{\zeta+k}) = 0$. However, $\mathbf{I}_R^T(t_0)\boldsymbol{\delta} \neq 0$, otherwise $\mathbf{I}_\zeta^T(0)\mathbf{P}_\delta = 0$, which directly contradicts (23). Therefore, $\sum_{j=1}^N(\alpha_j\chi_j^{\zeta+k}) = 0$ for $k = 1, \dots, N-2$, which directly implies that $\mathbf{I}_{\zeta+k}^T(\theta)\mathbf{P}_\delta = 0$ for all θ for $k = 1, \dots, N-2$. Intuitively, this result holds because the perturbations from (18) are of low rank relative to the overall dimension of the system. This property is exploited by choosing the basis (21) so that perturbations only modify a single mode from the set of asynchronous eigenvectors, allowing the others to be ignored. This property will be particularly useful in the full reduction to follow. Furthermore, this process can be repeated for all M sets of asynchronous eigenvectors $E^T(\boldsymbol{\mu}_i^A)$ to form a similar basis, where each contains $N-2$ elements for which the inner product of the resulting isostable response curves and the perturbation is always zero.

Using the bases described above in a full set of phase and isostable reduced equations of the system (18) and rearranging the indices on each isostable coordinate, we arrive at

$$\begin{aligned}
 \dot{\theta} &= \omega + \mathbf{Z}^T(\theta)\mathbf{P}_\delta u(t) + \sum_{k=1}^{2M-1} [\psi_k \mathbf{B}^{kT}(\theta)] \mathbf{P}_\delta u(t) + \sum_{k=2M}^{NM-1} [\psi_k \mathbf{B}^{kT}(\theta)] \mathbf{P}_\delta u(t), \\
 \dot{\psi}_i &= \kappa_i \psi_i + \mathbf{I}_i^T(\theta)\mathbf{P}_\delta u(t) + \sum_{k=1}^{2M-1} [\psi_k \mathbf{C}_i^{kT}(\theta)] \mathbf{P}_\delta u(t) + \sum_{k=2M}^{NM-1} [\psi_k \mathbf{C}_i^{kT}(\theta)] \mathbf{P}_\delta u(t) \\
 &\quad \text{for } i = 1, \dots, 2M-1, \\
 \dot{\psi}_j &= \kappa_j \psi_j + \sum_{k=1}^{2M-1} [\psi_k \mathbf{C}_j^{kT}(\theta)] \mathbf{P}_\delta u(t) + \sum_{k=2M}^{NM-1} [\psi_k \mathbf{C}_j^{kT}(\theta)] \mathbf{P}_\delta u(t) \\
 &\quad \text{for } j = 2M, \dots, NM-1.
 \end{aligned}
 \tag{27}$$

In the above equation, we have separated the isostable coordinates for which $\langle \mathbf{I}_i(\theta), \mathbf{P}_\delta \rangle = 0$. Recalling that $u(t) = \mathcal{O}(\epsilon)$, in Appendix C we show that for ϵ small enough relative to κ_i in the limit as time approaches infinity, $\psi_i = \mathcal{O}(\epsilon)$ for all i . Furthermore, because $\langle \mathbf{I}_j(\theta), \mathbf{P}_\delta \rangle = 0$ for some isostable coordinates, provided κ_j is large enough relative to ϵ , $\psi_j = \mathcal{O}(\epsilon^2)$ for $j = 2M, \dots, NM-1$. Additionally, in Appendix C, we show that on time scales of $1/\epsilon$, one can assume $\mathbf{C}_i^k(\theta) \equiv 0$ while limiting the error in the phase to $\mathcal{O}(\epsilon^3)$. Taking this information together, and truncating all $\mathcal{O}(\epsilon^3)$ terms, we arrive at the reduction

$$\begin{aligned}
 \dot{\theta} &= \omega + \mathcal{Z}(\theta)u(t) + \sum_{k=1}^{2M-1} [\psi_k \mathcal{B}^k(\theta)]u(t), \\
 \dot{\psi}_i &= \kappa_i \psi_i + \mathcal{I}_i(\theta)u(t) \text{ for } i = 1, \dots, 2M-1,
 \end{aligned}
 \tag{28}$$

where scalar functions of θ , $\mathcal{Z}(\theta) \equiv \mathbf{Z}^T(\theta)\mathbf{P}_\delta$, $\mathcal{B}^k(\theta) \equiv \mathbf{B}^{kT}(\theta)\mathbf{P}_\delta$, and $\mathcal{I}_i(\theta) \equiv \mathbf{I}_i^T(\theta)\mathbf{P}_\delta$ are defined for notational convenience. Here, the phase coordinate and $M-1$ isostable coordinates correspond to synchronous eigenvectors (19) and M isostable coordinates correspond to asynchronous (20) eigenvectors. Additionally, as mentioned in [69], for any κ_i large enough in magnitude (i.e., if their associated Floquet mul-

multipliers are close enough to zero), perturbations to ψ_k will quickly be forgotten and these terms can also be neglected from the reduction (28).

3.2. A Computationally Efficient Reduction Strategy. The reduction strategy detailed in section 3.1 is applicable for systems of arbitrary dimension. However, as the number of oscillators in the system increases, so too does the number of states, which can make calculation of the necessary functions difficult. As we will show here, if the periodic orbits of the individual oscillators have Floquet exponents which are large in magnitude, an intermediate reduction step can be performed in order to reduce the computational complexity of the overall reduction.

To do so, consider the periodic orbit γ_N of (18). As stated earlier, we assume that along the periodic orbit $\mathbf{x}_1 = \mathbf{x}_2 = \dots = \mathbf{x}_N$, so that individual oscillators are synchronized. For such solutions,

$$(29) \quad \begin{aligned} \frac{d}{dt} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_N \end{bmatrix} &= \begin{bmatrix} \mathbf{F}(\mathbf{x}_1) + \mathbf{G}(\bar{\mathbf{x}}, \mathbf{x}_1) \\ \mathbf{F}(\mathbf{x}_2) + \mathbf{G}(\bar{\mathbf{x}}, \mathbf{x}_2) \\ \vdots \\ \mathbf{F}(\mathbf{x}_N) + \mathbf{G}(\bar{\mathbf{x}}, \mathbf{x}_N) \end{bmatrix} + \begin{bmatrix} \alpha_1 \delta \\ \alpha_2 \delta \\ \vdots \\ \alpha_N \delta \end{bmatrix} u(t) \\ &= \begin{bmatrix} \mathbf{F}(\mathbf{x}_1) + \mathbf{G}(\mathbf{x}_1, \mathbf{x}_1) \\ \mathbf{F}(\mathbf{x}_2) + \mathbf{G}(\mathbf{x}_2, \mathbf{x}_2) \\ \vdots \\ \mathbf{F}(\mathbf{x}_N) + \mathbf{G}(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} + \begin{bmatrix} \alpha_1 \delta \\ \alpha_2 \delta \\ \vdots \\ \alpha_N \delta \end{bmatrix} u(t) \end{aligned}$$

and the system behaves as N uncoupled oscillators with identical states. Keeping this in mind, consider a phase reduction of the following system:

$$(30) \quad \dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x}, \mathbf{x}) + \mathbf{p}(t),$$

with $\mathbf{p}(t) \in \mathbb{R}^M$ being $\mathcal{O}(\epsilon)$ uniformly in time. Using a preliminary augmented phase reduction, (30) can be transformed to phase and isostable coordinates of the form (16)

$$(31) \quad \begin{aligned} \dot{\theta} &= \omega + \mathbf{z}^T(\theta) \mathbf{p}(t) + \sum_{k=1}^{\beta} [\mathbf{b}^{kT}(\theta) \psi_k] \mathbf{p}(t), \\ \dot{\psi}_j &= \kappa_j \psi_j + \mathbf{i}_j^T(\theta) \mathbf{p}(t) \end{aligned}$$

for $j = 1, \dots, \beta < M-1$. Here, θ and ψ_j are reduced phase and isostable coordinates, functions $\mathbf{z}$, $\mathbf{b}^j$, and $\mathbf{i}_j$ are calculated for the single oscillator system (30) which can be obtained numerically using isostable reduction techniques described in Appendix A, κ_j represent the Floquet exponents associated with each isostable coordinate, and ω is the natural frequency. Per the results of Appendix C, some of the terms of the isostable coordinates of (31) have been truncated to take a form similar to (C1). The dimensionality of (31) has also been reduced to $\beta+1$ compared to (30) assuming that some isostable coordinates decay rapidly so that they can be neglected (cf. [69]); if this is not the case, then there is no computational advantage to using an intermediate reduction and the methods detailed in section 3.1 should be employed. Note that in the formulation of the reduction (31), the dynamics need not be inherently oscillatory without coupling (i.e., when $\mathbf{G} \equiv 0$).

We use the intermediate reduction (31) by considering a rewritten version of the main equations (18):

$$(32) \quad \frac{d}{dt} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} \mathbf{F}(\mathbf{x}_1) + \mathbf{G}(\mathbf{x}_1, \mathbf{x}_1) \\ \mathbf{F}(\mathbf{x}_2) + \mathbf{G}(\mathbf{x}_2, \mathbf{x}_2) \\ \vdots \\ \mathbf{F}(\mathbf{x}_N) + \mathbf{G}(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} + \begin{bmatrix} \alpha_1 \delta u(t) + \mathbf{G}(\bar{\mathbf{x}}, \mathbf{x}_1) - \mathbf{G}(\mathbf{x}_1, \mathbf{x}_1) \\ \alpha_2 \delta u(t) + \mathbf{G}(\bar{\mathbf{x}}, \mathbf{x}_2) - \mathbf{G}(\mathbf{x}_2, \mathbf{x}_2) \\ \vdots \\ \alpha_N \delta u(t) + \mathbf{G}(\bar{\mathbf{x}}, \mathbf{x}_N) - \mathbf{G}(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix},$$

which can be reduced using the preliminary reduction (31) to

$$(33) \quad \frac{d}{dt} \begin{bmatrix} \theta_1 \\ \psi_{1,1} \\ \vdots \\ \psi_{1,\beta} \\ \vdots \\ \theta_N \\ \psi_{N,1} \\ \vdots \\ \psi_{N,\beta} \end{bmatrix} = \begin{bmatrix} \omega + \left(\mathbf{z}^T(\theta_1) + \sum_{k=1}^{\beta} [\mathbf{b}^{kT}(\theta_1) \psi_{1,k}] \right) \left(\mathbf{G}(\bar{\mathbf{x}}(\mathbf{s}), \mathbf{x}(\mathbf{s}_1)) - \mathbf{G}(\mathbf{x}(\mathbf{s}_1), \mathbf{x}(\mathbf{s}_1)) + \alpha_1 \delta u(t) \right) \\ \kappa_1 \psi_{1,1} + \mathbf{i}_1^T(\theta_1) (\mathbf{G}(\bar{\mathbf{x}}(\mathbf{s}), \mathbf{x}(\mathbf{s}_1)) - \mathbf{G}(\mathbf{x}(\mathbf{s}_1), \mathbf{x}(\mathbf{s}_1)) + \alpha_1 \delta u(t)) \\ \vdots \\ \kappa_{\beta} \psi_{1,\beta} + \mathbf{i}_{\beta}^T(\theta_1) (\mathbf{G}(\bar{\mathbf{x}}(\mathbf{s}), \mathbf{x}(\mathbf{s}_1)) - \mathbf{G}(\mathbf{x}(\mathbf{s}_1), \mathbf{x}(\mathbf{s}_1)) + \alpha_1 \delta u(t)) \\ \vdots \\ \omega + \left(\mathbf{z}^T(\theta_N) + \sum_{k=1}^{\beta} [\mathbf{b}^{kT}(\theta_N) \psi_{N,k}] \right) \left(\mathbf{G}(\bar{\mathbf{x}}(\mathbf{s}), \mathbf{x}(\mathbf{s}_N)) - \mathbf{G}(\mathbf{x}(\mathbf{s}_N), \mathbf{x}(\mathbf{s}_N)) + \alpha_N \delta u(t) \right) \\ \kappa_1 \psi_{N,1} + \mathbf{i}_1^T(\theta_N) (\mathbf{G}(\bar{\mathbf{x}}(\mathbf{s}), \mathbf{x}(\mathbf{s}_N)) - \mathbf{G}(\mathbf{x}(\mathbf{s}_N), \mathbf{x}(\mathbf{s}_N)) + \alpha_N \delta u(t)) \\ \vdots \\ \kappa_{\beta} \psi_{N,\beta} + \mathbf{i}_{\beta}^T(\theta_N) (\mathbf{G}(\bar{\mathbf{x}}(\mathbf{s}), \mathbf{x}(\mathbf{s}_N)) - \mathbf{G}(\mathbf{x}(\mathbf{s}_N), \mathbf{x}(\mathbf{s}_N)) + \alpha_N \delta u(t)) \end{bmatrix},$$

where $\mathbf{s}_i \equiv [\theta_1 \ \psi_{i,1} \ \dots \ \psi_{i,\beta}]^T$, $\bar{\mathbf{x}}(\mathbf{s}) \equiv \frac{1}{N} \sum_{k=1}^N \mathbf{x}(\mathbf{s}_k)$, and $\mathbf{x}(\mathbf{s}_k)$ gives the state of oscillator k as a function of its isostable coordinates (see, e.g., (13)).

