



## PAPER

## Inferring oscillator's phase and amplitude response from a scalar signal exploiting test stimulation

Rok Cestnik\*, Erik T K Mau and Michael Rosenblum

Department of Physics and Astronomy, University of Potsdam, Karl-Liebknecht-Str. 24/25, D-14476 Potsdam-Golm, Germany

\* Author to whom any correspondence should be addressed.

E-mail: [rokcestn@uni-potsdam.de](mailto:rokcestn@uni-potsdam.de)**Keywords:** phase response, amplitude response, phase-isostable reduction, inference

## OPEN ACCESS

## RECEIVED

18 June 2022

## REVISED

22 November 2022

## ACCEPTED FOR PUBLICATION

29 November 2022

## PUBLISHED

14 December 2022

Original Content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.**Abstract**

The phase sensitivity curve or phase response curve (PRC) quantifies the oscillator's reaction to stimulation at a specific phase and is a primary characteristic of a self-sustained oscillatory unit. Knowledge of this curve yields a phase dynamics description of the oscillator for arbitrary weak forcing. Similar, though much less studied characteristic, is the amplitude response that can be defined either using an ad hoc approach to amplitude estimation or via the isostable variables. Here, we discuss the problem of the phase and amplitude response inference from observations using test stimulation. Although PRC determination for noise-free neuronal-like oscillators perturbed by narrow pulses is a well-known task, the general case remains a challenging problem. Even more challenging is the inference of the amplitude response. This characteristic is crucial, e.g. for controlling the amplitude of the collective mode in a network of interacting units—a task relevant to neuroscience. Here, we compare the performance of different techniques suitable for inferring the phase and amplitude response, particularly with application to macroscopic oscillators. We suggest improvements to these techniques, e.g. demonstrating how to obtain the PRC in case of stimuli of arbitrary shape. Our main result is a novel technique denoted by IPID-1, based on the direct reconstruction of the Winfree equation and the analogous first-order equation for isostable dynamics. The technique works for signals with or without well-pronounced marker events and pulses of arbitrary shape; in particular, we consider charge-balanced pulses typical in neuroscience applications. Moreover, this technique is superior for noisy and high-dimensional systems. Additionally, we describe an error measure that can be computed solely from data and complements any inference technique.

**1. Introduction**

Analysis and control of real-world oscillatory dynamics require inference of oscillator's parameters from observations. In particular, one can explore the system by applying a specifically designed perturbation and measuring the reaction. This paper discusses experiments extracting a self-sustained unit's phase and amplitude response. We concentrate on oscillators with weakly-stable limit cycles and stimulation with pulses of arbitrary shape.

In the first approximation in perturbation's strength, the oscillator's phase response is quantified by the phase sensitivity function or phase response curve (PRC) [1–4], a crucial characteristic of a self-sustained oscillator. Knowing the PRC one describes the oscillator's phase dynamics via the Winfree equation

$$\dot{\varphi} = \omega + Z(\varphi)p(t), \quad (1)$$

where  $\varphi$  is the phase,  $\omega$  is the natural frequency,  $Z(\varphi)$  is the PRC, and  $p(t)$  is the external perturbation. Equation (1) predicts, e.g. the system's response to an arbitrary stimulus or synchronization of the oscillator by an external force. A description of oscillatory systems in terms of PRCs is useful in various fields, e.g. in computational neuroscience [5].

For known oscillator equations, one can obtain PRC by solving the adjoint problem [6]. Estimation of PRC in an experiment is a less trivial task. The standard approach is to apply weak, narrow pulses at different phases and measure the phase shifts caused by stimulation [2, 4]. Indeed, if the pulse at phase  $\varphi$  is Dirac's delta function, then the induced phase shift equals precisely  $Z(\varphi)$ . It is easy to implement this approach for neuronal oscillators, where the distance between the spikes gives the variation of the period and hence, the phase shift. It is not that easy to exploit this technique to investigate oscillators without well-pronounced marker events that can be assigned a specific phase value. Furthermore, for any oscillator, the applied pulses should reasonably approximate the Dirac's delta pulses, which is often impossible because stimulation of living tissue should fulfill specific criteria, namely to be charge-balanced.

A related, much more demanding problem is the inference of the amplitude response of an oscillator. This problem naturally arises, e.g. in model studies [7–27] of the clinical technique known as deep brain stimulation (DBS) [28–30]. DBS aims to modulate the brain rhythm by stimuli delivered through implanted microelectrodes. Thus, a question arises: stimulation at which phase results in the most significant change of the amplitude [25, 27, 31].

This paper critically assesses previously developed techniques and proposes two new approaches for estimating an oscillator's phase and amplitude response from observations, exploiting test pulses of arbitrary shapes. The first novel technique relies on a signal's instantaneous amplitude and phase and, hence, is in the spirit of traditional time series analysis. The second technique is model-based; it exploits recent advances in dynamical systems' description in terms of phases and isostable variables [32–34] and reconstructs the equations for phase and, for the first time, for the isostable variable. We exploit several test systems to probe our techniques' performance and compare them with those known in the literature. In particular, we concentrate on systems exhibiting signals without well-pronounced marker events, i.e. non-spiky signals, where the standard technique is inefficient.

The paper is organized as follows. In the rest of this section, we discuss the state-of-the-art and problem formulation in detail. In section 2 we present the tools for testing and comparing inference techniques. In section 3 we compare methods and their performance, including the novel reconstruction of the phase-isostable dynamics which is described in section 3.3. Section 4 summarizes and discusses the results. The appendix section introduces the phase-isostables representation and presents the details of the test systems and inference techniques.

### 1.1. State of the art and problems to be solved

Here, we specify the problems and introduce some notations.

#### 1.1.1. Estimating phase response in experiments

The traditional approach relies on measuring the phase shift  $\Delta\varphi$  evoked by a pulse applied at phase  $\varphi$ . However, in practice, the stimuli are not Dirac's delta pulses, so the inferred response function differs from  $Z(\varphi)$ . Therefore, we consider general pulses  $\mathcal{P}(t)$  of length  $\delta$ ,  $\mathcal{P}(t) = 0$  for  $t \notin [0, \delta]$ . In particular, we are interested in the charge-balanced stimuli that additionally fulfill the condition

$$\int_0^\delta \mathcal{P}(t) dt = 0. \quad (2)$$

For neuroscience applications, such stimuli are required to avoid charge accumulation in living tissue. We denote the response to non-Dirac stimuli as empirical PRC  $Z_{\mathcal{P}}(\varphi)$ . Thus, the first problem is to relate  $Z(\varphi)$  and  $Z_{\mathcal{P}}(\varphi)$  for a given  $\mathcal{P}(t)$ . In the following, we infer the oscillator's response by perturbing it with the pulse train

$$p(t) = \sum_k \mathcal{P}(t - t_k).$$

We specify the choice of stimulation times  $t_k$  in the test examples below. The phase shift caused by a stimulus also depends on the stimulus amplitude and hence, has to be normalized. For unipolar stimuli, it is natural to normalize by the integral  $f = \int_0^\delta \mathcal{P} dt$ , commonly referred to as the stimulus' 'action' [35]. For bipolar charge-balanced pulses the normalization is ambiguous; we choose  $f = \frac{1}{2} \int_0^\delta |\mathcal{P}| dt$ . Thus,  $Z_{\mathcal{P}} = \Delta\varphi/f$ .

The success of the standard technique is due to the presence of well-defined spikes, typical for neuronal oscillators. Consider now a system that exhibits nearly sinusoidal oscillation. Although no well-defined markers exist, threshold-crossing events can also determine the points of a constant phase value [36]. However, the error in phase estimation should be higher than in the case of neuron-like oscillators, especially in the presence of noise. Therefore, the natural idea is to avoid determining the marker events and estimate

the instantaneous phase. The first step in this direction has been done by Holt *et al* [25, 31]. They suggested fitting sine waves to the several oscillation periods before and after each stimulus. The phase shift caused by the stimulus is then easily obtained from these two sines. Duchet *et al* [27] approached this problem using the Hilbert transform (HT): they estimated the instantaneous phase before and after the stimulus and, in this way, computed the PRC [37]. Unfortunately, neither of the techniques published in [25, 27, 31] has been tested on an oscillatory model with known PRC; therefore, the performance of these techniques is unclear. Below, we test these techniques and propose a different, more precise approach.

### 1.1.2. Estimating amplitude response in experiments

The problem of amplitude response estimation appears, e.g. in the context of DBS, where the goal is to modulate the brain rhythm by weak pulses. Such modulation is possible only if the amplitude's perturbation decays slowly so that the effects of several pulses accumulate. For a dynamical model, it means that the oscillator's limit cycle is weakly stable, i.e. its largest Floquet multiplier is close to one (Floquet exponent close to 0). A popular model of brain rhythm generation is a large neuronal population exhibiting collective mode oscillation due to the synchronization of the population's elements. If the amplitude of the collective mode is not too large, i.e. the system is close to the synchronization transition point, then the limit cycle corresponding to the collective oscillation is weakly stable even if the cycles of individual neurons are strongly stable. This consideration motivates our analysis of the amplitude response using the models with a weakly stable cycle.

The first problem with the amplitude response is its definition. While the phase of a limit-cycle oscillator is unambiguously (up to an additive constant) defined via isochrons, there is no universal definition for the amplitude variable. A pragmatic approach [27] is to estimate the amplitude of an observed signal using the HT. Comparing the amplitude before and after the stimulus for different phases of stimulation, Duchet *et al* [27] in this way introduced the ARC.

Another approach to characterizing the amplitude variations exploits the notion of isostables [34, 38, 39]. In this framework, the Winfree equation (1) is complemented by an equation for the isostable variable  $\psi$  that can be interpreted as a deviation from the limit cycle:

$$\dot{\psi} = \kappa\psi + I(\varphi)p(t). \quad (3)$$

The real-valued parameter  $\kappa$  (Floquet exponent) determines the stability of the cycle and is negative for stable cycles. The function  $I(\varphi)$  quantifies the phase dependence of the stimulation's effect. We denote  $I(\varphi)$  as the isostable response curve, IRC. For an introduction to the phase-isostable representation, see appendix A. We remind that equations (1) and (3) represent the first-order approximation in the perturbation's strength.

This paper elaborates on the phase and amplitude response inference employing test stimulation. To evaluate the performance of different techniques, we test them on systems for which this response is known. First, we analyze the existing signal analysis techniques and propose some improvements. Next, we present the method for inference of the phase-isostable equations (1) and (3) for the first time.

## 2. Tools for testing and comparing inference techniques

In this section, we first suggest a simple approach for relating the response curve to stimuli of arbitrary shape  $Z_{\mathcal{P}}(\varphi)$  to the infinitesimal PRC  $Z(\varphi)$ . Then we design and present the test systems. And finally, we introduce the error measures we will be using.

### 2.1. Relation between empirical $Z_{\mathcal{P}}$ and infinitesimal $Z$ PRCs

From the infinitesimal PRC  $Z(\varphi)$ , one can compute the empirical PRC  $Z_{\mathcal{P}}(\varphi)$  as a response to an arbitrary stimulus  $\mathcal{P}$ . One simply has to evaluate the effective phase shift for the chosen stimulation  $\mathcal{P}$  at every phase by integrating the Winfree phase equation (1). For example, let us focus on evaluating the phase response for stimuli  $\mathcal{P}$  at a particular phase  $\varphi^*$ . One first needs to compute the phase in time for the duration of the pulse by solving the Winfree equation (1) using initial condition  $\varphi(t=0) = \varphi^*$ . Then the empirical PRC at the chosen phase corresponds to the phase shift that is induced by  $\mathcal{P}$  over the duration of the stimulation:

$$Z_{\mathcal{P}}(\varphi^*) = \frac{1}{f} \int_0^{\delta} Z(\varphi(t))\mathcal{P}(t) dt, \quad (4)$$

where  $f$  is the action of the pulse.

The inverse problem of how to determine the infinitesimal response  $Z(\varphi)$  when given the empirical one  $Z_{\mathcal{P}}(\varphi)$  and stimulation shape  $\mathcal{P}$ , is in general, much harder. However, for pulses with small action  $f$ , the

empirical response  $Z_{\mathcal{P}}(\varphi)$  is well approximated as the convolution of the infinitesimal curve  $Z(\varphi)$  and the pulse shape  $\mathcal{P}(t)$ . In Fourier space, convolution corresponds to multiplication, meaning the inverse operation is represented by division. One can, therefore, easily deconvolve an empirical response to the infinitesimal one by dividing the Fourier representations of  $Z_{\mathcal{P}}$  and  $\mathcal{P}$ ; we present the technique in more detail in appendix B. There we also present an algorithm that does not assume the smallness of  $f$ . The transformation  $Z_{\mathcal{P}} \rightarrow Z$  is generic and can be combined with any technique that infers an empirical response.

## 2.2. Test models

### 2.2.1. Two-dimensional test models

Our first test model is the Stuart–Landau (SL) oscillator. The reasons for using this system are threefold. First, the parameters of this system explicitly govern the limit cycle’s stability and shape of isochrons. Second, this system’s phase and amplitude response can be obtained analytically. Third, this system represents the normal form of the Hopf bifurcation and, therefore, serves as the model equation for the collective mode of an oscillator population close to the synchronization transition point. The system’s equations are:

$$\begin{aligned}\dot{x} &= \mu x - \eta y - (x^2 + y^2)(x - \alpha y) + \sigma \xi_x(t) + \cos \beta \cdot p(t), \\ \dot{y} &= \mu y + \eta x - (x^2 + y^2)(y + \alpha x) + \sigma \xi_y(t) + \sin \beta \cdot p(t),\end{aligned}\quad (5)$$

where the parameter  $\mu$  governs the limit cycle’s stability as it is directly related to its Floquet exponent by  $\kappa = -2\mu$  and  $\alpha$  is the non-isochronicity parameter. The frequency of the limit-cycle oscillation is given by  $\omega := \eta - \alpha\mu$ .  $\sigma$  is the noise strength, and  $\xi_{x,y}$  are two realizations of the Gaussian white noise with zero mean and unit variance.  $p(t)$  is the external perturbation specified in the test examples below, and parameter  $\beta$  determines the perturbation’s direction with respect to  $x$  and  $y$  variables. The solution of the noise-free and unperturbed SL system is a sine with the amplitude  $\sqrt{\mu}$ . For definiteness, throughout this paper, we assume that the perturbation is state-independent in the chosen Cartesian coordinates (5). Then, the SL system (5) has PRC (see, e.g. [40])

$$Z(\varphi) = -\frac{1}{\sqrt{\mu}} (\sin(\varphi - \beta) + \alpha \cos(\varphi - \beta)) \quad (6)$$

and IRC

$$I(\varphi) = \frac{2}{\sqrt{\mu}} \cos(\varphi - \beta). \quad (7)$$

For the SL system, we also define the empirical amplitude response  $A(\varphi) = R_e/R_s$ , where  $R_{s,e}$  are the values of the amplitude variable  $R$  immediately before and after the applied pulse. For a narrow unipolar pulse, this expression reads:

$$A(\varphi) = 1 + \frac{f \cos(\varphi - \beta)}{\sqrt{\mu}}, \quad (8)$$

where  $f$  is the pulse’s action. For the derivation of IRC and ARC for the SL model, and their interrelation, see appendix A.1.

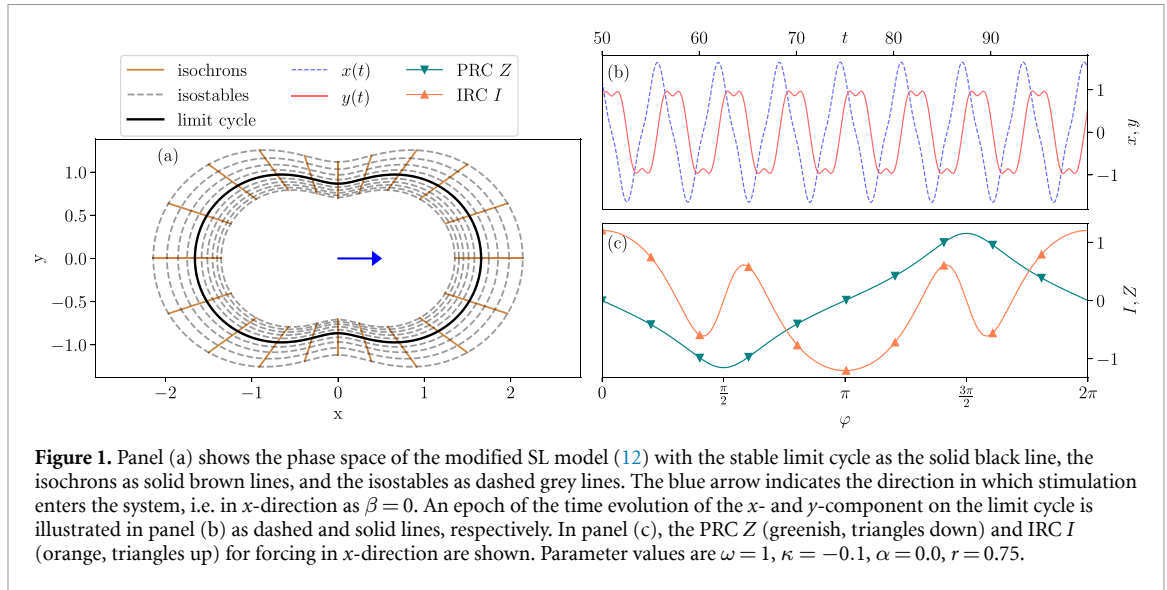
We modify the SL oscillator to obtain the second test model that provides a non-harmonic solution but is still analytically tractable. We denote this system as the modified SL oscillator (mSL). The system’s equations that we derive in appendix A.2 include the frequency of the limit cycle’s oscillation  $\omega$ , the Floquet exponent  $\kappa$ , and non-isochronicity  $\alpha$  as parameters. With perturbations, these equations read:

$$\begin{aligned}\dot{x} &= \omega [xC(x, y) - y] + \frac{\kappa}{2} [D(x, y) - 1] [x + \alpha (xC(x, y) - y)] + \sigma \xi_x(t) + \cos \beta \cdot p(t), \\ \dot{y} &= \omega [yC(x, y) + x] + \frac{\kappa}{2} [D(x, y) - 1] [y + \alpha (yC(x, y) + x)] + \sigma \xi_y(t) + \sin \beta \cdot p(t),\end{aligned}\quad (9)$$

where the functions  $C$  and  $D$  are:

$$C(x, y) = -\frac{2xy}{(r+2)x^2 + ry^2}, \quad (10)$$

$$D(x, y) = \frac{(x^2 + y^2)^2}{(r+2)x^2 + ry^2}, \quad (11)$$



and  $r$  is a positive parameter determining the shape of the limit cycle in state space. System (9) has IRC

$$I(\varphi) = 2 \frac{(r + 1) \cos(\varphi - \beta) + \cos(3\varphi - \beta)}{(r + 2 \cos^2(\varphi))^{\frac{3}{2}}} \quad (12)$$

and PRC

$$Z(\varphi) = -\frac{\sin(\varphi - \beta)}{\sqrt{r + 2 \cos^2(\varphi)}} - \alpha \frac{(r + 1) \cos(\varphi - \beta) + \cos(3\varphi - \beta)}{(r + 2 \cos^2(\varphi))^{\frac{3}{2}}}. \quad (13)$$

The system's dynamics are illustrated in figure 1.

