SIAM J. Applied Dynamical Systems
Vol. 22, No. 2, pp. 603–634

A Reduced Order Modeling Framework for Strongly Perturbed Nonlinear Dynamical Systems Near Arbitrary Trajectory Sets*

Dan Wilson†

Abstract. A reduced order modeling strategy is proposed that can accurately capture the behavior of strongly perturbed nonlinear dynamical systems, i.e., those that are subject to general, large magnitude inputs. In contrast to standard variational approaches which consider dynamics in the neighborhood of a single reference trajectory, the proposed methodology augments the dynamics with an additional variable that continuously selects from a family of reference trajectories in order to limit truncation errors that result from neglected nonlinear terms. Provided the reference trajectories contract sufficiently rapidly in some directions, a reduced order set of equations can be obtained by choosing an appropriate coordinate system. Direct numerical approaches for computation of the necessary terms of the associated reduced order equations are provided. Crucially, because the proposed reduction strategy does not require the existence of a persistent, stable fixed point or periodic orbit, it can be implemented in situations where external inputs cause the dynamics to transition through a bifurcation. Two examples with relevance to neural control are provided. In the first, the proposed reduction strategy allows for the formulation of a numerically tractable control problem to identify energy-optimal inputs that eliminate tonic firing in a neural model. In the second example, the proposed reduction framework accurately captures successive transitions between quiescence and tonic firing across the boundary of a saddle-node on invariant circle bifurcation.

Key words. model order reduction, Koopman operator, nonlinear dynamics, neuroscience, bifurcation, optimal control

MSC codes. 34E10, 37Cxx, 92Bxx, 34Hxx

DOI. 10.1137/21M1451154

1. Introduction.

Reduced order modeling techniques for nonlinear dynamical systems have seen a surge of interest in recent years. Much of this interest has been spurred by the tremendous promise of Koopman analysis, which can be used to represent a fully nonlinear dynamical system as a linear but possibly infinite-dimensional operator [7], [37], [38]. While the identification of a finite basis to represent an infinite-dimensional Koopman operator is generally a difficult task, strategies including dynamical mode decomposition (DMD) [51], [26], extended DMD [58], and deep learning approaches [32], [72] have been shown to be useful in various applications.

For nonlinear dynamical systems with states in the basin of attraction of a stable periodic orbit, the decay rates of eigenfunctions of the associated Koopman operator are directly related to Floquet exponents of the linearized system [38], [27]. Provided that most

*Received by the editors October 7, 2021; accepted for publication (in revised form) by K. Josic September 13, 2022; published electronically May 17, 2023.
https://doi.org/10.1137/21M1451154
Funding: This work was supported by National Science Foundation grant CMMI-2140527.
†Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996 USA (dwilso81@utk.edu.)

of these eigenfunctions decay sufficiently rapidly, a substantial reduction of dimensionality is possible. For instance, a well-established technique known as phase reduction [71], [24], [13] can be used to represent the dynamics of an N-dimensional system in terms of a 1-dimensional phase that encodes for oscillation timing. In situations where some of the Floquet eigenfunctions have slower decay rates, related phase-amplitude reduction approaches can be employed [17], [56], [29], [53], [66], [4] that also consider the nonnegligable dynamics of slowly decaying eigenmodes. In a similar manner, nonlinear dynamical systems in the basin of attraction of a stable fixed point can be analyzed using a reduced order basis of slowly decaying Koopman eigenmodes using the isostable coordinate framework [34], [33], [61], [65].

For dynamical systems in regimes that are not bound to a periodic or fixed point attractor, it can be difficult to apply general Koopman-based techniques to infer a low order basis to accurately represent the nonlinear dynamics. For such systems, other tools may be more appropriate. For instance, fast-slow analysis [19], [3], dynamical averaging [53], and contraction theory [30], [54] can be useful when dynamics that evolve on different timescales can be isolated. Techniques that use inertial manifolds [10], [15] or invariant submanifolds [24], [43], [1] can be useful in some situations, but such manifolds are not always easily identifiable. Other strategies have been proposed to represent the behavior along individual trajectories in a reduced coordinate basis [68], [70], [9], [49]; however, the associated reduction techniques suffer when the state is perturbed far from the reference trajectory. Alternative techniques have been developed that use a locally orthogonal moving coordinate frame to represent dynamical behaviors in reference to (and potentially far from) a base manifold [28], [29], [56]; however, when using these approaches it is not generally straightforward to subsequently identify a reduced order system of equations. Classical variational techniques can be used to approximate dynamical behavior in the vicinity of a reference solution by means of series expansion [22]; however, this method is only generally valid in a close neighborhood of the reference trajectory precluding implementation when large perturbations are required.

In this work, a reduced order modeling framework is proposed that can be used to capture the behavior of a strongly perturbed, nonlinear dynamical system in the vicinity of a continuous family of reference trajectories. Here, the term "strongly perturbed" is used to emphasize that the applied inputs are arbitrary and potentially of large magnitude. This is in direct contrast to methodologies that are valid in the "weakly perturbed" limit [71], [47], [13], which place restrictive limitations on the magnitudes of the allowable inputs. The proposed methodology augments the dynamics with an additional variable that continuously selects from this family of reference trajectories in order to limit truncation errors that result from neglected nonlinear terms. This reduction framework can generally be implemented provided that local linearizations of the individual trajectories contract sufficiently rapidly in some directions. Using this framework it is possible to overcome limitations of related reduction strategies [65], [64]; specifically, it allows for the consideration of applications for which external inputs cause the dynamics to transition through a bifurcation.

The organization of this paper is as follows: section 2 provides relevant background on recently proposed, Koopman-based reduction frameworks that have been used to represent the dynamics near a continuous collection of fixed points [65] or periodic orbits [64]. Section 3 provides a derivation of the proposed adaptive trajectory reduction framework and

discusses strategies for direct numerical computation of the associated reduced order equations. Section 4 provides two examples that implement the proposed reduction framework in the context of neural control applications. In the first application, the adaptive trajectory reduction framework allows for the formulation of a tractable optimal control problem to identify optimal inputs that transition a tonically firing neuron to a quiescent state. In the second application, direct numerical comparisons are given between the proposed adaptive trajectory reduction strategy and a recently proposed adaptive phase reduction framework from [64]. Section 5 provides concluding remarks and discusses opportunities for future extension.

2. Background on related phase-isostable-based adaptive reduction methods. Consider a general ordinary differential equation

$$\dot{x} = F(x, p_0) + U(t), \tag{1}$$

where $x \in \mathbb{R}^N$ is the state, $p = p_0 \in \mathbb{R}$ is a constant nominal parameter, F represents the unperturbed dynamics, and $U(t)$ is an external input. Previous work on adaptive reduction techniques has been limited to dynamical systems with a family of either stable periodic orbits [64] or stable fixed point attractors [65]. Here, a brief background of this strategy for systems with periodic orbits is provided. Supposing that $x^{\Omega}(t,p)$ is a stable periodic orbit of (1) that results when taking p to be a constant, the notion of isochrons [16], [71] can be used to define a phase coordinate that encodes for the timing of oscillations. Likewise, a set of the most slowly decaying isostable coordinates (i.e., level sets of the slowest decaying Koopman modes [34], [38]) can be used to encode for the transient decay of solutions towards the stable limit cycle. Transforming to phase and isostable coordinates, to a linear approximation, the dynamics near the periodic orbit can often be represented with a reduced order set of equations [69], [60]

$$\begin{aligned} \dot{\theta} &= \omega(p) + Z^T(\theta, p)U(t), \\ \dot{\psi}_j &= \kappa_j(p)\psi_j + I_j^T(\theta, p)U(t) \end{aligned} \tag{2}$$

for $j = 1, \ldots, \beta < N$, where state is related to the reduced coordinates according to

$$x(\theta, \psi_1, \ldots, \psi_\beta) = x^{\gamma}(\theta, p) + \sum_{k=1}^{\beta} g^k(\theta, p)\psi_k. \tag{3}$$

Here, θ is the phase (i.e., the isochron) associated with a given state, andψ_j is the jth isostable coordinate with associated Floquet exponent $\kappa_j(p)$ ordered so that $|\mathrm{Re}(\kappa_{\mathrm{j}})| \leq |\mathrm{Re}(\kappa_{\mathrm{j+1}})|$. The terms $Z(\theta,p)$ and $I_j(\theta,p)$ are the phase and isostable response curves, respectively, which capture how the effective input influences the reduced order coordinates, and g^k is a Floquet eigenfunction associated with the kth isostable coordinate. Rapidly decaying isostable coordinates (i.e., those with large magnitude, negative Floquet exponents) are truncated so that only $\beta \leq N-1$ of the slowest decaying isostable coordinates remain. Note here that the value of p is assumed to be constant. The interested reader is referred to [69], [66], [40], [38] for a more detailed background on isostable coordinates and their connection with the Koopman operator for dynamical systems with periodic oscillations.

As part of (2), the terms $Z(\theta,p)$ and $I_j(\theta,p)$ represent the gradient of the phase and isostable coordinates, respectively, evaluated on the periodic orbit. As such, when the state travels too far from the periodic orbit, the accuracy of (2) is degraded. One approach to mitigating this issue is to compute the gradients of the phase and isostable coordinates to higher orders of accuracy [66], [60], [9], [8]. These strategies can be used to extend the applicability of the phase-isostable reduction to include larger magnitude inputs, but they still usually break down when exceedingly large magnitude inputs are applied. Alternatively, one can consider a family of periodic orbits that emerge in (1) as the parameter p is changed. This approach was investigated in [25], [44] for use in situations where the input can be decomposed into the sum of a slowly varying component and another component that has sufficiently small magnitude. These ideas were augmented in [64] using the isostable coordinate paradigm to allow for more general inputs. Letting P be an allowable parameter set and considering a continuous set of stable periodic orbits $x^{\Omega}(t,p)$ that emerge when taking the parameter $p \in P$, one can rewrite (1) as

$$\dot{x} = F(x,p) + U_e(t,p,x), \tag{4}$$

where $U_e(t,p,x) = F(x,p_0) - F(x,p) + U(t)$. Notice that (4) is identical to (1) with the addition of a new variable p and with the difference $F(x,p_0) - F(x,p)$ being absorbed into the external forcing term. As shown in [64], by allowing the value of p to change over time, a reduced order representation of (4) can be obtained that accurately captures the behavior of the underlying system (1)

$$\begin{aligned} \dot{\theta} &= \omega(p) + Z(\theta,p)\cdot U_e(t,p,x) + D(\theta,p)\cdot \dot{p}, \\ \dot{\psi}_j &= \kappa_j(p)\psi_j + I_j(\theta,p)\cdot U_e(t,p,x) + Q_j(\theta,p)\cdot \dot{p}, \\ j &= 1,\ldots,\beta, \\ \dot{p} &= G_p(p,\theta,\psi_1,\ldots,\psi_\beta), \end{aligned} \tag{5}$$

where the system state can be approximated according to (3). Above, $Z(\theta,p)$ and $I_j(\theta,p)$ are the phase and isostable response curves associated with the periodic orbit $x^{\Omega}(t,p)$, $D(\theta,p)$ and each $Q_j(\theta,p)$ capture the change in the phase and isostable coordinates, respectively, in response to shifts in the adaptive parameter p, and G_p is a function that must be designed so that each ψ_j remains an $O(\epsilon)$ term, where $0 < \epsilon \ll 1$. In the present manuscript, reduced order models of the form (5) will be referred to as an adaptive phase reduction (APR). Intuitively, to implement the adaptive phase reduced equations (5) the function G_p must be designed to update the adaptive parameter p so that the state of the full system stays close to $x^{\gamma}(t,p)$; a general form G_p that often works well in practice is

$$G_p(p,\theta,\psi_1,\ldots,\psi_\beta) = -\alpha \sum_{i=1}^{\beta} \psi_i Q_i(\theta,p), \tag{6}$$

where α is a positive constant. Other formulations are also possible. The interested reader is referred to [64] for a more detailed background on the adaptive reduction strategy. The APR framework has been successfully used in control applications that require large magnitude

inputs [63] and has also been adapted in situations with fixed point attractors [65], [62] instead of periodic orbit attractors.

While the APR framework has been useful for analyzing the behavior of dynamical systems with oscillatory dynamics and those with stable fixed points, one of the key assumptions required for its implementation is that in the allowable range of p, each $\psi_j(x,p)$ is continuously differentiable with respect to both p and x. This condition is usually violated when the underlying attractor changes. As such, critical points of bifurcations are generally excluded from the allowable parameter sets when using the APR and other isostable-based reduction strategies.

3. Theoretical foundation of the adaptive trajectory reduction framework.

3.1. Technical approach. The goal of this work is to identify an adaptive reduction strategy that considers dynamical behavior on a low-dimensional manifold defined by a collection of reference trajectories. In direct contrast to the previously developed APR from (5), the reduction framework detailed here does not require the underlying ordinary differential equation (1) to have periodic orbits. As such, the reduction framework detailed here will be referred to as an adaptive trajectory reduction (ATR). To this end, consider a general dynamical system of the form (1). Letting P be some allowable parameter set for which $p_0 \in P$, consider an associated set of unperturbed reference trajectories, $x^\gamma(t,p)$, that are the solution to the ordinary differential equation

$$\begin{aligned} \dot{x} &= F(x,p), \\ x(0,p) &= \eta(p) \end{aligned} \tag{7}$$

on the interval $t=[0,T]$, where $\eta(p)$ gives the initial condition. Standard linearization techniques yield variational equations that can be used to represent the behavior in a close neighborhood of any single reference trajectory, $x^\gamma(t,p)$, according to

$$\frac{d\Delta x}{dt} = J(x^\gamma(t,p),p)\Delta x + O(||\Delta x||^2), \tag{8}$$

where $\Delta x = x(t,p_0) - x^\gamma(t,p)$ and $J(x^\gamma(t,p),p)$ is the Jacobian of F evaluated at $x^\gamma(t,p)$ and p. Provided Δx is small, truncation errors in (8) can be ignored. Practically, to guarantee that Δx is small, $x(0,p_0)$ must be close to $x^\gamma(0,p)$ and both $(p-p_0)$ and $U(t)$ must be sufficiently small. Asymptotic expansions that are valid to higher orders of accuracy can also be considered to accommodate larger magnitude inputs, but such techniques still generally require the state to remain in a close neighborhood of the reference trajectory. To circumvent this issue, the parameter p will be considered nonstatic, ultimately allowing for the truncation errors in the associated variational equations to be able to be actively managed. At the expense of adding an extra variable, the resulting equations can accommodate arbitrarily large inputs. Additionally, provided the reference trajectories contract sufficiently rapidly in some directions a reduced order set of equations can be obtained by choosing an appropriate moving coordinate system (illustrated in Figure 1) as detailed in the derivation to follow. Preliminary assumptions used in the derivation of this reduced order modeling strategy are listed below.