For $\eta \in [0, 2\pi)$, one can verify that when $\mathbf{s}_i = [\eta \ 0 \ \dots \ 0]$ and $u(t) = 0$, then $\dot{\mathbf{s}}_i = [\omega \ 0 \ \dots \ 0]$ for all i . This solution represents a periodic orbit of (33). Furthermore, we will define $\bar{\mathbf{s}} \in \mathbb{R}^{\beta+1}$ so that the first element of $\bar{\mathbf{s}}$ is equal to $\text{Arg}(\frac{1}{N} \sum_{j=1}^N e^{i\theta_j})$ and the k th element of $\bar{\mathbf{s}}$ is equal to $\frac{1}{N} \sum_{j=1}^N \psi_{j,k-1}$ for $k = 2, \dots, \beta + 1$. By using Taylor expansion, one can show that when $\mathbf{s}_k - \bar{\mathbf{s}}$ is small for all k , $\mathbf{x}(\mathbf{s})$ is well approximated by $\mathbf{x}(\bar{\mathbf{s}})$. This allows us to rewrite (33) with the same structure as (18), i.e., with unperturbed dynamics that settle to a limit cycle and an effective perturbation:

$$(34) \quad \frac{d}{dt} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_N \end{bmatrix} = \begin{bmatrix} \mathbf{F}^R(\mathbf{s}_1) + \mathbf{G}^R(\bar{\mathbf{s}}, \mathbf{s}_1) \\ \mathbf{F}^R(\mathbf{s}_2) + \mathbf{G}^R(\bar{\mathbf{s}}, \mathbf{s}_2) \\ \vdots \\ \mathbf{F}^R(\mathbf{s}_N) + \mathbf{G}^R(\bar{\mathbf{s}}, \mathbf{s}_N) \end{bmatrix} + \begin{bmatrix} \alpha_1 \delta_1^R(\mathbf{s}_1) \\ \alpha_2 \delta_2^R(\mathbf{s}_2) \\ \vdots \\ \alpha_N \delta_N^R(\mathbf{s}_N) \end{bmatrix} u(t),$$

where

$$\delta_i^R(\mathbf{s}_i) \equiv \begin{bmatrix} \left(\mathbf{z}^T(\theta_i) + \sum_{k=1}^{\beta} [\mathbf{b}^{kT}(\theta_i) \psi_{i,k}] \right) \delta \\ \mathbf{i}_1^T(\theta_i) \delta \\ \vdots \\ \mathbf{i}_{\beta}^T(\theta_i) \delta \end{bmatrix},$$

and $\mathbf{F}^R$ and $\mathbf{G}^R$ are found by collecting the remaining terms. Equation (34) separates all terms of (33) which are multiplied by $u(t)$ from the remaining terms, resulting in

a more familiar form. Furthermore, along the periodic orbit, where the states of each oscillator are synchronized, $\delta_1^R = \delta_2^R = \dots = \delta_N^R$ so that the reduction methodology outlined in section 3.1 can be applied to (34). Therefore, because $\mathbf{s}_i \in \mathbb{R}^{\beta+1}$ for all i , (34) can be reduced to a set of $2(\beta+1)$ equations of the form (28)

$$(35) \quad \begin{aligned} \dot{\Theta} &= \omega + \mathbf{Z}^T(\Theta) \mathbf{Q}(t) u(t) + \sum_{k=1}^{2\beta+1} [\Psi_k \mathbf{B}^k(\Theta)] \mathbf{Q}(t) u(t), \\ \dot{\Psi}_i &= K_i \Psi_i + \mathbf{I}_i^T(\Theta) \mathbf{Q}(t) u(t), \quad i = 1, \dots, 2\beta+1, \end{aligned}$$

where

$$\mathbf{Q}(t) \equiv \begin{bmatrix} \alpha_1 \delta_1^R(\mathbf{s}_1(t)) \\ \alpha_2 \delta_2^R(\mathbf{s}_2(t)) \\ \vdots \\ \alpha_N \delta_N^R(\mathbf{s}_N(t)) \end{bmatrix}.$$

Here, Θ and Ψ_i are population phase and isostable coordinates in the reduction of (34), the terms $\mathbf{Z}$, $\mathbf{B}^k$, $\mathbf{I}_k$ are functions of the reduction which can be calculated numerically using methods from Appendix A, and K_i is associated with the Floquet multiplier of the i^{th} isostable coordinate. Since (35) and (31) have the same natural frequency, when $\Psi_1 = \dots = \Psi_{2\beta+1} = 0$, $\theta_i = \Theta$ for all i . This fact allows us to Taylor expand $\delta_i^R(\mathbf{s}_i)$ in order to write it as a function of $\Theta, \Psi_1, \dots, \Psi_{2\beta+1}$ near the limit cycle:

$$(36) \quad \delta_i^R(\mathbf{s}_i) = \begin{bmatrix} \left(\mathbf{z}^T(\Theta) + \mathbf{z}'^T(\Theta) \sum_{k=1}^{2\beta+1} \left[\frac{\partial \theta_i}{\partial \Phi_k} \Phi_k \right] + \sum_{k=1}^{\beta} \left[\mathbf{b}^k(\Theta) \sum_{j=1}^{2\beta+1} \left(\frac{\partial \psi_{i,k}}{\partial \Psi_j} \Psi_j \right) \right] \right) \delta \\ \left(\mathbf{i}_1^T(\Theta) + \mathbf{i}_1'^T(\Theta) \sum_{k=1}^{2\beta+1} \left[\frac{\partial \theta_i}{\partial \Psi_k} \Psi_k \right] \right) \delta \\ \vdots \\ \left(\mathbf{i}_{\beta}^T(\Theta) + \mathbf{i}_{\beta}'^T(\Theta) \sum_{k=1}^{2\beta+1} \left[\frac{\partial \theta_i}{\partial \Psi_k} \Psi_k \right] \right) \delta \end{bmatrix} + \sum_{k=1}^{2\beta+1} \mathcal{O}((\Phi_k)^2) + \mathcal{O}(\Theta^2),$$

where $' \equiv \partial/\partial\Theta$. While (35) is composed of many terms, the only remaining variables are $\Theta, \Psi_1, \dots, \Psi_{2\beta+1}$ so that (35) can be simplified to

$$(37) \quad \begin{aligned} \dot{\Theta} &= \omega + \mathcal{Z}(\Theta) u(t) + \sum_{k=1}^{2\beta+1} [\Psi_k \mathcal{B}^k(\Theta)] u(t), \\ \dot{\Psi}_i &= K_i \Psi_i + \mathcal{I}_i(\Theta) u(t) \quad \text{for } i = 1, \dots, 2\beta+1, \end{aligned}$$

where functions $\mathcal{Z}$, $\mathcal{B}^k$, and $\mathcal{I}_i$ are scalar functions of Θ , found by appropriately grouping the terms of (35). Notice that the structures of both (37) and (28) are identical. Both equations provide the same information about (18), with the difference that (37) represents the reduced behavior of the intermediate reduction (34) while (28) is a direct reduction of (18).

While the computation of the reduced system detailed in this section requires an extra coordinate transformation and computation of many intermediate functions, this extra work can come at tremendous savings in computational efficiency provided β , the number of states kept in the intermediate reduction, is smaller than M , the number of states for each oscillator. This savings comes from the computation of the phase and isostable response curves, for which the computational effort to calculate and evaluate the Jacobians from (A5) grows quadratically with the number of states in the system.

4. Example Reduction of a Coupled Population of Circadian Oscillators. We now investigate the reduction strategy derived in the previous section applied to populations of circadian oscillators. In each example, we are able to significantly reduce the dimensionality of the perturbed equations while still retaining an accurate representation of the system behavior.

In nearly all organisms, circadian rhythms provide the adaptive advantage of being able to predict and respond to daily fluctuations in their environment [70], [16]. While the mechanism and complexity of circadian timekeeping can be vastly different between organisms, in mammals, the suprachiasmatic nucleus (SCN) is responsible for maintaining circadian time throughout the body [56], [47]. This “master clock” is comprised of roughly 10,000 coupled neurons and entrains to external cues such as light to regulate circadian timing. The individual cells within the SCN are capable of maintaining oscillations when removed from the larger population [63], [28], suggesting that the population level oscillations in the SCN are the result of synchronized oscillations of individual cells.

Here, we will consider a model for the behavior of coupled SCN neurons [17],

$$\begin{aligned}
 \dot{a}_i &= h_1 \frac{K_1^n}{K_1^n + c_i^n} - h_2 \frac{a_i}{K_2 + a_i} + h_c \frac{KF}{K_c + KF} + \alpha_i L(t), \\
 \dot{e}_i &= h_3 a_i - h_4 \frac{e_i}{K_4 + e_i}, \\
 \dot{c}_i &= h_5 e_i - h_6 \frac{c_i}{K_6 + c_i}, \\
 \dot{d}_i &= h_7 a_i - h_8 \frac{d_i}{K_8 + d_i}, \quad i = 1, \dots, N.
 \end{aligned}
 \tag{38}$$

Here, the variables a_i , e_i , and c_i represent concentrations of an mRNA of the clock gene, the resulting protein, and the nuclear form of the protein, respectively, in oscillator i , d_i is a neurotransmitter by which cells can communicate with each other, time is in units of hours, F is the average value of that neurotransmitter, i.e., $F \equiv (1/N) \sum_{k=1}^N d_k$, $L(t)$ a light perturbation, and α_i is a given oscillator’s sensitivity to light. In (38), it is assumed that the spatial transmission of neurotransmitters is fast relative to the approximately 24-hour time scale of oscillations so that oscillators are effectively coupled through their mean-field. In this example, $\alpha_1 = -0.5$, $\alpha_N = 1$, with the remaining coefficients evenly spaced so that $\alpha_i = -0.5 + 1.5 \frac{(i-1)}{(N-1)}$. With this choice of parameters, on average, light stimulation tends to activate the production of a_i , but in some oscillators, its production is moderately inhibited. Furthermore, to prevent negative concentrations in any individual oscillator when light stimulation is applied, $\dot{a}_i(t)$ is set to zero if failure to do so would produce a negative concentration. All parameters, with the exception of each α_i , are identical. We take $h_1 = 0.840$, $h_2 = 0.628$, $n = 8$, with all remaining parameters identical to those in the nominal parameter set given in Figure 1 of [17]. In the absence of light, the dynamics

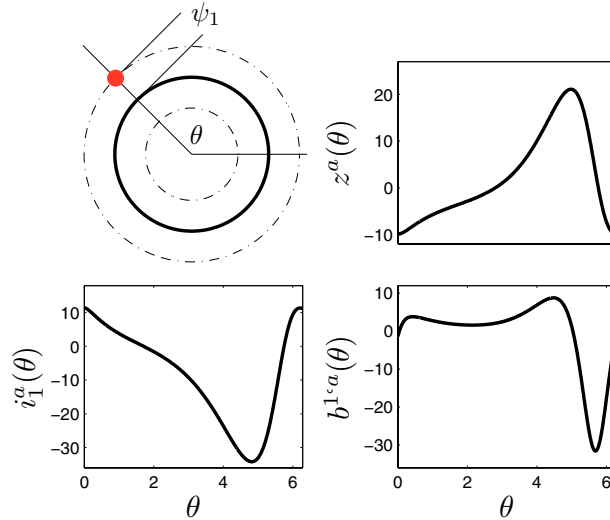


Fig. 5 In the intermediate reduction of (38), individual oscillators are reduced to an equation of the form (31). The resulting reduction has one phase coordinate, θ , and one isostable coordinate, ψ_1 . A representation of this reduced coordinate system is given in the top-left panel. The function $z^a(\theta)$, is shown in the top-right panel, which represents the gradient of the phase with respect to the variable a for locations on the limit cycle. According to this curve, light applied at $\theta \approx 3\pi/2$ will advance the phase, while the same light applied at $\theta \approx 0$ will delay the phase. The function $i_1^a(\theta)$ is shown in the bottom-left panel. According to this curve, light applied at $\theta \approx 3\pi/2$ (resp., $\theta = 0$) will decrease (resp., increase) the ψ_1 coordinate. The bottom-right panel shows $b^{1,a}(\theta)$. This function gives a correction for the local gradient of the phase when the state is perturbed from the limit cycle according to the relation $d\theta/da = z^a(\theta) + \psi_1 b^{1,a}(\theta)$.

of (38) settle to a limit cycle for which all oscillators are synchronized with period $T = 24.16$ hours.

In order to calculate the augmented isostable reduction, one could do so either directly using the methodology from section 3.1 or by first reducing the individual oscillators and then finding the population reduction using the methodology described in section 3.2. Here, we will illustrate the latter, as it provides more insight into the overall reduction methodology. As detailed in section 3.2, a preliminary reduction is performed on an individual oscillator from (38) taking $F = d_i$, i.e., replacing the mean-field (average) value of d_i with the value for a single oscillator as mandated by (32). An augmented phase reduction of the form (31) can be obtained for this intermediate system, where all necessary functions are found by calculating the appropriate solutions of the adjoint equation (A5) as detailed in Appendix A. While (38) has four variables, and hence three isostable coordinates, all but one of these coordinates decays rapidly, and we ignore the others. The resulting reduction requires two dimensions: one for the phase coordinate, and one for the slowly decaying isostable coordinate. In two dimensions, the limit cycle can be thought of as a circle, where the phase represents the location on the periodic orbit, and the isostable coordinate represents the location in the radial direction. The top-left panel of Figure 5 gives a representation of this reduced coordinate system. In (38), light perturbations only affect the a coordinate for a given oscillator; for this reason, the remaining panels of Figure 5 only show the resulting response functions for perturbations to the variable a for each oscillator in the intermediate reduction (31).

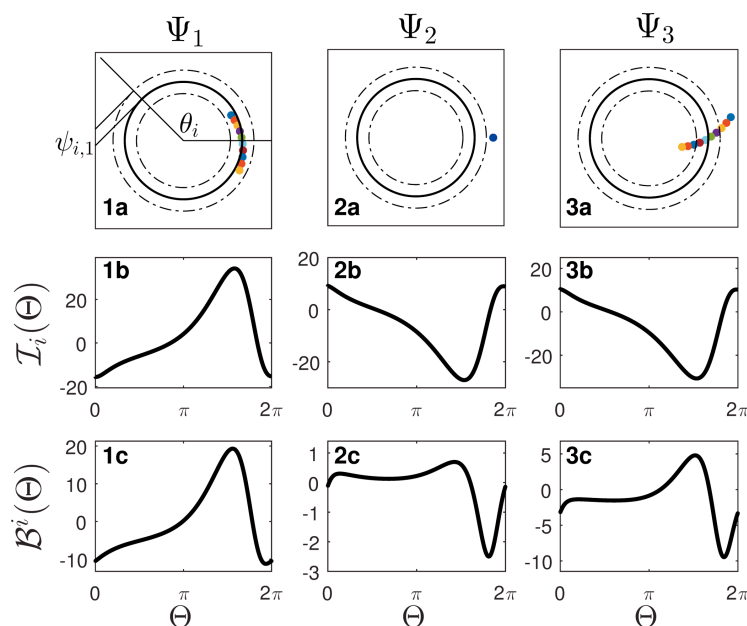


Fig. 6 Using the strategy detailed in section 3.2, the preliminary population of phase reduced oscillators (33) can be further reduced to an equation of the form (37). For an $N = 10$ system of oscillators, panels 1a, 2a, and 3a represent the population reduced coordinates $[\Theta, \Psi_1, \Psi_2, \Psi_3] = [0, 1, 0, 0]$, $[0, 0, 1, 0]$, and $[0, 0, 0, 1]$, respectively, where each dot represents $\psi_{i,1}$, θ_i for an individual oscillator. Note that in panel 2a the ψ_1 and θ coordinates of the individual oscillators are identical. The associated functions $\mathcal{I}_i(\Theta)$ and $\mathcal{B}^i(\Theta)$ are calculated numerically and shown in panels 1b–3b and 1c–3c, respectively.

