P H Y S I C A L   R E V I E W   L E T T E R S

Inferring Phase Equations from Multivariate Time Series

Isao T. Tokuda

School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa 923-1292, Japan

Swati Jain, István Z. Kiss, and John L. Hudson

Department of Chemical Engineering, University of Virginia, 102 Engineers' Way, Charlottesville, Virginia 22904, USA

An approach is presented for extracting phase equations from multivariate time series data recorded from a network of weakly coupled limit cycle oscillators. Our aim is to estimate important properties of the phase equations including natural frequencies and interaction functions between the oscillators. Our approach requires the measurement of an experimental observable of the oscillators; in contrast with previous methods it does not require measurements in isolated single or two-oscillator setups. This noninvasive technique can be advantageous in biological systems, where extraction of few oscillators may be a difficult task. The method is most efficient when data are taken from the nonsynchronized regime. Applicability to experimental systems is demonstrated by using a network of electrochemical oscillators; the obtained phase model is utilized to predict the synchronization diagram of the system.

Synchronization is a ubiquitous phenomenon of coupled nonlinear oscillators commonly found in both natural science and engineering. Significant developments have been made in theory and experiments on synchronization of coupled limit cycle or chaotic oscillators [1,2]. An important theoretical development in this field is the phase reduction theory of weakly coupled limit cycle oscillators [3]. This theory has become a standard technique with applications in a wide variety of systems such as electrochemical oscillators [4], neurons [5], brain activity [6,7], genetic networks [8], and circadian oscillations [9,10]. However, an important problem of construction of the phase models from experimental data still remains open. It is of special interest to infer phase equations that characterize the underlying coupled system simply in terms of natural frequencies of individual oscillators and interaction functions between the oscillatory elements. Up to now, several techniques have been developed to construct phase equations from nonlinear systems [3,5,11–13]. One of them is the perturbation method that estimates the interaction function from the phase response curve and the oscillation waveform [3,11,12]. Despite its reliability, the method has a limitation in use with experimental data, since it requires the application of external perturbations to a nonlinear oscillator in an isolated condition. Other methods also have restrictions such as applicability only to a system of two coupled oscillators [5,13].

For wider application of the phase models to experimental (most notably biological) systems, which often have a network structure composed of many oscillators, there is a need for additional techniques. A noninvasive approach is desired that does not destroy the original configuration of the biological network.

The aim of this Letter is to demonstrate an approach to the development of experiment-based phase models of rhythmic, weakly interacting systems. Our approach is based upon the construction of the model by the multiple-shooting method combined with the cross-validation technique. This approach has the important practical advantages that (1) it can be applied to a network of many oscillators, (2) it is a noninvasive approach, (3) the numerical techniques can be easily implemented, and (4) the parameter estimation has an excellent convergence property.

We start with the problem formulation. Consider a system of N weakly coupled nearly identical limit cycle oscillators:

$$\dot{x}_i = F_i(x_i) + \frac{C}{N} \sum_{j \neq i}^{N} G(x_i, x_j), \qquad (1)$$

where x_i and F_i ($i = 1, 2, \ldots, N$) represent state variables and dynamics of the ith oscillator, C and G represent the coupling constant and the interaction function between the ith and jth oscillators. Our assumption is that in isolated condition ($C = 0$) each oscillator F_i gives rise to a stable limit cycle with similar natural frequencies ω_i. Then the phase reduction theory [3] states that for weak uniform coupling C the network dynamics can be reduced to the phase equations: $\dot{\theta}_i = \omega_i + \frac{C}{N} \sum_{j=1}^{N} H(\theta_j - \theta_i)$ (θ_i, phase of ith oscillator; H, interaction function). As a recording condition, we assume that simultaneous measurement of all oscillators is made as $\{x_i(n\Delta t): n = 1, \ldots, M\}_{i=1}^{N}$ (Δt: sampling time).

Our goal is to infer the phase equations from the measurement data under the conditions that (i) the underlying dynamics (1) are unknown, (ii) the coupling constant C associated with the measured data is taken from a nonsynchronous regime, and (iii) the coupling type is known to be uniform and all-to-all connection. The estimation does

not require *a priori* knowledge of the specific value of the coupling constant (without loss of generality it can be taken to be unity).