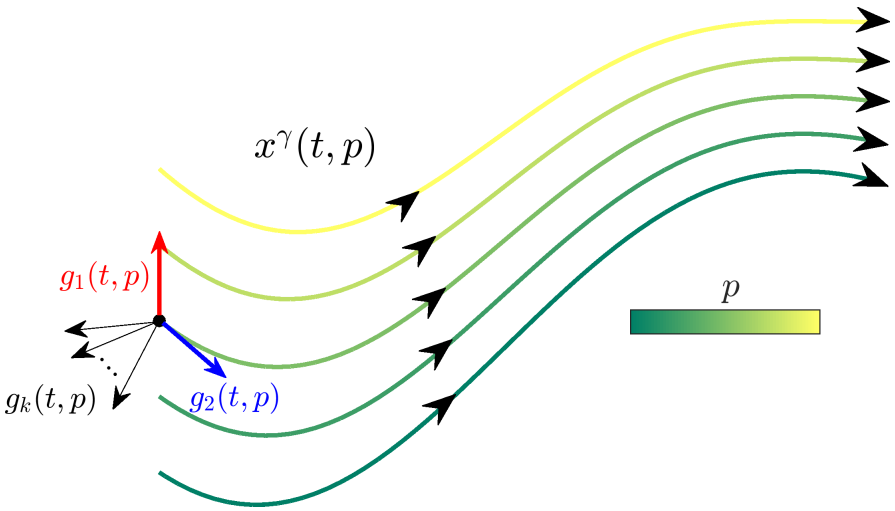

Figure 1. *Illustration of the proposed local coordinate system. For a continuous collection of reference trajectories $x^\gamma(t,p)$, in the vicinity of each trajectory it is assumed that the dynamics can be represented according to a local linearization of the form (8) and that all but two of the eigenvectors of the state transition matrix $\Phi_p(T,0)$ are small in magnitude so that these components of the solution decay rapidly. To capture the behavior of the remaining components of the solution, the coordinate basis defined by (12)–(14) will be used. For a given value of p, the basis element $g_1(t,p)$ points towards $\partial x^\gamma/\partial p$ and $g_2(t,p)$ points in the direction of the unperturbed flow.*

3.2. Sufficient assumptions to implement the proposed adaptive trajectory reduction (ATR) framework.

Assumption A. The function $F(x,p)$ is $\boldsymbol{C}^2$ differentiable.

Assumption A requires that (1) is sufficiently smooth so that the Jacobian, Hessian, and other second order partial derivatives exist and are continuous.

Assumption B. The reference trajectories $x^\gamma(t,p)$ are chosen so that both $\frac{\partial x^\gamma}{\partial p}$ and $\frac{\partial^2 x^\gamma}{\partial p^2}$ exist for all t and for all $p \in P$.

Like the APR strategy, a shadow system will be considered of the form

$$\dot{x} = F(x,p) + U_e(t,p,x), \tag{9}$$

where

$$U_e(t,p,x) = F(x,p_0) - F(x,p) + U(t), \tag{10}$$

and p is an adjustable parameter. Let $\Delta x(t,p) \equiv x - x^\gamma(t,p)$ represent a perturbed solution of (9), i.e., using the trajectory $x^\gamma(t,p)$ as a reference. When both $U_e \equiv 0$ and p is static, the dynamics of the variational equation are given by (8) to linear order. Let $\Phi_p(t,t_0)$ be the state transition matrix of the linearized system (8), i.e., with the property that $\Delta x(t,p) = \Phi_p(t,t_0)\Delta x(t_0,p)$. Let $v_1(p),\ldots,v_N(p)$ and $\lambda_1(p),\ldots\lambda_N(p)$ denote the eigenvectors and associated eigenvalues, respectively, of $\Phi_p(T,0)$ for some reference value of T chosen to be identical for all p. By convention, eigenvalues will be sorted so that $\max_p|\lambda_j(p)| > \max_p|\lambda_{j+1}(p)|$. Letting $\kappa_j(p) = \log(\lambda_j(p))/T$, the following will be assumed.

Assumption C. For $j > 2$, $\mathrm{Re}(\kappa_j(p)) < 0$ for all p and $\max_p(\mathrm{Re}(\kappa_j)) = O(1/\epsilon)$, where $0 < \epsilon \ll 1$.

Assumption C requires that for $j > 2$ the solutions of (8) associated with $\lambda_j(p)$ decay rapidly. As will be shown in the derivation to follow, this rapid decay ultimately allows for a reduction of the system order.

Assumption D. For $j > 2$, each $\lambda_j(p)$ is a simple eigenvalue for all p.

An additional constraint will be placed on the magnitude of $U_e(t,p,x)$ in the shadow system from (9).

Assumption E. For all time and allowable p, $||U_e(t,p,x)||_1 \leq M_U$, where $M_U = O(1)$ and $||\cdot||_1$ denotes the 1-norm.

Note that the inputs $U_e(t,p,x)$ and $U(t)$ are generally allowed to be large and are not constrained to be $O(\epsilon)$ terms. Towards stating a final assumption, let the matrix $A(t,p) \in \mathbb{R}^{N\times N}$ be defined such that the first column of $A(t,p)$ equals $\frac{\partial x^\gamma}{\partial p}$ evaluated at both t and p, the second column equals $F(x^\gamma(t,p),p)$, and the kth column is given by $\Phi_p(t,0)v_k(p)\exp(-\kappa_k t)$ for $k \geq 3$.

Assumption F. For any $p \in P$, the matrix inverse $A(t,p)^{-1}$ exists for all time and $||A(t,p)^{-1}||_1 = O(1)$.

Assumptions A–F represent a sufficient set of conditions in order to implement the proposed reduction framework. In section 5, modifications will be discussed that could be used to implement the proposed reduction framework in situations where some of these assumptions are not satisfied.

3.3. Derivation of the ATR equations. Consider a general dynamical system of the form (1) and an associated continuous set of reference trajectories as shown in Figure 1. Suppose that all assumptions from section 3.2 are satisfied. For the moment, solutions of (9) will be considered using a static choice of p and taking $U_e = 0$. Local linearization of these trajectories yields (8). The system state will be represented using the moving coordinate system

$$x(t) = x^\gamma(t,p) + \sum_{k=1}^{N} s_k(t,p)g_k(t,p), \tag{11}$$

where

$$g_1(t,p) = \left.\frac{\partial x^\gamma}{\partial p}\right|_{t,p}, \tag{12}$$

$$g_2(t,p) = F(x^\gamma(t,p),p), \tag{13}$$

$$g_k(t,p) = \exp(-\kappa_k(p)t)\Phi_p(t,0)v_k(p), \quad k = 3,\ldots,N. \tag{14}$$

Above, $\kappa_k(p) = \log(\lambda_k(p))/T$ and $s_1,\ldots,s_N$ represent coordinates in the moving basis. Intuitively, this moving coordinate system is defined so that g_1 points towards adjacent trajectories, g_2 points in the direction of the unperturbed flow, and each g_k for $k \geq 3$ captures a part of the solution that is rapidly decaying. Note that $g_k(T,p) = \frac{1}{\lambda_k(p)}\Phi_p(T,0)v_k(p) = v_k(p) = O(1)$. Drawing on Assumption F above, by taking

$$A(t,p) = \begin{bmatrix} g_1(t,p) & \ldots & g_N(t,p) \end{bmatrix}, \tag{15}$$

the inverse

$$A(t,p)^{-1} \equiv \begin{bmatrix} w_1(t,p) & \dots & w_N(t,p) \end{bmatrix}^T \tag{16}$$

exists where $w_1^T(t,p)$ are row vectors that comprise A^{-1} with ^{T} denoting the vector transpose. With these definitions the transformed coordinates are

$$s_j(t,p) = w_j^T(t,p)(x - x^\gamma(t,p)) \tag{17}$$

for all j. Towards defining a reduced order coordinate system, a time-like variable θ is introduced with the dynamics that follow:

$$\dot{\theta} = \begin{cases} 1 & \text{if } U_e = 0, \\ F_1(x,p,U_e) & \text{otherwise} \end{cases} \tag{18}$$

is introduced, where F_1 will be determined momentarily. Note that incorporating a time-like variable is a common analysis strategy for nonautonomous dynamical systems, particularly in conjunction with Poincaré maps [57]. Additionally, unlike in the phase-based reduced order modeling frameworks given by (2) and (5), θ in the ATR framework is not a direct function of the state. Given that each $x^\gamma(t,p)$, $g_k(t,p)$, and $w_k(t,p)$ is defined subject to $U_e = 0$, θ will be substituted for t in these expressions. In the derivation to follow, p will be allowed to change in time. To proceed, consider some initial condition $x(t=0)$ of (1) with corresponding values of θ and p chosen so that $s_k = O(\epsilon)$ for all k at $t = 0$. Intuitively, this initial condition must be close to the set of reference trajectories, i.e., so that $\min_{\theta,p} |x^\gamma(\theta,p) - x(0)| = O(\epsilon)$. Taking the time derivative of (17) yields

$$\begin{aligned} \dot{s}_k &= \dot{w}_k^T(x - x^\gamma) + w^T(\dot{x} - \dot{x}^\gamma) \\ &= \dot{w}_k^T \sum_{k=1}^N (s_k g_k) + w_k^T \left(F(x,p) + U_e(t,p,x) - \frac{\partial x^\gamma}{\partial \theta}\dot{\theta} - \frac{\partial x^\gamma}{\partial p}\dot{p} \right), \end{aligned} \tag{19}$$

where the second line is obtained by substituting both (9) and by taking the time derivative of $x^\gamma(\theta,p)$ using the chain rule. Here, all functions and partial derivatives are evaluated at θ and p. Continuing to manipulate (19), first by substituting (11) into $F(x,p)$ and Taylor expanding, yields

$$\begin{aligned} \dot{s}_k &= \dot{w}_k^T \sum_{k=1}^N (s_k g_k) + w_k^T \left(F(x^\gamma,p) + J(\sum_{k=1}^N s_k g_k) + U_e(t,p,x) - \frac{\partial x^\gamma}{\partial \theta}\dot{\theta} - \frac{\partial x^\gamma}{\partial p}\dot{p} \right) + O(\epsilon^2) \\ &= \dot{w}_k^T \sum_{k=1}^N (s_k g_k) + w_k^T \sum_{k=1}^N (s_k \dot{g}_k) \\ &\quad + w_k^T \left[F(x^\gamma,p) - \frac{\partial F}{\partial p} s_1 - \sum_{k=1}^N \left(s_k \frac{\partial g_k}{\partial p}\dot{p} \right) \right. \\ &\quad \left. + \sum_{k=3}^N (s_k \kappa_k g_k) + U_e(t,p,x) - \frac{\partial x^\gamma}{\partial \theta}\dot{\theta} - \frac{\partial x^\gamma}{\partial p}\dot{p} \right] + O(\epsilon^2) \end{aligned}$$

$$
\begin{aligned}
&= w_k^T \Bigg[F(x^\gamma, p) - \frac{\partial F}{\partial p} s_1 - \sum_{k=1}^{N} \left(s_k \frac{\partial g_k}{\partial p} \dot{p} \right) \\
&\quad + \sum_{k=3}^{N} (s_k \kappa_k g_k) + U_e(t,p,x) - \frac{\partial x^\gamma}{\partial \theta} \dot{\theta} - \frac{\partial x^\gamma}{\partial p} \dot{p} \Bigg] + O(\epsilon^2), \\
&= w_k^T \Bigg[g_2(\theta, p) - \frac{\partial F}{\partial p} s_1 - \sum_{k=1}^{N} \left(s_k \frac{\partial g_k}{\partial p} \dot{p} \right) \\
&\quad + \sum_{k=3}^{N} (s_k \kappa_k g_k) + U_e(t,p,x) - g_2(\theta,p)\dot{\theta} - g_1(\theta,p)\dot{p} \Bigg] + O(\epsilon^2),
\end{aligned} \tag{20}
$$

where the time derivatives $\dot{g}_k$ considered in the second line follow from the relations from (A4), the fourth line follows from the simplification $\dot{w}_j^T g_k + w_j^T \dot{g}_k = 0$ from (A5), and the final line follows from (12) and (13). Noting that $w_i^T g_j = 0$ for $j \neq i$ and 1 for $j = i$, evaluating (20) for specific coordinates of s_k at time $t = 0$ yields

$$
\begin{aligned}
\dot{s}_1 &= w_1^T(\theta,p) U_e(t,p,x) - \dot{p} + O(\epsilon), \\
\dot{s}_2 &= 1 + w_2^T(\theta,p) U_e(t,p,x) - \dot{\theta} + O(\epsilon), \\
\dot{s}_k &= \kappa_k(p) s_k + w_k^T(\theta,p) U_e(t,p,x) + O(\epsilon), \\
&\qquad k = 3, \ldots, N,
\end{aligned} \tag{21}
$$

where the additional $O(\epsilon)$ terms emerge because $s_k(t) = O(\epsilon)$ for all k at $t = 0$. The time interval on which each $s_k(t)$ remains an $O(\epsilon)$ term in (21) will be discussed momentarily. The full system can be approximated using (11) yielding

$$
x(t) = x^\gamma(\theta, p) + O(\epsilon). \tag{22}
$$

Letting θ and p evolve in time according to

$$
\begin{aligned}
\dot{\theta} &= 1 + w_2^T(\theta,p) U_e(t,p,\theta), \\
\dot{p} &= w_1^T(\theta,p) U_e(t,p,\theta),
\end{aligned} \tag{23}
$$

$\dot{s}_1$ and $\dot{s}_2$ become $O(\epsilon)$ terms. As such, with this choice for the θ and p dynamics, $s_1(t)$ and $s_2(t)$ remain $O(\epsilon)$ for $t \sim 1/\epsilon$. Additionally, in the above equation each $\mathrm{Re}(\kappa_k(p)) = O(1/\epsilon)$ for $k \geq 3$. Because of this, as shown in Appendix B, $s_k(t) = O(\epsilon)$ for $k \geq 3$ for all time. As such, the dynamics of each of the coordinates $s_1, \ldots, s_N$ can be neglected. Note that while $s_3, \ldots, s_N$ can generally take complex values, the modulus of the associated eigenvalues $\lambda_3, \ldots, \lambda_N$ is small because $\mathrm{Re}(\kappa_j) = O(1/\epsilon)$ for $k \geq 3$ by Assumption C. As such, there are no issues caused by resonances when eigenvalues are complex. Above, since x can be written as a function of θ and p, U_e can be written as a function of t, p, and θ, thereby making (23) self-contained. Note that $d\theta/dt$ in (23) has the same structure as mandated by (18).