The preliminary reduction, with numerically determined functions shown in Figure 5, can be used to approximate the full system (38) with the augmented phase reduction (33). Each oscillator from (38) has four variables, while the reduction (31) uses two; in this case the preliminary reduction decreases the number of states from $4N$ to $2N$. Equation (33) itself has a stable limit cycle which is used to calculate the secondary phase reduction. A naive choice of isostable coordinates would require us to keep track of all $2N - 1$ possible isostable coordinates in the subsequent reduction. However, by applying the reduction strategy from section 3.1, (33) can be reduced from a system of the form (37) with $2N$ variables to a system of four variables: Θ , the population phase coordinate, and Ψ_1 , Ψ_2 , and Ψ_3 , the population isostable coordinates. Intuitively, this second reduction is made possible by exploiting the low-dimensional nature of the allowed perturbations and ignoring the modes that are not altered by the perturbations (see, for example, the explanation given in Figure 4).

The population limit cycle is a solution for which all oscillators are synchronized and the location on the limit cycle can be inferred from the population phase Θ . The resulting isostable coordinates represent perturbations in directions away from the limit cycle. For $N = 10$ oscillators we calculate the necessary functions of the population reduction, with results shown in Figure 6. Panels 1a–1c of Figure 6 give a visual representation of each isostable coordinate and can be interpreted as follows: At $[\Theta, \Psi_1, \Psi_2, \Psi_3] = [0, 0, 0, 0]$, the state of each individual oscillator from (33) is $\theta_i = 0$, $\psi_{i,1} = 0$. Panel 1a, for instance, shows the state of each oscillator from (33)

when $[\Theta, \Psi_1, \Psi_2, \Psi_3] = [0, 1, 0, 0]$ for the population oscillation. The solid black circle represents the periodic orbit of each preliminary reduced oscillator (i.e., the curve with $\psi_1 = 0$) and the inner and outer dashed circles represent isostable coordinates of $\psi_{i,1} = \pm 0.2$. The population isostable coordinate Ψ_1 corresponds primarily to a spread in phases θ_i of the individual oscillators, but also causes a small spreading of $\psi_{i,1}$. In the adjacent panels, Ψ_2 corresponds to an identical perturbation in $\psi_{i,1}$ for each oscillator, and Ψ_3 primarily spreads the values of $\psi_{i,1}$ while slightly spreading the phases. Panels 1b–3b correspond to the population isostable response curves for each associated isostable coordinate, and panels 1c–3c show the correction to the phase response curve as the corresponding isostable coordinate grows. Note the similarity between the calculated functions in Figure 6 for the population reduction and those in Figure 5 for the individual oscillator reduction. For instance, $\mathcal{I}_2(\Theta)$ and $\mathcal{B}^2(\Theta)$ have the same shape as i_1^a and $b^{1,a}$, respectively. This is to be expected, as the coordinate Ψ_2 represents a synchronous shift in each individual isostable coordinate $\psi_{i,1}$.

Finally, to illustrate the advantage of using the augmented phase reduction over the standard phase reduction, we investigate the effect of pulsatile light perturbations for a population of $N = 20$ circadian oscillators. The left panels of Figure 7 show the change in phase resulting from 2-hour duration light perturbations of magnitude $L = 0.01$ (top-left panel) and $L = 0.02$ (bottom-left panel). Before each perturbation is applied, the system is allowed to relax to its limit cycle solution, and perturbations are applied starting at a known phase of oscillation. The resulting phase change, $\Delta\Theta$, is measured by comparing the phase of oscillation after returning to the limit cycle to the phase had the system not received the perturbation. Using this procedure, blue lines show numerically computed curves for the augmented reduced system (labeled AR) from (37) and red lines give the result using the standard phase reduction (labeled PR) from (2), which is obtained by calculating only the phase response curve for perturbations to (38) and neglecting all isostable coordinates. The black dots give numerical estimates for the full system of unreduced equations (labeled FS) given in (38). Much like for the coupled theta oscillators from section 2, the predicted phase change from augmented reduction is nearly identical to the phase change resulting in simulations of the full system. Furthermore, as the magnitude of perturbations increases, the standard phase reduction becomes a worse approximation of the full system. These differences are also apparent when investigating each system's ability to entrain to a periodic light perturbation of the form

$$(39) \quad L(t) = \begin{cases} L_p & \text{if } \text{mod}(t, T_p) \leq D_p, \\ 0 & \text{otherwise.} \end{cases}$$

Here, L_p is the magnitude of the pulse, T_p is the pulsing period, and D_p is the duration of each pulse. The right panel of Figure 7 shows Arnold tongues [64] for 1:1 locking to the periodic light stimulus of duration $D_p = 3$ hours; regions between lines of identical colors represent sets of parameters which result in stable locking to the periodic perturbation. These boundaries are found numerically by simulating the equations for the full system (38), augmented reduction (37), and standard phase reduction (2). The Arnold tongues for the full system (black curve) and the augmented reduction (blue curve) are nearly identical, with the right branches extending vertically from the natural period of 24.16 hours. The Arnold tongue for the standard phase reduced system (red curve) is more centered around the natural period and does not match the other two curves. The reason for this discrepancy becomes clear when we examine

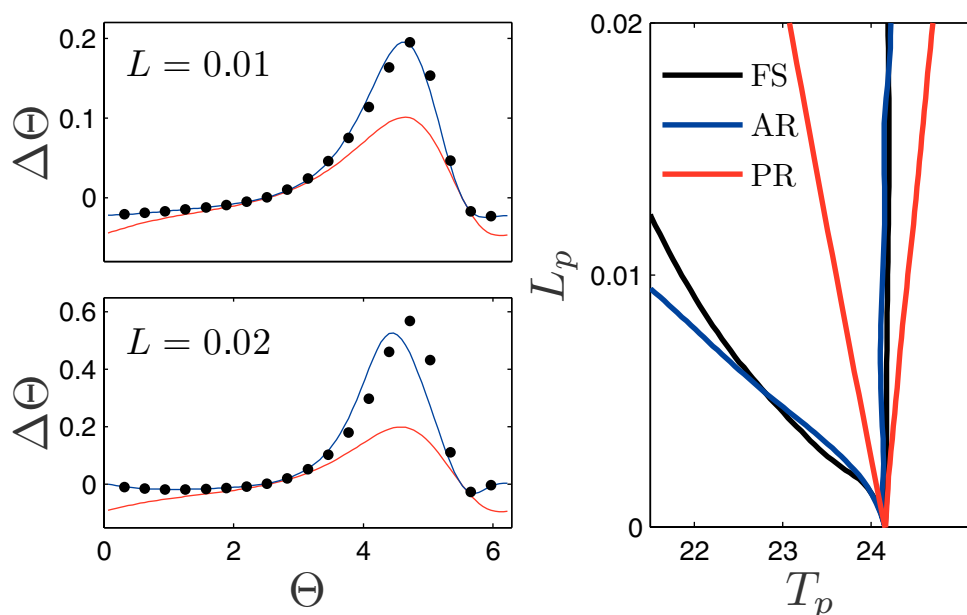


Fig. 7 For $N = 20$ circadian oscillators the full (38), augmented phase reduced (37), and standard phase reduced (2) models are simulated in response to pulsatile perturbations. Results for each model are shown in black, blue, and red, respectively. The left panels show numerically computed phase changes for 2-hour light perturbations beginning at phase Θ on the limit cycle at two different intensities. The right panel shows Arnold tongues computed for periodic perturbations of the form (39) applied to each model.

the left panels of Figure 7. When simulating the standard phase reduction, regardless of the magnitude of perturbations or pacing history, a significant portion of the response curve is negative, allowing for positive light perturbations to slow down the oscillation so that the standard reduced equations can entrain to perturbations with a larger period. For both the augmented phase reduction and full system simulations, as perturbations become larger in magnitude, the response curves are shifted to more positive values, helping entrainment to perturbations with smaller period, but hindering entrainment when the period becomes larger.

5. Direct Method for Augmented Phase Reduction. In the previous sections, we explored numerical strategies for reducing a large population of identical oscillators using both phase and isostable coordinates. In practice, however, the underlying equations of a given system may be unknown, as is the case in most experimental applications. Phase reduction is particularly useful in these contexts, as the iPRC can generally be measured experimentally using the direct method [29], [50], whereby short perturbations are applied when the system is at a known phase, the resulting phase change is inferred, and this process is repeated over multiple trials with the resulting data fit to a curve. This process has been used, for example, to calculate the phase-dependent sensitivity of human circadian rhythms to light [71], and to characterize the response of brain rhythms to electrical stimuli [4].

Here, we investigate a strategy for calculating an augmented isostable reduction of a large and heterogeneous population of oscillators of the form (1). As is usually the

case when employing the direct method for calculation of iPRCs, we will also assume that we can only measure a single observable from this system. As we illustrated in previous sections, while these large populations of oscillators may have a large number of isostable coordinates, it is often sufficient to consider a small subset in the augmented isostable reduction. For this reason, and because we have limited information about the state of the system, we will assume that a general system (1) can be well approximated by a reduction of the form

$$(40) \quad \dot{\theta} = \omega + \mathcal{Z}(\theta)u(t) + \psi\mathcal{B}(\theta)u(t),$$

$$(41) \quad \dot{\psi} = \kappa\psi + \mathcal{I}(\theta)u(t).$$

The above equation is identical to (28), except that we have dropped subscripted indices because we have a single isostable coordinate. For this system (1), consider some observable, $\text{obs}(\mathbf{x}(t))$, which is a function of the system's state. In a model of circadian oscillations, for example, this observable could be a measured concentration of a neurotransmitter. In order to define phase coordinates, one must decide where $\theta = 0$ on the periodic orbit which for the moment will be taken to be a local maximum or minimum of the observable. The case where $\theta = 0$ corresponds to the crossing of a predetermined threshold will be discussed next.

To begin, we must define isostable coordinates for the system (1). To this end, we will assume that (1) can be well approximated by an augmented reduction with a single isostable coordinate. Then, from (13), for small perturbations from the limit cycle,

$$(42) \quad \mathbf{x}(\theta, \psi) = \mathbf{x}^\gamma(\theta) + \mathbf{q}(\theta)\psi,$$

where $\mathbf{x}^\gamma(\theta)$ gives the state as a function of θ on the limit cycle, and $\mathbf{q}(\theta)$ gives the displacement from the periodic orbit as a function of both θ and ψ . For small values of ψ ,

$$(43) \quad \text{obs}(t) \approx \text{obs}(\mathbf{x}^\gamma(\theta)) + E\psi,$$

where $E \equiv \frac{\partial \text{obs}}{\partial \mathbf{x}} \big|_{\mathbf{x}^\gamma(\theta)} \mathbf{q}(\theta)$. In order to define isostable coordinates for this system, we must define some detectable event for which $\theta \approx 0$, for instance, a local maximum or minimum of the observable.

In a computational model, it was convenient to define isostable coordinates in reference to the eigenvectors of a Poincaré map. In the present case we assume that we do not have access to the full model equations, and therefore cannot define isostable coordinates in this way. Instead, we will define isostable coordinates as follows:

$$(44) \quad \psi(\mathbf{x}) = \lim_{j \rightarrow \infty} [(\text{obs}(m^j) - \text{obs}(m^\infty)) \exp(-\kappa m^j)].$$

Here, m^j corresponds to the j th time at which an event close to $\theta = 0$ is detected after the application of a perturbation and $\text{obs}(m^\infty)$ is the steady state value. We will explicitly assume that the system is at steady state before the application of the stimulus, i.e., $\text{obs}(m^0) = \text{obs}(m^\infty)$. See Panel A of Figure 9 for a visual explanation of these definitions. Here we take κ to be a representative rate at which solutions approach the steady state value. Notice that isostable coordinates in (44) are defined similarly to those from (8), i.e., the limit of an exponentially increasing function multiplied by a function decreasing exponentially at the same rate. Using this definition of isostables, we can estimate the functions in the augmented isostable reduction (40) and (41) with the steps presented in the section below.

5.1. A List of Steps to Infer the Functions and Parameters in the Augmented Phase Reduction.

Step 1. Calculate the average period T over multiple unperturbed oscillations. Take the natural frequency to be $\omega = 2\pi/T$.

Step 2. Using a protocol similar to the direct method, allow the unperturbed system to relax to its periodic orbit. Apply a small perturbation of magnitude $u(t) = \eta$ lasting for a duration of Δt units. For each perturbation, record $\text{obs}(m^0)$ (assumed to be the steady state value) and n additional data points $\text{obs}(m^1), \text{obs}(m^2), \dots, \text{obs}(m^n)$ as the system relaxes back to its periodic orbit (see panel A of Figure 9, for example). Also, record the times $m^0, \dots, m^n$ at which these events occur as well as the time t_p at which the perturbation is applied. The phase at which the perturbation is applied can be estimated as $\theta \approx (t_p - m^0)\omega$. Repeat this procedure multiple times, perturbing at various values for θ and for various strengths η . This data will be used to numerically estimate all required functions and parameters in the augmented phase reduction.