Our approach to the problem can be described as follows.

(1) Determine phases $\theta_i(t)$ from data $x_i(t)$. Among various definitions of phases [2], a simple formula is chosen, where the phase θ is increased by 2π at every local maximum of $x(t)$ and between the local maxima the phase grows proportionally in time.

(2) Fit the phases $\{\theta_i(t)\}$ to the phase equations:

$$\dot{\theta}_i = \omega_i + \frac{C}{N}\sum_{j=1}^{N}\tilde{H}(\theta_j - \theta_i), \qquad (2)$$

where the interaction function $\tilde{H}$, which is in general nonlinear and periodic with respect to 2π, is approximated by a Fourier expansion up to order of D as $\tilde{H}(\Delta\theta) = \sum_{j=1}^{D} a_j \sin j\Delta\theta + b_j(\cos j\Delta\theta - 1)$. [For simplicity, we consider difference coupling. Thus the interaction function is set to zero for zero phase difference, i.e., $\tilde{H}(0) = 0$].

The unknown parameters $p = \{\omega_i, a_j, b_j\}$ are estimated by the multiple-shooting method [14]. We denote the time evolution of the phase equations (2) with respect to the initial condition $\theta(0)$ by $\theta(t) = \phi^t(\theta(0), p)$. Then, at each sampling time $t = i\Delta t$, the phase equation must satisfy the boundary conditions: $\theta((n + 1)\Delta t) = \phi^{\Delta t}(\theta(n\Delta t), \mathbf{p})$. With respect to the unknown parameters p, we solve these nonlinear equations by the generalized Newton method. The evolution function ϕ^t is integrated numerically. For the computation of the gradients $\partial\phi/\partial p$ which are needed for the Newton method, variational equations of the phase equations (2) are also solved numerically.

A necessary condition to solve the nonlinear equations is that the number of the unknown parameters is less than the number of the equations, corresponding in this case to $N + 2D < N(M - 1)$. This always holds in the case of enough data M.

(3) To avoid overfitting problem, a cross-validation technique is utilized to determine the optimum number of higher harmonics in the interaction function D [15]. We divide the multivariate data into two parts. For the first half of the data, the parameter values p are estimated. Then we apply the estimated parameters to the latter half data and measure the error $E = \sum_n |\theta((n + 1)\Delta t) - \phi^{\Delta t}(\theta(n\Delta t), \mathbf{p})|^2$. The order number D providing the minimum error is considered to be the optimum.

We apply this technique to a prototypical example of weakly coupled limit cycle oscillators. We consider the following network of Rössler equations with diffusive coupling: $\dot{x}_i = -\alpha_i y_i - z_i$, $\dot{y}_i = \alpha_i x_i + 0.15 y_i + \frac{C}{N}\sum_{j=1}^{N}(y_j - y_i)$, $\dot{z}_i = 0.2 + z_i(x_i - 2)$ $(i = 1, \ldots, N)$. Each Rössler oscillator gives rise to a limit cycle attractor for the chosen parameter values without coupling $C = 0$. To consider an inhomogeneity in the network, parameter

values α_i, which determine rotation speed in the (x, y) space, are varied among the oscillators. The multivariate data are recorded as $\{y_i(t)\}_{i=1}^{N}$.

We start with the case of $N = 16$. Inhomogeneous parameter values are set as $\alpha_i = 1 + 0.0002(i - 8.5)$ $(i = 1, 2, \ldots, 16)$. The data $\{y_i(t)\}_{i=1}^{16}$ are recorded at the coupling strength $C = 0.002$, which gives rise to non-synchronized dynamics. The sampling interval was set to be $\Delta t = 0.08$ for the extraction of the phase $\{\theta_i(t)\}$. Then to apply the multiple-shooting method the data have been downsampled to $\Delta t = 1000 \times 0.08$ and a total of 200 data points have been collected for the parameter estimation. (As an initial condition, unknown parameter values are all set to be zero, i.e., $\omega_i = 0$, $a_j = b_j = 0$.) The convergence property of the multiple shooting was excellent; a single Newton procedure gives a good estimate.