A slight relaxation of Assumption F. For any choice of θ and p, Assumption F from section 3.2 ensures that the reduced order coordinates defined according to (17) exist and ensures that the neglected terms from (21) are of $O(\epsilon)$. However, some combinations of θ

and p may not be relevant for a given application. Assumption F can be relaxed slightly by focusing on specific solutions of (23) instead. In other words, any solution of (23) with associated state estimation (22) will be an accurate representation of the full order solution (1) provided that $\theta(t)$ and $p(t)$ remain bounded away from regions for which $w_1(\theta,p)$ and $w_2(\theta,p)$ are infinite. This point will be illustrated in the examples from section 4.

3.4. Numerical computation of the terms of the ATR equations. Equations (22) and (23) provide the foundation of the proposed ATR strategy. In order to implement this strategy, it is necessary to compute both $w_1(\theta,p)$ and $w_2(\theta,p)$. Of course, these can technically be computed by first computing $A(t,p)$ as defined in (15) (perhaps using jet transport [46] or other variational approaches) and then taking the matrix inverse as part of (16). However, accurate numerical evaluation of (16) is not generally feasible when the state transition matrix $\Phi_p(T,0)$ has near-zero eigenvalues, $\lambda_k(p)$, that subsequently make $A(t,p)$ ill-conditioned. As such, an alternative strategy detailed below must be used. This strategy is similar to the adjoint method for calculating phase response curves [6] and isostable response curves [69] for phase-amplitude reduced equations of the form (2).

To begin, first consider an initial condition $x = x^\gamma(\theta_1,p_1)$ with dynamics that follow (9). Here θ_1 and p_1 represent initial values of the phase and adaptive parameter, respectively. Let Δx represent an $O(\epsilon)$, arbitrary perturbation to that trajectory. Following this initial perturbation, suppose that $U_e(t,p,x)=0$ so that $\dot\theta = 1$. The evolution of $\Delta x(t)$ follows

$$\frac{d\Delta x}{dt} = J\Delta x + \frac{\partial F}{\partial p}\Delta p + O(\epsilon^2), \tag{24}$$

where all partial derivatives are evaluated at $x^\gamma(\theta_1+t,p_1)$. Here, Δp represents some constant shift in the adaptive parameter caused by the perturbation Δx. In the context of the reduced order equations (23), one can write

$$\Delta p = w_1^T(\theta_1,p_1)\Delta x, \tag{25}$$

so that (24) becomes

$$\frac{d\Delta x}{dt} = \left(J + \frac{\partial F}{\partial p}w_1^T\right)\Delta x. \tag{26}$$

With this in mind, consider the shift Δp from (25). Taking the time derivative yields

$$\frac{d\Delta p}{dt} = \frac{d\Delta x}{dt}^T w_1 + \Delta x^T \frac{dw_1}{dt}. \tag{27}$$

After the initial perturbation, the shift in p between the perturbed and unperturbed trajectories is zero since $U_e = 0$ (in other words, $d\Delta p/dt = 0$). Substituting this and (26) into (27) yields

$$0 = \Delta x^T\left(J^T w_1 + w_1\frac{\partial F}{\partial p}^T w_1 + \frac{dw_1}{dt}\right). \tag{28}$$

Because (28) is valid for any arbitrary Δx, one finds

$$\frac{dw_1}{dt} = -\left(J^T + w_1\frac{\partial F}{\partial p}^T\right)w_1, \tag{29}$$

where all functions and partial derivatives are evaluated at $\theta = \theta_1 + t$, $p = p_1$, and $x^\gamma(\theta_1 + t, p_1)$. Similarly, by considering the shift in the variable θ resulting from the initial perturbation Δx,

$$\Delta\theta = w_2^T(\theta_1, p_1)\Delta x, \tag{30}$$

and taking the time derivative of both sides (noting that $d\Delta\theta/dt = 0$ after the initial perturbation since $U_e = 0$), mirroring the arguments used to derive (27)–(29) yields

$$\frac{dw_2}{dt} = -\left(J^T + w_1 \frac{\partial F}{\partial p}^T\right) w_2, \tag{31}$$

where once again all functions and partial derivatives are evaluated at $\theta = \theta_1 + t$, $p = p_1$, and $x^\gamma(\theta_1 + t, p_1)$.

The derivation above yields (29) and (31), which can be used to evaluate w_1 and w_2 local to the reference trajectory $x^\gamma(\theta, p_1)$. As a matter of practical implementation, initial conditions for $w_1(T, p_1)$ and $w_2(T, p_1)$ can be obtained from (16). Subsequently, $w_1(\theta, p_1)$ and $w_2(\theta, p_1)$ can be obtained for all values of θ by evaluating (29) and (31) in backwards time. The necessity of this backwards time integration stems from the fact that for a given value of p_1 some solutions of the form $\Delta\dot{x} = J\Delta x$, by assumption, are rapidly contracting. As such, numerical solutions of equations that have the form $\Delta\dot{x} = -J^T\Delta x$ will have errors that are significantly amplified when evaluated in forward time. Solving such equations in backward time ameliorates this issue. Finally, w_1 and w_2 can be computed for different reference trajectories allowing the necessary terms $w_1(\theta, p)$ and $w_2(\theta, p)$ from (23) to be obtained using linear interpolation.

3.5. Explicit solutions for the terms of the ATR equations. As shown here, with knowledge of $w_1(0, p)$ and $w_2(0, p)$, one can write both $w_1(t, p)$ and $w_2(t, p)$ explicitly as functions of the state transition matrix $\Phi_p(t, t_0)$ and the terms $g_1(t, p)$ and $g_2(t, p)$ from (12) and (13). To begin, let

$$\begin{aligned} y(t, p) &= \Phi_p(t, 0)^T w_1(0, p), \\ z(t, p) &= \Phi_p(t, 0)^T w_2(0, p). \end{aligned} \tag{32}$$

Note here that $y(t)$ and $z(t)$ are solutions of the adjoint system of (8) so that

$$\begin{aligned} \frac{dy}{dt} &= -J^T(x, p)y, \\ \frac{dz}{dt} &= -J^T(x, p)z. \end{aligned} \tag{33}$$

With this in mind, one can show that w_1 and w_2 can be written explicitly as

$$w_1(t, p) = \frac{y(t, p)}{g_1^T(t, p)y(t, p)}, \tag{34}$$

$$w_2(t, p) = z(t, p) - \frac{g_1^T(t, p)z(t, p)}{g_1^T(t, p)y(t, p)} y(t, p). \tag{35}$$

To see that (34) is the solution to (29), first notice that when $t = 0$, the right-hand side of (34) simplifies to $w_1(0,p)$. Taking the time derivative of (34) yields

$$\begin{aligned}\dot{w}_1 &= \frac{1}{g_1^T y}\dot{y} - \left(\frac{\dot{g}_1^T y + g_1^T \dot{y}}{(g_1^T y)^2}\right) y \\ &= -\frac{1}{g_1^T y} J^T y - \frac{1}{(g_1^T y)^2}\left(\left(g_1^T J^T + \frac{\partial F}{\partial p}^T\right) y - g_1^T J^T y\right) y \\ &= -\frac{1}{g_1^T y} J^T y - \frac{1}{(g_1^T y)^2}\left(\frac{\partial F}{\partial p}^T y\right) y \\ &= -\left(J^T + w_1 \frac{\partial F}{\partial p}^T\right) w_1, \end{aligned} \tag{36}$$

which is identical to (29). The second line of (36) above is obtained, recalling that Δp is constant, by substituting in (A2) and (33), and the last line is obtained by substituting (35) and rearranging. Likewise, considering (35), the right-hand side simplifies to $w_2(0,p)$ when $t = 0$. To verify that (35) satisfies (31), taking the time derivative of (35) yields

$$\begin{aligned}\dot{w}_2 &= \dot{z} - \frac{g_1^T z}{g_1^T y}\dot{y} - \frac{1}{(g_1^T y)^2}((\dot{g}_1^T z + g_1^T \dot{z}) g_1^T y - g_1^T z(\dot{g}_1^T y + g_1^T \dot{y}))y \\ &= -J^T z + \frac{g_1^T z}{g_1^T y} J^T y \\ &\quad - \frac{1}{(g_1^T y)^2}\left[\left(\left(g_1^T J^T + \frac{\partial F}{\partial p}^T\right) z - g_1^T J^T z\right) g_1^T y - g_1^T z\left(\left(g_1^T J^T + \frac{\partial F}{\partial p}^T\right) y - g_1^T J^T y\right)\right] y, \end{aligned} \tag{37}$$

where the second line is obtained by substituting derivatives from (A2) and (33). Continuing to simplify, emphasizing that many of the terms of (37) are scalars for which the multiplication order can be rearranged as necessary, yields

$$\begin{aligned}\dot{w}_2 &= -J^T\left(z - \frac{g_1^T z}{g_1^T y} y\right) - \frac{1}{(g_1^T y)^2}\left[\frac{\partial F}{\partial p}^T z g_1^T y - g_1^T z \frac{\partial F}{\partial p}^T y\right] y \\ &= -J^T w_2 - \left[\frac{\partial F}{\partial p}^T z - \frac{g_1^T z}{g_1^T y}\frac{\partial F}{\partial p}^T y\right] w_1 \\ &= -J^T w_2 - \frac{\partial F}{\partial p}^T w_2 w_1 \\ &= -\left(J^T + w_1 \frac{\partial F}{\partial p}^T\right) w_2, \end{aligned} \tag{38}$$

which is indeed identical to (31). In most cases, it is easier to numerically compute $w_1(t,p)$ and $w_2(t,p)$ using (34) and (35) instead of solving (29) and (31) directly. Here, $y(t,p)$ and $z(t,p)$ can be obtained by solving equations (33) in backward time, and $g_1(t,p)$ can be found by solving (A2) in forward time. Equations (34) and (35) also emphasize the fact that both w_1 and w_2 tend towards infinity if there exists some t for which $g_1^T(t,p)y(t,p) = 0$.

3.6. An alternative reduced coordinate system for implementing the ATR strategy. The moving coordinate system defined by (11)–(14) allows the N-dimensional system (1) to be represented using only 2 dimensions. The implementation of this coordinate transformation, however, requires $A(t,p)^{-1}$ as defined in (16) to exist at all times (Assumption F). As discussed in the previous section, this assumption can be relaxed slightly by requiring θ and p in the reduced order equations (23) to be sufficiently bounded away from regions for which $w_1(\theta,p)$ and $w_2(\theta,p)$ are infinite. When this is not possible, an alternative reduced order coordinate system can be used that results in a reduced order system with 3 dimensions. This formulation is detailed below.

To begin, consider a moving coordinate system similar to (11) defined in reference to a continuous set of unperturbed reference trajectories of a shadow system (9) as shown in Figure 1. This alternative coordinate system is defined below for a given choice of p.

$$\begin{aligned} g_1(t,p) &= \Phi_p(t,0)\frac{\partial x^\gamma}{\partial p}\bigg|_{t=0}, \\ g_2(t,p) &= F(x^\gamma(t,p),p), \\ g_k(t,p) &= \exp(-\kappa_k t)\Phi_p(t,0)v_k(p), \\ &\qquad k=3,\dots,N. \end{aligned} \tag{39}$$

The above basis is identical to the one from (12)–(14), except for the change in the definition of g_1. Once again, the matrices $A(t,p)$ and $A(t,p)^{-1}$ will be defined identically to (15) and (16). As a consequence of (A3), one can write $g_2 = \Phi_p(t,0)g_2$. As such, all elements of the basis from (39) evolve according to the relationship

$$g_j(t,p) = \nu_j(t)\Phi_p(t,0)g_j(0,p) \tag{40}$$

for all j, where $\nu_j(t)$ is an appropriately defined, nonzero scalar. Note that because $\Phi_p(t,0)$ is always invertible, $A(t,p)$ is guaranteed to be invertible for all time provided that $A(0,p)$ is invertible. Therefore, the collection $w_1,\dots,w_n$ is guaranteed to exist at all times provided that it exists at $t=0$.