### 2.2.2. High-dimensional test model

Our third test system is a simple model of macroscopic neuronal oscillation. We take a system of  $N$  globally coupled Bonhoeffer–van der Pol oscillators:

$$\dot{x}_k = x_k - x_k^3/3 - y_k + J_k + \varepsilon X + p(t), \dot{y}_k = 0.1(x_k - 0.8y_k + 0.7), \quad (14)$$

where  $k$  is the oscillator index,  $k = 1, \dots, N$ , and the term  $\varepsilon X$  describes the mean-field coupling, where  $X = N^{-1} \sum_k x_k$ . In the following, we take  $N = 1000$ . The oscillators' frequencies are determined by the parameter  $J_k$  that is Gaussian-distributed with mean 0.6 and standard deviation 0.1. The coefficient  $\varepsilon$  explicitly describes the interaction between the ensemble elements. We choose  $\varepsilon = 0.023$ ; for this parameter's value, the system exhibits collective chaos so that the mean-field oscillation is amplitude-modulated. Figure 2 presents the two-dimensional trajectory in coordinates  $X$  and  $Y = N^{-1} \sum_k y_k$  in (a) and  $X(t)$  in (b). With this test example, we imitate the natural variability of real-world signals.

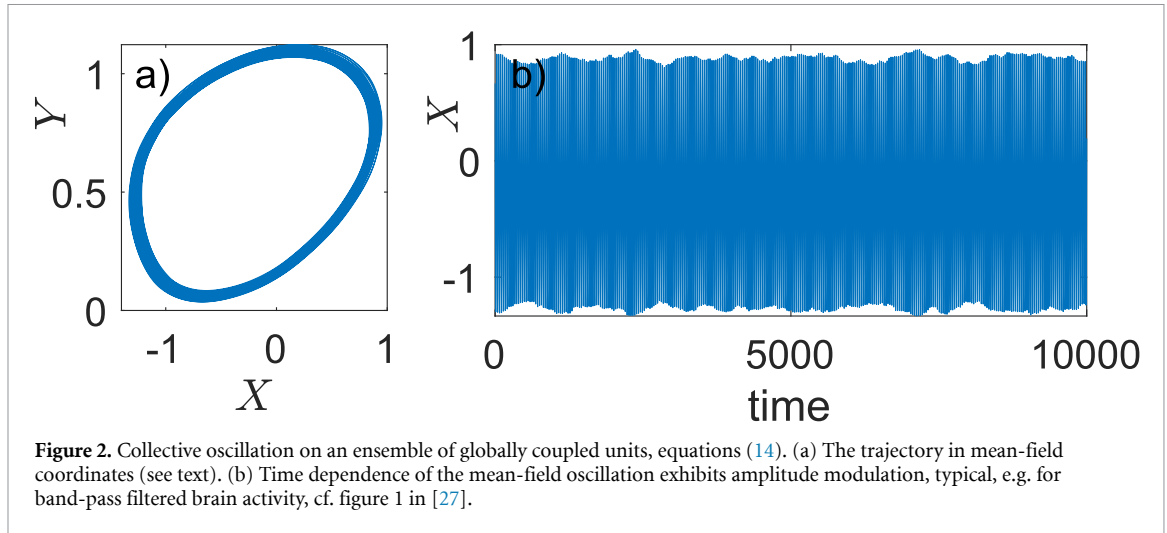
### 2.3. Error measures

We define two error measures with which we quantify the goodness of inference. If the true response curve is known from the theory, we evaluate the goodness of the inference by computing the normalized  $L^2$  distance between the inferred and true curves. We denote these errors with  $L_{Z,A,I}$ , where the subscript indicates which curves we are comparing (either PRC, ARC, or IRC). Thus, for error of, e.g. the PRC recovery, we obtain

$$L_Z = \sqrt{\frac{\langle (Z - Z^{\text{rec}})^2 \rangle}{\langle (Z - \langle Z \rangle)^2 \rangle}}, \quad (15)$$

where  $\langle \cdot \rangle = (2\pi)^{-1} \int_0^{2\pi} (\cdot) d\varphi$  denotes the average value of a function, and  $Z^{\text{rec}}$  is the inferred (recovered) PRC. The error value is  $L_{Z,A,I} = 0$  if the two curves coincide, it is of order 1 if the two curves are of the same order of magnitude but not similar and can also take higher values if one curve is significantly larger on average.

The other measure quantifies how well the inferred curve represents the dynamical model. Since the dynamics of phase and amplitude are distinct (see equations (1) and (3), respectively) the definition of the



corresponding error measures differs as well. We first introduce the error measure for the phase response and later in section 3.3.2 when introducing the method we also specify an analog for the isostable variable error. Note that the inferred phase response curve  $Z(\varphi)$  in conjunction with the natural frequency  $\omega$  can be used in equation (1) to reproduce  $\varphi(t)$  for a given realization of stimulation  $p(t)$  by means of numerical integration. In the first step, we threshold the observed signal to determine time events  $\tau_i$  that correspond to the same phase; for details, see appendix C. Without loss of generality, we set this phase to zero. Next, starting at  $\tau_i$  we reproduce the phase evolution up to the time  $\tau_{i+1}$ . In the ideal noise-free case, if the stimulation is weak and the inferred curve is exact, the reproduced phase  $\varphi(\tau_{i+1}) = \Phi_i$  equals  $2\pi$  (or, equivalently, zero since we consider the wrapped phase). In practice, this reproduced phase  $\Phi_i$  deviates from  $2\pi$ , and it is precisely this deviation that we use to quantify the inference's quality. We define the error of PRC reconstruction  $E_Z$  as the standard deviation of  $\Phi_i$  from  $2\pi$ :

$$E_Z = \langle (\Phi_i - 2\pi)^2 \rangle^{1/2}, \quad (16)$$

where  $\langle \cdot \rangle$  denotes averaging over the index  $i$ . The value of  $E_Z$  should be compared to a measure of the signal's irregularity. We evaluate the latter by a measure that is proportional to the standard deviation of the inter-event intervals  $T_i = \tau_{i+1} - \tau_i$  (instantaneous periods):

$$E_{Z0} = \frac{2\pi}{\langle T_i \rangle} \langle (T_i - \langle T_i \rangle)^2 \rangle^{1/2}. \quad (17)$$

A natural measure of the inference's quality is the ratio  $E_Z/E_{Z0}$  since it roughly describes how much of the signal's irregularity is explained by the inferred phase model. If one considers a PRC that is identically zero, this ratio equals one.

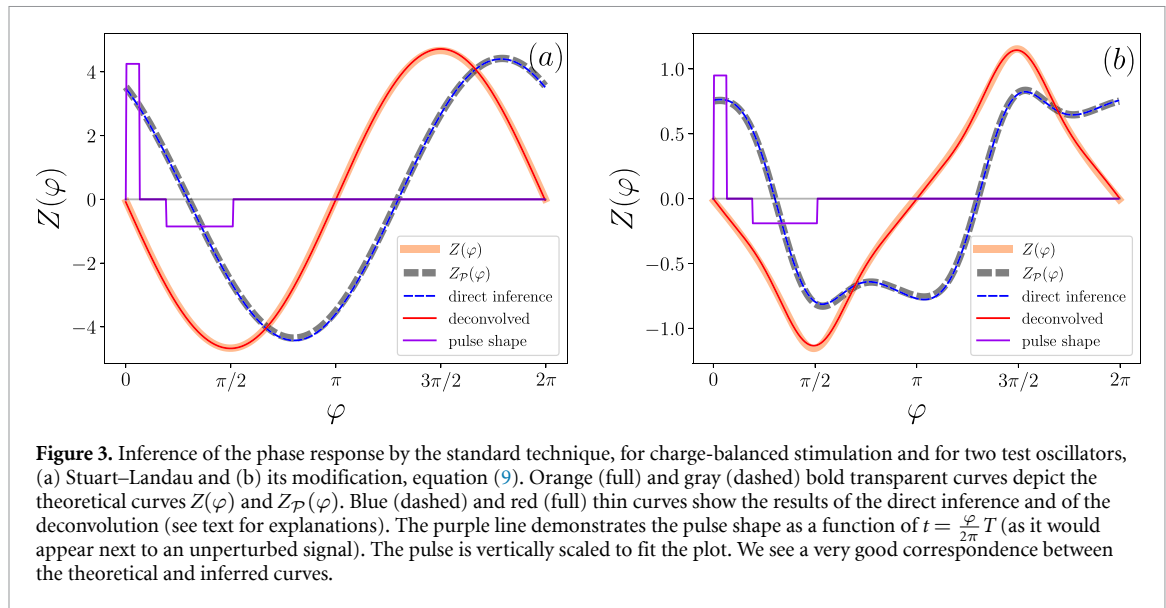
We emphasize an essential advantage of the error measure  $E_Z/E_{Z0}$ . The computation of the  $L^2$ -based measure equation (15) requires knowledge of the ground truth and, therefore, helps only in testing the techniques on model data from limit-cycle oscillators. In contrast, the error measure  $E_Z/E_{Z0}$  is obtained solely from our inferred model and the data itself. This makes it a helpful tool for experimental data analysis.

### 3. Methods and their performance

We formulate the inference problem as follows. Suppose we perturb the system by applying some known stimulation  $p(t)$  and measure the system's scalar output  $s(t)$ . We process  $s(t)$  to infer the PRC  $Z(\varphi)$  and IRC  $I(\varphi)$  or ARC  $A(\varphi)$  as specified below. Generally, we are free to construct  $p(t)$ , e.g. as a sequence of pulses. However, some constraints exist in specific settings, e.g. the pulses must be charge-balanced in neuroscience applications. In this section, we first review the standard PRC inference technique, followed by testing the performance of other methods in use. Here we also describe and test our approach to the problem.

#### 3.1. Standard approach (phase response only)

The standard technique is very efficient for neuronal oscillators exhibiting slow and fast motion. A spike corresponds to an epoch of fast motion, where the isochron density in the phase space is low. Therefore, since spike detection via threshold-crossing is weakly dependent on the threshold, the precision of the phase



**Figure 3.** Inference of the phase response by the standard technique, for charge-balanced stimulation and for two test oscillators, (a) Stuart–Landau and (b) its modification, equation (9). Orange (full) and gray (dashed) bold transparent curves depict the theoretical curves  $Z(\varphi)$  and  $Z_{\mathcal{P}}(\varphi)$ . Blue (dashed) and red (full) thin curves show the results of the direct inference and of the deconvolution (see text for explanations). The purple line demonstrates the pulse shape as a function of  $t = \frac{\varphi}{2\pi} T$  (as it would appear next to an unperturbed signal). The pulse is vertically scaled to fit the plot. We see a very good correspondence between the theoretical and inferred curves.

estimation is high. Consider now a system without slow-fast motion that exhibits nearly sinusoidal oscillation. To trace the variation of the period, we determine states with equal phases employing threshold-crossing [41], see [36]. For example, we apply a stimulus and look for the events with zero crossings from below. However, if the limit cycle is not strongly stable, observing such an event immediately after the stimulus' application at  $t_s$  does not suffice. Indeed, we have to wait until the system returns to the limit cycle. Thus, we take the instant  $\tau_n$  of the  $n$ th threshold-crossing event and compute

$$Z_{\mathcal{P}}(\varphi) = \frac{2\pi}{f} \frac{nT - (\tau_n - \tau_0)}{T}, \quad (18)$$

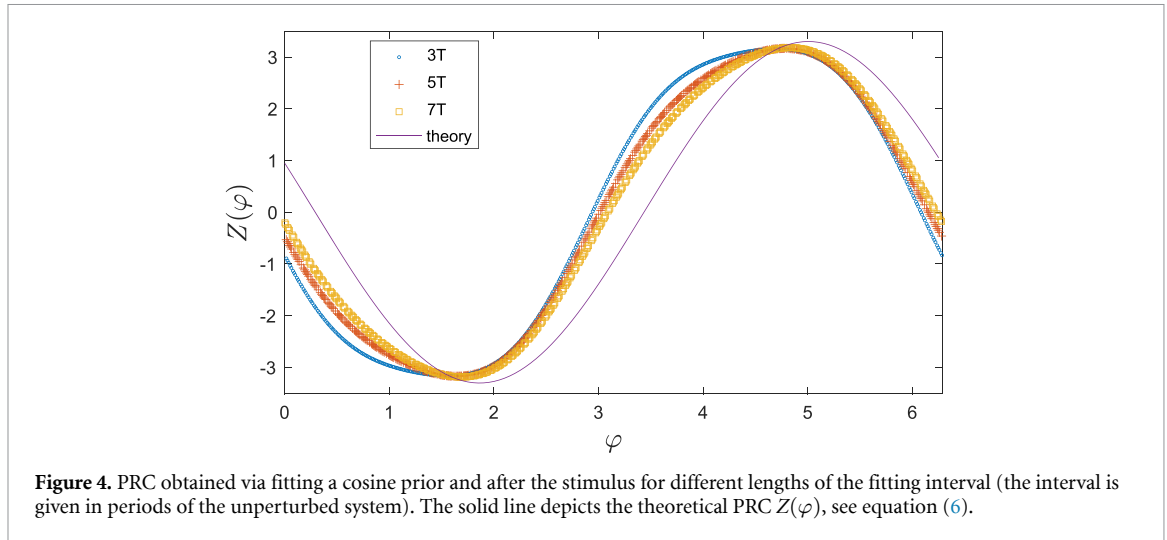
where  $\tau_0$  is the threshold-crossing event preceding the stimulus, and the phase of the stimulus application is  $\varphi = 2\pi(t_s - \tau_0)/T$  where  $T$  is the natural period. The choice of  $n$  depends on the stability of the limit cycle, quantified with the Floquet exponent  $\kappa$ ; as a rule of thumb, we suggest  $n\kappa T \gg 1$ . However, in practice,  $\kappa$  is unknown, and  $n$  shall be chosen by trial and error. To obtain the function  $Z_{\mathcal{P}}(\varphi)$ , we repeat the perturbation  $\mathcal{P}$  for different  $\varphi$ . To this end, we either choose random intervals between stimuli or stimulate periodically with the period incommensurate with  $T$ . In any case, the stimulation interval shall be large enough to ensure observation of  $n$  threshold-crossing events after each applied pulse.

We illustrate the inference via the standard technique in figure 3 where we consider a charge-balanced stimulation and infer the response for the two test models presented earlier, see equations (5) and (9). For the SL system (5) the parameter values are:  $\omega = 1$ ,  $\kappa = -0.1$ ,  $\alpha = -0.3$  and for the mSL system (9) they are  $\omega = 1$ ,  $\kappa = -0.1$ ,  $\alpha = 0$ ,  $r = 0.75$ . There is no noise,  $\sigma = 0$ . For the observable, we take the  $x$  variable, and  $n = 3$  threshold-crossing events after stimulation were considered for the phase shift determination. The charge-balanced pulse used has a short interval of positive stimulation (duration 0.2), followed by an interval without stimulation (duration 0.4), and then a longer interval of negative stimulation (duration 1.0). The amplitude of stimulation depends on the oscillator, the action  $f = \frac{1}{2} \int |\mathcal{P}| dt$  is 0.01 for SL and 0.07 for mSL, see appendix section D for details on perturbation generation. We show the following curves: (1) theoretical PRC  $Z(\varphi)$  according to equations (6) and (13), (2) theoretical response to charge-balanced pulses  $Z_{\mathcal{P}}$ —the ground truth, (3) inferred empirical PRC  $Z_{\mathcal{P}}$  obtained via equation (18) ('direct inference') [42], and (4) the result of the deconvolution of the empirical PRC  $Z_{\mathcal{P}}$  as described in appendix B ('deconvolved'). For this simple case, we see a nearly perfect reconstruction of the effective PRC  $Z_{\mathcal{P}}$  (compare blue and gray curves in figure 3) as well as a successful deconvolution to the infinitesimal curve (compare orange and red). Since we used bipolar pulses, the empirical PRC resembles the negative derivative of the infinitesimal curve (compare gray with orange).

We emphasize that there is no analog to the standard technique for inferring the amplitude response.