Step 3. Estimate κ , the rate of decay of the isostable coordinate. To obtain a single estimate for κ , for a single perturbation from Step 2, determine the coefficients a and b for the exponential fit $\text{obs}(m_i) - \text{obs}(m_0) = a \exp(b(i-1))$. The decay rate, κ , is approximately equal to b/T . Take κ to be the average over multiple perturbations to mitigate errors introduced by noise in the system.

Step 4. Estimate $\mathcal{I}(\theta)$, the isostable response curve using the data from Step 2. Note that in (44), isostables are defined in the limit as the system approaches the periodic orbit. However, precise measurements of the convergence to the periodic orbit can be difficult in the presence of noise and other uncertainties. For this reason, for each trial from Step 2, we approximate (44), the isostable shift caused by a perturbation with

$$(45) \quad \Delta\psi \approx (\text{obs}(m^2) - \text{obs}(m^0)) \exp(-\kappa(m^2 - t_p)),$$

where the second event is chosen so that the system has had a chance to recover after the perturbation, but is still sufficiently far from the periodic orbit so that the left-hand side is not too small. Measurements $\text{obs}(m^k)$ with $k > 2$ could also be used to approximate (44) provided $(\text{obs}(m^k) - \text{obs}(m^0))$ is not strongly influenced by noise as the system dynamics approach the limit cycle. With this estimate, and the approximation for the initial phase $\theta \approx (t_p - m^0)\omega$, $\mathcal{I}(\theta) = \Delta\psi/\eta\Delta t$. Using multiple perturbations at different phases an estimation for the entire isostable response curve can be obtained.

Step 5. Estimate both $\mathcal{Z}(\theta)$ and $\mathcal{B}(\theta)$. To do so, for a given perturbation from Step 2, suppose η and Δt are small enough that θ does not change much while the perturbation is applied, i.e., $\mathcal{I}(\theta(t)) \approx \mathcal{I}(\theta(t_p))$, $\mathcal{Z}(\theta(t)) \approx \mathcal{Z}(\theta(t_p))$, and $\mathcal{B}(\theta(t)) \approx \mathcal{B}(\theta(t_p))$. Furthermore, if η is large relative to κ ,

$$(46) \quad \begin{aligned} \psi(t) &= \int_{t_p}^t [\kappa\psi + \mathcal{I}(\theta)\eta] dt \\ &\approx \mathcal{I}(\theta(t_p))\eta(t - t_p) \quad \text{for } t_p \leq t \leq t_p + \Delta t. \end{aligned}$$

With this information, we can infer the phase as a function of time,

$$\begin{aligned}
 \theta(t + \Delta t) &= \theta(t_p) + \int_{t_p}^{t_p + \Delta t} [\omega + \mathcal{Z}(\theta)\eta + \psi\mathcal{B}(\theta)\eta] \\
 &\approx \theta(t_p) + \int_{t_p}^{t_p + \Delta t} [\omega + \mathcal{Z}(\theta(t_p))\eta + \mathcal{I}(\theta(t_p))\mathcal{B}(\theta(t_p))\eta^2(t - t_p)] \\
 (47) \quad &= \theta(t_p) + \omega\Delta t + \mathcal{Z}(\theta(t_p))\eta\Delta t + \frac{1}{2}\mathcal{I}(\theta(t_p))\mathcal{B}(\theta(t_p))\eta^2\Delta t^2.
 \end{aligned}$$

Comparing the perturbed phase, $\theta_p(t + \Delta t)$, to the phase had the perturbation not occurred, $\theta_u(t + \Delta t)$, we have

$$(48) \quad \Delta\theta \equiv \theta_p(t + \Delta t) - \theta_u(t + \Delta t) = \mathcal{Z}(\theta(t_p))\eta\Delta t + \frac{1}{2}\mathcal{I}(\theta(t_p))\mathcal{B}(\theta(t_p))\eta^2\Delta t^2.$$

In the above equation for each perturbation, the resulting $\Delta\theta$ can be inferred from the observable once the system reaches the periodic orbit, as is commonly done for the direct method (see also Figure 9). This leaves $\mathcal{Z}(\theta)$ and $\mathcal{I}(\theta)\mathcal{B}(\theta)$ as the remaining unknowns in (48). Over multiple iterations using different values of θ and η , both unknown functions can be obtained by finding a least squares fit to the data using an appropriate basis (i.e., polynomials or sinusoids). The function $\mathcal{B}(\theta)$ can then be found by dividing $\mathcal{I}(\theta)\mathcal{B}(\theta)$ by the $\mathcal{I}(\theta)$ that was calculated in Step 4. For this reason, when fitting the functions to the data, given $\mathcal{I}(\theta)$ it is useful to constrain $\mathcal{I}(\theta)\mathcal{B}(\theta)$ so that every root of the former is also a root of the latter.

A Remark About the Choice of $\theta = 0$. Above, it is assumed that $\theta = 0$ corresponds to a local minimum or maximum of the observable. Practically, this would require an accurate measurement of the moment that the local minimum or maximum is achieved and requires E from (43) to be sufficiently large that $\text{obs}(m^2) - \text{obs}(m^0)$ can provide a useful signal in a noisy environment. If either of these requirements is not fulfilled, defining $\theta = 0$ to correspond to a threshold crossing (e.g., when the voltage of a neuron crosses 0 mV in neurological applications) may be more appropriate. In this case, $\Delta\theta$ can be inferred in the usual way by measuring timing differences between threshold crossings. However, in this setting the measurement of an isostable response curve will be different. In this setting, as illustrated in [65], a principle isostable coordinate $\psi(\mathbf{x})$ can be defined according to (8) and the change in isostable coordinate resulting from a perturbation can be determined according to

$$(49) \quad \Delta\psi = \frac{\kappa[2\pi(n-1) - \omega(t_n - t_1)]}{c[\exp(\kappa(t_n - t_p)) - \exp(\kappa(t_1 - t_p))]}.$$

Here, $n \geq 2$ denotes the n th threshold crossing after the perturbation, t_n is the timing of the n th threshold crossing, t_p denotes the time that the perturbation is applied, and c is a scaling factor that can be taken to be 1. Equation (49) can be used to estimate $\mathcal{I}(\theta)$ as in Step 4 of the current section, and κ can be inferred using methods described in [65]. After these terms are determined, $\mathcal{B}(\theta)$ and $\mathcal{Z}(\theta)$ can be found using the strategy detailed in Step 5.

5.2. Example of the Direct Method Applied to a Circadian System. A significant benefit of using the above procedure for calculating the augmented phase reduction is that it does not require knowledge of the right-hand side of a general sys-

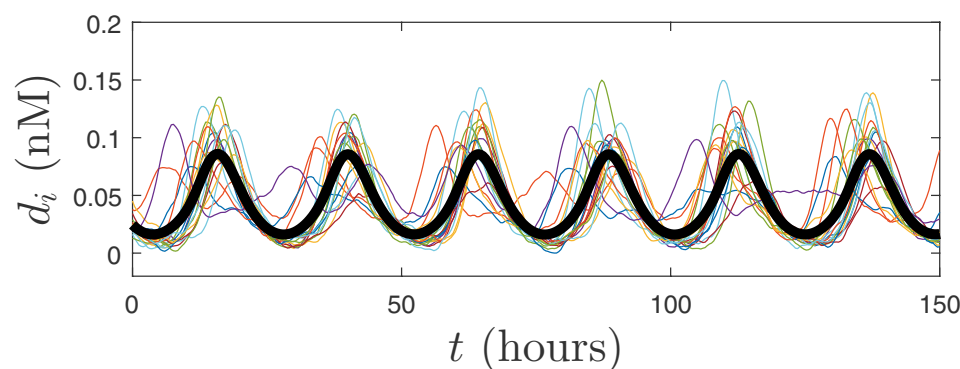


Fig. 8 Representative traces of $d_i(t)$ after initial transient behavior has died out in a heterogeneous and noisy version of (38). The system dynamics are nearly periodic, where the aggregate signal $F(t) = \frac{1}{N} \sum_i^N d_i$ achieves a local maximum every 24.15 hours on average.

tem (1). Therefore, it can be applied to an experimental system or to a system which is too large to calculate response curves numerically. Here we apply the direct method to a large, noisy, and heterogeneous population of 3000 oscillators from (38). In this example, nominal parameters are chosen as $h_1 = 0.840$, $h_2 = 0.605$, $h_c = 1.2$, $n = 6$, and all other nominal parameters are taken to be identical to those given in Figure 1 of [17]. To incorporate heterogeneity into the model, the parameters h_1, h_2, h_3, h_4, h_5 , and h_6 are drawn from a normal distribution where the mean is the nominal parameter value and the standard deviation is equal to 0.03. The sensitivity of each oscillator to light is also chosen randomly as $\alpha_i = \max(0.5 + 0.4\mathcal{N}(0, 1), 0)$, where $\mathcal{N}(0, 1)$ is a normal distribution with mean zero and unit variance. We also add independent and identically distributed zero mean white noise with intensity 1.6×10^{-4} to the variable a_i for each oscillator. Model code is provided in supplementary material.

Figure 8 shows the behavior of this system in the absence of external perturbations. Due to noise, the system behavior is not perfectly periodic; however, the aggregate signal $F(t)$ appears to be nearly periodic, with an average period of $T = 24.15$ hours so that $\omega = 0.260$, which is in line with estimates of the free running period of the human circadian pacemaker [10]. In this example, we will assume that $F(t)$ is an aggregate signal that can be measured from the system, and we will use this as our observable in order to apply the direct method in order to determine the necessary terms and functions in the augmented phase reduction (40) and (41).

Following the steps outlined in the previous section, for a general system (1) where the right-hand side is unknown, one can calculate all necessary functions in the augmented phase reduction (40) and (41). Panel B of Figure 9 gives a plot of the change in phase scaled by the stimulation strength for light perturbations lasting one hour. This data is used to calculate the functions in the augmented phase reduction, shown in panels C through F. Notice in panel B that when $\theta \approx 3$, $\Delta\theta/\eta$ tends to decrease as η becomes larger. This is reflected in the estimated curve $\mathcal{I}(\theta)\mathcal{B}(\theta)$ being negative and large in magnitude at $\theta \approx 3$. In panel F, the function $\mathcal{B}(\theta)$ has a large local maximum at $\theta \approx 2.5$, implying that as the isostable coordinate becomes more positive, the effect of a perturbation at $\theta \approx 2.5$ will be shifted toward more positive values. Conversely, $\mathcal{B}(\theta)$ has zeros at $\theta \approx 0.5$ and 4. At these phases, the effect of a perturbation does not depend on the isostable coordinate.

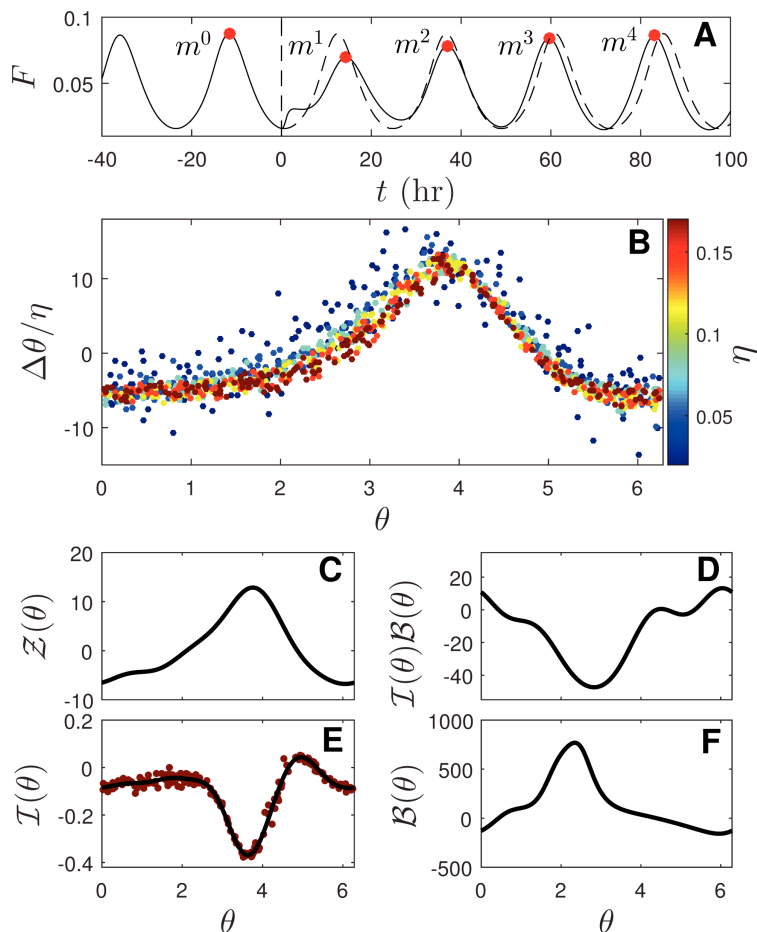


Fig. 9 Illustration of the direct method for calculating the terms in the augmented isostable reduction. Panel A gives an example of a set of measurements required from Step 2 of section 5.1. A perturbation is given at $t_p = 0$ and the times m^i and magnitudes $\text{obs}(m^i)$ at which F achieves a local maximum are recorded. The dashed line represents the expected signal (based on the natural frequency) that would be observed if the perturbation had not been given. Panel B shows $\Delta\theta/\eta$ for $\Delta t = 1$ hour perturbations of various strengths. These datapoints are used to fit functions of the form $\sum_{n=0}^4 [u_i \sin(n\theta) + v_i \cos(n\theta)]$ to both $Z(\theta)$ and $I(\theta)B(\theta)$. Panel E shows individual datapoints for estimation of the isostable response curve, and a curve fit to this data using the same basis. Panels C and D show fits of both $Z(\theta)$ and $I(\theta)B(\theta)$ to the data from panel B. Panel F shows $B(\theta)$ as calculated from the fits in panels D and E.