Once again defining a transformed coordinate system according to (17), taking time derivatives and simplifying following the steps used to obtain (19) and (20) yields

$$\dot{s}_k = w_k^T\left(F(x^\gamma,p) - \sum_{k=1}^{N}\left(s_k\frac{\partial g_k}{\partial p}\dot{p}\right) + \sum_{k=3}^{N}(s_k\kappa_k g_k) + U_e(t,p,x) - \frac{\partial x^\gamma}{\partial\theta}\dot{\theta} - \frac{\partial x^\gamma}{\partial p}\dot{p}\right) + O(\epsilon^2). \tag{41}$$

Assuming that $s_k = O(\epsilon)$ for all k, evaluating (41) at specific values of k gives

$$\dot{s}_1 = w_1^T(\theta,p)U_e(t,p,x) - w_1^T(\theta,p)\frac{\partial x^\gamma}{\partial p}\dot{p} + O(\epsilon), \tag{42}$$

$$\dot{s}_2 = 1 + w_2^T(\theta,p)U_e(t,p,x) - \dot{\theta} - w_2^T(\theta,p)\frac{\partial x^\gamma}{\partial p}\dot{p} + O(\epsilon), \tag{43}$$

$$\dot{s}_k = \kappa(p)s_k + w_k^T(\theta,p)U_e(t,p,x) - w_k^T(\theta,p)\frac{\partial x^\gamma}{\partial p}\dot{p} + O(\epsilon). \tag{44}$$

Equations (42)–(44) are similar to those from (21); however, fewer cancellations occur due to the fact that $g_1(t,p)$ is no longer identical to $\partial x^\gamma/\partial p$. Provided that each s_k remains an order ϵ term for all time, the full system (9) can be approximated according to

$$x(t) = x^\gamma(\theta,p) + O(\epsilon). \tag{45}$$

Noting that (44) is in a form similar to the equation from (B1) and that each $\kappa_k(p) = O(1/\epsilon)$, provided that $|\dot{p}|$ and $||\partial x^\gamma/\partial p||_1$ are both bounded uniformly in time by $O(1)$ constants, using a line of reasoning similar to that of Appendix B, one can show that $s_k(t) = O(\epsilon)$ for all time for $k \geq 3$. Consequently, the following reduction can then be implemented to represent the perturbed behavior of (9):

$$\dot{\theta} = 1 + w_2^T(\theta,p)U_e(t,p,\theta) - \alpha s_1 \left(w_2^T(\theta,p)\frac{\partial x^\gamma}{\partial p}\right)\left(w_1^T(\theta,p)\frac{\partial x^\gamma}{\partial p}\right), \tag{46}$$

$$\dot{p} = \alpha s_1 w_1^T(\theta,p)\frac{\partial x^\gamma}{\partial p}, \tag{47}$$

$$\dot{s}_1 = w_1^T(\theta,p)U_e(t,p,\theta) - \alpha s_1 \left(w_1^T(\theta,p)\frac{\partial x^\gamma}{\partial p}\right)^2, \tag{48}$$

where α is a positive constant. Above, the parameter update in (47) is chosen so that the resulting term $-\alpha s_1(w_1^T(\theta,p)\partial x^\gamma/\partial p)^2$ in (48) serves to counteract the effect of $w_1^T(\theta,p)U_e(t,p,\theta)$ so that s_1 stays an $O(\epsilon)$ term for all time. The update for (46) is chosen so that $\dot{s}_2 = O(\epsilon)$, allowing this variable to be ignored from the reduced order equations. As in (23), U_e can be written as a function of t,p, and θ using (45) so that (46)–(48) are self-contained. Note that the set of reduced order equations of the form (46)–(48) is similar in form to the set from (5) for use in systems with periodic oscillations.

The three-dimensional ATR represented by (45)–(48) constitute an alternative reduced order modeling framework to the two-dimensional ATR represented by (22) and (23). It is generally preferable to use the two-dimensional ATR (23) when possible because it requires fewer reduced order variables. Nevertheless, the three-dimensional ATR can be applied in situations where the two-dimensional ATR cannot. Applications to follow will consider these points further.

3.7. Reduction of time-scaled trajectories. The proposed dimension reduction strategy explicitly assumes that the end time, T, is constant between reference trajectories $x^\gamma(t,p)$. In some cases, it may be useful to compare trajectories for which the end time T is a function of p. For example, when some $x^\gamma(t,p)$ are periodic, it may be desirable to take $T(p)$ to be the period of oscillation. This can be handled straightforwardly by applying a preliminary time scaling $\tau(p) = t/T(p)$ to the model (1). With this scaling, (9) becomes

$$\frac{dx}{d\tau} = T(p)F(x,p) + T(p)U_e(t,p,\theta). \tag{49}$$

Applying the ATR strategy detailed in section 3.3 to the time-scaled model (49) yields

$$\begin{aligned} \frac{d\theta}{d\tau} &= 1 + T(p) w_{2,\tau}^T(\theta,p) U_e(t,p,\theta), \\ \frac{dp}{d\tau} &= T(p) w_{1,\tau}^T(\theta,p) U_e(t,p,\theta), \end{aligned} \tag{50}$$

where the terms $w_{1,\tau}$ and $w_{2,\tau}$ are computed using (49) and not (9). Undoing the time scaling in (50) yields

$$\begin{aligned} \dot{\theta} &= \frac{1}{T(p)} + w_{2,\tau}^T(\theta,p) U_e(t,p,\theta), \\ \dot{p} &= w_{1,\tau}^T(\theta,p) U_e(t,p,\theta). \end{aligned} \tag{51}$$

An analogous strategy can also be implemented for the three-dimensional ATR described by (46)–(48) yielding the time-scaled equations

$$\begin{aligned} \frac{d\theta}{d\tau} &= 1 + T(p) w_{2,\tau}^T(\theta,p) U_e(t,p,\theta) - \hat{\alpha} s_1 \left(w_{2,\tau}^T(\theta,p) \frac{\partial x^\gamma}{\partial p} \right) \left(w_{1,\tau}^T(\theta,p) \frac{\partial x^\gamma}{\partial p} \right), \\ \frac{dp}{d\tau} &= \hat{\alpha} s_1 w_{1,\tau}^T(\theta,p) \frac{\partial x^\gamma}{\partial p}, \\ \frac{ds_1}{d\tau} &= T(p) w_{1,\tau}^T(\theta,p) U_e(t,p,\theta) - \hat{\alpha} s_1 \left(w_{1,\tau}^T(\theta,p) \frac{\partial x^\gamma}{\partial p} \right)^2, \end{aligned} \tag{52}$$

where $\hat{\alpha} > 0$ can be chosen arbitrarily. Once again, the terms $w_{1,\tau}$ and $w_{2,\tau}$ are computed using (49) and not (9). Taking $\hat{\alpha} = \alpha T(p)$ and undoing the time scaling in (52) yields

$$\begin{aligned} \dot{\theta} &= 1/T(p) + w_{2,\tau}^T(\theta,p) U_e(t,p,\theta) - \alpha s_1 \left(w_{2,\tau}^T(\theta,p) \frac{\partial x^\gamma}{\partial p} \right) \left(w_{1,\tau}^T(\theta,p) \frac{\partial x^\gamma}{\partial p} \right), \\ \dot{p} &= \alpha s_1 w_{1,\tau}^T(\theta,p) \frac{\partial x^\gamma}{\partial p}, \\ \dot{s}_1 &= w_{1,\tau}^T(\theta,p) U_e(t,p,\theta) - \alpha s_1 \left(w_{1,\tau}^T(\theta,p) \frac{\partial x^\gamma}{\partial p} \right)^2. \end{aligned} \tag{53}$$

The time scaling discussed above can sometimes be used to compute necessary terms of the ATR equations more efficiently. For instance, when $x^\gamma(t,p)$ is periodic, using the preliminary scaling $\tau(p) = t/T(p)$, one can compute the necessary terms of the ATR over a single period. Simulations can subsequently be performed considering $\theta \in [0,1)$ and resetting θ to 0 every time it reaches 1. Without time scaling, such resetting would not be possible and one would need to compute the terms of the ATR equations over the full time interval considered. Appropriate time scaling can also be numerically beneficial. For instance, considering the trajectories plotted in panel A in the example from Figure 5, the trajectories are time-scaled so that the action potentials are aligned. If the action potentials were not aligned, both $w_1(\theta,p)$ and $w_2(\theta,p)$ would change more sharply across trajectories requiring a finer discretization in order to obtain an accurate interpolation between trajectories. The time-scaled forms of the reduced order equations (51) and (53) will be used in the examples to follow.

3.8. Choosing an adaptive parameter set and reference trajectory set. In general, any parameter p from (1) can be used as an adaptive parameter for the purposes of implementing the ATR strategy. It often works well in practice to consider an adaptive parameter that mimics the effect of the applied input $U(t)$. For instance, if it is possible to write (1) in the form $\dot{x} = F(x, p_0) + U(t) = F(x) + U(p_0) + U(t)$, one can consider the shadow system $\dot{x} = F(x) + U(p) + U_e(t, p)$ with $U_e(t, p) = U(p_0) - U(p) + U(t)$. In this case, U_e no longer depends explicitly on the state x, which greatly simplifies the implementation and analysis of the resulting reduced order equations. In the applications provided in section 4, the adaptive parameter is chosen in this manner.

In the implementation of the ATR strategy, there is significant freedom in the choice of the reference trajectory set $x^\gamma(t, p)$. In general one can start with a single, trajectory $x^\gamma(t, p_0)$ associated with the unperturbed flow of (1) and add additional trajectories that are relevant to the specific application. Considering Assumption B from section 3.2, these additional trajectories must be chosen so that $\frac{\partial x^\gamma}{\partial p}$ and $\frac{\partial^2 x^\gamma}{\partial p^2}$ exist for all t and p. This can be accomplished by choosing a set of initial conditions $x^\gamma(0, p)$ that satisfy these differentiability constraints and integrate (9) forward in time to obtain the full trajectory set. Examples of the implementation of this process are given in the examples provided in section 4.

4. Examples. Two applications are provided here with relevance to neural spiking behavior. In the first example, the problem of identifying an energy optimal transmembrane current input to transition a periodically firing neuron to a quiescent state is considered. In the second example, the adaptive trajectory reduction framework is used to characterize transitions between tonic firing and quiescence in response to external inputs.

4.1. Optimal inputs to eliminate tonic firing in a neural model. Control strategies for neural spiking have received a great deal of interest in recent years, primarily motivated by the advent of deep brain stimulation as a treatment for various neurological disorders [31], [5]. One class of neuron control problems seeks to identify an optimal input that will yield an action potential in an otherwise quiescent neuron [2], [14], [21], [42]. These problems generally restrict the dynamics to a neighborhood of the quiescent state, and consider an action potential to have occurred if the transmembrane voltage surpasses some threshold value, but do not consider the full transmembrane voltage upstroke and subsequent recovery. Another large class of neuron control problems considers optimal modification of action potential timing in tonically firing neurons [39], [67], [11], [50], [48], [41], [20]. In these applications, the hyperbolicity of an underlying periodic orbit generally allows for the implementation of phase-based reduction strategies.

A significant difficulty associated with neural control problems is that the dynamical equations associated with excitable neurons are highly nonlinear. As such, in order to formulate a tractable control problem, it is generally necessary to restrict the state to be near a stable attractor, for instance to a neighborhood of a stable fixed point when considering optimal inputs to produce action potentials or to the neighborhood of a stable periodic orbit when considering the modification of spike timing. As such, it is difficult to formulate control problems that require a transition through a bifurcation, for example, from tonic spiking to quiescence. In the following example, the ATR strategy is used to circumvent this limitation. Using the reduced order equations from (51), a tractable optimal control problem is formulated and

solved to identify minimum energy inputs required to drive a neuron from a tonically firing state to a quiescent state.

To approach this problem, consider a dynamical model from [52] which replicates firing behavior in thalamic neurons

$$\begin{aligned} C\dot{V} &= -I_{\mathrm{L}}(V) - I_{\mathrm{Na}}(V,h) - I_{\mathrm{K}}(V,h) - I_{\mathrm{T}}(V,r) + I_{\mathrm{SM}} + u(t), \\ \dot{h} &= (h_\infty(V) - h)/\tau_h(V), \\ \dot{r} &= (r_\infty(V) - r)/\tau_r(V), \end{aligned} \tag{54}$$

where V represents the transmembrane voltage, h and r are gating variables, $u(t)$ is an externally applied transmembrane current, $C = 1\mu\mathrm{F/cm}^2$ is the membrane capacitance, and $I_{\mathrm{SM}} = 3\mu\mathrm{A/cm}^2$. The reader is referred to [52] for a full description of the remaining functions that determine the ionic currents I_{L}, I_{Na}, I_{K}, and I_{T}. For this choice of I_{SM}, the neural model has a stable periodic orbit with period 11.89 ms corresponding to tonic firing.

As stated above, the goal here is to identify an energy-optimal input (i.e., that minimizes the L^2 norm) that can eliminate tonic firing in favor of quiescence. To proceed, a shadow system will be considered with the same h and r dynamics as (54), but where the transmembrane voltage dynamics follow

$$C\dot{V} = -I_{\mathrm{L}}(V) - I_{\mathrm{Na}}(V,h) - I_{\mathrm{K}}(V,h) - I_{\mathrm{T}}(V,r) + p + U_e(t,p), \tag{55}$$

where p is an adaptive parameter that behaves like a baseline current and

$$U_e(t,p) = \begin{bmatrix} I_{\mathrm{SM}} - p + u(t) \\ 0 \\ 0 \end{bmatrix} \tag{56}$$

is the effective input. When $U_e(t,p) = 0$, the shadow model (55) displays qualitatively different steady state behavior for different values of p as shown in Figure 2. For instance in Panel E taking $p > 0.33\ \mu\mathrm{A/cm}^2$, tonic firing emerges in steady state. As p decreases, the steady state behavior transitions through quiescence, tonic firing, and bursting (Panels B–D) before settling to a quiescent state when $p < -0.61\ \mu\mathrm{A/cm}^2$. Notice that while the steady state behavior is substantially different for this range of values of p, the trajectories in panels A–E of Figure 2 are qualitatively similar over the first 60 ms of simulation.