### 3.2. Inferring response measuring instantaneous phase and amplitude

The idea of the approach is straightforward. Suppose we perturb the oscillator by a sequence of pulses at instants  $t_k$  and measure the system's output  $s(t)$ . Computing the instantaneous phase and amplitude of  $s(t)$  before and after each pulse, we obtain the phase shift and amplitude variation as functions of the



**Figure 4.** PRC obtained via fitting a cosine prior and after the stimulus for different lengths of the fitting interval (the interval is given in periods of the unperturbed system). The solid line depicts the theoretical PRC  $Z(\varphi)$ , see equation (6).

stimulation's phase. To obtain a reasonable estimation of these functions, we have to apply the pulses at different phases—we ensure this by either using stimulation with a period incommensurate with that of the unperturbed oscillator or by choosing random  $t_k$ .

Let the instantaneous phase and amplitude immediately before and after a finite-width stimulus be  $\varphi_{s,e}$  and  $a_{s,e}$ , respectively, where indices  $s$  and  $e$  stand for 'start' and 'end'. From these quantities, we obtain the empirical PRC  $Z_{\mathcal{P}}(\varphi_s) = (\varphi_e - \varphi_s - \omega\delta)/f$ , where  $\delta$  is the pulse's width,  $\omega$  is the frequency of the unperturbed oscillation, and  $f$  is the normalization factor. Next, following Duchet *et al* [27], we introduce the empirical amplitude response curve (ARC) as  $A(\varphi) = a_e/a_s$ , to be distinguished from IRC. Below, we explore the performance of three different techniques.

### 3.2.1. Fitting the signal by a sine (phase response only)

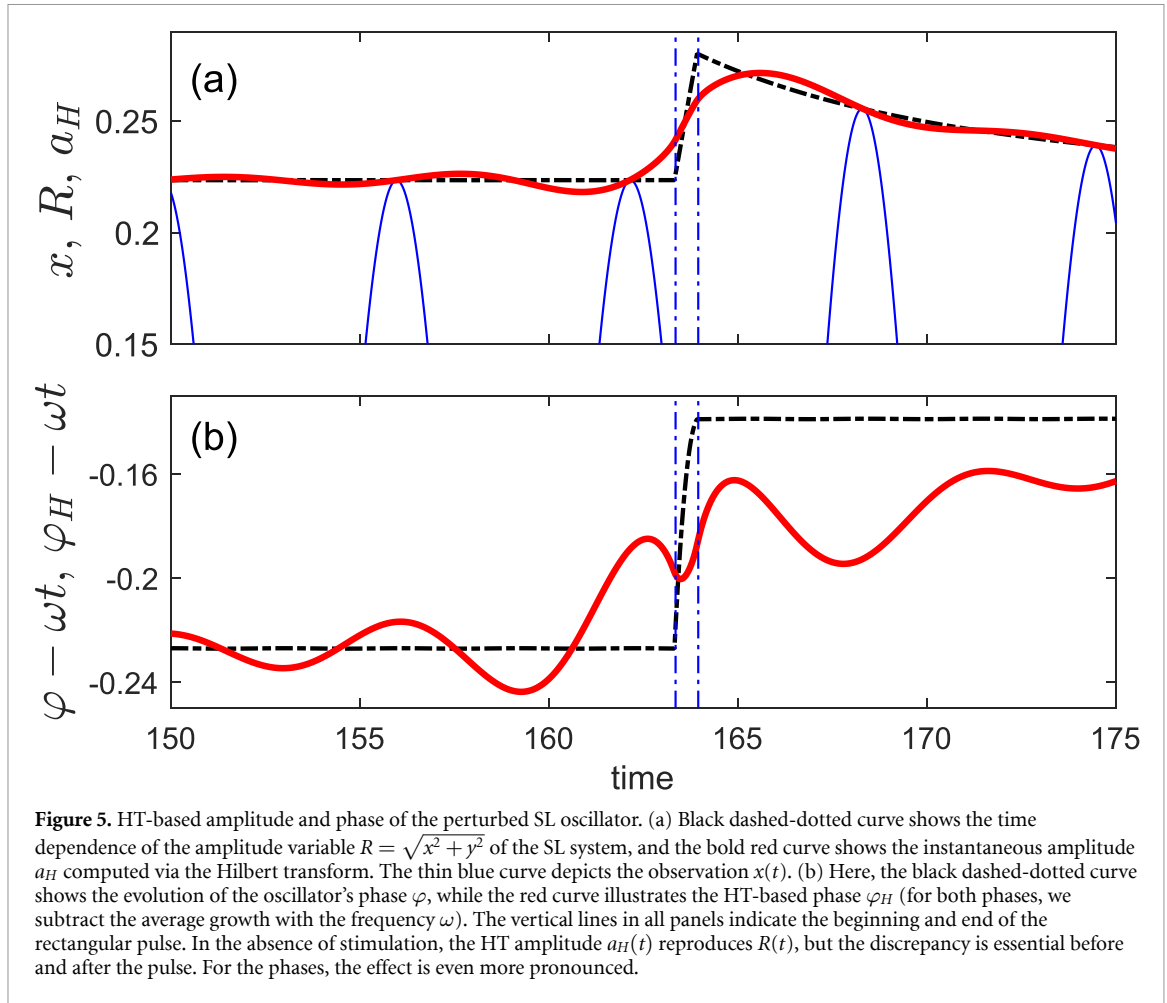
Holt *et al* [25, 31] suggested fitting a harmonic to several periods of the signal  $s(t)$  before the stimulus and another harmonic to an interval of the same length  $T_{\text{fit}}$  after the stimulus. Then, they exploited the Fourier Transform of both fitted functions to find the phase shift evoked by the stimulus. Obviously, this technique does not provide information on the amplitude response. Indeed, the perturbation in the amplitude normally decays rather quickly and cannot be captured by a fit over several oscillation periods. We illustrate the performance of this technique on the test model (5) in figure 4, computing the phase shift in the time domain. Suppose the pulse of length  $\delta$  occurs at  $t_s$ . Let the fitted functions be  $s(t) \approx a_1 \cos[\omega(t - (t_s - T_{\text{fit}})) + \chi_1]$  for  $t_s - T_{\text{fit}} < t < t_s$  and  $s(t) \approx a_2 \cos[\omega(t - (t_s - T_{\text{fit}})) + \chi_2]$  for  $t_s + \delta < t < t_s + \delta + T_{\text{fit}}$ . The phase of the second cosine at  $t = t_s + \delta$  is  $\varphi = \omega(T_{\text{fit}} + \delta) + \chi_2$ . If there were no stimulus, the phase of the first cosine at this point would be  $\varphi = \omega(\delta + T_{\text{fit}}) + \chi_1$ . Hence, the phase shift is  $\Delta\varphi = \chi_2 - \chi_1$  [43].

Naturally, this technique applies to the signals close to harmonic oscillation. Therefore, we test it on the SL oscillator only. We use the same parameters as used to illustrate the standard technique, see figure 3. For simplicity, we use unipolar rectangular pulses with amplitude 0.1 and width  $\delta = 0.03$ , therefore having action  $f = 0.003$ . Such a stimulus is a good approximation for the delta-pulse, hence, we can expect a good correspondence of the inferred PRC with the theoretical one. However, the results summarized in figure 4 demonstrate a limited precision of the approach. As another drawback of the technique, we mention that, since fitting requires several periods before and after the stimulus, the stimulation pulses shall be relatively rare, and hence, the time interval required for PRC estimation is rather long.

### 3.2.2. Estimating phase and amplitude using the Hilbert transform

A typical way to obtain the narrow-band signal's instantaneous phase and amplitude is to use the HT. Duchet *et al* [27] exploited HT to compute the phase and amplitude response of tremor oscillation to DBS of the thalamus. They took the instantaneous phase and amplitude immediately before and after the stimulus for this computation [44].

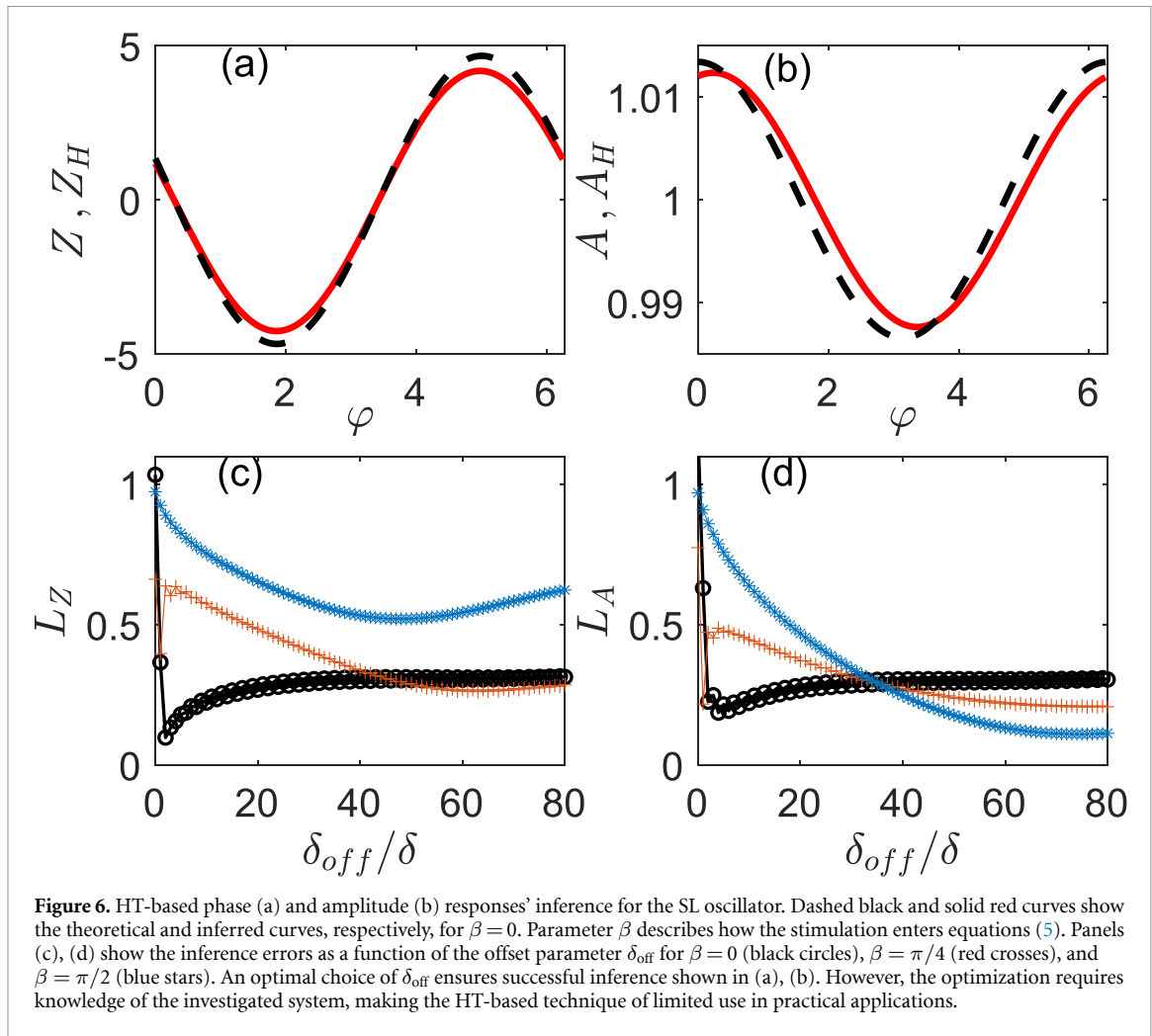
However, being a nonlocal operation, the HT is not suitable for measuring response to a pulse, e.g. the HT shows the systems' reaction even *before* the stimulus begins. We illustrate this property of the HT by perturbing the SL oscillator with a rectangular pulse. For a better visibility, we choose a wide pulse (amplitude 0.1,  $\delta = 0.6$ ). The parameters of the SL system are the same as before. Figure 5 illustrates the results, see also figure 3.5 in [45]. We see that the instantaneous Hilbert-based amplitude immediately before the pulse and immediately after it strongly deviates from the actual value. This deviation is of the same order



of magnitude as the amplitude's change due to the pulse. We observe similar behavior for the instantaneous phases. In summary, we shall interpret the HT-based phase and amplitude with caution in this context.

We suggest improving the performance of the HT-based inference in the following way. We neglect the instantaneous phases in the intervals  $(t_s - \delta_{\text{off}}, t_s)$  and  $(t_s + \delta, t_s + \delta + \delta_{\text{off}})$ , where  $t_s$  is the instant of the stimulus's onset and  $\delta_{\text{off}}$  is the offset time. Next, we obtain the phase  $\varphi_s$  by extrapolating the linear fit of  $\varphi(t)$  over the interval  $(t_s - \delta_{\text{off}} - \delta_{\text{fit}}, t_s - \delta_{\text{off}})$  to the instant  $t_s$ ; the parameter  $\delta_{\text{fit}}$  is the length of the fitting interval. Note that fitting requires the unwrapped phase. Similarly, extrapolating the fit over the interval  $(t_s + \delta + \delta_{\text{off}}, t_s + \delta + \delta_{\text{off}} + \delta_{\text{fit}})$  to  $t_s + \delta$  yields  $\varphi_e$ . Moreover, linear fit provides the frequency  $\omega$ . Hence, we compute  $Z_{\mathcal{P}}(\varphi_s)$ . The amplitude response (ARC) we compute as  $A(\varphi_s) = a_H(t_s + \delta + \delta_{\text{off}}) / a_H(t_s - \delta_{\text{off}})$ . We emphasize that the algorithm has two parameters,  $\delta_{\text{off}}$  and  $\delta_{\text{fit}}$ .

We demonstrate the performance of the described algorithm perturbing the SL system (5) by rectangular pulses, taking the inter-pulse interval incommensurate with the natural period  $T$ —in this way, we ensure that stimulation occurs at different phases. We set the observable  $s(t) = x(t)$  and first choose  $\beta = 0$ . Other parameters are the same as in the illustration of the standard and sine-fitting techniques. We illustrate the results in figures 6(a) and (b). Here, we show the theoretical curves for  $Z(\varphi)$ ,  $A(\varphi)$  according to equations (7, 8) along with inferred characteristics  $Z_H(\varphi)$ ,  $A_H(\varphi)$ . To compute the latter, we choose  $\delta_{\text{fit}} = 50\delta$  and  $\delta_{\text{off}} = 2\delta$  and perform an 8th-order Fourier fit of obtained points. Numerical tests show that although the results are not very sensitive to the choice of  $\delta_{\text{fit}}$ , the choice of  $\delta_{\text{off}}$  is crucial. We demonstrate this in figures 6(c) and (d) by showing the dependence of the error of inference on  $\delta_{\text{off}}$ . We compute this error according to equation (15). We see, that a proper choice of the offset  $\delta_{\text{off}}$  essentially affects the inference; the reasonable results in figures 6(a) and (b) are due to the optimal value  $\delta_{\text{off}}/\delta = 2$ . Unfortunately, we do not see any practical way to choose  $\delta_{\text{off}}$  when the true curve is unknown [46]. Moreover, the results strongly depend on how the stimulation enters the system's equations. To demonstrate this, we plot in figures 6(c) and (d) the corresponding curves for  $\beta = \pi/4$  and  $\beta = \pi/2$ . We see that, generally, the error of inference is not small. Summarizing this example, we say that the HT-based phase and amplitude response inference generally yields imprecise results. Though one can use this approach to obtain some empirical measures of the response, see [27], these inference results may be loosely related to the theory.