Finally, we can use the augmented reduction (40) and (41) to investigate entrainment in the heterogeneous 3000 oscillator system to a 24-hour light-dark cycle. Here, the external light perturbations are

$$(50) \quad L(t) = \begin{cases} 0.07 & \text{if } \text{mod}(t, 24) \leq 12, \\ 0 & \text{otherwise.} \end{cases}$$

Here, we specifically investigate the amount of time it takes to reentrain following an abrupt time shift to simulate the effect of transcontinental flight on the circadian pacemaker (i.e., jet lag). Using the full, unreduced system, the augmented phase

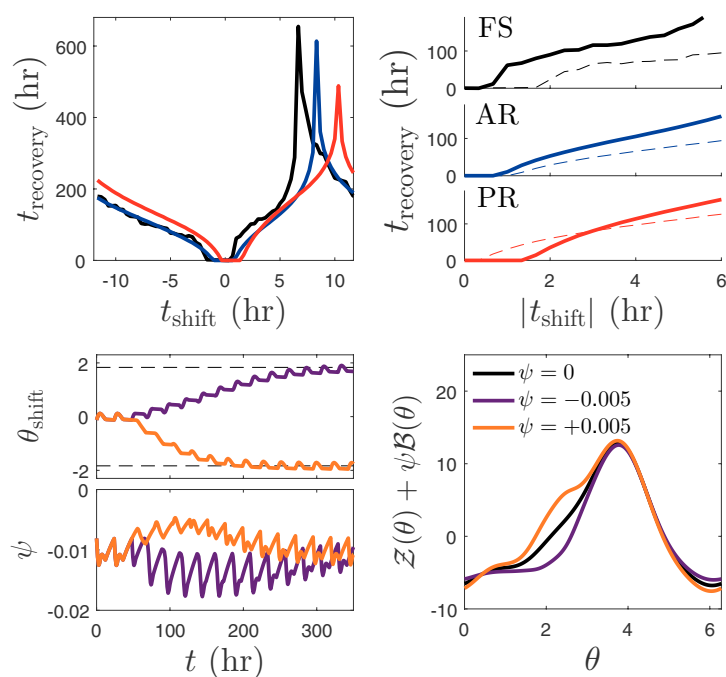


Fig. 10 The top-left panel shows the recovery time as a function of the time shift for the full (FS), phase reduced (PR), and augmented phase reduced (AR) models as black, red, and blue lines, respectively. Top-right panels compare the effect of positive (solid lines) and negative (dashed lines) values of t_{shift} on the recovery time, illustrating the ability of the augmented phase reduction to reproduce the asynchrony of recovery times in each direction. The bottom-left panels show the isostable and phase coordinates of the augmented phase reduction for $t_{\text{shift}} = +7$ and -7 (purple and orange lines, respectively) occurring at $t = 48$ hours. In the bottom-right panel, negative (resp., positive) isostable coordinates shift the effect of perturbations toward more negative (resp., positive) values.

reduced system (40) and (41), and the standard phase reduced system (i.e., where $\dot{\theta} = \omega + \mathcal{Z}(\theta)L(t)$), each system is entrained to the 24-hour cycle long enough for the transient behavior to decay. Then at some time for which $\text{mod}(t, 24) = 0$, we abruptly shift t so that the 24-hour cycle is advanced by an amount t_{shift} . Finally, the recovery time t_{recovery} is defined as the time it takes for the phase to return to within one hour of its steady state behavior. The following results do not change qualitatively if the initial time shift occurs at a different time in the circadian day. The top-left panels of Figure 10 show the result of these simulations. Here, the full system (black line) has a significant peak for $t_{\text{shift}} \approx 7$ hours. For shifts of this magnitude, the system delays its phase by approximately 17 hours, rather than advancing it by 7 hours. In the standard phase reduced model, this transition occurs for $t_{\text{shift}} \approx 10.5$ hours, a significant mismatch. Using the augmented phase reduction, the behavior is closer to the full model, with the peak occurring at $t_{\text{shift}} \approx 8.5$ hours. Furthermore, as shown in the top-right panels, both the full and augmented phase reduced systems show a marked asymmetry between shifting forward and shifting backward in time. Using the standard phase reduction, however, recovery times do not depend as much on the direction of the shift. The reason for this asymmetry becomes clear in the bottom panels of Figure 10. In the bottom-left panels, we show

the time course of ψ and $\theta_{\text{shift}}(t) \equiv \theta(t) - \theta(0) - 24t/2\pi$ for $t_{\text{shift}} = +7$ hours (purple curve) and $t_{\text{shift}} = -7$ hours (orange curve). Here, for a positive (resp., negative) value of t_{shift} , ψ decreases significantly (resp., increases slightly) during reentrainment. In the bottom-right panel, we find that for negative values of ψ , the response to perturbations, $\mathcal{Z}(\theta) + \psi\mathcal{B}(\theta)$, predominantly becomes more negative, hindering the phase advancement required for reentrainment. This feature of the augmented phase reduction drives the asymmetry in recovery time for phase advancements and phase delays.

6. Discussion and Conclusions. This paper investigates the applicability of augmented phase reduction to oscillatory populations of coupled oscillators. In applications where phase reduction alone is insufficient to understand the behavior of a population oscillation, if oscillators are identical and coupled through a mean-field, only a small number of isostable coordinates is necessary to accurately characterize the behavior of the system in response to larger perturbations. Using a strategy akin to the direct method [29], [50] for computation of phase response curves, an augmented isostable reduction could be obtained in experimental systems, or in systems for which the underlying model equations are not known. Most notably, the necessary number of isostable coordinates does not grow as the number of oscillators in the system grows; rather, it depends on the dimension and complexity of the underlying dynamics for the individual oscillators. Furthermore, we have developed more effective strategies for obtaining the required functions in the augmented reduction than were proposed in [66], with a discussion on this issue given in Appendix A.

The augmented phase reduction presented in this work (28) predicts system behavior more accurately than standard phase reduction (2), particularly as perturbations become larger. Indeed, while the standard phase reduction is based on a linear approximation of isochrons near the limit cycle, the augmented phase reduction employs a second order approximation while still yielding a closed form set of equations. However, the augmented phase reduction still truncates higher order terms in its approximation of a system's isochrons and is not valid for perturbations of arbitrary magnitude. In the examples presented in this work, this can be seen in Figure 7, where predictions between the augmented phase reduced and full system models begin to diverge as the applied perturbations become larger. Additionally, light perturbations of relatively small magnitude were used to investigate reentrainment in section 5; for larger perturbations the augmented reduction did not accurately capture the model behavior. The resulting recovery times from Figure 10 are larger than those observed physiologically [57]. Related work [59] could be viewed as a first step toward extending the accuracy of the augmented reduction presented here; however, it is not obvious how to use this framework to arrive at a closed set of reduced equations. It could be particularly useful to develop reduction strategies with arbitrarily high orders of accuracy in future work.

We illustrate the utility of this reduction strategy in a model with applications to circadian oscillations. In this example, when simulating the full circadian model (38), each additional oscillator adds four state variables, quickly increasing the dimensionality and complexity of the model. However, regardless of the number of oscillators used, only one population phase coordinate and three population isostable coordinates are required to accurately predict the effect of perturbations of moderate magnitude. We emphasize that the model (38) used in this study is relatively simple, with only four variables per oscillator, and assumes that light only affects the expression of mRNA in a manner strictly proportional to the magnitude of the applied

perturbation. In reality, much more complicated mechanisms are thought to give rise to mammalian circadian oscillations (cf., [39], [61]). While the specific model used in this study is not as physiologically detailed as others, application of augmented isostable reduction can help elucidate the mechanisms behind its observed behavior with possible applications to living systems. Specifically, in humans, jet-lag is known to last longer after eastward travel, as compared to westward travel [51], [62], [3]. This asymmetry has been attributed to factors such as a mismatch between daylight hours and the body's corresponding phase response curve at the time of arrival in a new time zone [57] and to the slightly greater than 24 hour free running period of the circadian pacemaker [42]. The present study points to a third potential source of asymmetry, namely, that during reentrainment, perturbations shift not only the phase but also the arrangement of the oscillators. Interpreting the results in this fashion, it may be useful to develop not only jet-lag reduction treatments that work to adjust the timing of the body's internal circadian pacemaker, but also strategies that attempt to restore the steady state arrangement of the oscillators once reentrainment has occurred. By understanding the time-dependent effect of zeitgebers such as light or melatonin on the arrangement of the oscillators in the body's circadian clock, more effective strategies could be developed to speed reentrainment and mitigate the effects of jet-lag. Previous experimental and detailed computational studies have hinted at such a hypothesis. For instance, [2], [22] observed that when SCN cells are first desynchronized before entrainment, they tend to reentrain much more rapidly than populations that were not initially desynchronized. The authors of [2] coined this phenomenon "phase tumbling"; in the context of our study, particularly in the results from section 5, we believe that phase tumbling can be precisely understood using the notion of an augmented phase reduction. Specifically, desynchronization modifies the system's isostable coordinates resulting in a change in its response to light perturbations in such a way that resynchronization to a new light schedule is hastened.

Phase reduction has been particularly useful in experimental systems where the full dynamical equations are unknown. For instance, in vitro control applications have been successfully applied to oscillatory neurons [49], [67]. Also, phase response curves have been calculated for melatonin [41] and light [32] in order to determine the phase dependent effect of perturbations to circadian oscillations. Because the direct method for calculation of an augmented phase reduction requires similar data to the direct method for the calculation of phase response curves, we believe that this strategy could be successfully applied experimentally. Furthermore, the direct method proposed here is not limited to systems of coupled oscillators, and could be applied to other oscillatory systems to determine how the phase response curve changes as the system is perturbed from the limit cycle.

While phase reduction is an invaluable tool in the study of nonlinear limit cycle oscillators, its assumptions break down as strong perturbations drive the system farther from its stable limit cycle. The augmented phase reduction described in this work provides a convenient means of addressing this fundamental limitation, and its application in both experimental and numerical settings can predict and explain behaviors that standard phase reduction is unable to address.

Appendix A. Computation of the iPRC, iIRC, and All Required Functions in the Augmented Phase Reduction. In the augmented phase reduction (16) for a general system (1), each of the functions $\mathbf{Z}(\theta) \equiv \frac{\partial \theta}{\partial \mathbf{x}}|_{\mathbf{x}^\gamma(\theta)}$, $\mathbf{B}^k(\theta) \equiv (\frac{\partial}{\partial \psi_k} \frac{\partial \theta}{\partial \mathbf{x}})|_{\mathbf{x}^\gamma(\theta)}$, $\mathbf{I}_k(\theta) \equiv \frac{\partial \psi_k}{\partial \mathbf{x}}|_{\mathbf{x}^\gamma(\theta)}$, and $\mathbf{C}_j^k(\theta) \equiv (\frac{\partial}{\partial \psi_k} \frac{\partial \psi_j}{\partial \mathbf{x}})|_{\mathbf{x}^\gamma(\theta)}$ needs to be calculated. Recall here

that θ is the phase coordinate, ψ_j is the j th isostable coordinate, and $\mathbf{x}^\gamma(\theta)$ is the location on the periodic orbit, γ , as a function of θ . When the system is small, as is the case for the theta model example in section 2, the most straightforward way to do this is by calculating phase and isostable coordinates over a discretized grid of initial conditions and interpolating to obtain the desired functions. However, for larger systems, such a method is not computationally feasible, as the number of interpolated values grows exponentially with the dimensionality of the system.

The most common way of calculating the functions $\mathbf{Z}$ and $\mathbf{I}_j$ is by using the adjoint equation [14], [7], the solution to which gives the necessary derivatives along the periodic orbit. To derive the adjoint equation, consider (1) with $\mathbf{P} \equiv 0$, and suppose we apply a small perturbation $\Delta\mathbf{x}$ at time $t = 0$ to any trajectory $\mathbf{x}(t)$ which is a solution to (1). For $\Delta\mathbf{x}$ small enough, its evolution is well approximated by the linearization

$$(A1) \quad \frac{d\Delta\mathbf{x}}{dt} = J(\mathbf{x}(t))\Delta\mathbf{x}(t) + \mathcal{O}(\|\Delta\mathbf{x}\|^2),$$

where J is the Jacobian of (1) evaluated at $\mathbf{x}(t)$. The corresponding phase shift $\Delta\theta \equiv \theta(\mathbf{x}(t) + \Delta\mathbf{x}(t)) - \theta(\mathbf{x}(t))$ and isostable shift $\Delta\psi_k \equiv \psi_k(\mathbf{x}(t) + \Delta\mathbf{x}(t)) - \psi_k(\mathbf{x}(t))$ can also be approximated by the linearization

$$(A2) \quad \Delta\Lambda = \nabla_{\mathbf{x}(t)}\Lambda \cdot \Delta\mathbf{x}(t) + \mathcal{O}(\|\Delta\mathbf{x}\|^2)$$

for $\Lambda = \theta$ and ψ_k , where $\nabla_{\mathbf{x}(t)}\Lambda$ is the gradient of Λ evaluated at $\mathbf{x}(t)$. As inspired by [7], by taking the time derivative of (A2) to lowest order in $\|\Delta\mathbf{x}\|$,

$$(A3) \quad \frac{d\Delta\Lambda}{dt} = \left\langle \frac{d\nabla_{\mathbf{x}(t)}\Lambda}{dt}, \Delta\mathbf{x}(t) \right\rangle + \left\langle \nabla_{\mathbf{x}(t)}\Lambda, \frac{d\Delta\mathbf{x}(t)}{dt} \right\rangle,$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product. Equation (A3) can be simplified by noting that in the absence of perturbation, $d\Delta\theta/dt = 0$ after the initial perturbation. Furthermore, because isostable coordinates converge exponentially as their associated trajectories approach the periodic orbit, $d\Delta\psi_k/dt = \kappa_k\Delta\psi_k = \kappa_k\nabla_{\mathbf{x}(t)}\psi_k \cdot \Delta\mathbf{x}$, where κ_k is defined as in (16).