Figure 1 shows the estimated natural frequencies of the uncoupled oscillators and the interaction function with the Fourier order of $D = 2$, which is optimized by the cross-validation test. The statistical errors indicated by the error bars were computed from the inverse of the Hessian matrix of the squared error function, based on the assumption that the phase data contain uncorrelated observational noise [16]. The estimated natural frequencies are distributed on a diagonal line with the original frequencies computed from each of the Rössler oscillators. Moreover, the estimated interaction function $\tilde{H}(\Delta\theta)$ is in very good agreement with that estimated by applying the perturbation method to a single isolated Rössler oscillator [3,11]. (For the comparison, we utilized the coupling constant as prior knowledge.) The error bars, which overlap with the perturbation method, imply that the slight deviation is due to the fitting error to the phase data.

It is important to note that the estimation results depend on the coupling strength C used to generate the time series. Figure 2(a) shows dependence of estimation error on the coupling strength. The estimation error e is evaluated as the normalized deviation of the estimated interaction function $\tilde{H}_s(\Delta\theta)$ from $\tilde{H}_p(\Delta\theta)$ estimated by the perturbation method; i.e., $e = 100 \int_0^{2\pi} |\tilde{H}_s(\Delta\theta) - \tilde{H}_p(\Delta\theta)| d\Delta\theta / \int_0^{2\pi} |\tilde{H}_p(\Delta\theta)| d\Delta\theta$ [%]. When the coupling
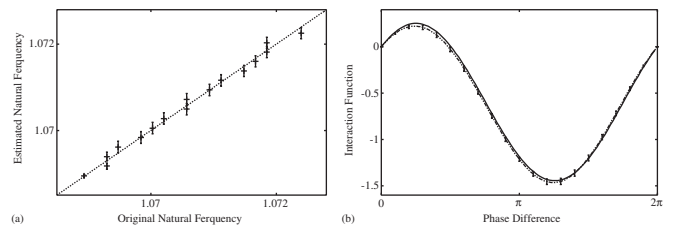

FIG. 1. (a) The estimated natural frequencies (vertical axis) $\{\omega_i\}_{i=1}^{16}$ of 16 Rössler oscillators versus the natural frequencies of the noncoupling simulation (horizontal axis). (b) Interaction function $\tilde{H}(\Delta\theta)$ estimated by the present method (dotted line) is compared with that estimated by the perturbation method (solid line).

strength is too small, the interaction effect is weak. This makes it difficult to recover the precise structure of the interaction function, resulting in a relatively large estimation error. On the other hand, when the coupling strength is very close to the onset of synchronization, the estimation error increases significantly. This can be understood as follows. When the network is synchronized, phase differences between all the oscillators become nearly constant $\Delta\theta = $ const. This provides no information at all about the interaction function. A similar result has been reported from modeling of coupled chaotic oscillators [17]. Hence, the most efficient condition for precise estimation of phase equations is to utilize an intermediate coupling strength between the two extreme situations. One possibility to overcome the problem of synchronized data is to apply an external perturbation to destroy the network synchronization and to utilize the relaxation process for the parameter estimation.

Another important dependence of the model estimation is on the network size. As shown in Fig. 2(b), the estimation error increases almost linearly as the network size N is increased. This might be due to the fact that the procedure becomes more sensitive to the estimate of interaction effect from each individual oscillator, when interactions from many other oscillators are summed up in the term $\sum_{j=1}^{N} H(\theta_j - \theta_i)$ of Eq. (2). This difficulty applied to large systems with all-to-all connection. However, for non-all-to-all coupled systems the limitation is not expected to be so severe because a typical node may only be connected to a small number of other elements. Such small number of local connectivities in complex networks is important in many biological systems [18].

Let us apply our technique to experimental data. We use an electrochemical oscillatory system in which the interaction function was recently calculated [12] with the perturbation method and thus direct comparison between the two approaches can be made.