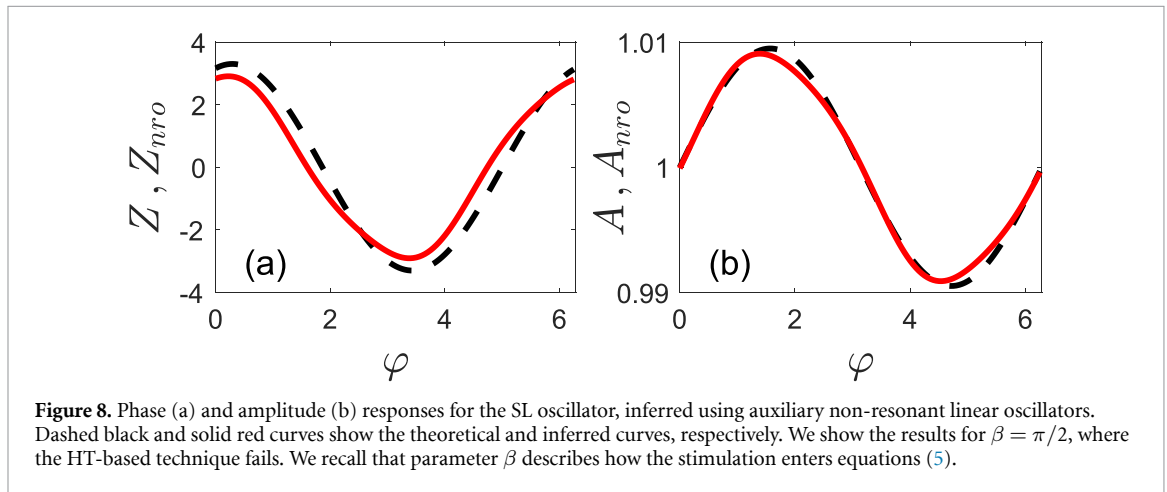
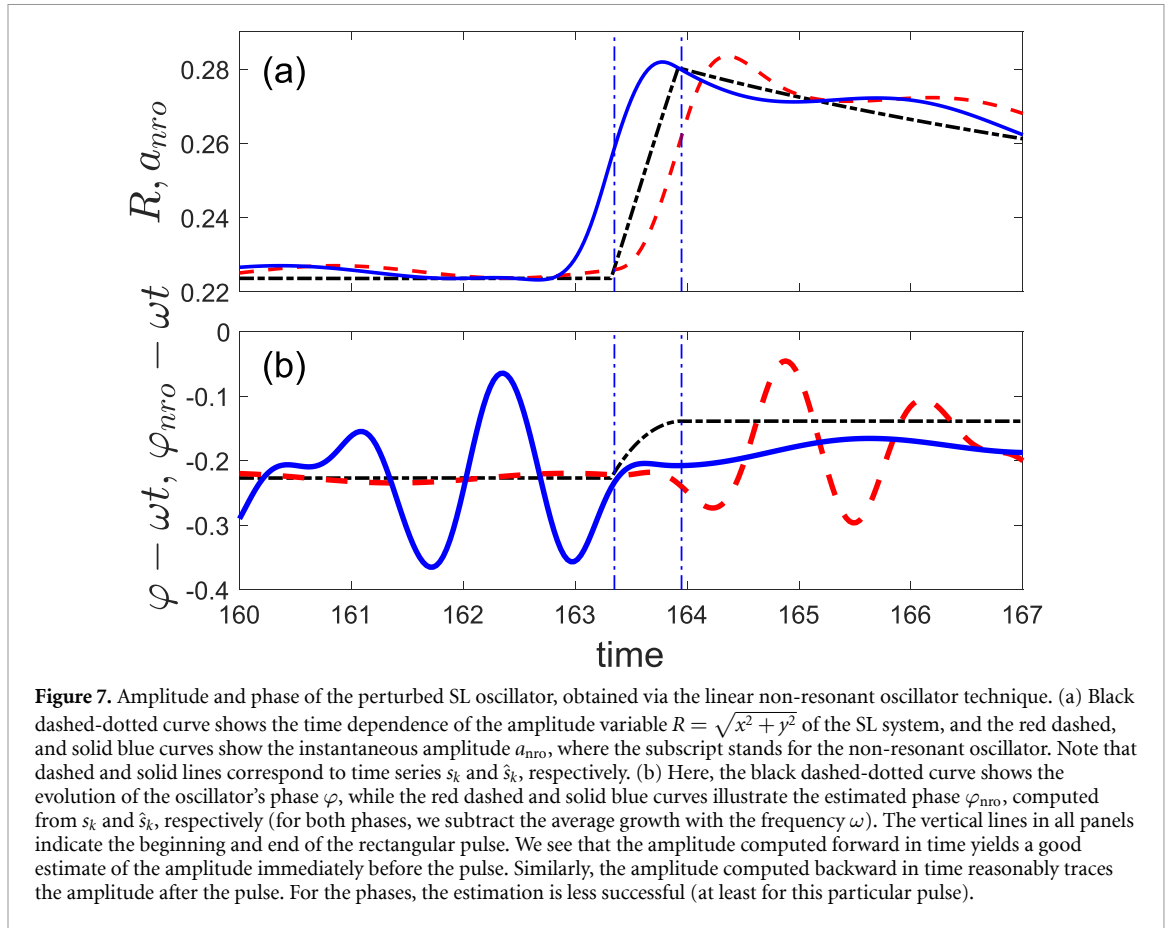


Additionally, we mention that HT serves merely as a signal embedding technique, yielding the angle variable or protophase; see [47, 48] for a discussion. A protophase generally depends on the embedding; so, e.g. the Hilbert-based protophase does not necessarily coincide with the angle variable in the  $x, y$  plane. Although protophases and the true (asymptotic) phase provide the same average frequencies, they generally differ microscopically, i.e. on a time scale smaller than the period. This difference is due to the non-uniform rotational velocity of the protophase.

### 3.2.3. Estimating phase and amplitude using a virtual auxiliary oscillator

Finally, we adapted the technique [49] for real-time estimation of phase and amplitude for our purpose. This technique exploits two virtual linear oscillators—one for phase and one for amplitude determination—to yield a causal estimation. Namely, we use the signal  $s(t)$  as an input to a damped oscillator  $\ddot{x} + \alpha_{a,\varphi}\dot{x} + \eta^2x = s(t)$ . We choose the oscillators' frequencies  $\eta$  to be much larger than the characteristic frequency  $\nu$  of  $s(t)$  so that the systems are far from resonance and the response weakly depends on  $\nu$ . Next, we take the damping parameters  $\alpha_\varphi$  and  $\alpha_a$  for the phase and amplitude measurement, to ensure a simple relation between the phase  $\varphi(t) = \arctan(-\dot{x}/\nu x)$  and amplitude  $\sqrt{x^2 + (\dot{x}/\nu)^2}$  of the forced oscillation and those of the investigated signal. The implementation is simple and boils down to the numerical integration of the linear oscillator's equation driven by a signal given at discrete time points. For details, see appendix E.

We expect this causal approach to yield a precise estimation of the amplitude and phase *before* the stimulus. Immediately after the stimulus, the estimation is poor due to transients. Indeed, the approach implies that the oscillation with the oscillator's frequency  $\eta$  decays, and only the oscillation with the frequency of the input  $s(t)$  remains. We suggest the following solution to this problem. We compute the amplitude/phase twice, first for the original signal and then for the time-inverted one. Namely, we flip the time series  $s_k$  so that the time 'runs' backward, i.e. from the  $N$ -point time series  $s_k, k = 1, 2, \dots, N$ , we construct the flipped series  $\hat{s}_1 = s_N, \hat{s}_2 = s_{N-1}, \dots, \hat{s}_N = s_1$ . Next, we compute the amplitude of  $\hat{s}_k$  via the same



algorithm and flip it in time to obtain the flipped amplitude denoted by  $\hat{a}(t)$ .  $\hat{a}(t)$  is a precise estimation for the time interval after the stimulus, while the transients corrupt the estimation before the stimulus. However, since the phase changes its sign with the time inversion, we must reflect the obtained phase,  $\hat{\varphi}(t) \rightarrow -\hat{\varphi}(t)$ . We illustrate this algorithm in figure 7; parameters of the virtual oscillators are  $\eta = 5$ ,  $\alpha_a = 6$ ,  $\alpha_\varphi = 0.2$ .

A remark is in order. The linear oscillator used to compute the phase has a smaller damping parameter than the oscillator for the amplitude estimation. Correspondingly the transients in the phase measurement are essentially longer than in the case of the amplitude measurement, which makes the phase measurement less precise. We present the inference's results for the SL model in figure 8. Additionally, just like any approach relying on estimating small differences in the angle variable, it generally suffers significant errors because it does not include information on local isochrons.

### 3.3. Inferring response curves by reconstructing the first order phase-isostable dynamics: the IPID-1 technique

Here, we describe a method for inferring the first-order phase-isostable dynamics from observations by fitting the model equations (1), (3). We name the method IPID-1 standing for ‘Inferring Phase–Isostable Dynamics of order 1’. We carry out the procedure in two steps. First, we infer the PRC and instantaneous phase by adapting the algorithm introduced in [50] for the case of pulse stimulation. Next, we use the inferred phase to reconstruct the isostable dynamics. As in the previous sections, we assume that a scalar signal  $s(t)$  and the perturbation  $p(t)$  are known.

#### 3.3.1. Inference of the phase response

The key step in the inference is determining time instants  $\tau_i$  corresponding to the same asymptotic phase  $\varphi$ ; these instants must be extracted from the observed signal  $s(t)$ . Sometimes the choice of such events is obvious, e.g. in the case of a spiky signal where spikes indicate the same phase. In general one considers threshold-crossing events,  $s(\tau_i) = s_{\text{thr}}$ , choosing one crossing per period, e.g. with an additional condition  $\frac{d}{dt}s(\tau_i) > 0$ . The choice of the threshold  $s_{\text{thr}}$  is important. The proper thresholding should closely match the crossing of a local isochron, see appendix C for details. Since isochrons make a full rotation over a limit cycle, at least two such thresholds exist for any scalar signal—we find appropriate thresholds with a direct search, seeing which threshold yields the best fit in terms of the error (16).

The core of the inference is fitting the Winfree phase equation (1) integrated over individual periods, determined as the time interval between two threshold-crossing events,  $[\tau_i, \tau_{i+1}]$ :

$$2\pi = \omega(\tau_{i+1} - \tau_i) + \int_{\tau_i}^{\tau_{i+1}} Z(\varphi(t))p(t) dt. \quad (19)$$

The left-hand side equals  $2\pi$  due to the definition of events  $\tau_i$  having the same phase. On the right-hand side, we approximate the PRC by a finite Fourier series of order  $N_F$ :  $Z(\varphi) = \sum_{n=0}^{N_F} [z_n^{\cos} \cos(n\varphi) + z_n^{\sin} \sin(n\varphi)]$ . By interchanging the order of integration and summation, we obtain from equation (19) a linear system for the unknown Fourier coefficients  $z_n$  and frequency  $\omega$ . Since we have as many equation (19) as there are periods of the observed signal, for a sufficiently long data set we have more equations than unknowns. Thus, we can solve the linear system, e.g. by least-squares minimization. The coefficients  $\int \cos(n\varphi)p(t) dt$ ,  $\int \sin(n\varphi)p(t) dt$ , however, cannot be evaluated yet because the phase  $\varphi(t)$  is not known to us *a priori*. We overcome this with an iterative procedure; by first approximating the phase to obtain an approximate solution, and then exploiting this solution to improve the phase estimate [51].

Approximating the phase initially as linearly growing between the events,  $\varphi^{(0)}(t) = 2\pi \frac{t - \tau_i}{\tau_{i+1} - \tau_i}$ , we obtain the first-approximation solution of equation (19), namely  $\omega^{(1)}$  and  $Z^{(1)}(\varphi)$ . (The superscripts denote the iteration number.) Then, we obtain the next approximation of the phase by integrating Winfree equation (1) between events. In general, the  $m$ th-order approximation for the phase is obtained as:

$$\varphi^{(m)}(t) = \int_{\tau_i}^t \left[ \omega^{(m)} + Z^{(m)}(\varphi^{(m)}(t')) \right] dt'. \quad (20)$$

Estimated in this way, the phase at the end of a period generally differs from  $2\pi$ :  $\Phi_i^{(m)} = \lim_{t \uparrow \tau_{i+1}} \varphi^{(m)}(t) \neq 2\pi$ .

Thus, we additionally re-scale the phase with the factor  $\Phi_i^{(m)}$  to ensure  $\varphi^{(m)}(\tau_{i+1}) \equiv 2\pi$ . As a result, the approximations gradually improve through iterations. Estimated phases at the end of periods  $\Phi_i$  indicate how well our inference fits the observations, see error measure (16). For each iteration ( $m$ ) we can compute the error and monitor the convergence. For further details of the technique, we refer to [50].

#### 3.3.2. Inference of the isostable variable response

This section extends our approach to cover the isostable dynamics reconstruction. We still need the signal  $s(t)$  and perturbation  $p(t)$ , and since we have already inferred the PRC, we also have the instantaneous phase  $\varphi(t)$ .

Like the PRC technique, the isostable inference relies on time events of equal phase  $\tau_i$ , and additionally, on estimating the isostable variable at those events,  $\psi_i \equiv \psi(\tau_i)$ . The time events  $\tau_i$  are straightforwardly obtained from the instantaneous phase  $\varphi(t)$  [52], while the isostable variable needs to be estimated from the observed signal  $s(t)$ . The IRC function describes a linear response and is operable in the close vicinity of the limit cycle, where straight lines can approximate isochrons. Thus in the first-order approximation, the isostable variable linearly depends on signal  $s(t)$ :

$$\psi_i = c(s(\tau_i) - s_0) + \mathcal{O}((s(\tau_i) - s_0)^2), \quad (21)$$

where factors  $c$  and  $s_0$  are generally phase-dependent. However, since all crossing events  $\tau_i$  correspond to the same phase,  $c$  and  $s_0$  are constant in this context. Additionally, the isostable variable is inherently determined up to a constant factor. Therefore, without loss of generality, we can set  $c \equiv 1$ . Note that approximation via equation (21) is also justified for high-dimensional systems given if one Floquet exponent is significantly smaller in absolute value than the rest, i.e. there is a slow manifold (see discussion in appendix A).

We now consider the isostable dynamics (3) integrated over periods determined by the events of the same phase  $[\tau_i, \tau_{i+1}]$ :

$$\psi_{i+1} - \psi_i = \kappa \int_{\tau_i}^{\tau_{i+1}} \psi(t) dt + \int_{\tau_i}^{\tau_{i+1}} I(\varphi)p(t) dt. \quad (22)$$

This system of equations is similar to equation (19) in that it can be approximated with a linear system by expanding the unknown function in a Fourier series,  $I(\varphi) = \sum_{n=0}^{N_F} [u_n^{\cos} \cos(n\varphi) + u_n^{\sin} \sin(n\varphi)]$ , and interchanging the order of integration and summation. Furthermore, since we know the phase as a function of time, the  $i$ -dependent factors of Fourier coefficients can be computed directly. Using approximation (21) we evaluate the left-hand side as  $\psi_{i+1} - \psi_i = s(\tau_{i+1}) - s(\tau_i)$  because the still unknown constant  $s_0$  cancels in the difference. The integral  $\int \psi(t) dt$  is challenging since it includes a time-dependent isostable variable, which we do not have.

There are two issues with evaluating the isostable variable integral  $\int_{\tau_i}^{\tau_{i+1}} \psi(t) dt$ : (a) we do not know the constant  $s_0$  *a priori*, and (b) we want to use approximation (21) only in discrete events  $\tau_i$  (otherwise, we would have to consider  $c$  and  $s_0$  as phase-dependent). We tackle the first issue by splitting the integral in two:

$$\kappa \int_{\tau_i}^{\tau_{i+1}} \psi(t) dt = -\kappa s_0 (\tau_{i+1} - \tau_i) + \kappa \int_{\tau_i}^{\tau_{i+1}} (\psi(t) + s_0) dt. \quad (23)$$

In this representation  $\kappa s_0$  becomes just another unknown variable that will be determined while solving the linear system (22), which together with determining  $\kappa$  allows expressing  $s_0$  as the ratio  $s_0 = \frac{\kappa s_0}{\kappa}$ . The remaining integral involves the continuous quantity  $\psi(t) + s_0$ , which coincides with our observable  $s(t)$  at events  $\tau_i$  (generally this quantity differs from the signal for  $t \neq \tau_i$ :  $s(t) \neq \psi(t) + s_0$ ). The property  $s(\tau_i) = \psi(\tau_i) + s_0$  also helps us resolve the second issue of only using discrete values  $s(\tau_i)$  to estimate the integral of the continuous isostable variable. Just as in the phase response method, we approach this iteratively. First, we linearly interpolate  $\psi(t) + s_0$  between known events so that we can evaluate the last term:  $\int_{\tau_i}^{\tau_{i+1}} (\psi(t) + s_0) dt \approx [s(\tau_i) + s(\tau_{i+1})](\tau_{i+1} - \tau_i)/2$ . Then we solve the linear system equation (22) and obtain the first-approximation solution  $s_0^{(1)}$ ,  $\kappa^{(1)}$ ,  $I^{(1)}(\varphi)$ . Next, we use this solution to improve the estimate of the continuous isostable variable by integrating the underlying dynamical equation (3). In general, from the approximate solution of order  $m$ , we obtain the subsequent isostable variable estimation via integration [53]:

$$\psi^{(m)}(t) = s(\tau_i) - s_0^{(m)} + \int_{\tau_i}^t [\kappa^{(m)} \psi^{(m)}(t') + I^{(m)}(\varphi)p(t')] dt'. \quad (24)$$

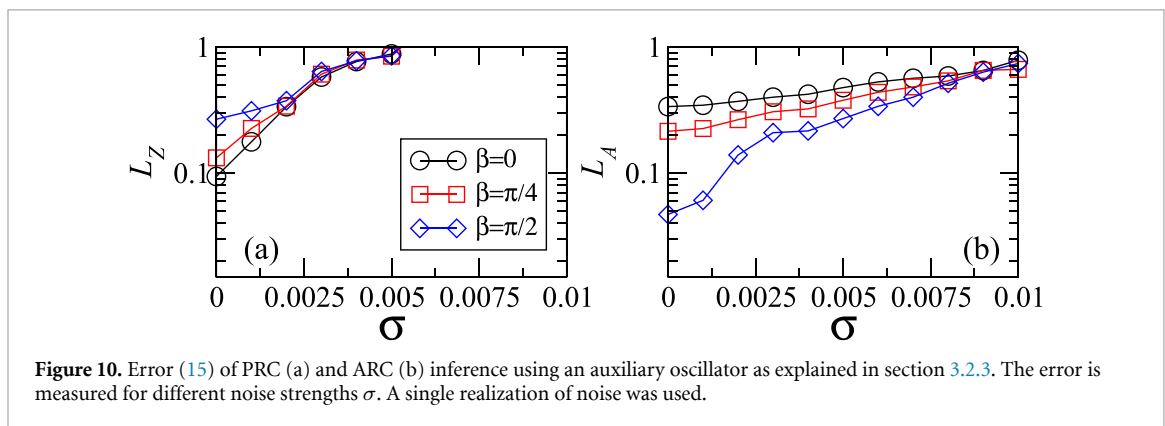
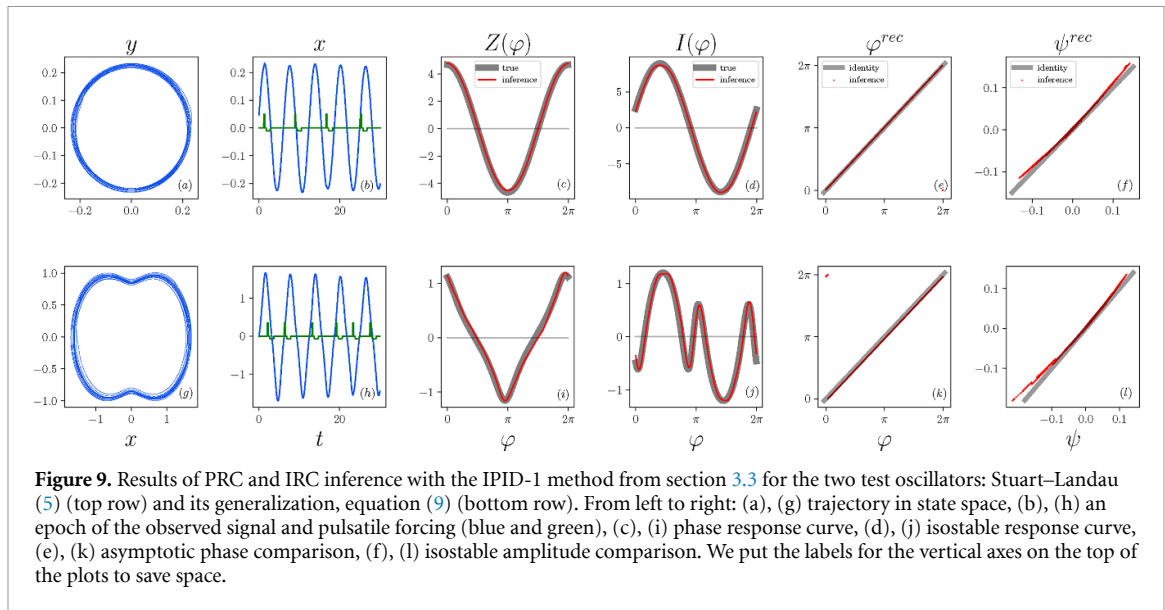
Estimated in this way, the isostable variable  $\psi^{(m)}$  starts the interval  $[\tau_i, \tau_{i+1}]$  as a direct estimation from the signal:  $s(\tau_i) - s_0^{(m)}$ , but further along the interval it can deviate due to the approximation in integrands  $\kappa^{(m)}$ ,  $I^{(m)}$ . Thus, at the end of the interval this quantity does not exactly correspond to  $s(\tau_{i+1}) - s_0^{(m)}$ , which is the value that we take to start the next interval  $[\tau_{i+1}, \tau_{i+2}]$ . The estimated time series of the isostable variable obtained in this way is thus discontinuous at the events  $\tau_i$ . If the inference were perfect, this discontinuity would disappear, which means we can use the magnitude of the discontinuity to quantify the quality of the model as follows. Let us denote the isostable variable estimated at the end of a period as  $\Psi_i^{(m)} = \lim_{t \uparrow \tau_{i+1}} \psi^{(m)}(t) \neq s(\tau_{i+1}) - s_0^{(m)}$ . We define the error of the fit as the standard deviation of the difference:

$$E_I^{(m)} = \left\langle \left( \Psi_i^{(m)} - \left( s(\tau_{i+1}) - s_0^{(m)} \right) \right)^2 \right\rangle^{1/2}. \quad (25)$$