The ATR framework is implemented by choosing trajectories $x^\gamma(t,p)$ as follows: first the model (54) (i.e., with nominal parameters in the tonically firing regime) is simulated until it reaches its limit cycle and an initial condition x_{init} is chosen on the upstroke of the action potential 0.6 seconds before the peak voltage is achieved. Using this initial condition, the shadow system (55) is then simulated for 0.6 seconds for various choices of p. The resulting state after this 0.6 s simulation is chosen to correspond to $x^\gamma(0,p)$. Trajectories of the ATR are time-scaled using the strategy detailed in section 3.7 taking $T(p) = 13.5 - 1.5p + 0.25(6-p)^2$. This scaling is used to efficiently cover the domain of interest for the proposed control problem of eliminating action potentials. Here action potentials do not occur on the interval $\theta = [0,1]$ for any choice of p. Resulting trajectories are shown in panel A of Figure 3 with corresponding time scalings shown in panel B. For each $x^\gamma(p,\theta)$, the associated eigenvalues

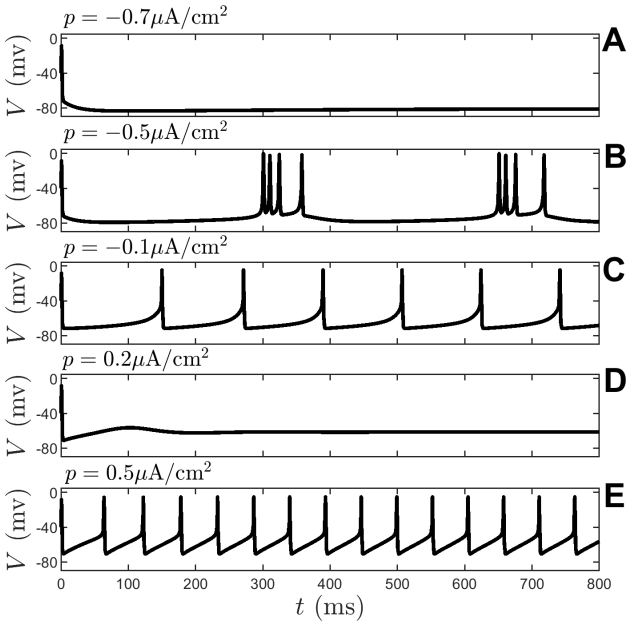

Figure 2. *The neural model (55) transitions from tonic firing for $p > 0.33\,\mu A/cm^2$ (panel E) to quiescence when $p < -0.61\,\mu A/cm^2$ (panel A). The model is quiescent for $p \in (0, .33]$ (panel D), tonic firing occurs when $p \in (-.4, 0]$ (panel C), and bursting occurs when $p \in (-.61, -.4]$. Because the steady state behavior changes on this interval, methods such as the previously proposed APR from (5) cannot be used in this application and the ATR must be used instead.*

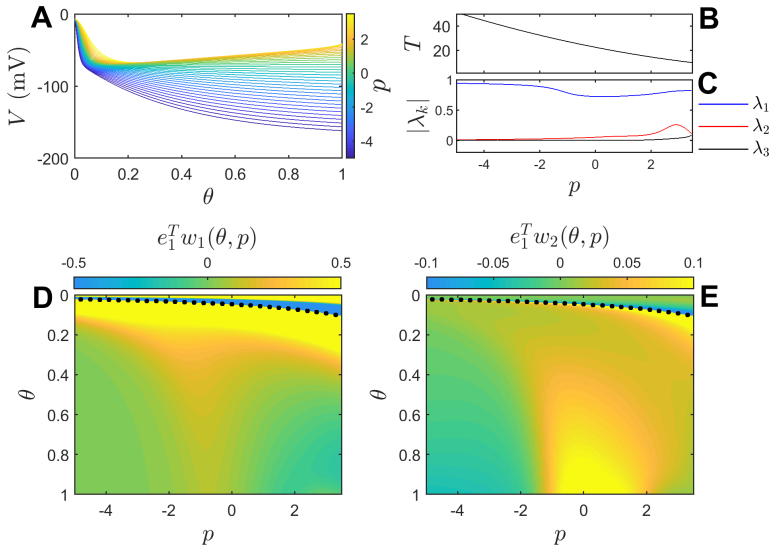

Figure 3. *Individual trajectories of the shadow system (55) are shown in panel A for various values of the adaptive parameter p. Each trajectory is time-scaled using the strategy described in section 3.7 so that no action potentials occur for any of the considered trajectories. Panel B shows the magnitude of the time scaling as a function of p. Resulting eigenvalues of the state transition matrix are shown in panel C. Panels D and E show terms of the reduced order equations $e_1^T w_1(\theta, p)$ and $e_1^T w_1(\theta, p)$, respectively. Dotted lines highlight a boundary for which both $w_1(\theta, p)$ and $w_2(\theta, p)$ are infinite.*

of the state transition matrix $\Phi_p(T(p),0)$ are shown in panel C. In this case, λ_3 is nearly zero for all values of p allowing for the application of the ATR strategy. For each value of p, the terms $w_1(\theta,p)$ and $w_2(\theta,p)$ are computed numerically according to the relations (34) and (35). Colormaps of $e_1^T w_1(\theta,p)$ and $e_1^T w_2(\theta,p)$, i.e., components of the reduced order equations corresponding to perturbations in the transmembrane voltage variable, are shown in panels D and E, respectively. The dotted lines in these panels correspond to locations for which $w_1(\theta,p)$ and $w_2(\theta,p)$ are infinite—the state of the reduced order model must remain sufficiently far from this set to avoid compromising the accuracy of the reduced order equations as explained at the end of section 3.3.

Upon implementing the proposed adaptive trajectory reduction strategy, the resulting ATR equations are

$$\begin{aligned} \dot{\theta} &= 1/T(p) + R(\theta,p)(I_{\mathrm{SM}} - p + u), \\ \dot{p} &= Y(\theta,p)(I_{\mathrm{SM}} - p + u), \end{aligned} \tag{57}$$

where $R(\theta,p) = e_1^T w_2(\theta,p)$ and $Y(\theta,p) = e_1^T w_1(\theta,p)$. The problem of identifying an energy-optimal input to eliminate tonic firing in favor of quiescence will be approached using a calculus of variations framework [23] to identify the control input $u(t)$ that minimizes the cost functional

$$C = \int_0^{t_{\mathrm{end}}} u^2(t)dt \tag{58}$$

subject to the initial conditions conditions $p(0) = I_{\mathrm{SM}} = 3\mu\mathrm{A/cm}^2$ and $\theta(0) = \theta_0$. End point conditions are chosen so that $p(t_{\mathrm{end}}) = p_{\mathrm{targ}}$ and $\theta(t_{\mathrm{end}})$ is free. The initial conditions are chosen to correspond to a location on the neuron's stable limit cycle $x^\gamma(t, I_{\mathrm{SM}})$ that the system relaxes to when $u(t) = 0$. Provided $p_{\mathrm{targ}} < -0.61\ \mu\mathrm{A/cm}^2$ the neuron will be on a trajectory that does not result in an action potential (input will constantly need to be applied to maintain this quiescent state, however). The term $u^2(t)$ is chosen in the integrand of the cost functional (58) which is proportional to the power consumed by the stimulus provided the system obeys Ohm's law. Inputs that minimize the cost functional from (58) can be identified by defining the associated Hamiltonian

$$H(\Psi,u,L,t) = u^2 + L_1[1/T(p) + R(\theta,p)(I_{\mathrm{SM}} - p + u)] + L_2[Y(\theta,p)(I_{\mathrm{SM}} - p + u)], \tag{59}$$

where $\Psi \equiv \begin{bmatrix}\theta & p\end{bmatrix}^T$ is a vector of state variables and $L \equiv \begin{bmatrix}L_1 & L_2\end{bmatrix}^T$ are Lagrange multipliers that force the dynamics to satisfy the reduced order equations (57). The Euler–Lagrange equations associated with the Hamiltonian (59) are [23]

$$\dot{\Psi} = \frac{\partial H}{\partial L}, \tag{60}$$

$$\dot{L} = -\frac{\partial H}{\partial \Psi}, \tag{61}$$

$$0 = \frac{\partial H}{\partial u}. \tag{62}$$

Evaluation of (60) returns the reduced order equations from (57). Evaluation of (61) yields

$$\begin{aligned}\dot{L}_1 &= -L_1 R_\theta(\theta,p)(I_{\mathrm{SM}}-p+u) - L_2 Y_\theta(\theta,p)(I_{\mathrm{SM}}-p+u),\\ \dot{L}_2 &= -(L_1 R_p(\theta,p)+L_2 Y_p(\theta,p))(I_{\mathrm{SM}}-p+u) + L_1 R(\theta,p) + L_2 Y(\theta,p) + \frac{L_1 T_p}{T^2(p)},\end{aligned} \tag{63}$$

where the subscripts θ and p are used to denote the appropriate partial derivatives. Finally, evaluation of (62) yields

$$u = \frac{-L_1 R(\theta,p) - L_2 Y(\theta,p)}{2}, \tag{64}$$

so that (57), (63), and (64) comprise a set of Euler–Lagrange equations that must be satisfied by control inputs $u(t)$ that locally minimize the cost functional (58). Boundary conditions $p(0) = I_{\mathrm{SM}}$, $\theta(0) = \theta_0$, and $p(t_{\mathrm{end}}) = p_{\mathrm{targ}}$ are mandated by the problem formulation. The final boundary condition is $L_1(t_{\mathrm{end}}) = 0$ and follows from the fact that $\theta(t_{\mathrm{end}})$ is free. Solutions of this two-point boundary value problem can be found by identifying initial values of $L_1(0)$ and $L_2(0)$ that give the correct final conditions. This is accomplished by first noticing that when $p_{\mathrm{targ}} = I_{\mathrm{SM}}$ the optimal solution can be found by taking $L_1(0) = L_2(0) = 0$. Solutions for nearby values of p_{targ} can be obtained by updating the initial values for $L_1(0)$ and $L_2(0)$ using a Newton iteration until convergence is achieved. This process can be performed iteratively until the solution associated with the desired value of p_{targ} is obtained.

Using $\theta_0 = 0.2$ and $p_{\mathrm{targ}} = -2$, the resulting optimal inputs are shown in panel A of Figure 4 for various choices of t_{end}. For smaller values of t_{end}, the optimal input resembles a linear ramp. For larger values of t_{end}, the optimal input plateaus near $u(t) = -2\mu\mathrm{A/cm}^2$ before sharply decreasing as time approaches t_{end}. Corresponding values of $\theta(t)$ and $p(t)$ are shown in panels B and C. Panel D gives plots of p versus θ for various values of t_{end} illustrating that these trajectories are sufficiently bounded away from the threshold for which the functions $R(\theta,p)$ and $Y(\theta,p)$ from the reduced order equations (57) are infinite. The optimal solutions are compared to the input $u(t) = -5\mu\mathrm{A/cm}^2$. This comparison is chosen as a simple input for which $p(t)$ approaches the target value of $-2\mu\mathrm{A/cm}^2$ over time (recall that the baseline current $I_{\mathrm{sm}} = 3\mu\mathrm{A/cm}^2$). The overall energy consumption computed according to (58) is shown in panel E for the optimal and the comparison stimuli. For larger values of t_{end}, the optimal stimulus is approximately three times more efficient than the comparison stimulus. Finally, the optimal stimulus associated with $t_{\mathrm{end}} = 60$ ms is applied to the full order model, with results shown in panel F. The black line shows the first 2.4 ms after an action potential where no inputs are applied as θ increases from 0 to 0.2. The blue trace shows transmembrane voltage in response to the optimal input. The dashed line shows the transmembrane voltage in response to the optimal input applied to the reduced order model (57) where (22) is used to determine the associated state estimation. Near-perfect agreement is observed between the full and reduced order models. The comparison input $u(t) = -5\mu\mathrm{A/cm}^2$ is also applied to the full order model, with results shown in red. Note that the resulting $\theta(t_{\mathrm{end}})$ resulting from the application of the optimal stimulus is different for $\theta(t_{\mathrm{end}})$ resulting from application of the comparison stimulus. As such, the final states are also different.