The experiments were carried out in a standard electrochemical cell containing 3 mol/dm^3 sulfuric acid kept at a temperature of 11 °C with an 8×8 array of 1 mm diameter nickel, a Hg/Hg$_2$SO$_4$/K$_2$SO$_4$ reference, and a Pt counterelectrode. (For the experiments with 32 electrodes only half of the electrodes were connected in the array.) The currents, proportional to the rate of electrodissolution,
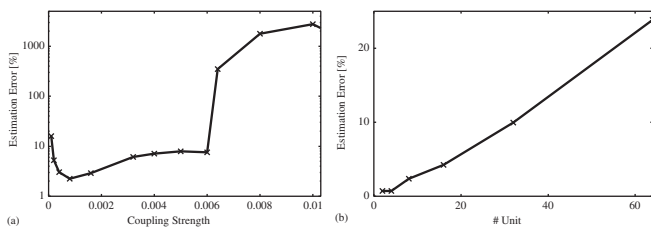
were measured at a frequency of 100 Hz using zero resistance ammeters. The electrodes, held at a potential of $V = 1.105$ V, were connected to a potentiostat through one series (collective) resistor R_s and through N parallel resistors (N is the number electrodes). The dimensionless interaction strength $K = NR_s/R_p$ was controlled through the external resistors by keeping the equivalent resistance $R_{eq} = R_p + NR_s$ constant [4]. [The physical coupling strength is obtained as $C = K/(R_{eq}C_d)$, $C_d = 1.3$ mS as measured by impedance spectroscopy.]

First, we consider a network of 32 electrochemical oscillators. Multivariate data $\{y_i(t)\}_{i=1}^{32}$ were measured from all $N = 32$ oscillators with a coupling strength of $C = 0.018$ 1/s, which gives rise to nonsynchronized dynamics. The recording interval of 5 s was utilized for the parameter estimation. Figure 3 shows the estimation results based on the optimum $D = 1$. Natural frequencies are well estimated with slightly higher values than those obtained with $C = 0$. Moreover, the shape of the estimated interaction function $\tilde{H}(\Delta\theta)$ is in a good agreement with that estimated by applying the perturbation method to a single isolated electrochemical oscillator [12]. The interaction function consists of strong $\sin(\Delta\phi)$ and $1 - \cos(\Delta\phi)$ terms and higher harmonics are very weak; this functional form is consistent with theoretical prediction of Stuart-Landau oscillators close to a Hopf bifurcation [3]. The difference in amplitude between the interaction functions obtained with the two different methodologies is $e = 23.7\%$ in Fig. 3(b).

We found that the estimation procedure was efficient in extracting the phase model for sufficiently strong coupling. With very weak coupling strengths, small drifts of the natural frequencies of oscillators caused the estimation of the relatively weak coupling effects to be rather inaccurate.

Finally, we demonstrate the capability of our approach to predict the synchronization structure. For another network of 64 globally coupled electrochemical oscillators a synchronization diagram (dependence of the order parameter Φ on the coupling strength in the range of $C \in [0, 0.17]$) has been previously obtained [19]. We examine the power of the extracted experiment-based phase model

FIG. 2. (a) Dependence of the estimation error e on the coupling strength C used to generate multivariate data. (b) Dependence of the estimation error e on the network size N.
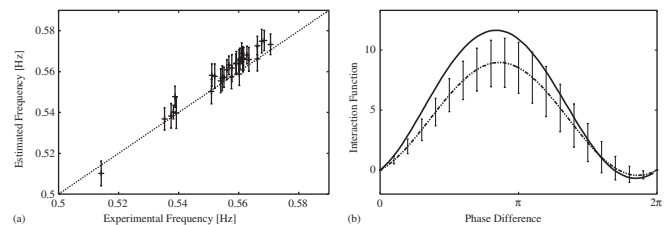

FIG. 3. (a) Estimated natural frequencies (vertical axis) $\{\omega_i\}_{i=1}^{32}$ of 32 electrochemical oscillators versus the measured natural frequencies (horizontal axis). (b) Interaction function $H(\Delta\theta)$ estimated by the present method (dotted line) is compared with that estimated by applying the perturbation to a single isolated electrochemical oscillator (solid line) [12].
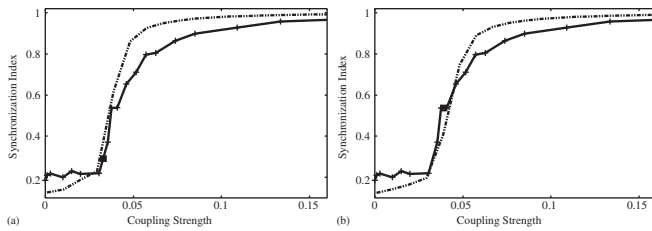
FIG. 4. Synchronization diagram of a network of 64 electro-chemical oscillators. The order parameter Φ is plotted against the coupling strength K. The model prediction (dotted line) is compared with the experimental curve (solid line with crosses). (a),(b) correspond to the modeling data measured at $K = 0.033$ and $K = 0.039$ (boxes), respectively.