Just like the PRC error was compared to the irregularity of inter-event intervals, see equation (17), this error should be compared to the irregularity of the isostable variable at the events:

$$E_{I0} = \left\langle \left( s(\tau_i) - \langle s(\tau_i) \rangle \right)^2 \right\rangle^{1/2}. \quad (26)$$

The entire procedure of inferring the isostable variable response is condensed into step-by-step instructions in appendix F.

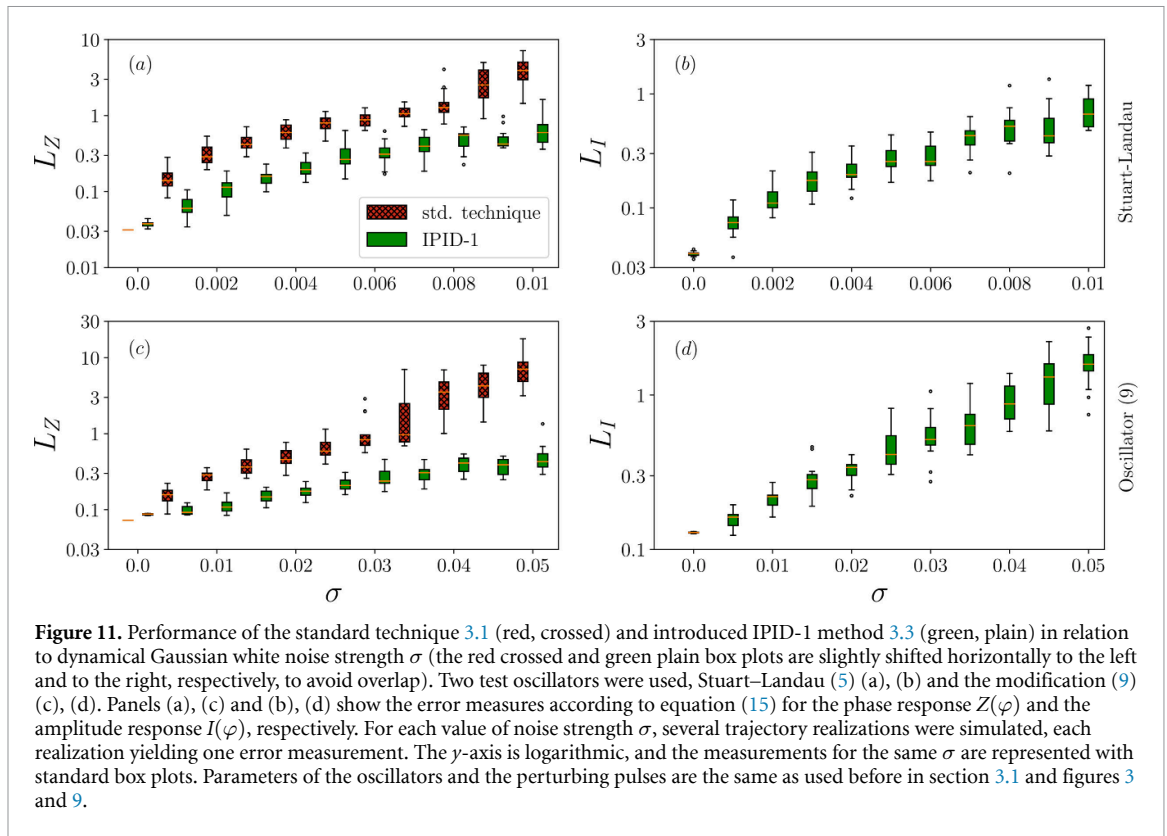


### 3.3.3. Performance of the IPID-1 phase-isostable reconstruction

We test the performance of the introduced IPID-1 method on the two example oscillators (5) and (9). We infer the PRC  $Z(\varphi)$  and IRC  $I(\varphi)$  with the approaches explained in sections 3.3.1 and 3.3.2 respectively. The results can be seen in figure 9. Parameters are the same as used in section 3.1 and figure 3. We stimulate with a bipolar charge-balanced pulse with a period of positive stimulation lasting 0.2, a period of no stimulation lasting 0.4, and a period of negative stimulation lasting 1.0 (pulse shape can be seen compared to the signal in figures 9(b) and (h)). We used a relatively long time series (1500 periods), and there was no noise or other unknown inputs (see appendix D for details). Additionally, we determined events of equal phase with the optimal threshold [54] by performing a direct search over possible threshold values, minimizing the error (16). As a result, the inferred curves (red) accurately reflect the true ones (thick gray). The slight deviations are mostly due to using a finite Fourier representation ( $N_F = 10$ ). As a byproduct of the inference, we also obtain the asymptotic phase and isostable amplitude as functions of time, and plot them against their true counterparts in figures 9(e),(f),(k) and (l). The obtained isostable time series is also a good representation of the signal envelope if shifted to match the signal maxima, as shown in a later figure 13.

### 3.4. Inference in the presence of noise

Real-world oscillators are inevitably noisy, and therefore, we test the performance of introduced techniques in the presence of dynamical noise. For this purpose, we simulate the same two systems (5) and (9) with different strengths  $\sigma$  of Gaussian white noise. We start with the SL oscillator and test the sine-fitting, Hilbert-based, and auxiliary oscillator techniques; again we use the unipolar pulses. In figure 10 we illustrate the performance of the latter method. (Since the sine-fitting and Hilbert-based techniques are less efficient for noise-free systems, we do not illustrate them here.) We show the inference errors as a function of the noise intensity  $\sigma$ , for three different values of  $\beta$ , cf. equation (5). As expected, for this method the PRC inference is more sensitive to noise than the ARC reconstruction.



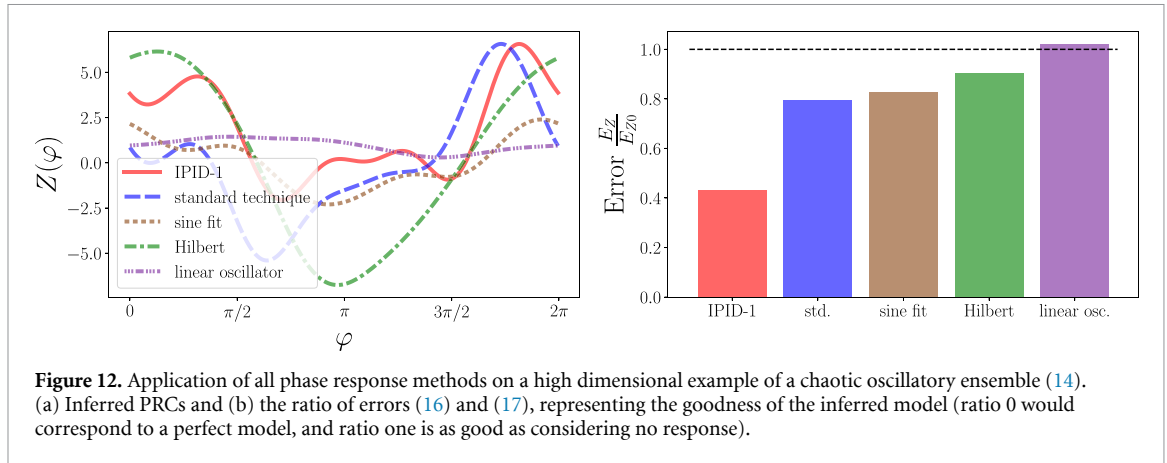
Next, we proceed with testing the most promising IPID-1 technique, based on the phase-isostable dynamics reconstruction. We keep the forcing pulse action constant, and then for each  $\sigma$  generate several trajectory realizations, and for each one, we compute the  $L$  error measure using equation (15). These errors are then depicted with standard box plots in figures 11(a) and (b). For comparison, we also show the results for the standard technique (PRC only). We conclude that the PRC inference with the technique introduced in section 3.3.1 is most stable to noise; in particular, it outperforms the standard approach. The direct comparison of the amplitude response reconstruction is not that easy since the auxiliary oscillator approach yields the empirical amplitude and, correspondingly, the ARC, while the isostable reconstruction provides the IRC. However, the results presented in figures 10(b) and 11(b) indicate approximately equal performance. We underline that stronger forcing would imply lower errors in the strong noise regime since the forcing-to-noise ratio would increase. On the other hand, the errors in the weak noise regime would increase since the linear approximation works worse for strongly forced systems. We also mention that the standard technique (red) has stronger limitations on the observations, namely, the pulses have to be rare (one pulse per several periods), which is why in the case of no noise, the standard technique has a marginally smaller error. For our technique, we considered a more realistic case of more than one pulse per period on average. It means our technique requires a much shorter time series, see appendix section D for details.

Finally, we test our phase-isostable reconstruction technique on the complicated case of the modified SL oscillator (9) that provides a non-sinusoidal observable, cf. figure 1(b). The results in figures 11(c) and (d) confirm the advantage of this technique, namely, it is not restricted to sinusoidal signals. Moreover, the inference errors  $L_Z$ ,  $L_I$  for the modified SL are not much different from those for the standard SL that provides a sinusoidal signal. The results in panel (c) clearly demonstrate that our technique is more stable with respect to noise than the standard one.

### 3.5. Response of a high-dimensional system

As a final test we apply the studied and newly developed techniques to a signal generated by a globally coupled oscillatory ensemble, see equation (14). The individual units are governed by the Bonhoeffer–van der Pol equations. This system can be treated as a simple model of neuronal rhythmical activity.

We consider a large number of oscillators and choose the parameters such that the system exhibits weak collective chaos, see test oscillators section 2 for more details on the system and figure 2 for a depiction of the mean-field orbit. We assume that we observe collective dynamics. This example is a hard test. In this case, not only is the ground truth phase model unknown to us, but since the system is chaotic, we know that there exists no phase description that can exactly describe the local deviations [55]. We, therefore, estimate the



**Figure 12.** Application of all phase response methods on a high dimensional example of a chaotic oscillatory ensemble (14). (a) Inferred PRCs and (b) the ratio of errors (16) and (17), representing the goodness of the inferred model (ratio 0 would correspond to a perfect model, and ratio one is as good as considering no response).

goodness of inference solely based on how well the inferred model reproduces observations; namely, we consider the ratio of errors (16) and (17):  $E_Z/E_{Z_0}$ . Our IPID-1 method performs best and yields an error ratio of less than half, while other methods perform rather poorly. Figure 12 shows all the corresponding phase response curves and the values of errors in a bar plot. We inferred the empirical PRCs using the mean field in the  $x$  variable as our observable,  $s = X$ , and exploiting bipolar pulses as before. Next, we deconvolved all curves to obtain the theoretical phase response curves. We also performed the inference on a different observable  $s = X + 2Y$  of the same system and obtained similar results.

Only three techniques can infer the amplitude response: the Hilbert-based, auxiliary oscillator-based, and IPID-1. The first two methods yield the effective ARC. In contrast, the IPID-1 technique estimates the infinitesimal isostable curve  $I(\varphi)$ . Hence, for comparison, we recompute  $I(\varphi)$  into the ARC in the following way: first, using the infinitesimal curve  $I(\varphi)$  and pulse shape  $\mathcal{P}(t)$  we evaluate the isostable shift  $\Delta\psi$ . Then we relate it to the ARC. For a particular phase  $\varphi^*$ , we integrate the first-order dynamics (1, 3) with the considered pulse shape  $\mathcal{P}(t)$  to obtain phase as a function of time,  $\varphi(t)$ , by solving the Winfree equation with the initial condition  $\varphi(t=0) = \varphi^*$ , and then compute the effective isostable shift as:

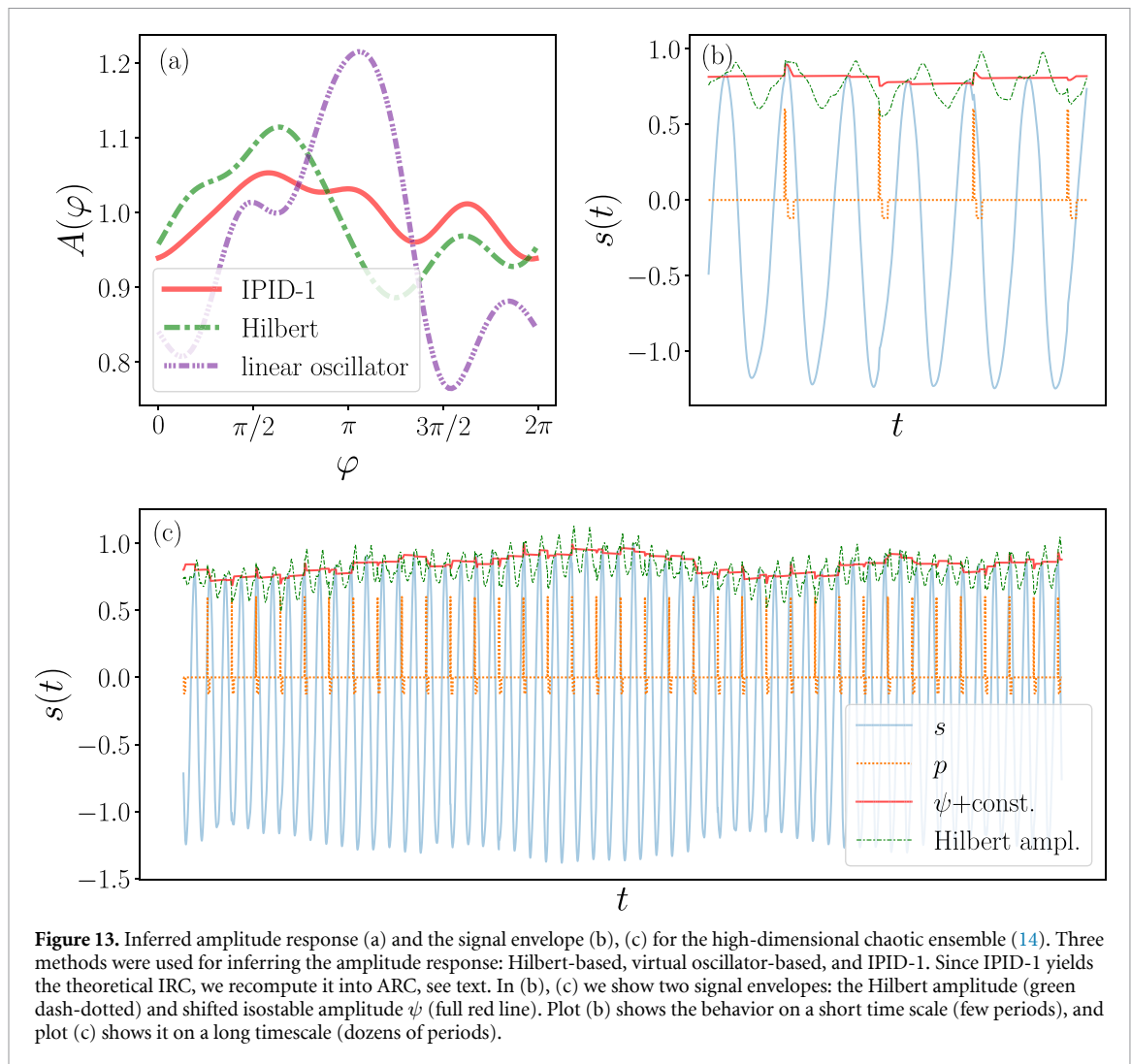
$$\Delta\psi(\varphi^*) = \int_0^{\delta} (\kappa\psi(t) + I(\varphi(t))\mathcal{P}(t)) dt. \quad (27)$$

Next, we have to relate the isostable variable  $\psi$  to amplitude  $a$ . As already mentioned, locally, the isostable linearly depends on the distance from the limit cycle, see equation (21):  $\psi \approx c(\varphi)(R - R_0(\varphi))$ . Here  $R$  represents the distance from the origin, while  $c(\varphi)$  and  $R_0(\varphi)$  are phase-dependent and hold the information of the isostable structure and parametrization of the limit cycle, respectively. If amplitude is simply defined as  $a = R$ , then  $a = R_0(\varphi) + \psi/c(\varphi)$ . Recalling the definition of ARC as a ratio of amplitudes before and after a pulse, we write:

$$A(\varphi) = \frac{a_e}{a_s} = 1 + \frac{\Delta\psi(\varphi)}{c(\varphi)R_0(\varphi)}. \quad (28)$$

We cannot estimate  $c(\varphi)$  nor  $R_0(\varphi)$  from a scalar signal—therefore, here we approximate them by constants. We choose  $c = 1$  since this value was used in the inference, and for  $R_0$  we use the maximal value of the signal:  $R_0 = s_{\max}$  [56]. We stress that difficulties of recomputing IRC into ARC are due to an ad hoc definition of the amplitude. Indeed, defined as the distance from the origin, it depends on the projection space (depends on both the observable and embedding technique). In contrast,  $I(\varphi)$  is an invariant characteristic of a limit cycle.