With this information, starting with (A3), we can write

$$(A4) \quad \begin{aligned} \left\langle \frac{d\nabla_{\mathbf{x}(t)}\Lambda}{dt}, \Delta\mathbf{x}(t) \right\rangle &= - \left\langle \nabla_{\mathbf{x}(t)}\Lambda, \frac{d\Delta\mathbf{x}(t)}{dt} \right\rangle + \frac{d\Delta\Lambda}{dt} \\ &= - \left\langle \nabla_{\mathbf{x}(t)}\Lambda, J(\mathbf{x}(t))\Delta\mathbf{x}(t) \right\rangle + \left\langle \kappa_\Lambda \nabla_{\mathbf{x}(t)}\Lambda, \Delta\mathbf{x}(t) \right\rangle \\ &= - \left\langle J(\mathbf{x}(t))^T \nabla_{\mathbf{x}(t)}\Lambda - \kappa_\Lambda \nabla_{\mathbf{x}(t)}\Lambda, \Delta\mathbf{x}(t) \right\rangle. \end{aligned}$$

In the above equation, the transpose is the adjoint of the real-valued matrix $J(\mathbf{x}(t))$, $\kappa_\Lambda = 0$ for $\Lambda = \theta$ and $\kappa_\Lambda = \kappa_k$ for $\Lambda = \psi_k$. Equation (A4) is true for any perturbation $\Delta\mathbf{x}$, so that the adjoint equation follows immediately from (A4),

$$(A5) \quad \frac{d\nabla_{\mathbf{x}(t)}\Lambda}{dt} = (\kappa_\Lambda I - J(\mathbf{x}(t))^T) \nabla_{\mathbf{x}(t)}\Lambda,$$

where I is the identity matrix.

Equation (A5) can be used to solve for $\mathbf{Z}$ and $\mathbf{I}_j$ by taking $\mathbf{x}(t)$ to be the periodic orbit. For this choice, the periodic solution of (A5) with one additional boundary condition gives the desired functions. Specifically, when calculating $\mathbf{Z}$, we require

$\mathbf{Z}^T(\theta)\mathbf{F}(\mathbf{x}(\theta(t))) = \omega$ as a normalizing condition [14] (i.e., in the absence of perturbations, the phase increases at the natural frequency). When calculating $\mathbf{I}_k$, we require that when $\theta = 0$, $\langle \mathbf{I}_k^T(\theta), \mathbf{v}_k \rangle = 1$ as the normalizing condition [69], which follows directly from the definition of isostable coordinates (8) (cf. [69]).

A strategy to calculate the functions $\mathbf{B}^k(\theta)$ and $\mathbf{C}_j^k(\theta)$ was detailed in [66], whereby the Hessian matrix of second derivatives of the phase and isostable coordinates with respect to the periodic orbit were calculated using an equation similar to the adjoint equation (A5). Here, we give an alternative strategy which is computationally easier to implement in high-dimensional systems and uses the adjoint equation itself.

To do so, consider the trajectory

$$(A6) \quad \mathbf{x}_{\epsilon,k}(t) = \phi(t, \mathbf{x}_0 + \epsilon \mathbf{v}_k),$$

where ϕ represents the unperturbed flow of (1), $\mathbf{x}_0$ is the location for which $\theta = 0$ on the periodic orbit, $\mathbf{v}_k$ is the eigenvector which defines the k th isostable coordinate per the definition of isostables in (8), and ϵ is a small, positive constant. Notice that with the definition in (A6), $\mathbf{x}_{0,k}(t) = \mathbf{x}^\gamma(t)$, i.e., the stable periodic orbit.

Using the definition of isostable coordinates (8), the solution (A6) represents a trajectory which has been perturbed slightly from the periodic orbit to increase the ψ_k coordinate. Therefore, through Taylor expansion, one can show that to leading order in ψ_k ,

$$(A7) \quad \left. \frac{\partial \theta}{\partial \mathbf{x}} \right|_{\mathbf{x}_{\epsilon,k}(t)} - \left. \frac{\partial \theta}{\partial \mathbf{x}} \right|_{\mathbf{x}^\gamma(t)} = \left(\frac{\partial}{\partial \psi_k} \frac{\partial \theta}{\partial \mathbf{x}} \right) \Big|_{\mathbf{x}^\gamma(t)} \psi_k(\mathbf{x}_{\epsilon,k}(t)).$$

Recalling that in the absence of perturbations, $\dot{\psi}_k = \kappa_k \psi_k$, where κ_k is defined in (11), we know that $\psi_k(\mathbf{x}_{\epsilon,k}(t)) = \psi_k(\mathbf{x}_{\epsilon,k}(0)) \exp(\kappa_k t)$. Furthermore, based on the definition of isostables in (8), $\psi_k(\mathbf{x}_{\epsilon,k}(0)) = \epsilon$. Substituting this information into (A7) and rearranging, we have

$$(A8) \quad \left(\frac{\partial}{\partial \psi_k} \frac{\partial \theta}{\partial \mathbf{x}} \right) \Big|_{\mathbf{x}^\gamma(t)} = \frac{\exp(-\kappa_k t)}{\epsilon} \left(\left. \frac{\partial \theta}{\partial \mathbf{x}} \right|_{\mathbf{x}_{\epsilon,k}(t)} - \left. \frac{\partial \theta}{\partial \mathbf{x}} \right|_{\mathbf{x}^\gamma(t)} \right).$$

All that remains is to calculate the gradient of θ along the trajectory $\mathbf{x}_{\epsilon,k}(t)$. To do so, notice that in the right-hand side of (A7), $\psi_k(\mathbf{x}_{\epsilon,k}(T)) = \exp(\kappa_k T) \psi_k(\mathbf{x}_{\epsilon,k}(0))$. Therefore, $\left. \frac{\partial \theta}{\partial \mathbf{x}} \right|_{\mathbf{x}_{\epsilon,k}(t)}$ can be computed as the solution to

$$(A9) \quad \frac{d \nabla_{\mathbf{x}_{\epsilon,k}(t)} \theta}{dt} = -J(\mathbf{x}_{\epsilon,k}(t))^T \nabla_{\mathbf{x}_{\epsilon,k}(t)} \theta$$

that satisfies the boundary conditions

$$(A10) \quad \left. \frac{\partial \theta}{\partial \mathbf{x}} \right|_{\mathbf{x}_{\epsilon,k}(T)} - \left. \frac{\partial \theta}{\partial \mathbf{x}} \right|_{\mathbf{x}^\gamma(T)} = \exp(\kappa_k T) \left(\left. \frac{\partial \theta}{\partial \mathbf{x}} \right|_{\mathbf{x}_{\epsilon,k}(0)} - \left. \frac{\partial \theta}{\partial \mathbf{x}} \right|_{\mathbf{x}^\gamma(0)} \right).$$

The solution of (A9) which satisfies the above boundary conditions is relatively straightforward to calculate, for instance, with a Newton iteration. Once this solution is computed numerically, all terms in the right-hand side of (A8) are known (the left-hand side is $\mathbf{B}^k(\theta)$). This process can be used to solve for all functions $\mathbf{B}^k(\theta)$

necessary in the augmented phase reduction (16). This strategy outlined above is significantly more efficient than solving for the Hessian matrices of the phase and isostable coordinates, whose computational effort grows rapidly as the dimensionality of the system increases (cf. [66]). In principle, the equations $C_j^k(\theta)$ could also be found in an analogous manner. However, practically this strategy does not work well for the calculation of these functions for reasons not investigated here.

Appendix B. General Structure of the Eigenvectors of the Linearization to the Poincaré Map for a Population of Identical Oscillators. Consider a population of N identical oscillators with mean-field coupling from (18) in the main text. Also consider a Poincaré map of the form (6) and its corresponding linearization J_ϕ , which are used to define isostable coordinates. In order to determine an adequate reduction for the system (18) it is necessary to understand the structure of the eigenvectors of and multiplicity of the corresponding eigenvalues of J_ϕ . To this end, consider a small perturbation $\mathbf{x}_\epsilon = \mathbf{x}_0 + \Delta\mathbf{x}$. Through Taylor expansion of (18), to leading order we write

$$(B1) \quad \frac{d}{dt}\Delta\mathbf{x}(t) = \begin{bmatrix} A(t) + B(t) & B(t) & \dots & B(t) \\ B(t) & A(t) + B(t) & \dots & B(t) \\ \vdots & \vdots & \ddots & \vdots \\ B(t) & B(t) & \dots & A(t) + B(t) \end{bmatrix} \begin{bmatrix} \Delta\mathbf{x}_1 \\ \Delta\mathbf{x}_2 \\ \vdots \\ \Delta\mathbf{x}_N \end{bmatrix} + \mathcal{O}(\|\Delta\mathbf{x}\|^2),$$

where $A(t) \equiv (\frac{\partial F}{\partial \mathbf{x}} + \frac{\partial G}{\partial \mathbf{x}})|_{\gamma(t)}$, $B(t) \equiv \frac{1}{N} \frac{\partial G}{\partial \mathbf{x}}|_{\gamma(t)}$, and $\Delta\mathbf{x} \equiv [\Delta\mathbf{x}_1^T \dots \Delta\mathbf{x}_N^T]^T$. Here, we will show that the eigenvectors of J_ϕ have a special structure of the form

$$(B2) \quad \mathbf{v}_i^S = \begin{bmatrix} \boldsymbol{\mu}_i^S \\ \boldsymbol{\mu}_i^S \\ \vdots \\ \boldsymbol{\mu}_i^S \end{bmatrix} \in \mathbb{R}^{MN}, \quad \mathbf{v}_i^A = \begin{bmatrix} \boldsymbol{\mu}_i^A \\ -\boldsymbol{\mu}_i^A/(N-1) \\ \vdots \\ -\boldsymbol{\mu}_i^A/(N-1) \end{bmatrix} \in \mathbb{R}^{MN}, \quad 1 \leq i \leq M,$$

where $\boldsymbol{\mu}_i^S$ and $\boldsymbol{\mu}_i^A \in \mathbb{R}^M$. To begin, consider any initial condition to (B1) $\Delta\mathbf{x}$ of the form $\mathbf{v}_i^S$ or $\mathbf{v}_i^A$. Matrix multiplication and simplification of (B1) yields

$$(B3) \quad \begin{aligned} \dot{\boldsymbol{\mu}}_i^S &= (A(t) + NB(t))\boldsymbol{\mu}_i^S, \\ \dot{\boldsymbol{\mu}}_i^A &= A(t)\boldsymbol{\mu}_i^A. \end{aligned}$$

The solutions to (B3) are [25]

$$(B4) \quad \begin{aligned} \boldsymbol{\mu}_i^S(T) &= \Phi^S(T, 0)\boldsymbol{\mu}_i^S(0), \\ \boldsymbol{\mu}_i^A(T) &= \Phi^A(T, 0)\boldsymbol{\mu}_i^A(0), \end{aligned}$$

where $\Phi^S(T, 0)$ and $\Phi^A(T, 0)$ are the state transition matrices of their respective time-dependent linear systems. Furthermore, eigenvector solutions

$$(B5) \quad \begin{aligned} \Phi^S(T, 0)\boldsymbol{\mu}_i^S(0) &= \lambda_i\boldsymbol{\mu}_i^S(0), \\ \Phi^A(T, 0)\boldsymbol{\mu}_i^A(0) &= \lambda_i\boldsymbol{\mu}_i^A(0) \end{aligned}$$

must correspond to eigenvector solutions of $J_\phi\mathbf{v}_i^S = \lambda_i\mathbf{v}_i^S$ and $J_\phi\mathbf{v}_i^A = \lambda_i\mathbf{v}_i^A$.

One can show that $\Phi^S(T, 0)$ and $\Phi^A(T, 0)$ are both diagonalizable so that each matrix has M linearly independent eigenvalues. To see this, toward contradiction,

suppose either $\Phi^S(T, 0)$ or $\Phi^A(T, 0)$ is not diagonalizable. This implies that one of these matrices contains a Jordan block of size larger than 1, so that solutions of (B3) can admit solutions that decay as an exponential multiplied by a polynomial function of time [25]. However, this is not possible because J_ϕ is diagonalizable admitting no such solutions, and thus $\Phi^S(T, 0)$ and $\Phi^A(T, 0)$ are both diagonalizable. Recalling that $\Phi^S(T, 0)$ and $\Phi^A(T, 0) \in \mathbb{R}^{M \times M}$ and that the eigenvectors of $\Phi^S(T, 0)$ and $\Phi^A(T, 0)$ are related to eigenvectors of J_ϕ , this implies that there are M linearly independent eigenvectors of J_ϕ of the form $\mathbf{v}_i^A$, and an additional M linearly independent eigenvectors of the form $\mathbf{v}_i^S$. Furthermore, due to symmetry in (B1), for any μ_i^A from $1 \leq i \leq M$ (corresponding to the eigenvector $\mathbf{v}_i^A$ of J_ϕ), one can verify that the $N - 1$ columns of the matrix

$$(B6) \quad E(\mu_i^A) = \begin{bmatrix} \mu_i & -\mu_i/(N-1) & & -\mu_i/(N-1) \\ -\mu_i/(N-1) & \mu_i & & -\mu_i/(N-1) \\ -\mu_i/(N-1) & -\mu_i/(N-1) & & -\mu_i/(N-1) \\ \vdots & \vdots & \dots & \vdots \\ -\mu_i/(N-1) & -\mu_i/(N-1) & & \mu_i \\ -\mu_i/(N-1) & -\mu_i/(N-1) & & -\mu_i/(N-1) \end{bmatrix}$$

with $E(\mu_i^A) \in \mathbb{R}^{MN \times (N-1)}$ are also linearly independent eigenvectors of J_ϕ , each corresponding to the same eigenvalue. Linear independence can be shown by considering a subset of the rows of $E(\mu_i^A)$, specifically the k th, $M + k$ th, $2M + k$ th, $\dots$, and $(N-1)M + k$ th rows for $k < M + 1$, to create a matrix of the form

$$(B7) \quad E_{\text{sub}}^{(k)}(\mu_i^A) = \mu_{i,k}^A \begin{bmatrix} 1 & -1/(N-1) & \dots & -1/(N-1) \\ -1/(N-1) & 1 & \dots & -1/(N-1) \\ \vdots & \vdots & \ddots & \vdots \\ -1/(N-1) & -1/(N-1) & \dots & 1 \\ -1/(N-1) & -1/(N-1) & \dots & -1/(N-1) \end{bmatrix},$$

where $\mu_{i,k}^A$ is the k th element of μ_i^A and $E_{\text{sub}}^{(k)}(\mu_i^A) \in \mathbb{R}^{N \times (N-1)}$. By adding a column of ones to the end of $E_{\text{sub}}^{(k)}(\mu_i^A)$ and using elementary row operations on the resulting matrix, an upper triangular matrix can be obtained with nonzero elements on the diagonal, thus implying $E_{\text{sub}}^{(k)}(\mu_i^A)$ and consequently that $E(\mu_i^A)$ has full rank.