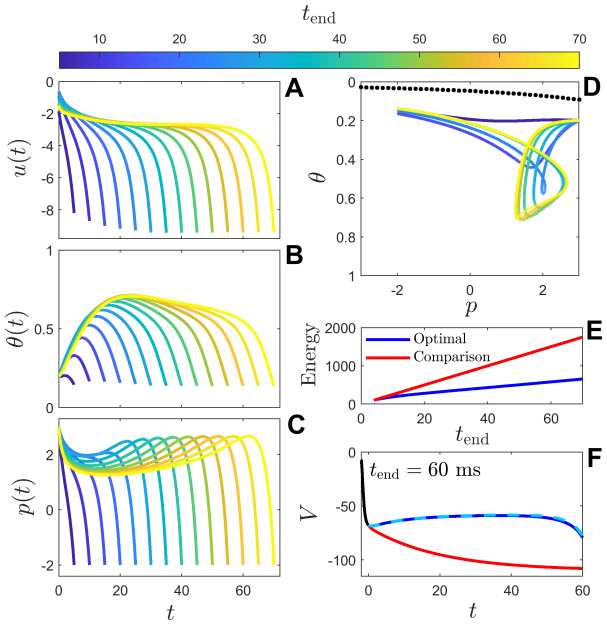

Figure 4. *Optimal control computation to eliminate tonic firing in the thalamic neural model from (54). Panel* A *shows the control inputs (in units of $\mu A/cm^2$) obtained according to (64) that satisfy the Euler–Lagrange equations (60)-(62) for various values of t_{end}. Panels* B *and* C *show associated values of θ and p during the application of these inputs. Panel* D *shows plots of p versus θ for these trajectories. The dotted black line denotes a boundary for which both $R(\theta, p)$ and $Y(\theta, p)$ from the reduced order equations (57) are infinite. As such, all valid trajectories of the reduced order model must remain sufficiently far from this boundary. Panel* E *gives an energy comparison, computed according to (58), for the optimal stimuli and the input $u(t) = -5\mu A/cm^2$. Panel* F *shows the resulting transmembrane voltage (in units of mV) in response to the optimal and comparison input applied to the full order model. The dashed line shows the output of the reduced order model for the optimal input and is in near-perfect agreement with the full model simulations. Recall that $\theta(0) = 0.2$ in the optimal control formulation; the black line in panel* F *represents the first* 2.4 *ms after an action potential occurs during which the optimal phase transitions from $\theta = 0$ to $\theta = 0.2$.*

4.2. Adaptive trajectory reduction to capture transitions between quiescence and tonic firing.

As a second example, neural spiking of the Wang–Buzsáki [55] neuron is considered:

$$\begin{aligned} C\dot{V} &= -I_{\text{L}}(V) - I_{\text{Na}}(V,h) - I_{\text{K}}(V,n) + I_{\text{SM}} + u(t), \\ \dot{h} &= \phi(\alpha_h(1-h) - \beta_h h), \\ \dot{n} &= \phi(\alpha_n(1-n) - \beta_n n), \end{aligned} \tag{65}$$

where V represents the transmembrane voltage, h and n are gating variables, $u(t)$ is an externally applied transmembrane current, $C = 1\mu\text{F/cm}^2$ is the membrane capacitance, and $I_{\text{SM}} = 10\mu\text{A/cm}^2$ is the baseline current. The reader is referred to [55] for a full description of the remaining functions that determine the ionic currents I_{L}, I_{Na}, and I_{K}. For this choice of I_{SM}, when taking $u(t) = 0$ the model (65) has a stable limit cycle with period 3.5 ms. This periodic orbit exists for $I_{SM} > 0.16$ $\mu\text{A/cm}^2$; when I_{SM} falls below this threshold, stable oscillations disappear due to a saddle node on an invariant circle (SNIC) bifurcation.

The goal here is to represent the model (65) using the proposed adaptive trajectory reduction framework, particularly for situations where the system transitions between tonic firing and quiescence across the SNIC bifurcation. To this end, a shadow system will be considered with the same h and n dynamics as (65), but where the transmembrane voltage dynamics follow

$$C\dot{V} = -I_{\mathrm{L}}(V) - I_{\mathrm{Na}}(V,h) - I_{\mathrm{K}}(V,n) + p + U_e(t,p), \tag{66}$$

where p is an adaptive parameter that behaves like a baseline current and $U_e(t,p)$ is an effective input identical to the one given in (56).

The ATR framework is implemented by choosing trajectories $x^\gamma(\theta,p)$ as follows: for values of $p \in [4.6,16]$, the periodic orbit is taken to be $x^\gamma(\theta,p)$, with each of these trajectories aligned and time-scaled so that $\theta = 0$ and $\theta = 1$ both correspond to the peak value of the transmembrane voltage. While additional periodic orbits exist for values of $4.6 > p > 0.16$, the accuracy of the resulting reduced order model is diminished when using reference trajectories that are close to the SNIC bifurcation. For values of $p < 4.6$, the initial conditions for the reference trajectories $x^\gamma(0,p)$ are chosen so that both $\frac{\partial x^\gamma}{\partial p}|_{\theta=0}$ and $\frac{\partial^2 x^\gamma}{\partial p^2}|_{\theta=0}$ are continuous. Likewise, the time scaling for $p < 4.6$ is taken to be $T(p) = T(4.6) + (p-4.6)\frac{dT}{dp}|_{p=4.6}$ and is chosen so that $T(p)$ is also continuously differentiable. Voltage traces of the resulting trajectories and the corresponding time scaling $T(p)$ are shown in panels A and B of Figure 5. Panel C shows the relative magnitudes of the eigenvalues of the state transition matrix. Because λ_3 is always

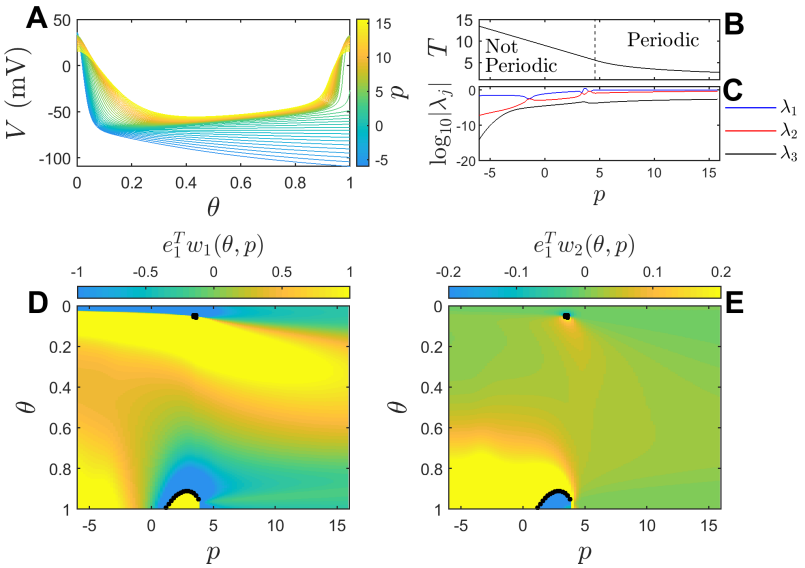

Figure 5. *Panel* A *shows individual trajectories of the shadow system* (66) *for various values of* p *obtained using the procedure described in the text. Trajectories with* $p \geq 4$ *represent limit cycle solutions. Other trajectories are chosen so that the terms of the resulting ATR of the form* (51) *are continuously differentiable. Panel* B *shows the associated time scaling of each trajectory. Panel* C *shows the resulting eigenvalues of the state transition matrix associated with the linearized trajectories. Panels* D *and* E *show terms of the reduced order equations* $e_1^T w_1(\theta,p)$ *and* $e_1^T w_2(\theta,p)$*, respectively. Dotted lines highlight a boundary for which both* $w_1(\theta,p)$ *and* $w_2(\theta,p)$ *are infinite, necessitating the use of the three-dimensional ATR of the form* (53) *in some situations.*

near-zero, the reduction framework can be applied to eliminate dynamics associated with this rapid decay. Colormaps of $e_1^T w_1(\theta,p)$ and $e_1^T w_1(\theta,p)$, i.e., components of the reduced order equations corresponding to perturbations in the transmembrane voltage variable, are shown in panels D and E, respectively. The dotted lines in these panels correspond to locations for which $w_1(\theta,p)$ and $w_2(\theta,p)$ are infinite. As such, the state of the reduced order model must remain sufficiently far from this set to avoid compromising the accuracy of the reduced order equations as explained at the end of section 3.3. In this example, the three-dimensional ATR with transformed equations of the form (53) will also be investigated. For this reduction, $\alpha = 10$ is used. Note that while the three-dimensional ATR does not result in a reduction in dimension for this application, this strategy is readily generalizable to higher dimensional models.

The APR strategy is also considered to represent the dynamics of (65). To implement the APR, stable periodic orbits are considered taking $p \in [0.6, 16]$. Note that the SNIC bifurcation occurs at $p = 0.16$, but accuracy of the APR degrades substantially for values of p that are close to this critical value. When using the APR framework, the update function G_p is chosen to be of the form (6) taking $\alpha = 1000$.

Simulations of the full model (65), the proposed ATR (both the two- and the three-dimensional versions from (51) and (53), respectively), and the APR from (5) are shown in Figure 6. Initial conditions in each simulation are chosen to correspond to the moment that

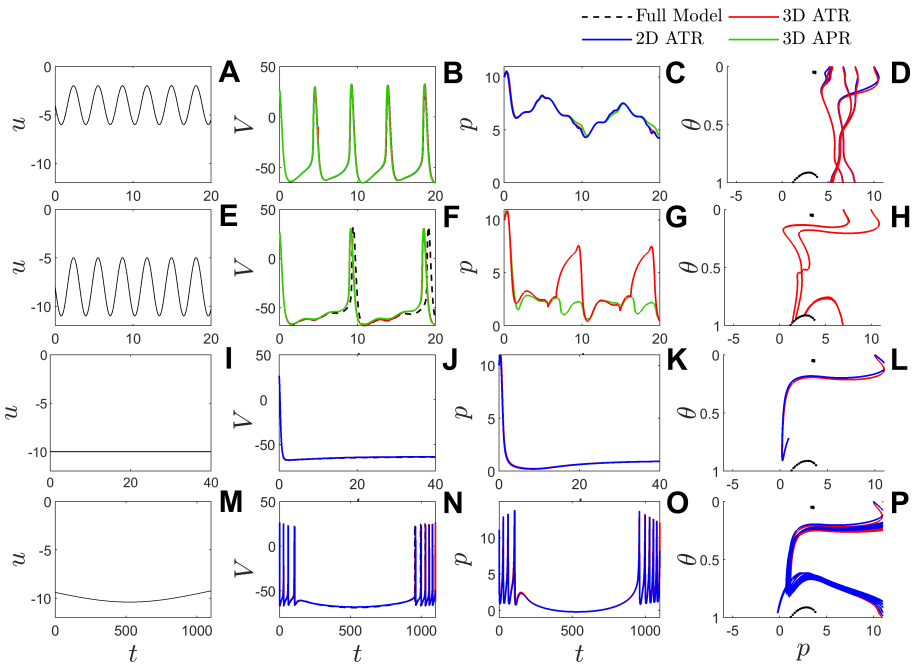

Figure 6. *Results comparing simulations of the full order model (65) to the proposed 2- and 3-dimensional ATR equations ((51) and (53), respectively). Solutions using the the previously proposed APR framework from (5) are also considered. Panels* A, E, I, *and* M *show four different stimuli applied to each model. As explained in the text, not every input can be considered by all of the reduced order models. Panels* B, F, J, *and* N *show the resulting transmembrane voltage output. Panels* C, G, K, *and* O *show associated traces of the adaptive parameter* p *and panels* D, H, L, *and* P *show traces of the variables* p *versus* θ *when using the proposed ATR equations.*

the neuron reaches its peak value of the transmembrane voltage on the limit cycle that results when $u(t) = 0$. When using the ATR strategy, when θ reaches 1 with a value of $p \geq 4.6$, the phase is reset to 0 since these trajectories are periodic.

The response to a purely sinusoidal input of moderate magnitude is shown in panels A–D of Figure 6. For these inputs, all reduced order models are able to faithfully reproduce the full model behavior. Panels E–H show the response to a larger magnitude sinusoid. Here, the two-dimensional ATR cannot be used because the input drives the system to values of p and θ for which the reduced order terms are infinite. The three-dimensional ATR is able to circumvent this limitation. The APR strategy can also be used in this regime. Larger magnitude, negative inputs cannot be considered for the APR because they cause p to approach values that are near the SNIC bifurcation. A constant input $u(t) = -10\mu\text{A}/\text{cm}^2$ is considered in panels I–L. In this case, the APR from (5) cannot be considered because the input drives the neuron past the SNIC bifurcation to a quiescent state. Both the proposed two- and three-dimensional ATR strategies yield results that are indistinguishable from the full model simulations when using this constant input. Finally, the input $u(t) = -9.4 - \sin(0.03t)$ is considered in panels M–P. This input is chosen to highlight the ability of the ATR strategy to accurately capture transitions between the quiescent and tonically firing regimes. Once again, the APR strategy cannot be considered for this input because it drives the system past the SNIC bifurcation.

5. Conclusion and future directions.

In this work, a general reduced order modeling strategy is developed that can accurately capture the behavior of a strongly perturbed, nonlinear dynamical system in the vicinity of a continuous collection of reference trajectories. To implement this strategy, it is assumed that the local linearization of each trajectory is rapidly contracting in some directions (as gauged by the magnitudes of the eigenvalues of the associated state transition matrix). To capture the behavior of the nonrapidly contracting directions, a local coordinate transformation is proposed that explicitly considers the behavior of other nearby trajectories. As illustrated in the examples from section 4, the ATR framework can be used to accurately capture the behavior of systems for which external inputs are of sufficient strength to precipitate bifurcations.

There are many opportunities for extension of the proposed ATR strategy, particularly for situations where some of the mandated assumptions from section 3.2 are not satisfied. Foremost, Assumption A requires that $F(x, p)$ from (9) is sufficiently smooth. This constraint precludes the use of the ATR for nonsmooth dynamical systems. Recent work in [45], [59] has considered phase-based reduced order modeling strategies that are valid for piecewise smooth dynamical systems subject to a transverse flow condition. For such piecewise smooth dynamical systems, the infinitesimal phase and isostable response curves (i.e., the gradients of the phase and isostable coordinates evaluated on the periodic orbit) are generally discontinuous across boundaries for which the corresponding vector field is discontinuous. It is likely that methods similar to those used in [45], [59] could be used to identify the necessary terms of the ATR for a piecewise smooth dynamical system.