We compare the curves in figure 13(a). Similarly to the characterization of the phase response, we estimate the goodness of the IRC inference by quantifying how well the dynamical model fits the observation. Namely, we compute the ratio of errors (25) and (26). We obtain  $E_I/E_{I_0} = 0.28$ , which indicates a decent inference (we remind that this measure is 0 for a perfect fit and of order one for a completely wrong curve). Note that the two other techniques do not yield dynamical equations; therefore, no errors were computed. As a byproduct, our technique IPID-1 also yields an estimation of  $\psi$  as a function of time. In figures 13(b) and (c), we demonstrate that the isostable variable provides a good envelope if shifted to match the signal's maxima (or minima). We depict such an envelope along with the Hilbert amplitude as the commonly used alternative. Notice how Hilbert amplitude significantly fluctuates on the timescale of one period, while the isostable envelope changes mainly in times of stimulation and otherwise slowly follows the signal.



**Figure 13.** Inferred amplitude response (a) and the signal envelope (b), (c) for the high-dimensional chaotic ensemble (14). Three methods were used for inferring the amplitude response: Hilbert-based, virtual oscillator-based, and IPID-1. Since IPID-1 yields the theoretical IRC, we recompute it into ARC, see text. In (b), (c) we show two signal envelopes: the Hilbert amplitude (green dash-dotted) and shifted isostable amplitude  $\psi$  (full red line). Plot (b) shows the behavior on a short time scale (few periods), and plot (c) shows it on a long timescale (dozens of periods).

## 4. Discussion and conclusions

This paper addressed the problem of inferring an oscillator's phase and amplitude response from observations using test stimulation. Below, we discuss these two tasks of phase and amplitude inference separately.

In addition to critically testing several techniques presented in the literature, we also developed a new method for fitting first-order phase-isostable dynamics to observations, denoted as the IPID-1 method. We concentrated on the case where we apply a specially designed stimulation. However, this technique can be exploited for almost any input, as long as we can observe it and it does not entrain the oscillator [57].

### 4.1. Phase response

PRC determination for noise-free neuron-like oscillators has been known for decades. We here concentrated on noisy signals without well-defined marker events, such as spikes. As our first result, we mention the technique for recomputing the empirical PRC obtained in response to stimuli of arbitrary, but known shape, into the theoretical PRC,  $Z_{\mathcal{P}}(\varphi) \rightarrow Z(\varphi)$ . This technique can be used along side with any approach for PRC inference, in particular, it can enhance the applicability of the standard approach. This addition is especially beneficial for analyzing biological systems where some restrictions on the stimulus' shape may apply.

Next, we suggested a simple approach to inference error estimation. Without knowing the ground truth and using only the observations, we quantify the inferred PRC's accuracy in reproducing the observations. Namely, we quantify how much of the signal's variability can be explained by the inferred PRC. Again, this approach can complement any PRC estimation technique. We strongly recommend using this error estimation tool in experiments. Indeed, blind application of an inference technique always provides some response curve, but in the case of noisy or chaotic systems, the obtained curve may not describe the underlying dynamics.

We critically considered previously developed techniques that were introduced without being tested on examples where the ground truth is known. We performed tests and highlighted the potential weaknesses and drawbacks. In particular, we suggested improving the inference approach based on the popular analytical signal technique exploiting the HT by ignoring the data points immediately before and after stimuli. Furthermore, we compared the performance of the known and newly developed techniques in the noise-free and noisy cases using specially designed test data with the known ground truth. We tested the dependence of the results on the noise level and the observable used for data analysis. In particular, we have shown that the HT-based technique is generally unreliable. As a result, we have demonstrated the essential advantage of our IPID-1 approach based on the direct reconstruction of the Winfree equation: it is not restricted to sine-like signals, it immediately provides the theoretical PRC, and is robust against noise. We confirmed this conclusion by applying different inference techniques to a high-dimensional system representing the simplest model of brain rhythm generation. Estimating the error, we have demonstrated that IPID-1 is the only approach to yield a good result for this challenging test.

We conclude the discussion of the phase response inference with a remark. While testing the algorithms on the model systems, we ensured that the stimulation was weak. In practice, one must perform stimulation with stimuli of different strengths and check whether the inferred PRC depends on this strength. No dependence means validity of the linear approximation and, hence, of the PRC description. Otherwise, the revealed curve is not the theoretical (infinitesimal) PRC but quantifies the effect of a finite-strength perturbation and is amplitude-dependent.

#### 4.2. Amplitude response

The main problem with the amplitude response's inference is the definition of the amplitude. While phase can be uniquely defined for any point in the basin of attraction of a limit cycle, the definition of the amplitude is ambiguous. An operational approach introduces the (empirical) amplitude variable as an envelope to the observable and computes the envelope employing, e.g. the HT. However, this definition obviously depends on the observable. Moreover, for non-harmonic signals, HT and other techniques generally do not provide a 'good' envelope since they show variation for perfectly periodic signals.

Another approach involves equating the amplitude with the slowest decaying isostable variable, which provides a universal description independent of the choice of observable. The IPID-1 approach developed here reconstructs the first-order isostable dynamics directly from a scalar time series. The advantage of this technique is that inferred equation (3) can be further exploited to predict the effect of stimuli of a different shape. Furthermore, like the PRC inference, the technique provides an error measure that can be computed solely from data. We also mention the limitations of the procedure. It seems that inferring the isostable response is generally harder than inferring the phase response, and this is reflected in the performance of the IPID-1 technique by, e.g. comparing the inference where the ground truth is known. We also stress that equation (3) suffices for describing the dynamics only if the oscillator is two-dimensional or generally high-dimensional, but perturbations' decay in one direction is much slower than in other directions. The reduction of limit-cycle oscillators to phase-isostable dynamics is rather recent [38] and the physical interpretation of the isostable structure is still a matter of discussion. Another problem is to relate the isostable variables to envelopes of observed signals. Generally, inference and quantification of the amplitude response remain a subject of further studies.

#### 4.3. Relevance for oscillatory dynamics control

Knowledge of the phase and amplitude response allows efficient control of an oscillatory system. Indeed, probing the system by pulses of a known shape, we infer the Winfree equation (1) and then can exploit this equation to optimize the stimulus' shape, e.g. aiming to maximize the response.

In many cases, the goal of the control is to suppress or enhance the oscillation. The most illustrative example is DBS, aiming to quench the Parkinsonian tremor. The design of efficient control schemes requires the determination of vulnerable phases where the stimulation is most efficient. A possible approach relies on an adaptive control scheme [24, 26]. The alternative idea is to use first a test stimulation to infer the response properties and then exploit the corresponding curves to determine the proper stimulation phases [27]. Maxima and minima of the IRC or ARC provide the optimal stimulation phases for correspondingly enhancing and suppressing the oscillation. We emphasize that PRC alone does not yield the required information. Although a relationship between PRC and ARC has been demonstrated for some examples [27], generally, this relationship does not hold, see appendix A.3 for details. Another idea is stimulating around zeros of  $Z(\varphi)$ . If the phase response is zero, the stimulus acts along the isochron and modifies only the amplitude. This idea might be a good starting point, but it does not guarantee that this phase is optimal - the final amplitude variation depends on isostables density. We believe that the reconstruction of the isostable equation from data provides a means to design an optimal stimulation. Just as

in the case of phase response, using the reconstructed equation (3) we obtain the amplitude responses to arbitrary stimuli and can exploit this equation to optimize the stimulus' shape.

### Data availability statement

No new data were created or analysed in this study.

### Acknowledgments

R C and E T K M acknowledge financial support from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project-IDs PI 220/21-1 and 424778381 - TRR 295, respectively.

### Appendix A. Phase-isostable coordinates

Isostable coordinates extend the phase description of a limit cycle by adding variables  $\psi_i$  that describe the decay or growth of deviations from the limit cycle in the isochronal hyperplane. In the vicinity of a limit cycle, any  $N$ -dimensional system can be represented in the isostable coordinate system

$$\dot{\varphi} = \omega, \quad (\text{A1})$$

$$\dot{\psi}_i = \kappa_i \psi_i, \quad i = 1, \dots, N-1, \quad (\text{A2})$$

where  $\omega$  is the frequency, and  $\kappa_i$  are Floquet exponents of the limit cycle. For a system with a stable limit cycle, all Floquet exponents have a negative real part,  $\Re(\kappa_i) < 0$ . However, in many cases, it is sufficient to consider only one isostable variable  $\psi_{\max}$  corresponding to the largest real part of its Floquet exponent. This two-dimensional approximation is justified if there is a clear separation of time scales and all neglected  $\psi_i$  decay considerably faster than  $\psi_{\max}$ . Restricting our consideration to only one isostable coordinate  $\psi_{\max}$  complementing  $\varphi$ , we also assume the corresponding Floquet exponent to be real. It means that we do not address the case of the slowest mode described by a complex-conjugate pair of Floquet exponents. Thus  $\kappa := \max_i \Re(\kappa_i)$  and below we omit the subscript max in the notation  $\psi_{\max}$ . For a detailed discussion, in particular, on how to interpret a general  $N$ -dimensional system of phase-isostable coordinates with pairs of complex conjugate Floquet exponents, see [34, 38, 39].

As one can check from the autonomous dynamical equation (A2), all variables  $\psi_i$  are invariant to scaling with some non-zero factor  $c$ —the dynamical equations remain unchanged after a change of variables  $\tilde{\psi} := c\psi$ . As soon as external perturbations  $p(t)$  are introduced to the system via

$$\dot{\varphi} = \omega + G_\varphi(\varphi, \psi)p(t), \quad (\text{A3})$$

$$\dot{\psi} = \kappa\psi + G_\psi(\varphi, \psi)p(t), \quad (\text{A4})$$

the scaling of isostable variable leads to the adjustment in the stimulation function

$$G_{\tilde{\psi}}(\varphi, \tilde{\psi}) = cG_\psi\left(\varphi, \frac{\tilde{\psi}}{c}\right). \quad (\text{A5})$$

Thus, the IRC (defined as the isostable amplitude response evaluated at the limit cycle) is also scaled as a result of the scaling of  $\psi$ :

$$I_{\tilde{\psi}}(\varphi) = G_{\tilde{\psi}}(\varphi, 0) = cI_\psi(\varphi). \quad (\text{A6})$$

The PRC is invariant under rescaling of the isostable  $\psi$ . In the following derivations, we keep the scaling factor  $c$  to show where it appears in the equations.

### A.1. Isostable reduction for the SL system

The SL system, equation (5), is the analytically solvable normal form of a Hopf bifurcation: For  $\mu > 0$ , a radial limit cycle at  $R_0 = \sqrt{\mu}$  establishes with a basin of attraction being the entire phase space except for the fixed point at the origin. Thus, each point in the phase space, here given in polar coordinates  $R$  and  $\theta$ , can be assigned a phase by the well-known formula

$$\varphi(R, \theta) = \theta - \alpha \ln \left( \frac{R}{\sqrt{\mu}} \right), \quad (\text{A7})$$

for which one can verify that  $\dot{\varphi} = \eta - \alpha\mu =: \omega$ . Since the SL system is two-dimensional, one additional isostable  $\psi$  suffices to describe the dynamics in the vicinity of the limit cycle completely. Hence, for the SL example, the isostable coordinate system is a one-to-one transformation for  $\mathbb{R} \setminus \{0\}$ . For the case of radial isochrons ( $\alpha = 0$ ),  $\psi$  has already been derived in [33] as

$$\psi(R, \theta) = c \left( 1 - \frac{\mu}{R^2} \right). \quad (\text{A8})$$

Moreover, this coordinate transformation yields  $\dot{\psi} = -2\mu\psi := \kappa\psi$  also in the case of non-radial isochrons ( $\alpha \neq 0$ ). From this expression, one can check the system's dynamical properties: the roots of  $\psi$  at  $R = \sqrt{\mu}$  yield the coordinates of the limit cycle, and the divergence at  $R = 0$  indicates the border of the limit cycle's basin of attraction. Also, as  $\psi$  converges to a finite value as  $R \rightarrow \infty$ , we deduce that it takes a finite time to reach the vicinity of the limit cycle for initial conditions with arbitrarily large  $R$ .

The SL system can be re-written using the dynamical parameters  $\omega$  and  $\kappa$  instead of  $\eta$  and  $\mu$  as

$$\begin{aligned} \dot{x} &= -\omega y - \left( x^2 + y^2 + \frac{\kappa}{2} \right) (x - \alpha y) + G_x(x, y)p(t), \\ \dot{y} &= \omega x - \left( x^2 + y^2 + \frac{\kappa}{2} \right) (y + \alpha x) + G_y(x, y)p(t), \end{aligned} \quad (\text{A9})$$

where the functions  $G_{x,y}(x, y)$  specify how the perturbation acts on the oscillator.

The knowledge of the analytical form of the coordinate transformation of  $\psi$  and  $\varphi$  allows us to compute the PRC  $Z(\varphi)$  and IRC  $I(\varphi)$  also in a closed analytical form. Those are, by definition, the response curves of phase and isostable variables evaluated at the limit cycle, where  $\psi = 0$ . In general, the response curves are computed via

$$\begin{pmatrix} I(\varphi) \\ Z(\varphi) \end{pmatrix} = \begin{pmatrix} G_\psi \\ G_\varphi \end{pmatrix} \Big|_{\psi=0} = J_{\text{isostable} \leftarrow \text{polar}} \Big|_{\psi=0} \cdot J_{\text{polar} \leftarrow \text{Cartesian}} \Big|_{\psi=0} \cdot \begin{pmatrix} G_x \\ G_y \end{pmatrix} \Big|_{\psi=0}, \quad (\text{A10})$$

where  $J_{\text{isostable} \leftarrow \text{polar}}$  is the Jacobian of the coordinate transformation from polar to isostable coordinates, and  $J_{\text{polar} \leftarrow \text{Cartesian}}$  is the well-known Jacobian of the coordinate transformation from Cartesian to polar coordinates:

$$J_{\text{isostable} \leftarrow \text{polar}} = \begin{pmatrix} \partial_R \psi & \partial_\theta \psi \\ \partial_R \varphi & \partial_\theta \varphi \end{pmatrix} = \begin{pmatrix} \frac{2\mu c}{R^3} & 0 \\ -\frac{\alpha}{R} & 1 \end{pmatrix}, \quad J_{\text{polar} \leftarrow \text{Cartesian}} = \begin{pmatrix} \partial_x R & \partial_y R \\ \partial_x \theta & \partial_y \theta \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\frac{\sin(\theta)}{R} & \frac{\cos(\theta)}{R} \end{pmatrix}. \quad (\text{A11})$$

Throughout this paper, we assume a state-independent stimulation, namely,  $G_x = \cos(\beta)$  and  $G_y = \sin(\beta)$ . Hence we arrive at the following phase and isostable response curves for the SL system:

$$I(\varphi) = \frac{2c}{\sqrt{-\frac{\kappa}{2}}} \cos(\varphi - \beta), \quad (\text{A12})$$

$$Z(\varphi) = -\frac{1}{\sqrt{-\frac{\kappa}{2}}} (\sin(\varphi - \beta) + \alpha \cos(\varphi - \beta)). \quad (\text{A13})$$

For the rotationally invariant SL system, it is convenient to interpret the polar radius  $R$  as the amplitude of the limit cycle oscillation. Thus, from the response curve of  $R$ , denoted as  $G_R$ , we infer the ARC as

$$A(\varphi) = \frac{R_e}{R_s} \approx \frac{R_s + fG_R(\varphi)}{R_s} = 1 + \frac{fG_R(\varphi)}{R_s}, \quad (\text{A14})$$

approximating a short pulse with action  $f$  by a Dirac delta function, hitting the system at phase  $\varphi$ . For the SL system, the initial amplitude  $R_s = \sqrt{\mu}$  is independent of  $\varphi$ . For a stimulation that is state-independent in