To summarize, for the M eigenvalue and eigenvector pairs (λ_i, μ_i^S) of $\Phi^S(T, 0)$, there is exactly one corresponding eigenvalue and eigenvector pair $(\lambda_i, \mathbf{v}_i^S)$ of J_ϕ . Furthermore, for each of the M eigenvalue and eigenvector pairs (λ_i, μ_i^A) of $\Phi^A(T, 0)$, there are exactly $N - 1$ linearly independent eigenvectors of J_ϕ formed by the columns of $E(\mu_i^A)$ with corresponding eigenvalue λ_i . This accounts for all NM eigenvectors of J_ϕ . We will refer to the eigenvectors which comprise the columns of $E(\mu_i^A)$ as a set of asynchronous eigenvectors, and the eigenvectors of the form $\mathbf{v}_i^S$ will be referred to as synchronous eigenvectors.

It will be important to note that all eigenvectors which comprise the columns of $E(\mu_i^A)$ are orthogonal to all eigenvectors of the form $\mathbf{v}_i^S$. Also, the columns of $E(\mu_i^A)$ sum to zero. Furthermore, because the periodic orbit of (18) is asymptotically stable, exactly one eigenvalue of J_ϕ must be equal to one. Since all $\mathbf{v}_i^A$ are repeated, this eigenvector must be of the form $\mathbf{v}_i^S$. For all other eigenvectors of J_ϕ , the corresponding eigenvalues must have an absolute value of less than 1.

Appendix C. An Approximation to the Augmented Phase Reduction in the Limit of Small Forcing. In previous work [66] an augmented isostable reduction of the form (16) was derived. Here, we will consider the equation

$$(C1) \quad \begin{aligned} \dot{\theta}_\epsilon &= \omega + \mathbf{Z}^T(\theta_\epsilon) \mathbf{P}(t) + \sum_{k=1}^{n-1} [\mathbf{B}^{kT}(\theta_\epsilon) \psi_{k,\epsilon}] \mathbf{P}(t), \\ \dot{\psi}_{j,\epsilon} &= \kappa_j \psi_{j,\epsilon} + \mathbf{I}_j^T(\theta_\epsilon) \mathbf{P}(t), \end{aligned}$$

which is identical to (16) with all functions $\mathbf{C}_j^k = 0$. We will show that provided $\mathbf{P}(t) = \mathcal{O}(\epsilon)$ uniformly in time and ϵ is sufficiently small in relation to the minimum of $|\kappa_j|$ over all j , then when comparing the solutions of (C1) and (16) with the same initial conditions, $\theta(t) - \theta_\epsilon(t) \sim \mathcal{O}(\epsilon^3)$ for $t \sim \mathcal{O}(1/\epsilon)$.

To begin, let $\mathbf{x}$ be some initial condition on the periodic orbit (i.e., with $\psi_j(0) = 0$ for all j). By taking the absolute value of each term of $\dot{\psi}_j$ for any j from (16), one can show

$$(C2) \quad |\dot{\psi}_j| \leq \kappa_j |\psi_j| + \max_{\theta,t} |\mathbf{I}_j^T(\theta) \mathbf{P}(t)| + (n-1) \max_{\theta,t,k} |\mathbf{C}_j^{kT}(\theta) \mathbf{P}(t)| \max_j |\psi_j|.$$

Equation (C2) uses the fact that $\kappa_j < 0$ for all j . Setting the left-hand side of (C2) to zero implies

$$(C3) \quad -\kappa_j |\psi_j| \leq \max_{\theta,t} |\mathbf{I}_j^T(\theta) \mathbf{P}(t)| + (n-1) \max_{\theta,t,k} |\mathbf{C}_j^{kT}(\theta) \mathbf{P}(t)| \max_j |\psi_j|$$

for $t \geq 0$. Equation (C3) is valid for any isostable coordinate so that

$$(C4) \quad -\kappa_\xi \max_j |\psi_j| \leq \max_{\theta,t,j} |\mathbf{I}_j^T(\theta) \mathbf{P}(t)| + (n-1) \max_{\theta,t,k,j} |\mathbf{C}_j^{kT}(\theta) \mathbf{P}(t)| \max_j |\psi_j|,$$

where $\xi = \arg \max_j |\psi_j|$. Provided $\min_j [-\kappa_j]$ is large relative to ϵ (recall that $\mathbf{P}(t) = \mathcal{O}(\epsilon)$), it follows that $(\min_j [-\kappa_j] - (n-1) \max_{\theta,t,k,j} |\mathbf{C}_j^{kT}(\theta) \mathbf{P}(t)|) > 0$ and the following bound is obtained:

$$(C5) \quad \max_j |\psi_j| \leq \frac{\max_{\theta,t,j} |\mathbf{I}_j^T(\theta) \mathbf{P}(t)|}{\min_j [-\kappa_j] - (n-1) \max_{\theta,t,k,j} |\mathbf{C}_j^{kT}(\theta) \mathbf{P}(t)|}.$$

Furthermore, provided ϵ is small enough so that the denominator of (C5) is order 1, the above equation implies that $\max_j |\psi_j| = \mathcal{O}(\epsilon)$.

Next, let $\Delta\theta \equiv \theta_\epsilon - \theta$ and $\Delta\psi_j \equiv \psi_{j,\epsilon} - \psi_k$, i.e., the difference between the solutions of (C1) and (16). Using these definitions, one can derive

$$(C6) \quad \begin{aligned} \Delta\dot{\theta} &= \mathbf{Z}'^T(\theta) \mathbf{P}(t) \Delta\theta + \sum_{k=1}^{n-1} \left[\mathbf{B}^{k'T}(\theta) \Delta\theta \psi_k + \mathbf{B}^{kT}(\theta) \Delta\psi_k \right] \mathbf{P}(t) \\ &\quad + \mathcal{O}(\Delta\theta^2) + \sum_{k=1}^{n-1} \mathcal{O}(\Delta\psi_k^2), \end{aligned}$$

$$(C7) \quad \Delta\dot{\psi}_j = \kappa_j \Delta\psi_j + \mathbf{I}_j^T(\theta) \mathbf{P}(t) \Delta\theta - \sum_{k=1}^{n-1} [\mathbf{C}_j^{kT}(\theta) \psi_k] \mathbf{P}(t) + \mathcal{O}(\Delta\theta^2).$$

Equations (C6) and (C7) are derived by subtracting (16) from (C1) and Taylor expanding the result, with $' \equiv \partial/\partial\theta$. Using the above equations, we can bound the growth of $\Delta\theta$ as follows:

$$\begin{aligned}
 \frac{d}{dt}|\Delta\theta| &\leq \underbrace{\max_{\theta,t} \left| \mathbf{Z}'^T(\theta) \mathbf{P}(t) \right| + \max_{\theta,t} \left| \sum_{k=1}^{n-1} \left[\mathbf{B}^{k'}{}^T(\theta) \right] \mathbf{P}(t) \right|}_{K_1} \max_{k,t} |\psi_k| |\Delta\theta| \\
 &+ \underbrace{\max_{\theta,t} \left| \sum_{k=1}^{n-1} \left[\mathbf{B}^{kT}(\theta) \right] \mathbf{P}(t) \right|}_{K_2} \max_{k,t} |\Delta\psi_k| \\
 &= K_1 |\Delta\theta| + K_2.
 \end{aligned}
 \tag{C8}$$

Recall from (C5) that $\max_{k,t} |\psi_k| = \mathcal{O}(\epsilon)$, which implies that $\max_{k,t} |\Delta\psi_k|$ can be no larger than $\mathcal{O}(\epsilon)$. Therefore, since $\mathbf{P}(t) = \mathcal{O}(\epsilon)$ uniformly in time, K_1 is no larger than $\mathcal{O}(\epsilon)$ and K_2 is no larger than $\mathcal{O}(\epsilon^2)$. Using (C8), this allows us to write

$$\begin{aligned}
 |\Delta\theta(t)| &\leq \frac{K_2}{K_1} (e^{K_1 t} - 1) \\
 &= K_2 t + \mathcal{O}((K_2 t)^2).
 \end{aligned}
 \tag{C9}$$

Therefore, on the time scale of $t \sim 1/\epsilon$, the above equation implies $\Delta\theta(t)$ is no larger than $\mathcal{O}(\epsilon^2)$. Now, using (C7), we can also bound the growth of $\Delta\psi_j$ as

$$\begin{aligned}
 \frac{d}{dt}|\Delta\psi_j| &\leq \kappa_j |\Delta\psi_j| + \max_{\theta,t} \left| \mathbf{I}'^T(\theta) \mathbf{P}(t) \right| \max_t |\Delta\theta| + \max_{\theta,t} \left| \sum_{k=1}^{n-1} \left[\mathbf{C}_j^{kT}(\theta) \right] \mathbf{P}(t) \right| \max_k |\psi_k| \\
 &\leq \underbrace{\max_{\theta,t} \left| \mathbf{I}'^T(\theta) \mathbf{P}(t) \right| \max_t |\Delta\theta|}_{K_3} + \underbrace{\max_{\theta,t} \left| \sum_{k=1}^{n-1} \left[\mathbf{C}_j^{kT}(\theta) \right] \mathbf{P}(t) \right| \max_{k,t} |\psi_k|}_{K_4},
 \end{aligned}
 \tag{C10}$$

where the second line is obtained by noting that $\kappa_j < 0$. Using (C9), on the time scale of $t \sim 1/\epsilon$, $\max_t |\Delta\theta|$ is at most $\mathcal{O}(\epsilon^2)$. Therefore on this time scale, K_3 is at most $\mathcal{O}(\epsilon^2)$. Furthermore, since $\max_k |\psi_k|$ is at most $\mathcal{O}(\epsilon)$, K_4 is at most $\mathcal{O}(\epsilon^2)$. Using (C10) on the time scale of $t \sim 1/\epsilon$, $\max_k |\Delta\psi_k|$ can be no larger than $\mathcal{O}(\epsilon^2)$. We can now use this revised bound to show that K_2 from (C8) is no larger than $\mathcal{O}(\epsilon^3)$ on $t \sim 1/\epsilon$ which, using (C9) implies that

$$|\Delta\theta(t)| \leq \mathcal{O}(\epsilon^3) \text{ for } t \sim 1/\epsilon.
 \tag{C11}$$

To summarize, for any initial condition that is sufficiently close to the periodic orbit γ , for perturbations of $\mathcal{O}(\epsilon)$ uniformly in time and small enough relative to the minimum of $|\kappa_j|$ over all j , the difference $\theta(t) - \theta_\epsilon(t)$ between solutions of (16) and (C1) remain $\mathcal{O}(\epsilon^3)$ on the time scale of $t \sim 1/\epsilon$. Because we are often only practically interested in the phase of a system, we primarily use (C1) for applications in this work.

REFERENCES

- [1] M. AINSWORTH, S. LEE, M. O. CUNNINGHAM, R. D. TRAUB, N. J. KOPELL, AND M. A. WHITTINGTON, *Rates and rhythms: A synergistic view of frequency and temporal coding in neuronal networks*, *Neuron*, 75 (2012), pp. 572–583. (Cited on p. 278)
- [2] S. AN, R. HARANG, K. MEEKER, D. GRANADOS FUENTES, C. A. TSAI, C. MAZUSKI, J. KIM, F. J. DOYLE, L. R. PETZOLD, AND E. D. HERZOG, *A neuropeptide speeds circadian entrainment by reducing intercellular synchrony*, *Proc. Natl. Acad. Sci. USA*, 110 (2013), pp. E4355–E4361. (Cited on p. 306)
- [3] J. ASCHOFF, K. HOFFMANN, H. POHL, AND R. WEVER, *Re-entrainment of circadian rhythms after phase-shifts of the Zeitgeber*, *Chronobiologia*, 2 (1974), pp. 23–78. (Cited on p. 306)
- [4] R. AZODI-AVVAL AND A. GHARABAGHI, *Phase-dependent modulation as a novel approach for therapeutic brain stimulation*, *Frontiers Comput. Neurosci.*, 9 (2015), art. 26. (Cited on pp. 279, 298)
- [5] M. BONNIN, *Amplitude and phase dynamics of noisy oscillators*, *Internat. J. Circuit Theory Appl.*, 45 (2017), pp. 636–659. (Cited on pp. 279, 285)
- [6] 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, <https://doi.org/10.1137/17M1155235>. (Cited on p. 285)
- [7] 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. (Cited on pp. 285, 307)
- [8] G. BUZSÁKI AND A. DRAGUHN, *Neuronal oscillations in cortical networks*, *Science*, 304 (2004), pp. 1926–1929. (Cited on p. 278)
- [9] O. CASTEJÓN, A. GUILLAMON, AND G. HUGUET, *Phase-amplitude response functions for transient-state stimuli*, *J. Math. Neurosci.*, 3 (2013), art. 13. (Cited on p. 279)
- [10] C. A. CZEISLER, J. F. DUFFY, T. L. SHANAHAN, E. N. BROWN, J. F. MITCHELL, D. W. RIMMER, J. M. RONDA, E. J. SILVA, J. S. ALLAN, J. S. EMENS, D. J. DIJK, AND R. E. KRONAUER, *Stability, precision, and near-24-hour period of the human circadian pacemaker*, *Science*, 284 (1999), pp. 2177–2181. (Cited on p. 302)
- [11] M. DETRIXHE, M. DOUBECK, J. MOEHLIS, AND F. GIBOU, *A fast Eulerian approach for computation of global isochrons in high dimensions*, *SIAM J. Appl. Dyn. Syst.*, 15 (2016), pp. 1501–1527, <https://doi.org/10.1137/140998615>. (Cited on p. 279)
- [12] P. EMERY, W. V. SO, M. KANEKO, J. C. HALL, AND M. ROSBASH, *CRY, a drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity*, *Cell*, 95 (1998), pp. 669–679. (Cited on p. 279)
- [13] G. B. ERMENTROUT AND N. KOPELL, *Parabolic bursting in an excitable system coupled with a slow oscillation*, *SIAM J. Appl. Math.*, 46 (1986), pp. 233–253, <https://doi.org/10.1137/0146017>. (Cited on p. 280)
- [14] G. B. ERMENTROUT AND N. KOPELL, *Multiple pulse interactions and averaging in systems of coupled neural oscillators*, *J. Math. Biol.*, 29 (1991), pp. 195–217. (Cited on pp. 285, 307, 308)
- [15] G. B. ERMENTROUT AND D. H. TERMAN, *Mathematical Foundations of Neuroscience*, *Interdiscip. Appl. Math.* 35, Springer, New York, 2010. (Cited on p. 278)
- [16] D. A. GOLOMBEK AND R. E. ROSENSTEIN, *Physiology of circadian entrainment*, *Physiol. Rev.*, 90 (2010), pp. 1063–1102. (Cited on pp. 278, 294)
- [17] D. GONZE, S. BERNARD, C. WALTERMANN, A. KRAMER, AND H. HERZEL, *Spontaneous synchronization of coupled circadian oscillators*, *Biophys. J.*, 89 (2005), pp. 120–129. (Cited on pp. 294, 302)
- [18] 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. (Cited on p. 285)
- [19] R. GRIMSHAW, *Nonlinear Ordinary Differential Equations*, *Appl. Math. Engrg. Sci. Texts* 2, CRC Press, Boca Raton, FL, 1993. (Cited on p. 284)
- [20] J. GUCKENHEIMER, *Isochrons and phaseless sets*, *J. Math. Biol.*, 1 (1975), pp. 259–273. (Cited on pp. 278, 280)
- [21] J. GUCKENHEIMER AND P. HOLMES, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, *Appl. Math. Sci.* 42, Springer-Verlag, New York, 1983. (Cited on p. 282)
- [22] M. L. GUERRIERO, A. POKHILKO, A. P. FERNÁNDEZ, K. J. HALLIDAY, A. J. MILLAR, AND J. HILLSTON, *Stochastic properties of the plant circadian clock*, *J. Roy. Soc. Interface*, 9 (2012), pp. 744–756. (Cited on p. 306)