Assumption C, i.e., the requirement that all but two of the eigenvalues of $\Phi_p(T, 0)$ are small for all p, is one of the most restrictive conditions required for the implementation of the proposed ATR strategy. It would be of interest to investigate possible extensions valid for use when this constraint is not satisfied. To approach this problem, one could start with

a modified version of (39) that includes the dynamics of $\zeta > 2$ slowly decaying components of the linearized solution. This would ultimately yield a reduced order set of coordinates $s_1, s_2, \ldots, s_{\zeta-1}$ in APR equations from (53) in order to account for these additional slowly decaying directions. Subsequently, similar to the APR strategy from [64], one would need to identify a parameter update rule that keeps the magnitude of these reduced order coordinates small in order to arrive at a reduced order set of equations. Depending on the number of nontruncated reduced order coordinates required, it may also be necessary to consider adaptive parameter sets, p, with multiple dimensions.

As part of Assumption D, the eigenvalues $\lambda_k(p)$ of $\Phi_p(T,0)$ are required to be simple for $k \geq 3$. This assumption ensures that the time derivatives of the coordinate basis $g_1, \ldots, g_N$ as defined in section 3.3 exist. It may be possible to explicitly extend this derivation for application in situations where Assumption D is not satisfied, for instance by using the notion of generalized eigenvectors [35]. A similar strategy was employed in [62] for applications of the APR framework to dynamical systems with defective linearizations.

For the proposed ATR framework, a set of reference trajectories is required to perform the reduction; however, there is no set strategy for constructing this trajectory set. In the examples considered in this work, the reference trajectory set is constructed in an ad hoc manner to capture the salient features of the full model behavior. It would be of interest to identify a concrete metric that gives a sense of how the accuracy of the resulting reduced order equations is influenced by the choice of the reference trajectory set. It is worth noting that for the examples considered in this work, small changes in the choice of the reference trajectory sets do not substantially influence the accuracy of the resulting reduced order equations.

In contrast to reduced order modeling techniques based on the use of asymptotic phase and isostable coordinates, the proposed ATR framework can be applied to a set of arbitrary reference trajectories. As such, it can be used to investigate problems where the application of external input causes the dynamics to transition through a bifurcation. Future work will consider the possibility of using the proposed ATR strategy in conjunction with other phase and isostable-based reduction frameworks with the goal of creating a unified framework that can represent the behavior of a general dynamical system in multiple operating regimes.

Appendix A. Dynamical relationships between terms of the proposed moving coordinate basis. Here, dynamical relationships associated with the terms $g_k(t,p)$ that comprise the moving coordinate basis from (11) are considered. To begin, consider an initial trajectory $x_1 = x^\gamma(t,p)$ and an adjacent trajectory $x_2 = x^\gamma(t,p+\Delta p)$ where Δp is a small, constant shift in the parameter p. Asymptotically expanding in reference to $x^\gamma(t,p)$ yields

$$\begin{aligned}
\dot{x}_1 &= F(x^\gamma, p),\\
\dot{x}_2 &= F(x^\gamma + \Delta x, p + \Delta p)\\
&= F(x^\gamma, p) + J\Delta x + \frac{\partial F}{\partial p}\Delta p + O(||\Delta x||^2) + O(||\Delta p||^2)\\
&= F(x^\gamma, p) + J\frac{\partial x^\gamma}{\partial p}\Delta p + \frac{\partial F}{\partial p}\Delta p + O(||\Delta p||^2),
\end{aligned} \tag{A1}$$

where the relationship $\Delta x = \frac{\partial x^\gamma}{\partial p}\Delta p$ is used in the fourth line and all partial derivatives are evaluated at $x^\gamma(t,p)$ using the parameter set p. Recalling that $g_1 = \frac{\partial x^\gamma}{\partial p}$, by considering $\lim_{\Delta p \to 0}[(\dot{x}_2 - \dot{x}_1)/\Delta p]$, one finds

$$\dot{g}_1 = Jg_1 + \frac{\partial F}{\partial p}. \tag{A2}$$

Furthermore, noting that $\frac{\partial}{\partial t}\Phi_p(t,0) = J\Phi_p(t,0)$, directly differentiating (13) and (14) with respect to time (recalling that p is constant) yields

$$\begin{aligned} \dot{g}_2 &= Jg_2, \\ \dot{g}_k &= (J - \kappa_k \mathrm{Id})g_k, \\ & \qquad k = 3, \ldots, N, \end{aligned} \tag{A3}$$

where the partial derivatives are calculated at $x^\gamma(t,p)$ using the parameter set p and Id is an appropriately sized identity matrix. Equations (A2) and (A3) are valid, provided p is constant, and are used in the derivations presented in section 3.5. If p is allowed to change in time, these total derivatives become

$$\begin{aligned} \dot{g}_1 &= Jg_1 + \frac{\partial F}{\partial p} + \frac{\partial g_1}{\partial p}\dot{p}, \\ \dot{g}_2 &= Jg_2 + \frac{\partial g_2}{\partial p}\dot{p}, \\ \dot{g}_k &= (J - \kappa_k \mathrm{Id})g_k + \frac{\partial g_k}{\partial p}\dot{p}, \\ & \qquad k = 3, \ldots, N. \end{aligned} \tag{A4}$$

Equations (A4) are applied in the derivation from section 3.3.

Finally, recalling the definitions of w_k^T from (16), for any choice of j and k one can write

$$w_j^T(t,p)g_k(t,p) = c, \tag{A5}$$

where c is a constant that equals 1 if $j = k$ and 0 otherwise. As such, taking the time derivative of (A5) yields

$$\dot{w}_j^T g_k + w_j^T \dot{g}_k = 0 \tag{A6}$$

for any choice of k and j.

Appendix B. Bounding the magnitude of rapidly decaying terms of the moving coordinate system. Here, we consider the dynamics of the rapidly decaying terms from the moving coordinate system from (21):

$$\dot{s}_k = \kappa_k(p)s_k + w_k^T(t,p)U_e(t,p,x) + O(\epsilon) \tag{B1}$$

for $k \geq 3$. As in the derivation from section 3.3, it will be assumed that each $s_k = O(\epsilon)$ at time $t = 0$. Noting that s_k may be complex valued, let $|s_k|^2 = s_k s_k^*$, where * denotes the complex conjugate. Taking the time derivative of both sides of this relation yields

$$2|s_k||\dot{s}_k| = s_k\dot{s}_k^* + \dot{s}_k s_k^*. \tag{B2}$$

Considering that $\kappa_k(p)$ and $w_k(t,p)$ may be complex valued, substituting (B1) into (B2) yields

$$\begin{aligned} 2|s_k||\dot{s}_k| &= s_k(\kappa_k^*(p)s_k^* + w_k^{T^*}(t,p)U_e(t,p,x)) + s_k^*(\kappa_k(p)s_k + w_k^T(t,p)U_e(t,p,x)) + O(\epsilon^2) \\ &= 2\mathrm{Re}(\kappa_k(p))|s_k|^2 + s_k(w_k^{T^*}(t,p)U_e(t,p,x)) + s_k^*(w_k^T(t,p)U_e(t,p,x)) + O(\epsilon^2) \\ &\le 2\mathrm{Re}(\kappa_k(p))|s_k|^2 + 2|s_k||w_k^T(t,p)U_e(t,p,x)| + O(\epsilon^2). \end{aligned} \tag{B3}$$

Rearranging (B3), invoking Assumption E from section 3.2, noting that each $\mathrm{Re}(\kappa_k(p)) < 0$, and truncating $O(\epsilon^2)$ terms, one finds

$$\begin{aligned} |\dot{s}_k| &\le \mathrm{Re}(\kappa_k(p))|s_k| + \max_{t,p,x}|w_k^T(t,p)U_e(t,p,x)| \\ &\le \max_p(\mathrm{Re}(\kappa_k(p)))|s_k| + M_U \max_{t,p}||w_k(t,p)||_1. \end{aligned} \tag{B4}$$

Once again, recalling that $\max_p(\mathrm{Re}(\kappa_k(p))) < 0$, considering the relationship from (B4), if

$$|s_k| > \frac{M_U \max_{t,p}||w_k(t,p)||_1}{-\max_p(\mathrm{Re}(\kappa_k(p)))} \tag{B5}$$

is satisfied, then $|\dot{s}_k| < 0$. As such,

$$|s_k| \le \frac{M_U \max_{t,p}||w_k(t,p)||_1}{-\max_p(\mathrm{Re}(\kappa_k(p)))} = O(\epsilon) \tag{B6}$$

is an upper bound for $|s_k|$. This upper bound is an $O(\epsilon)$ term because $\max_p(\mathrm{Re}(\kappa_k(p))) = O(1/\epsilon)$ (by Assumption C) and because M_U and w_k are $O(1)$ terms as a consequence of Assumptions E and F, respectively.

Appendix C. Partial derivatives of the state transition matrix with respect to adjacent trajectories. Here, it is shown that for the state transition matrix $\Phi_p(t,0)$ as defined below (8), $\frac{\partial \Phi_p}{\partial p}$ exists provided the assumptions from section 3.2 are satisfied. This will be shown through direct computation. To begin, for a given trajectory $x^\gamma(t,p)$, using the Peano–Baker series representation for the state transition matrix [18] associated with solutions of (8), one finds

$$\Phi_p(t,0) = \mathrm{Id} + \int_0^t J_p(\sigma_1)d\sigma_1 + \int_0^t J_p(\sigma_1)\int_0^{\sigma_1} J_p(\sigma_2)\sigma_2 d\sigma_1 + \cdots, \tag{C1}$$

where Id is the identity matrix, and $J_p(t)$ is a shorthand representation of the Jacobian from (8). Next consider the nearby trajectory $x^\gamma(t,p+\Delta p) = x^\gamma(t,p) + \frac{\partial x^\gamma}{\partial p}\Delta p + O(\Delta p^2)$. Using equation (28) from [60], through asymptotic expansion one finds

$$J_{p+\Delta p} = J_p + \begin{bmatrix} a_1 & \cdots & a_N \end{bmatrix}^T + \frac{\partial}{\partial p}\left(\frac{\partial F}{\partial x}\right)\Delta p + O(\Delta p^2), \tag{C2}$$

where each a_i is a column vector given by

$$a_i = \Delta p\left((\partial x^\gamma/\partial p)^T \otimes \mathrm{Id}\right)\mathrm{vec}(H_i), \tag{C3}$$

H_i is the Hessian matrix of partial derivatives associated with the ith entry of F, $\otimes$ denotes the Kronecker product, vec$(\cdot)$ is an operator that stacks each column of a matrix to form a single column vector, and Id is an appropriately sized identity matrix. In (C2) and (C3), all partial derivatives are evaluated at $x(t)$ and p. Considering (C2), one can write

$$J_{p+\Delta p}(t) = J_p(t) + M(t)\Delta p + O(\Delta p^2), \tag{C4}$$

where $M(t)$ is defined appropriately. With this in mind, using the Peano–Baker series representation from (C1) for both $\Phi_p(t,0)$ and $\Phi_{p+\Delta p}(t,0)$ and simplifying appropriately yields

$$\begin{aligned}\Phi_p(t,0) - \Phi_{p+\Delta p}(t,0) &= \Delta p \int_0^t M(\sigma_1)d\sigma_1 \\ &+ \Delta p \int_0^t \int_0^{\sigma_1} \Big[J_p(\sigma_1)M(\sigma_2) + M(\sigma_1)J_p(\sigma_2)\Big]d\sigma_2 d\sigma_1 \\ &+ \Delta p \int_0^t \int_0^{\sigma_1} \int_0^{\sigma_2} \Big[J_p(\sigma_1)J_p(\sigma_2)M(\sigma_3) + J_p(\sigma_1)M(\sigma_2)J_p(\sigma_3) \\ &+ M(\sigma_1)J_p(\sigma_2)J_p(\sigma_3)\Big]d\sigma_3 d\sigma_2 d\sigma_1 + \cdots + O(\Delta p^2).\end{aligned} \tag{C5}$$

As such, the required partial derivatives exist and can be written as a series representation according to

$$\begin{aligned}\lim_{\Delta p \to 0} \frac{\Phi_p(t,0) - \Phi_{p+\Delta p}(t,0)}{\Delta p} &= \int_0^t M(\sigma_1)d\sigma_1 \\ &+ \int_0^t \int_0^{\sigma_1} \Big[J_p(\sigma_1)M(\sigma_2) + M(\sigma_1)J_p(\sigma_2)\Big]d\sigma_2 d\sigma_1 + \cdots .\end{aligned} \tag{C6}$$

Appendix D. Existence of time derivatives of basis elements of the proposed moving coordinate system. In the derivations from section 3.3, it is necessary for the time derivatives of $g_1(t,p),\ldots,g_N(t,p)$ and $w_1(t,p),\ldots,w_N(t,p)$ to exist. Considering the set of equations from (A4), the terms $\dot{g}_1(t,p),\ldots,\dot{g}_N(t,p)$ exist provided all the terms of the right- hand side of (A4) exist. The Jacobian and $\frac{\partial F}{\partial p}$ exist as a consequence of Assumption A from section 3.2. The term $\frac{\partial g_1}{\partial p} = \frac{\partial^2 x^\gamma}{\partial p^2}$ exists according to Assumption B. The term $\frac{\partial g_2}{\partial p} = \frac{\partial x^\gamma}{\partial p}$ also exists according to Assumption B.