Cartesian coordinates, the response in  $R$  is given by  $G_R(\varphi) = \cos(\varphi - \beta)$  and thus the ARC simplifies to the expression

$$A(\varphi) \approx 1 + \frac{f \cos(\varphi - \beta)}{\sqrt{\mu}}. \tag{A15}$$

**A.2. Test model with a non-sinusoidal solution and known phase and isostable response**

In order to construct more complex two-dimensional test models with known properties, we use the isostable coordinate system. Instead of trying to derive the phase and isostable variable from known systems, we give an explicit analytical expression of  $\varphi$  and  $\psi$  in the first place. Since the roots of  $\psi(\mathbf{x})$  are the location of the stable invariant set for  $\kappa < 0$ , we can control the position and shape of the limit cycle. In particular, the coordinate transformation

$$\psi(R, \theta) = c \left( 1 - \frac{q(\theta)}{R^2} \right) \tag{A16}$$

describes a limit cycle at  $R_0 = \sqrt{q(\theta)}$  for any strictly positive,  $2\pi$ -periodic function  $q$  of  $\theta$ . Together with the phase defined as

$$\varphi(R, \theta) = \theta - \alpha \ln \left( \frac{R}{\sqrt{q(\theta)}} \right), \tag{A17}$$

such that  $\varphi = \theta$  on the limit cycle, we obtain a generalized SL system. For  $q(\theta) = -\frac{\kappa}{2} = \mu$  we recover the original SL system. Since the Jacobian  $J_{\text{isostable} \leftarrow \text{polar}}$  can be computed in the same straightforward manner as for the SL system as

$$J_{\text{isostable} \leftarrow \text{polar}} \Big|_{\psi=0} = \begin{pmatrix} \partial_R \psi & \partial_\theta \psi \\ \partial_R \varphi & \partial_\theta \varphi \end{pmatrix} \Big|_{\psi=0} = \begin{pmatrix} \frac{2cq(\theta)}{R^3} & -\frac{cq'(\theta)}{R^2} \\ -\frac{\alpha}{R} & 1 + \frac{\alpha q'(\theta)}{2q(\theta)} \end{pmatrix} \Big|_{\psi=0} = \begin{pmatrix} \frac{2c}{\sqrt{q(\theta)}} & -\frac{cq'(\theta)}{q(\theta)} \\ -\frac{\alpha}{\sqrt{q(\theta)}} & 1 + \frac{\alpha q'(\theta)}{2q(\theta)} \end{pmatrix}, \tag{A18}$$

we can explicitly write the dynamical equations, both in polar

$$\dot{R} = \frac{R}{2q(\theta)} \left( \kappa \left( 1 + \frac{\alpha q'(\theta)}{2q(\theta)} \right) (R^2 - q(\theta)) + \omega_0 q'(\theta) \right), \tag{A19}$$

$$\dot{\theta} = \omega_0 + \frac{\alpha \kappa}{2q(\theta)} (R^2 - q(\theta)), \tag{A20}$$

and in Cartesian coordinates (where  $q$  and  $q'$  have to be read as functions of  $\theta(x, y)$ )

$$\dot{x} = \omega \left( \frac{xq'}{2q} - y \right) + \frac{\kappa}{2} \left( \frac{x^2 + y^2}{q} - 1 \right) \left( x + \alpha \left( \frac{xq'}{2q} - y \right) \right), \tag{A21}$$

$$\dot{y} = \omega \left( \frac{yq'}{2q} + x \right) + \frac{\kappa}{2} \left( \frac{x^2 + y^2}{q} - 1 \right) \left( y + \alpha \left( \frac{yq'}{2q} + x \right) \right). \tag{A22}$$

Having the explicit Jacobians for the coordinate transformation, we obtain the PRC and IRC generally given by the response curves in polar angle  $G_\theta$  and polar radius  $G_R$  as

$$I(\varphi) = 2c\mathcal{I}(\varphi), \tag{A23}$$

$$Z(\varphi) = G_\theta \Big|_{\psi=0}(\varphi) - \alpha \mathcal{I}(\varphi), \tag{A24}$$

with the function  $\mathcal{I}$  defined as

$$\mathcal{I}(\varphi) = q^{-\frac{1}{2}}(\varphi) G_R \Big|_{\psi=0}(\varphi) - \frac{q'(\varphi)}{2q(\varphi)} G_\theta \Big|_{\psi=0}(\varphi). \tag{A25}$$

The polar angle  $\theta$  serves as a protophase in this model since, on the limit cycle,  $\theta$  and  $\varphi$  coincide despite not being equal in general. Thus, on the limit cycle where  $\psi = 0$ , we have  $\theta = \varphi$  and we obtain the shape of the limit cycle  $R_0 = \sqrt{q(\varphi)}$  as a function of phase  $\varphi$ .

Note the relation from equation (A24), stating that the IRC is proportional to the difference of PRC and the response curve of the polar angle  $G_\theta|_{\psi=0}$  for non-zero non-isochronicity parameter  $\alpha$ :

$$Z(\varphi) - G_\theta|_{\psi=0}(\varphi) = -\frac{\alpha}{2c}I(\varphi). \quad (\text{A26})$$

We emphasize this relation between PRC and IRC since it exists without further assumptions made about  $G_\theta$  and  $G_R$ , i.e. how external stimulation enters the system, but is exclusively due to the autonomous dynamics of the system. However, despite the generalizing extension this model has added to the simple SL model, it still holds strong assumptions that lead to this conclusion, e.g. that phase  $\varphi$  and polar angle  $\theta$  are equal on the limit cycle.

Assuming that the external stimulation enters state independently in Cartesian coordinates, we compute the polar angle response function as

$$G_\theta|_{\psi=0}(\varphi) = -\frac{\sin(\varphi - \beta)}{\sqrt{q(\varphi)}} \quad (\text{A27})$$

and  $\mathcal{I}$  as

$$\mathcal{I}(\varphi) = \frac{2\cos(\varphi - \beta)q(\varphi) + \sin(\varphi - \beta)q'(\varphi)}{2q^{\frac{3}{2}}(\varphi)}. \quad (\text{A28})$$

For the particular example in section 2, we use  $q(\theta) = r + 2\cos^2(\theta)$ , where  $r$  is a positive parameter. We denote this particular model as the modified SL (mSL) model since it does not contain the original SL model as a special case. Its polar angle response curve is given by

$$G_\theta^{\text{mSL}}|_{\psi=0}(\varphi) = -\frac{\sin(\varphi - \beta)}{\sqrt{r + 2\cos^2(\varphi)}} \quad (\text{A29})$$

and the function  $\mathcal{I}$  as

$$\mathcal{I}(\varphi) = \frac{(r+1)\cos(\varphi - \beta) + \cos(3\varphi - \beta)}{(r + 2\cos^2(\varphi))^{\frac{3}{2}}}. \quad (\text{A30})$$

Having these two function, the PRC in equation (13) and IRC in equation (12) are derived using the relations given in equation (A24) for  $c = 1$ .

### A.3. Relationship of phase and ARCs

Duchet *et al* [27] hypothesized that the amplitude response is proportional to  $\partial_\varphi Z(\varphi)$ . Indeed, figure 14(a) demonstrates that this hypothesis holds for the isochronous SL oscillator. In the following, we sketch how geometrical assumptions about an oscillating system can lead to this hypothesis on its own: Assume (as we also did throughout this paper) that in a two-dimensional system, the external stimulation is state-independent in Cartesian coordinates  $x$  and  $y$ . Then, without loss of generality, the response functions can be written as  $G_x = \rho\cos(\beta)$  and  $G_y = \rho\sin(\beta)$ , where the angle  $\beta$  determines the direction of the stimulation in the  $x$ - $y$ -plane and the non-negative value  $\rho$  determines the strength. Without loss of generality, we can set  $\rho = 1$  by absorbing the stimulation strength as a factor into the stimulation shape function  $p(t)$ . Using a coordinate transformation from Cartesian to polar coordinates, see equation (A11), the response functions in the polar angle  $\theta$  and the polar radius  $R$  follow as

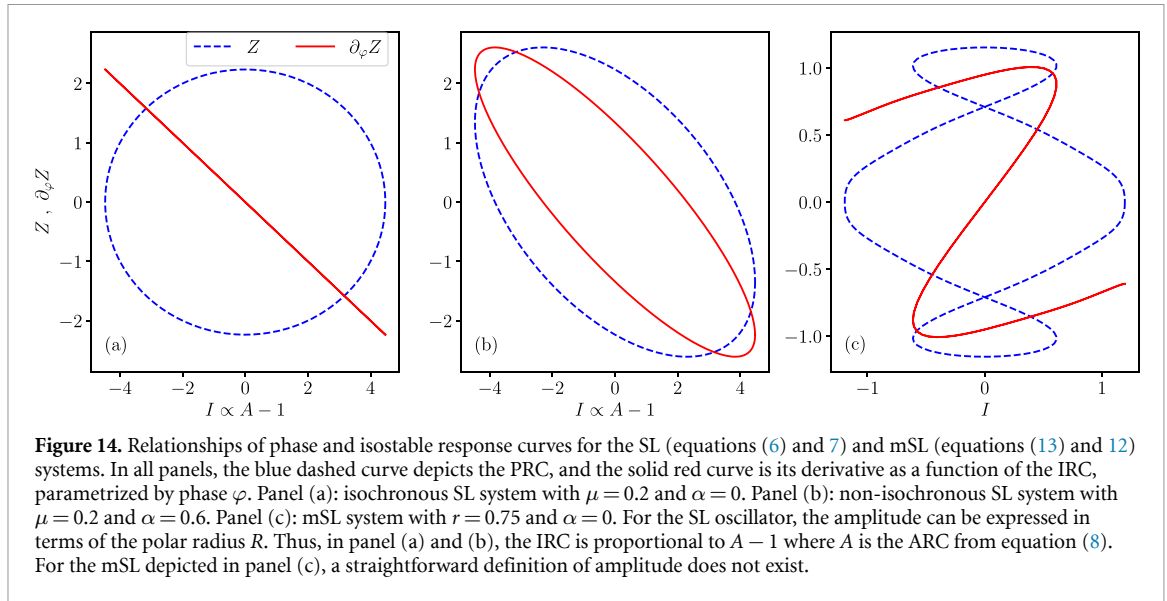
$$G_\theta = -\frac{1}{R}\sin(\theta - \beta) \quad \text{and} \quad G_R = \cos(\theta - \beta). \quad (\text{A31})$$

These two functions establish several relations, e.g.

$$G_R = -R \cdot \partial_\theta G_\theta \quad \text{or} \quad R^2 G_\theta^2 + G_R^2 = 1. \quad (\text{A32})$$

Thus, we obtain the relation that  $G_R$  is proportional to the derivative of  $G_\theta$ . However, we emphasize that up to this point, the considerations are purely geometrical and are valid regardless of the system's dynamics.

If the limit cycle rotates around the origin, one sometimes tries to approximate the system's phase  $\varphi$  by the polar angle  $\theta$  and the amplitude by the polar radius  $R$ . Indeed, for the standard SL system with its circular limit cycle, the polar radius  $R$  can serve as a measure of amplitude, leaving the isostable response curve proportional to the response in  $R$  and proportional to the ARC shifted by one:  $I \propto G_R \propto A - 1$ . In addition, for the isochronous SL system ( $\alpha = 0$ , figure 14(a)) the phase  $\varphi$  equals the polar angle  $\theta$  and the geometric relation (A32) directly translates into a relation between PRC and IRC:  $I \propto \partial_\varphi Z$ .



**Figure 14.** Relationships of phase and isostable response curves for the SL (equations (6) and (7)) and mSL (equations (13) and (12)) systems. In all panels, the blue dashed curve depicts the PRC, and the solid red curve is its derivative as a function of the IRC, parametrized by phase  $\varphi$ . Panel (a): isochronous SL system with  $\mu = 0.2$  and  $\alpha = 0$ . Panel (b): non-isochronous SL system with  $\mu = 0.2$  and  $\alpha = 0.6$ . Panel (c): mSL system with  $r = 0.75$  and  $\alpha = 0$ . For the SL oscillator, the amplitude can be expressed in terms of the polar radius  $R$ . Thus, in panel (a) and (b), the IRC is proportional to  $A - 1$  where  $A$  is the ARC from equation (8). For the mSL depicted in panel (c), a straightforward definition of amplitude does not exist.

For the non-isochronous SL system illustrated in figure 14(b), this relation is no longer strict since phase  $\varphi$  and polar angle  $\theta$  differ outside the limit cycle. Consequently,  $Z \neq G_\theta$  and  $\partial_\varphi Z$  is only correlated instead of strictly proportional to  $I$ . The modified SL system (figure 14(c)) has no circular limit cycle and hence no straightforward representation of amplitude by  $R$ , thus providing a counter-example to the hypothesis.

## Appendix B. Relationship between empirical and infinitesimal PRC

Here, we discuss how to recompute the empirical PRC  $Z_{\mathcal{P}}$  into theoretical infinitesimal PRC  $Z$ . The phase shift evoked by a pulse of finite width  $\delta$  applied at the instant  $t_s$  is  $\Delta\varphi = \varphi(t_e) - \varphi(t_s) - \omega\delta$ , where  $t_e = t_s + \delta$ . Hence, the Winfree equation (1) yields  $\Delta\varphi = \int_{t_s}^{t_e} Z(\varphi(t))\mathcal{P}(t) dt$ . If forcing is weak:  $Z(\varphi)p(t) \ll \omega$ , then in the first approximation we take  $\varphi \approx \omega(t - t_s)$  and obtain:

$$Z_{\mathcal{P}}(\varphi(t_s)) = \frac{1}{f} \int_{t_s}^{t_e} Z(\varphi(t_s) + \omega(t - t_s))\mathcal{P}(t) dt, \quad (\text{B1})$$

where  $f$  is the action of the pulse, in this paper either  $f = \int_0^\delta \mathcal{P} dt$  for unipolar pulses, or  $f = \frac{1}{2} \int_0^\delta |\mathcal{P}| dt$  for charge balanced pulses. This relation is a convolution of the true PRC  $Z$  and the pulse's waveform  $\mathcal{P}$ :  $Z_{\mathcal{P}} = Z \star \mathcal{P}$ . The convolution operation can be expressed with a product relation in Fourier space:  $\mathcal{F}(Z_{\mathcal{P}}) = \mathcal{F}(Z) \cdot \mathcal{F}(\mathcal{P})$ , which we utilize for expressing  $Z(\varphi)$ :

$$Z(\varphi) = \mathcal{F}^{-1}\left(\mathcal{F}(Z_{\mathcal{P}})/\mathcal{F}(\mathcal{P})\right), \quad (\text{B2})$$

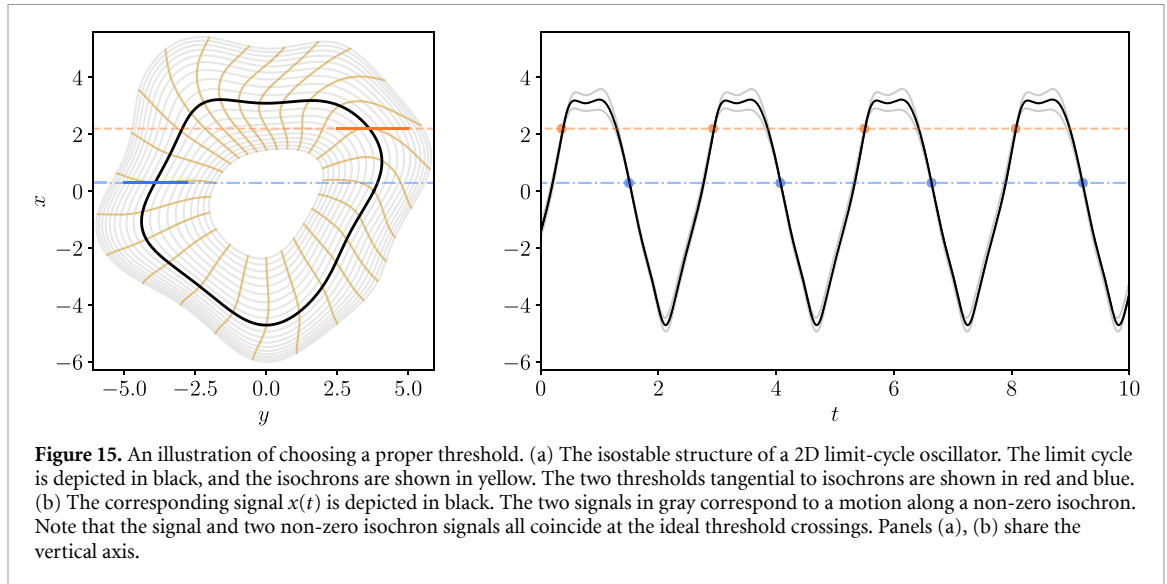
where  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  denote the direct and inverse Fourier Transform, respectively.

We mention that for a charge-balanced pulse we do not recover the constant term in  $Z$ . Indeed, substituting in equation (B1)  $Z$  by  $Z + C$ , where  $C$  is an arbitrary constant, we obtain the same function  $Z_{\mathcal{P}}$  due to the charge balance condition  $\int_{t_s}^{t_e} \mathcal{P} dt = 0$ . To solve this problem, we set the mean value of  $Z$  equal to that of  $Z_{\mathcal{P}}$ .

An alternative solution for the inverse problem that does not assume  $\varphi \approx \omega(t - t_s)$  is as follows. We represent the yet unknown function  $Z(\varphi)$  as a finite Fourier series with coefficients  $z_n^{\cos}, z_n^{\sin}$ . Then using  $\mathcal{P}(t)$  we express  $Z_{\mathcal{P}}(\varphi)$  in terms of  $z_n^{\cos}, z_n^{\sin}$  by numerically solving the Winfree equation. The nonlinear problem of equating the empirical curve with its symbolic representation:  $Z_{\mathcal{P}} = Z_{\mathcal{P}}(\varphi; z_n^{\cos}, z_n^{\sin})$ , is then solved using the Levenberg–Marquardt technique.

## Appendix C. Thresholding a signal

Here, we discuss how to determine events corresponding to the same asymptotic phase from the observed signal  $s(t)$ . This task is easy if the system moves on the limit cycle; then the signal  $s(t)$  is a projection of a one-dimensional trajectory, and any defining event that occurs once per period will correspond to the same state and, therefore, the same phase. Generally, defining events can be threshold crossings, extrema, and inflection points.