in recovering this diagram. The estimated model is simulated to produce the synchronization diagram. Figure 4(a) shows the dependence of the order parameter Φ on the coupling strength when only a single data set with $K = 0.033$ close to the Kuramoto transition point was used to obtain the phase model. (The order parameter can be computed according to $\mathrm{Re}^{i\Phi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\phi_j}$ [3,4].) In spite of the drift, relative shortness of the time series, and the larger (64) population size, the phase model gives an excellent prediction of the order versus coupling strength curve. We note that the modeling accuracy has a dependency on the data location. For example, when a relatively short data set with only a few phase slips among the oscillators at $K = 0.022$ was analyzed, the reconstruction of the synchronization diagram was not successful. However, when enough phase slips are present, the synchronization diagram can be well recovered as demonstrated at $K = 0.039$ in Fig. 4(b).

To conclude, an approach has been presented for inferring phase equations from multivariate time series. Analysis of measurement data from electrochemical oscillators demonstrated practical applicability to experimental systems. The present approach is rather general and can be extended to more complex situations such as the case of nonuniform coupling $C_{i,j}$. Such extended cases will be investigated in forthcoming studies. One of the future challenges is to apply the present approach to a network of oscillators in biological systems. Recent technological advances made simultaneous (optical or electrical) measurements of the rhythmic cells possible in networks; examples include circadian gene expressions of suprachiasmatic nucleus [9] or the synchronous electrical activity in neuronal pacemakers [20]. Such recording conditions would facilitate the construction of phase models from experimental data that could be used for the simple yet accurate description of a large number of rhythmic, interacting cells.

[1] A. T. Winfree, *The Geometry of Biological Time* (Springer, New York, 1980).
[2] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization—A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2001).
[3] Y. Kuramoto, *Chemical Oscillations, Waves and Turbulence* (Springer, Berlin, 1984).
[4] I. Z. Kiss, Y. M. Zhai, and J. L. Hudson, Science **296**, 1676 (2002).
[5] R. F. Galan, G. B. Ermentrout, and N. N. Urban, Phys. Rev. Lett. **94**, 158101 (2005).
[6] P. Tass *et al.*, Phys. Rev. Lett. **81**, 3291 (1998).
[7] L. Angelini *et al.*, Phys. Rev. Lett. **93**, 038103 (2004).
[8] D. Battogtokh, K. Aihara, and J. J. Tyson, Phys. Rev. Lett. **96**, 148102 (2006).
[9] S. Yamaguchi *et al.*, Science **302**, 1408 (2003).
[10] H. Kori and A. S. Mikhailov, Phys. Rev. Lett. **93**, 254101 (2004).
[11] H. Sakaguchi, S. Shinomoto, and Y. Kuramoto, Prog. Theor. Phys. **77**, 1005 (1987).
[12] I. Z. Kiss, Y. M. Zhai, and J. L. Hudson, Phys. Rev. Lett. **94**, 248301 (2005).
[13] J. Miyazaki and S. Kinoshita, Phys. Rev. Lett. **96**, 194101 (2006).
[14] E. Baake *et al.*, Phys. Rev. A **45**, 5524 (1992).
[15] M. Stone, J. R. Stat. Soc. Ser. B. Methodol. **36**, 111 (1974).
[16] W. Horbelt *et al.*, Phys. Rev. E **64**, 016222 (2001).
[17] I. Tokuda, J. Kurths, and E. Rosa, Jr., Phys. Rev. Lett. **88**, 014101 (2001).
[18] D. Watts and S. Strogatz, Nature (London) **393**, 440 (1998).
[19] Y. M. Zhai *et al.*, Physica (Amsterdam) **205D**, 57 (2005).
[20] E. E. Verheijck *et al.*, J. Gen. Physiol. **111**, 95 (1998).