- [23] A. GUILLAMON AND G. HUGUET, *A computational and geometric approach to phase resetting curves and surfaces*, SIAM J. Appl. Dyn. Syst., 8 (2009), pp. 1005–1042, <https://doi.org/10.1137/080737666>. (Cited on p. 279)
- [24] K. S. HENRY AND M. G. HEINZ, *Diminished temporal coding with sensorineural hearing loss emerges in background noise*, Nature Neurosci., 15 (2012), pp. 1362–1364. (Cited on p. 278)
- [25] J. P. HESPAÑHA, *Linear Systems Theory*, Princeton University Press, Princeton, NJ, 2009. (Cited on pp. 309, 310)
- [26] A. B. HOLT AND T. I. NETOFF, *Origins and suppression of oscillations in a computational model of Parkinson's disease*, J. Comput. Neurosci., 37 (2014), pp. 505–521. (Cited on p. 279)
- [27] A. B. HOLT, D. WILSON, M. SHINN, J. MOEHLIS, AND T. I. NETOFF, *Phasic burst stimulation: A closed-loop approach to tuning deep brain stimulation parameters for Parkinson's disease*, PLoS Comput. Biol., 12 (2016), art. e1005011. (Cited on p. 279)
- [28] S. HONMA, T. SHIRAKAWA, Y. KATSUNO, M. NAMIHIRA, AND K. I. HONMA, *Circadian periods of single suprachiasmatic neurons in rats*, Neurosci. Lett., 250 (1998), pp. 157–160. (Cited on p. 294)
- [29] E. M. IZHKEVICH, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*, MIT Press, London, 2007. (Cited on pp. 278, 280, 298, 305)
- [30] D. JORDAN AND P. SMITH, *Nonlinear Ordinary Differential Equations: An Introduction for Scientists and Engineers*, Oxford University Press, Oxford, 2007. (Cited on pp. 279, 284)
- [31] Y. KAWAMURA, H. NAKAO, K. ARAI, H. KORI, AND Y. KURAMOTO, *Collective phase sensitivity*, Phys. Rev. Lett., 101 (2008), art. 024101. (Cited on p. 279)
- [32] S. B. S. KHALSA, M. E. JEWETT, C. CAJOCHEN, AND C. A. CZEISLER, *A phase response curve to single bright light pulses in human subjects*, J. Physiol., 549 (2003), pp. 945–952. (Cited on p. 306)
- [33] T. W. KO AND G. B. ERMENTROUT, *Phase-response curves of coupled oscillators*, Phys. Rev. E, 79 (2009), art. 016211. (Cited on p. 279)
- [34] H. KOEPL, M. HAFNER, A. GANGULY, AND A. MEHROTRA, *Deterministic characterization of phase noise in biomolecular oscillators*, Phys. Biol., 8 (2011), art. 055008. (Cited on pp. 279, 285)
- [35] K. KOTANI, I. YAMAGUCHI, L. YOSHIDA, Y. JIMBO, AND G. B. ERMENTROUT, *Population dynamics of the modified theta model: Macroscopic phase reduction and bifurcation analysis link microscopic neuronal interactions to macroscopic gamma oscillation*, J. Roy. Soc. Interface, 11 (2014), art. 20140058. (Cited on p. 279)
- [36] Y. KURAMOTO, *Chemical Oscillations, Waves, and Turbulence*, Springer-Verlag, Berlin, 1984. (Cited on p. 278)
- [37] W. KUREBAYASHI, S. SHIRASAKA, AND H. NAKAO, *Phase reduction method for strongly perturbed limit cycle oscillators*, Phys. Rev. Lett., 111 (2013), art. 214101. (Cited on p. 279)
- [38] P. LAKATOS, G. KARLOS, A. D. MEHTA, I. ULBERT, AND C. E. SCHROEDER, *Entrainment of neuronal oscillations as a mechanism of attentional selection*, Science, 320 (2008), pp. 110–113. (Cited on p. 278)
- [39] J. C. LELOUP AND A. GOLDBETER, *Toward a detailed computational model for the mammalian circadian clock*, Proc. Natl. Acad. Sci. USA, 100 (2003), pp. 7051–7056. (Cited on p. 306)
- [40] Z. LEVNAJIC AND A. PIKOVSKY, *Phase resetting of collective rhythm in ensembles of oscillators*, Phys. Rev. E, 82 (2010), art. 056202. (Cited on p. 279)
- [41] A. J. LEWY, S. AHMED, J. M. L. JACKSON, AND R. L. SACK, *Melatonin shifts human circadian rhythms according to a phase-response curve*, Chronobiol. Internat., 9 (1992), pp. 380–392. (Cited on p. 306)
- [42] Z. LU, K. KLEIN CARDEÑA, S. LEE, T. M. ANTONSEN, M. GIRVAN, AND E. OTT, *Resynchronization of circadian oscillators and the east-west asymmetry of jet-lag*, Chaos, 26 (2016), art. 094811. (Cited on p. 306)
- [43] A. MAUROY, I. MEZIC, AND J. MOEHLIS, *Isostables, isochrons, and Koopman spectrum for the action-angle representation of stable fixed point dynamics*, Phys. D, 261 (2013), pp. 19–30. (Cited on pp. 280, 283)
- [44] C. C. MCINTYRE, W. M. GRILL, D. L. SHERMAN, AND N. V. THAKOR, *Cellular effects of deep brain stimulation: Model-based analysis of activation and inhibition*, J. Neurophysiol., 91 (2004), pp. 1457–1469. (Cited on p. 287)
- [45] J. P. MCKENNA, R. DHUMPA, N. MUKHITOV, M. G. ROPER, AND R. BERTRAM, *Glucose oscillations can activate an endogenous oscillator in pancreatic islets*, PLoS Comput. Biol., 12 (2016), art. e1005143. (Cited on p. 278)
- [46] S. MCKENNOCH, T. VOEGTLIN, AND L. BUSHNELL, *Spike-timing error backpropagation in theta neuron networks*, Neural Comput., 21 (2009), pp. 9–45. (Cited on p. 280)
- [47] R. Y. MOORE, J. C. SPEH, AND R. K. LEAK, *Suprachiasmatic nucleus organization*, Cell and Tissue Res., 309 (2002), pp. 89–98. (Cited on p. 294)

- [48] A. NABI, M. MIRZADEH, F. GIBOU, AND J. MOEHLIS, *Minimum energy desynchronizing control for coupled neurons*, J. Comput. Neurosci., 34 (2013), pp. 259–271. (Cited on p. 279)
- [49] A. NABI, T. STIGEN, J. MOEHLIS, AND T. NETOFF, *Minimum energy control for in vitro neurons*, J. Neural Engrg., 10 (2013), art. 036005. (Cited on p. 306)
- [50] T. NETOFF, M. A. SCHWEMMER, AND T. J. LEWIS, *Experimentally estimating phase response curves of neurons: Theoretical and practical issues*, in Phase Response Curves in Neuroscience, Springer, 2012, pp. 95–129. (Cited on pp. 280, 298, 305)
- [51] A. N. NICHOLSON, M. B. SPENCER, P. A. PASCOE, B. M. STONE, T. ROEHRS, AND T. ROTH, *Sleep after transmeridian flights*, The Lancet, 328 (1986), pp. 1205–1208. (Cited on pp. 279, 306)
- [52] H. M. OSINGA AND J. MOEHLIS, *Continuation-based computation of global isochrons*, SIAM J. Appl. Dyn. Syst., 9 (2010), pp. 1201–1228, <https://doi.org/10.1137/090777244>. (Cited on pp. 279, 281)
- [53] Y. PARK AND B. ERMENTROUT, *Weakly coupled oscillators in a slowly varying world*, J. Comput. Neurosci., 40 (2016), pp. 269–281. (Cited on p. 279)
- [54] M. G. PEDERSEN, E. MOSEKILDE, K. S. POLONSKY, AND D. S. LUCIANI, *Complex patterns of metabolic and Ca^{2+} entrainment in pancreatic islets by oscillatory glucose*, Biophys. J., 105 (2013), pp. 29–39. (Cited on p. 278)
- [55] K. PYRAGAS AND V. NOVIČENKO, *Phase reduction of a limit cycle oscillator perturbed by a strong amplitude-modulated high-frequency force*, Phys. Rev. E, 92 (2015), art. 012910. (Cited on p. 279)
- [56] S. M. REPPERT AND D. R. WEAVER, *Coordination of circadian timing in mammals*, Nature, 418 (2002), pp. 935–941. (Cited on p. 294)
- [57] R. L. SACK, D. AUCKLEY, R. R. AUGER, M. A. CARSKADON, K. P. WRIGHT, M. V. VITIELLO, AND I. V. ZHADONOVA, *Circadian rhythm sleep disorders: Part I, basic principles, shift work and jet lag disorders*, Sleep, 30 (2007), pp. 1460–1483. (Cited on pp. 305, 306)
- [58] S. SHIRASAKA, W. KUREBAYASHI, AND H. NAKAO, *Phase-amplitude reduction of transient dynamics far from attractors for limit-cycling systems*, Chaos, 27 (2017), art. 023119. (Cited on p. 279)
- [59] D. TAKESHITA AND R. FERES, *Higher order approximation of isochrons*, Nonlinearity, 23 (2010), pp. 1303–1323. (Cited on p. 305)
- [60] 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), art. 194102. (Cited on p. 285)
- [61] T. L. TO, M. A. HENSON, E. D. HERZOG, AND F. J. DOYLE, *A molecular model for intercellular synchronization in the mammalian circadian clock*, Biophys. J., 92 (2007), pp. 3792–3803. (Cited on p. 306)
- [62] J. WATERHOUSE, T. REILLY, G. ATKINSON, AND B. EDWARDS, *Jet lag: Trends and coping strategies*, The Lancet, 369 (2007), pp. 1117–1129. (Cited on pp. 279, 306)
- [63] D. K. WELSH, D. E. LOGOTHETIS, M. MEISTER, AND S. M. REPPERT, *Individual neurons dissociated from rat suprachiasmatic nucleus express independently phased circadian firing rhythms*, Neuron, 14 (1995), pp. 697–706. (Cited on p. 294)
- [64] S. WIGGINS, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Texts Appl. Math. 2, Springer, 2003. (Cited on pp. 282, 297)
- [65] D. WILSON AND B. ERMENTROUT, *An operational definition of phase characterizes the transient response of perturbed limit cycle oscillators*, SIAM J. Appl. Dyn. Syst., 17 (2018), pp. 2516–2543, <https://doi.org/10.1137/17M1153261>. (Cited on p. 301)
- [66] D. WILSON AND B. ERMENTROUT, *Greater accuracy and broadened applicability of phase reduction using isostable coordinates*, J. Math. Biol., 76 (2018), pp. 37–66. (Cited on pp. 280, 282, 283, 284, 285, 287, 305, 308, 309, 311)
- [67] D. WILSON, A. B. HOLT, T. I. NETOFF, AND J. MOEHLIS, *Optimal entrainment of heterogeneous noisy neurons*, Frontiers in Neuroscience, 9 (2015), art. 192. (Cited on p. 306)
- [68] D. WILSON AND J. MOEHLIS, *Determining individual phase response curves from aggregate population data*, Phys. Rev. E, 92 (2015), art. 022902. (Cited on p. 279)
- [69] D. WILSON AND J. MOEHLIS, *Isostable reduction of periodic orbits*, Phys. Rev. E, 94 (2016), art. 052213. (Cited on pp. 279, 280, 282, 283, 284, 290, 291, 308)
- [70] A. WINFREE, *The Geometry of Biological Time*, 2nd ed., Springer-Verlag, New York, 2001. (Cited on pp. 278, 280, 294)
- [71] J. M. ZEITZER, D. J. DIJK, R. E. KRONAUER, E. N. BROWN, AND C. A. CZEISLER, *Sensitivity of the human circadian pacemaker to nocturnal light: Melatonin phase resetting and suppression*, J. Physiol., 526 (2000), pp. 695–702. (Cited on pp. 279, 298)