**Figure 15.** An illustration of choosing a proper threshold. (a) The isostable structure of a 2D limit-cycle oscillator. The limit cycle is depicted in black, and the isochrons are shown in yellow. The two thresholds tangential to isochrons are shown in red and blue. (b) The corresponding signal  $x(t)$  is depicted in black. The two signals in gray correspond to a motion along a non-zero isochron. Note that the signal and two non-zero isochron signals all coincide at the ideal threshold crossings. Panels (a), (b) share the vertical axis.

If the system is not on the limit cycle, determining the asymptotic phase is challenging. Generally, one would have to know the isochronal structure and then determine events as crossings of a particular isochron. (We remind that the isochrons are the manifolds of the constant phase.) However, even with *a priori* knowledge of isochrons, one cannot determine the crossing of a potentially high-dimensional manifold while observing a scalar signal  $s(t)$  only. Therefore, we restrict our consideration to 2D oscillators or higher dimensional oscillators with one Floquet exponent much smaller in absolute value than the rest (having a slow 2D manifold). In general, for higher dimensional systems the approximation will be less accurate. Additionally, we mention that one does not need to find the events exactly, but just their good approximation.

Let us consider threshold crossings such as, e.g.  $s(t) = s_{\text{thr}}, \frac{d}{dt}s(t) > 0$ . Figure 15 illustrates an example. Since we are thresholding a scalar signal  $s(t)$  we only have one parameter: the threshold value  $s_{\text{thr}}$ . Different thresholds correspond to different parallel straight lines (see two examples in figure 15 depicted in red and blue). The best approximation in this situation is when the line is tangential to a particular isochron at the limit cycle. Then, the corresponding threshold provides a first-order approximation for the states with the same phase. While we cannot change the inclination of the line, we can choose its vertical position. For 2D systems, there always exist at least two such values, one when the signal is growing (see red dashed line in figure 15) and one when it is falling (blue dash-dotted). From this illustration, one can deduce that extremal points are generally poor approximations for equal-phase states since the corresponding lines are (close to) tangential to the limit cycle and thus unlikely to be tangential to any isochron.

## Appendix D. Test perturbation used in our experiments

Here, we discuss the test signals we apply to the test oscillators in our experiments. Except for the cases where we mention it explicitly, we consistently use a bipolar charge-balanced pulse with a period of positive forcing lasting 0.2, no forcing lasting 0.4, and a period of negative forcing lasting 1.0. The overall duration of the pulse is therefore 1.6, and since both our test oscillators (5) and (9) are chosen to have a period  $2\pi$  each pulse lasts approximately  $1/4$  of a period, see figure 3 for a depiction of the pulse relative to the oscillators phase response curves. We use different forcing action  $f = \frac{1}{2} \int_0^{\delta} |\mathcal{P}| dt$  for different oscillators: for the Stuart–Landau system (5) we used  $f = 0.01$  while for the generalization (9) we used  $f = 0.07$  since it has a larger limit cycle and thus needs a stronger perturbation for roughly the same effect.

For the test with the standard inference technique, we introduced the pulses periodically with a much larger period (roughly five times) than the period of oscillators  $2\pi$ , ensuring that the two periods are incommensurate; hence the oscillator would eventually be stimulated at almost every phase. Such perturbation is ideal for inferring the phase response since the pulses appear rarely, giving the oscillator time to relax to the limit cycle where the phase shift is easy to read. In contrast, for our introduced IPID-1 method 3.3 we devised a more realistic experimental example where the pulses come randomly and often. We sample the inter-pulse times from a Poisson distribution such that, on average, one oscillator period receives 1.6 stimulating pulses, see figures 9(b), (h) for a short signal vs. forcing example.

## Appendix E. Causal phase and amplitude estimation exploiting virtual oscillators

Here, we show how to estimate phase and amplitude through a virtual oscillator. We explain how to obtain the amplitude first. We choose the oscillator's frequency  $\eta$  to be much larger than the characteristic frequency  $\nu$  of  $s(t)$  and the damping parameter  $\alpha_a$  sufficiently large so that the amplitude response is practically independent of  $\nu$ . Then, we recompute the amplitude of the virtual oscillator (that we can monitor) into the amplitude of the input signal, using the well-known expression for the resonance curve. Namely, we obtain the input's amplitude as

$$a(t) = \sqrt{x^2 + (\dot{x}/\nu)^2} \sqrt{(\eta^2 - \nu^2)^2 + (\alpha_a \nu)^2}.$$

For phase estimation, we choose the damping parameter  $\alpha_\varphi$  to be small so that the phase shift  $\beta = \arctan[-\alpha_\varphi \nu / (\eta^2 - \nu^2)]$  between the virtual oscillator and its input is nearly zero and weakly dependent on  $\nu$  for  $\nu \ll \eta$ . Then, we compute input's phase as  $\varphi(t) = \arctan(-\dot{x}/\nu x) - \beta$ . For further detail and an efficient numerical scheme, we refer to [49].

## Appendix F. Isostable inference step-by-step summary

- Determine events  $\tau_i$  by thresholding the phase, e.g.  $\varphi(\tau_i) = \varphi_{\text{thr}}$ ,  $\frac{d}{dt}s(\tau_i) > 0$ . The choice of threshold  $\varphi_{\text{thr}}$  is not crucial but generally should be picked such that the signal in events  $s(\tau_i)$ , has a wide range of values.
- Write system (22) by linearly approximating the isostable integral  $\int_{\tau_i}^{\tau_{i+1}} (\psi(t) + s_0) dt \approx [s(\tau_i) + s(\tau_{i+1})] (\tau_{i+1} - \tau_i) / 2$ :

$$s(\tau_{i+1}) - s(\tau_i) = -\kappa s_0 (\tau_{i+1} - \tau_i) + \kappa \frac{s(\tau_i) + s(\tau_{i+1})}{2} (\tau_{i+1} - \tau_i) + \int_{\tau_i}^{\tau_{i+1}} I(\varphi(t)) p(t) dt.$$

- Express the response function  $I(\varphi)$  as a Fourier series:  $I(\varphi) = \sum_{n=0}^{N_F} [u_n^{\cos} \cos(n\varphi) + u_n^{\sin} \sin(n\varphi)]$ , and replace the order of integration and summation in the last term to express it as a series of computable integrals:

$$\int_{\tau_i}^{\tau_{i+1}} I(\varphi(t)) p(t) dt = \sum_{n=0}^{N_F} [u_n^{\cos} P_{n,i}^{\cos} + u_n^{\sin} P_{n,i}^{\sin}],$$

where we denoted the Fourier integrals with  $P_{n,i}^{\cos} = \int_{\tau_i}^{\tau_{i+1}} \cos(n\varphi(t)) p(t) dt$  and  $P_{n,i}^{\sin} = \int_{\tau_i}^{\tau_{i+1}} \sin(n\varphi(t)) p(t) dt$ . Numerically compute the integrals for each inter-event interval.

- Combine the approximations to express system (22) as a linear system with known coefficients:

$$s(\tau_{i+1}) - s(\tau_i) = -\kappa s_0 (\tau_{i+1} - \tau_i) + \kappa \frac{s(\tau_i) + s(\tau_{i+1})}{2} (\tau_{i+1} - \tau_i) + \sum_{n=0}^{N_F} [u_n^{\cos} P_{n,i}^{\cos} + u_n^{\sin} P_{n,i}^{\sin}].$$

There are as many equations as there are inter-event intervals. The known coefficients are:

$s(\tau_{i+1}) - s(\tau_i)$ ,  $\tau_{i+1} - \tau_i$ ,  $(s(\tau_i) + s(\tau_{i+1})) / 2$ , and the Fourier integrals  $P_{n,i}$ . The unknown quantities are:  $\kappa s_0$ ,  $\kappa$  and the Fourier modes  $u_n^{\cos}$ ,  $u_n^{\sin}$ , representing the response  $I(\varphi)$ . Thus proceed to minimize this system via least squares (or a similar method) to obtain the first approximated solution:  $s_0^{(1)}$ ,  $\kappa^{(1)}$ ,  $I^{(1)}(\varphi)$ .

- Iterate the following: approximate the continuous amplitude  $\psi(t)$  by integrating the dynamics according to equation (24) using the current approximation  $s_0^{(m)}$ ,  $\kappa^{(m)}$ ,  $I^{(m)}(\varphi)$ . Numerically compute the integral  $S_i = \int_{\tau_i}^{\tau_{i+1}} (\psi(t) + s_0) dt$  with the approximated amplitude. Solve system (22) with the better approximated integral:

$$s(\tau_{i+1}) - s(\tau_i) = -\kappa s_0 (\tau_{i+1} - \tau_i) + \kappa S_i + \sum_{n=0}^{N_F} [u_n^{\cos} P_{n,i}^{\cos} + u_n^{\sin} P_{n,i}^{\sin}]$$

to obtain the next approximate solution  $s_0^{(m+1)}$ ,  $\kappa^{(m+1)}$ ,  $I^{(m+1)}(\varphi)$ . The deviation of the reconstructed amplitude from the true amplitude estimated at events  $\tau_i$  can be used as an error measure.

## References

- [1] Winfree A T 1980 *The Geometry of Biological Time* (Berlin: Springer)
- [2] Mackey M C and Glass L 1988 *From Clock to Chaos: The Rhythms of Life* (Princeton, NJ: Princeton University Press)
- [3] Rinzel J and Ermentrout B 1998 *Methods of Neuronal Modeling*, ed C Koch and I Segev (Cambridge: MIT Press) pp 251–92
- [4] Canavier C C 2006 *Scholarpedia* **1** 1332
- [5] Achuthan S and Canavier C C 2007 *J. Neurosci.* **29** 5218
- [6] Ermentrout B 2002 *Simulating, Analyzing and Animating Dynamical Systems: A Guide To Xppaut for Researchers and Students* (Philadelphia, PA: Society for Industrial and Applied Mathematics)
- [7] Tass P A 1999 *Phase Resetting in Medicine and Biology. Stochastic Modelling and Data Analysis* (Berlin: Springer) (<https://doi.org/10.1007/978-3-540-38161-7>)
- [8] Tass P A 2000 *Prog. Theor. Phys. Suppl.* **139** 301–13
- [9] Tass P A 2001a *Europhys. Lett.* **53** 15
- [10] Tass P 2001b *Europhys. Lett.* **55** 171–7
- [11] Tass P 2002 *Phys. Rev. E* **66** 036226
- [12] Rosenblum M G and Pikovsky A S 2004a *Phys. Rev. Lett.* **92** 114102
- [13] Rosenblum M G and Pikovsky A S 2004b *Phys. Rev. E* **70** 041904
- [14] Popovych O, Hauptmann C and Tass P A 2005 *Phys. Rev. Lett.* **94** 164102
- [15] Tukhlina N, Rosenblum M, Pikovsky A and Kurths J 2007 *Phys. Rev. E* **75** 011019
- [16] Hauptmann C and Tass P A 2009 *J. Neural Eng.* **6** 016004
- [17] Wilson C J, Beverlin B and Netoff T 2011 *Front. Syst. Neurosci.* **5** 50
- [18] Popovych O V and Tass P A 2012 *Front. Hum. Neurosci.* **6** 58
- [19] Lin W, Pu Y, Guo Y and Kurths J 2013 *Europhys. Lett.* **102** 20003
- [20] Zhou S, Ji P, Zhou Q, Feng J, Kurths J and Lin W 2017 *New J. Phys.* **19** 083004
- [21] Wilson D and Moehlis J 2016 *PLoS Comput. Biol.* **11** 1
- [22] Popovych O, Lysyansky B, Rosenblum M, Pikovsky A and Tass P 2017 *PLoS One* **12** e0173363
- [23] Krylov D, Dylow D and Rosenblum M 2020 *Chaos* **30** 033126
- [24] Montaseri G, Javad Yazdanpanah M, Pikovsky A and Rosenblum M 2013 *Chaos* **23** 033122
- [25] Holt A, Wilson D, Shinn M, Moehlis J and Netoff T 2016 *PLoS Comput. Biol.* **12** e1005011
- [26] Rosenblum M 2020 *Chaos* **30** 093131
- [27] Duchet B, Weerasinghe G, Cagnan H, Brown P, Bick C and Bogacz R 2020 *J. Math. Neurosci.* **10** 4
- [28] Benabid A, Pollak P, Gervason C, Hoffmann D, Gao D, Hommel M, Perret J and De Rougemont J 1991 *Lancet* **337** 403
- [29] Benabid A, Chabardes S, Mitrofanis J and Pollak P 2009 *Lancet Neurol.* **8** 67
- [30] Kühn A and Volkman J 2017 *Mov. Disorders* **32** 11
- [31] Holt A and Netoff T 2014 *J. Comput. Neurosci.* **37** 505
- [32] Wilson D and Moehlis J 2016b *Phys. Rev. E* **94** 012211
- [33] Wilson D and Ermentrout B 2018a *J. Math. Biol.* **76** 37
- [34] Wilson D 2019 *Phys. Rev. E* **99** 022210
- [35] We highlight that in the context of electrical stimulation, e.g., in neuroscience applications,  $p(t)$  is a current or a voltage and the introduced action has the physical meaning of electrical charge
- [36] Callenbach L, Hänggi P, Linz S J, Freund J A and Schimansky-Geier L 2002 *Phys. Rev. E* **65** 051110
- [37] Instantaneous phase and amplitude of a real-valued signal  $s(t)$  are the argument and absolute value of the complex-valued analytical signal  $s(t) + is_H(t)$ , where  $s_H$  is the Hilbert transform of  $s$ , see, e.g., [45, 58, 59]
- [38] Wilson D and Moehlis J 2016c *Phys. Rev. E* **94** 052213
- [39] Wilson D and Ermentrout B 2018b *SIAM J. Appl. Dyn. Syst.* **17** 2516
- [40] Rosenblum M and Pikovsky A 2019 *Phil. Trans. R. Soc. A* **377** 20190093
- [41] Alternatively, one can detect the signal's maxima; this approach corresponds to threshold-crossing for the derivative
- [42] The empirical curve is an 8th-order Fourier fit of experimental points
- [43] Note that the oscillation frequency  $\omega$  shall be either determined by fit or set to be equal to the central frequency of the bandpass filter, see [25, 31]
- [44] In their experiments, Duchet et al [27] used a complex stimulus consisting of 25 bursts of high-frequency pulses of a total length of 5s
- [45] Feldman M 2011 *Hilbert Transform Applications in Mechanical Vibration* (New York: Wiley)
- [46] For PRC estimation, a possible approach would be computing the error measure  $E_Z$  according to equation (16) as a function of  $\delta_{\text{off}}$  and searching for the minimum. However, this approach does not apply to ARC estimation since the Hilbert technique does not yield the Floquet exponent. While we can use the same  $\delta_{\text{off}}$  value to infer both curves, there is no guarantee that the optimal value is the same.
- [47] Kraleman B, Cimponeriu L, Rosenblum M, Pikovsky A and Mrowka R 2007 *Phys. Rev. E* **76** 055201
- [48] Kraleman B, Cimponeriu L, Rosenblum M, Pikovsky A and Mrowka R 2008 *Phys. Rev. E* **77** 066205
- [49] Rosenblum M, Pikovsky A, Kühn A and Busch J 2021 *Sci. Rep.* **11** 18037
- [50] Cestnik R and Rosenblum M 2018 *Sci. Rep.* **8** 13606
- [51] Each phase approximation  $\varphi^{(m)}$  is used to compute the integrals that represent coefficients of the linear system:  $\int \cos(n\varphi^{(m)})p(t) dt$ ,  $\int \sin(n\varphi^{(m)})p(t) dt$ . Then when solving the linear system and obtaining a better PRC approximation  $Z^{(m+1)}(\varphi)$ , it is used to recompute the next approximation of the phase  $\varphi^{(m+1)}$  by integrating the Winfree equation (1)
- [52] Note that the events  $\tau_i$  used for isostable inference generally differ from those used for the PRC reconstruction, though we keep the same notation. Practically, to determine  $\tau_i$  we use linear interpolation of the monotonically growing function  $\varphi(t)$  given in a discrete set of points
- [53] Since the forcing term is just a function of time, the isostable dynamics can be integrated via variation of constant, in which case the time integral can be computed explicitly:  $\psi^{(m)}(t) = (s(\tau_i) - s_0^{(m)})e^{i\kappa^{(m)}(t-\tau_i)} + \int_{\tau_i}^t I^{(m)}(\varphi(t'))p(t')e^{i\kappa^{(m)}(t-t')} dt'$
- [54] The optimal threshold corresponds to thresholding with a Poincaré section that is tangential to a local isochron, see figure 15. We find it by performing a direct search over a reasonable range of threshold values and choosing the one that corresponds to the lowest error value (16), as explained in the method paper [50]
- [55] Schwabedal J T C, Pikovsky A, Kraleman B and Rosenblum M 2012 *Phys. Rev. E* **85** 026216

- [56] Ignoring the phase dependence of  $c(\varphi)$  is similar to ignoring the information of local isochrons in the phase response, as most techniques do, e.g., the Hilbert-based one
- [57] Another example of a problematic perturbation is periodic stimulation with a frequency faster than the observed oscillator. In the latter case, there is more than one stimulus per period, and the phase of subsequent stimulations  $\varphi^{*(k)}$  fully depends on the phase of the first one,  $\varphi^{*(1)}$ . This makes distinguishing the phase-shift contributions of individual pulses  $Z(\varphi^{*(k)})$  impossible
- [58] Pikovsky A, Rosenblum M, and Kurths J 2001 Synchronization: a universal concept in nonlinear sciences The Cambridge nonlinear science series No. 12 (Cambridge: Cambridge University Press)
- [59] King F 2009 *Hilbert Transforms, vol 1,2 of Encyclopedia of Mathematics and its Applications* (Cambridge: Cambridge University Press)