To establish the existence of $\frac{\partial g_k}{\partial p}$ for $k \geq 3$, consider the following definition from (14):

$$g_k(t,p) = \exp(-\kappa_k(p)t)\Phi_p(t,0)v_k(p) \tag{D1}$$

for $k = 3\ldots N$. Define $\lambda_k(p) = f_{\lambda_k}(\Phi_p(T,0))$ and $v_k(p) = f_{v_k}(\Phi_p(T,0))$, where each $v_k(p)$ is scaled such that $||v_k(p)||_2 = 1$, where $||\cdot||_2$ is the two-norm. Appendix C illustrates that $\frac{\partial \Phi_p(t,0)}{\partial p}$ exists. By assumption D, each $\lambda_k(p)$ is a simple eigenvalue for $k \geq 3$. For these simple eigenvalues, considering the associated sensitivity to perturbations of eigenvectors and

eigenvalues established in [12], [36], $\frac{\partial \lambda_k}{\partial p}$ and $\frac{\partial v_k}{\partial p}$ also exist. Note that $\lambda_k \neq 0$ for any k since the state transition matrix must be invertible. As such, $\frac{\partial \kappa_k}{\partial p}$ exists for $k \geq 3$. All partial derivatives with respect to p of the terms that comprise the right-hand side of (D1) exist so that $\frac{g_k}{\partial p}$ also exists. Finally, because all terms of the right-hand sides of the equations that comprise (A4) exist, $\dot{g}_1(t,p), \ldots, \dot{g}_N(t,p)$ also exist.

As defined in (16) each $w_k(t,p)$ corresponds to a row of $A(t,p)^{-1}$ which must exist by Assumption F. One can also write

$$\begin{aligned} \frac{dA^{-1}}{dt} &= \frac{\partial A^{-1}}{\partial t} + \frac{\partial A^{-1}}{\partial p}\dot{p} \\ &= A^{-1}\frac{\partial A}{\partial t}A^{-1} + A^{-1}\frac{\partial A}{\partial p}A^{-1}\dot{p}. \end{aligned} \tag{D2}$$

The matrix $A(t,p)$ is comprised of the vectors $g_1(t,p), \ldots, g_N(t,p)$ with time derivatives and partial derivatives that exist. As such, $\frac{dA^{-1}}{dt}$ exists and consequently each $\dot{w}_1(t,p), \ldots, \dot{w}_N(t,p)$ exists.

A nearly identical argument can be applied to show that the required time derivatives exist when using the alternative reduced coordinate system defined in section 3.6.

REFERENCES

[1] D. M. Abrams and S. H. Strogatz, *Chimera states for coupled oscillators*, Phys. Rev. Lett., 93 (2004), 174102.

[2] Y. Ahmadian, A. M. Packer, R. Yuste, and L. Paninski, *Designing optimal stimuli to control neuronal spike timing*, J. Neurophysiol., 106 (2011), pp. 1038–1053.

[3] R. E. Bertram and J. E. Rubin, *Multi-timescale systems and fast-slow analysis*, Math. Biosci., 287 (2017), pp. 105–121.

[4] 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.

[5] J. M. Bronstein, M. Tagliati, R. L. Alterman, A. M. Lozano, J. Volkmann, A. Stefani, F. B. Horak, M. S. Okun, K. D. Foote, P. Krack, R. Pahwa, J. M. Henderson, M. I. Hariz, R. A. Bakay, A. Rezai, W. J. Marks, Jr., E. Moro, J. L. Vitek, F. M. Weaver, R. E. Gross, and M. R. DeLong, *Deep brain stimulation for Parkinson disease: An expert consensus and review of key issues*, Arch. Neurol., 68 (2011), pp. 165–165.

[6] 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.

[7] M. Budišić, R.Mohr, and I.Mezić, *Applied Koopmanism*, Chaos, 22 (2012), 047510.

[8] O. Castejón and A. Guillamon, *Phase-amplitude dynamics in terms of extended response functions: Invariant curves and Arnold tongues*, Commun. Nonlinear Sci. Numer. Simul., 81 (2020), 105008.

[9] O. Castejón, A. Guillamon, and G. Huguet, *Phase-amplitude response functions for transient-state stimuli*, J. Math. Neurosci., 3 (2013), 13.

[10] P. Constantin, C. Foias, B. Nicolaenko, and R. Temam, *Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations*, Appl. Math. Sci. 70, Springer, New York, 1989.

[11] I. Dasanayake and J. S. Li, *Optimal design of minimum-power stimuli for phase models of neuron oscillators*, Phys. Rev. E (3), 83 (2011), 061916.

[12] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997, https://doi.org/10.1137/1.9781611971446.

[13] G. B. Ermentrout and D. H. Terman, *Mathematical Foundations of Neuroscience*, Interdiscip. Appl. Math. 35, Springer, New York, 2010.

[14] J. Feng and H. C. Tuckwell, *Optimal control of neuronal activity*, Phys. Rev. Lett., 91 (2003), 018101.

[15] C. Foias, G. R. Sell, and R. Temam, *Inertial manifolds for nonlinear evolutionary equations*, J. Differential Equations, 73 (1988), pp. 309–353.
[16] J. Guckenheimer, *Isochrons and phaseless sets*, J. Math. Biol., 1 (1975), pp. 259–273.
[17] 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.
[18] J. P. Hespanha, *Linear Systems Theory*, Princeton University Press, Princeton, NJ, 2018.
[19] M. H. Holmes, *Introduction to Perturbation Methods*, Texts Appl. Math. 20, Springer, New York, 2013.
[20] 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), e1005011.
[21] A. Iolov, S. Ditlevsen, and A. Longtin, *Stochastic optimal control of single neuron spike trains*, J. Neural Eng., 11 (2014), 046004.
[22] H. K. Khalil, *Nonlinear Systems*, Prentice Hall, Upper Saddle River, NJ, 2002.
[23] D. Kirk, *Optimal Control Theory*, Dover, New York, 1998.
[24] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, Springer-Verlag, Berlin, 1984.
[25] W. Kurebayashi, S. Shirasaka, and H. Nakao, *Phase reduction method for strongly perturbed limit cycle oscillators*, Phys. Rev. Lett., 111 (2013), 214101.
[26] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*, SIAM, Philadelphia, 2016, https://doi.org/10.1137/1.9781611974508.
[27] M. D. Kvalheim and S. Revzen, *Existence and uniqueness of global Koopman eigenfunctions for stable fixed points and periodic orbits*, Phys. D, 425 (2021), 132959.
[28] B. Letson and J. E. Rubin, *A new frame for an old (phase) portrait: Finding rivers and other flow features in the plane*, SIAM J. Appl. Dyn. Syst., 17 (2018), pp. 2414–2445, https://doi.org/10.1137/18M1186617.
[29] B. Letson and J. E. Rubin, *LOR for analysis of periodic dynamics: A one-stop shop approach*, SIAM J. Appl. Dyn. Syst., 19 (2020), pp. 58–84, https://doi.org/10.1137/19M1258529.
[30] W. Lohmiller and J. J. E. Slotine, *On contraction analysis for non-linear systems*, Automatica, 34 (1998), pp. 683–696.
[31] A. M. Lozano and N. Lipsman, *Probing and regulating dysfunctional circuits using deep brain stimulation*, Neuron, 77 (2013), pp. 406–424.
[32] B. Lusch, N. J. Kutz, and S. L. Brunton, *Deep learning for universal linear embeddings of nonlinear dynamics*, Nat. Commun., 9 (2018), pp. 1–10.
[33] A. Mauroy and I. Mezić, *Global stability analysis using the eigenfunctions of the Koopman operator*, IEEE Trans. Automat. Control, 61 (2016), pp. 3356–3369.
[34] A. Mauroy, I. Mezić, 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.
[35] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.
[36] C. D. Meyer and G. W. Stewart, *Derivatives and perturbations of eigenvectors*, SIAM J. Numer. Anal., 25 (1988), pp. 679–691, https://doi.org/10.1137/0725041.
[37] I. Mezić, *Analysis of fluid flows via spectral properties of the Koopman operator*, Annu. Rev. Fluid Mech., 45 (2013), pp. 357–378.
[38] I. Mezić, *Spectrum of the Koopman operator, spectral expansions in functional spaces, and state-space geometry*, J. Nonlinear Sci., 30 (2020), pp. 2091–2145.
[39] J. Moehlis, E. Shea-Brown, and H. Rabitz, *Optimal inputs for phase models of spiking neurons*, ASME J. Comput. Nonlinear Dyn., 1 (2006), pp. 358–367.
[40] B. Monga, D. Wilson, T. Matchen, and J. Moehlis, *Phase reduction and phase-based optimal control for biological systems: A tutorial*, Biol. Cybernet., 113 (2019), pp. 11–46.
[41] A. Nabi and J. Moehlis, *Time optimal control of spiking neurons*, J. Math. Biol., 64 (2012), pp. 981–1004.
[42] A. Nandi, H. Schättler, J. T. Ritt, and S. Ching, *Fundamental limits of forced asynchronous spiking with integrate and fire dynamics*, J. Math. Neurosci., 7 (2017), pp. 1–37.
[43] E. Ott and T. M. Antonsen, *Low dimensional behavior of large systems of globally coupled oscillators*, Chaos, 18 (2008), 037113.

[44] Y. Park and B. Ermentrout, *Weakly coupled oscillators in a slowly varying world*, J. Comput. Neurosci., 40 (2016), pp. 269–281.
[45] Y. Park, K. M. Shaw, H. J. Chiel, and P. J. Thomas, *The infinitesimal phase response curves of oscillators in piecewise smooth dynamical systems*, European J. Appl. Math., 29 (2018), pp. 905–940.
[46] D. Pérez-Palau, J. J. Masdemont, and G. Gómez, *Tools to detect structures in dynamical systems using Jet transport*, Celestial Mech. Dynam. Astronom, 123 (2015), pp. 239–262.
[47] B. Pietras and A. Daffertshofer, *Network dynamics of coupled oscillators and phase reduction techniques*, Phys. Rep., 819 (2019), pp. 1–105.
[48] K. Pyragas, O. Popovych, and P. Tass, *Controlling synchrony in oscillatory networks with a separate stimulation-registration setup*, Europhys. Lett., 80 (2007), 40002.
[49] A. J. Roberts, *Appropriate initial conditions for asymptotic descriptions of the long term evolution of dynamical systems*, J. Austral. Math. Soc. Ser. B., 31 (1989), pp. 48–75.
[50] M. Rosenblum and A. Pikovsky, *Controlling synchronization in an ensemble of globally coupled oscillators*, Phys. Rev. Lett., 92 (2004), 114102.
[51] C. W. Rowley, I. Mezic, S. Bagheri, P. Schlatter, and D. S. Henningson, *Spectral analysis of nonlinear flows*, J. Fluid Mech., 641 (2009), pp. 115–127.
[52] J. Rubin and D. Terman, *High frequency stimulation of the subthalamic nucleus eliminates pathological thalamic rhythmicity in a computational model*, J. Comput. Neurosci., 16 (2004), pp. 211–235.
[53] J. A. Sanders, F. Verhulst, and J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems*, 2nd ed., Springer-Verlag, New York, 2007.
[54] D. D. Vecchio and J. J. E. Slotine, *A contraction theory approach to singularly perturbed systems*, IEEE Trans. Automat. Control, 58 (2012), pp. 752–757.
[55] X. J. Wang and G. Buzsáki, *Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model*, J. Neurosci., 16 (1996), pp. 6402–6413.
[56] K. C. A. Wedgwood, K. K. Lin, R. Thul, and S. Coombes, *Phase-amplitude descriptions of neural oscillator models*, J. Math. Neurosci., 3 (2013), 2.
[57] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, 2nd ed., Springer, New York, 2003.
[58] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, *A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition*, J. Nonlinear Sci., 25 (2015), pp. 1307–1346.
[59] D. Wilson, *Isostable reduction of oscillators with piecewise smooth dynamics and complex Floquet multipliers*, Phys. Rev. E (3), 99 (2019), 022210.
[60] D. Wilson, *Phase-amplitude reduction far beyond the weakly perturbed paradigm*, Phys. Rev. E (3), 101 (2020), 022220.
[61] D. Wilson, *Data-driven inference of high-accuracy isostable-based dynamical models in response to external inputs*, Chaos, 31 (2021), 063137.
[62] D. Wilson, *Degenerate isostable reduction for fixed-point and limit-cycle attractors with defective linearizations*, Phys. Rev. E (3), 103 (2021), 022211.
[63] D. Wilson, *Optimal control of oscillation timing and entrainment using large magnitude inputs: An adaptive phase-amplitude-coordinate-based approach*, SIAM J. Appl. Dyn. Syst., 20 (2021), pp. 1814–1843, https://doi.org/10.1137/20M1373165.
[64] D. Wilson, *An adaptive phase-amplitude reduction framework without $O(\epsilon)$ constraints on inputs*, SIAM J. Appl. Dyn. Syst., 21 (2022), pp. 204–230, https://doi.org/10.1137/21M1391791.
[65] D. Wilson and S. M. Djouadi, *Adaptive isostable reduction of nonlinear PDEs with time varying parameters*, IEEE Control Syst. Lett., 5 (2021), pp. 187–192.
[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.
[67] D. Wilson and J. Moehlis, *Optimal chaotic desynchronization for neural populations*, SIAM J. Appl. Dyn. Syst., 13 (2014), pp. 276–305, https://doi.org/10.1137/120901702.
[68] D. Wilson and J. Moehlis, *Extending phase reduction to excitable media: Theory and applications*, SIAM Rev., 57 (2015), pp. 201–222, https://doi.org/10.1137/140952478.
[69] D. Wilson and J. Moehlis, *Isostable reduction of periodic orbits*, Phys. Rev. E (3), 94 (2016), 052213.

[70] D. Wilson and J. Moehlis, *Isostable reduction with applications to time-dependent partial differential equations*, Phys. Rev. E (3), 94 (2016), 012211.
[71] A. Winfree, *The Geometry of Biological Time*, 2nd ed., Springer-Verlag, New York, 2001.
[72] E. Yeung, S. Kundu, and N. Hodas, *Learning deep neural network representations for Koopman operators of nonlinear dynamical systems*, in Proceedings of the 2019 IEEE American Control Conference, 2019, pp. 4832–4839.