Chemical Engineering Communications

ISSN: 0098-6445 (Print) 1563-5201 (Online) Journal homepage: http://www.tandfonline.com/loi/gcec20

DISCRETE- vs. CONTINUOUS-TIME NONLINEAR SIGNAL PROCESSING OF Cu ELECTRODISSOLUTION DATA

R. RICO-MARTÍNEZ , K. KRISCHER , I.G. KEVREKIDIS , M.C. KUBE & J.L. HUDSON

To cite this article: R. RICO-MARTÍNEZ , K. KRISCHER , I.G. KEVREKIDIS , M.C. KUBE & J.L. HUDSON (1992) DISCRETE- vs. CONTINUOUS-TIME NONLINEAR SIGNAL PROCESSING OF Cu ELECTRODISSOLUTION DATA, Chemical Engineering Communications, 118:1, 25-48, DOI: 10.1080/00986449208936084

To link to this article: https://doi.org/10.1080/00986449208936084

Published online: 30 Mar 2007.

Chem. Eng. Comm. 1992, Vol. 118, pp. 25–48
Reprints available directly from the publisher.
Photocopying permitted by license only.

Printed in the United States of America

DISCRETE- vs. CONTINUOUS-TIME NONLINEAR SIGNAL PROCESSING OF Cu ELECTRODISSOLUTION DATA

R. RICO-MARTÍNEZ, K. KRISCHER and I.G. KEVREKIDIS

Department of Chemical Engineering, Princeton University, Princeton, NJ 08544-5263.

and

M.C. KUBE and J.L. HUDSON

Department of Chemical Engineering, University of Virginia, Charlottesville, VA 22093-2442.

(Submitted August 20, 1991; in revised form January 22, 1992)

Artificial neural networks (ANNs) are often used for short term discrete time predictions of experimental data. In this paper we focus on the capability of such nets to correctly identify long term behavior and, in particular, observed bifurcations. As we show, the usual discrete time mapping approach is (precisely because of its discrete nature) often incapable of reproducing observed bifurcation sequences. If the interest is only in periodic or temporally more complicated behavior, a Poincaré map extracted from the experimental time series can be used to circumvent this problem. A complete dynamic picture including bifurcations of steady states can, however, only be captured by a continuous-time model. We present an ANN configuration which couples a "nonlinear principal component" network for data processing (Kramer, 1991, Usui *et al.*, 1990) with a composite ANN based on a simple integrator scheme. This ANN is able to correctly reconstruct the bifurcation diagram of our experimental data. All time series we process stem from the potentiostatic electrodissolution of Cu in phosphoric acid solution. As the applied potential is varied, the electrodissolution rate changes from steady behavior to periodic oscillations, followed by a sequence of period doublings to apparently chaotic motion, and then returns to simple oscillations via a reverse cascade of period doublings.
KEYWORDS Neural networks Time-series Electrodissolution Bifurcation.

I. INTRODUCTION

Artificial neural networks (ANNs) are currently extensively used in time series processing and nonlinear system identification tasks; this use is widespread over a range of fields from artificial intelligence to fluid mechanics (Rummelhart and McClelland, 1986). We are interested in the characterization of time series stemming from chemical reaction processes (McAvoy *et al.*, 1989, Hudson *et al.*, 1990). Such processes are inherently nonlinear (due, for example, to the Arrhenius dependence of the reaction rate on temperature, or to kinetic nonlinearities); as a result, measurements of the reaction rate or other important variables (e.g. reactant concentrations) may be steady or oscillatory in time, and the oscillations can vary in nature from simply periodic to complicated ("chaotic") as the operating parameters of the process (e.g. flow rates, pressure,

temperature) are varied. Models based on first principles, when available, can lead to qualitative understanding of the dynamic behavior of the process and its dependence on operating parameters. However, such models may not exist, or may not be robust enough to quantitatively predict the system behavior in real time (e.g. they may not take into account slow changes of a catalytic surface under reaction conditions). In such cases, the *ad hoc* models obtained by using experimental measurements to train ANNs can be used to characterize and predict the system dynamics. Accurate short term prediction is the main objective for real time control applications. In our work we focus rather on using the long term solutions (attractors) of such ANN-based models to understand and characterize the dependence of the system behavior on parameters, and more specifically the instabilities and bifurcations it exhibits.

Usually when analyzing time series the input to the ANN used consists of variable measurements at the current and possibly previous sampling times as well as operating parameter values; the ANN output is then a prediction of the state of the system at the next sampling time. When this output is fed back into the ANN and the procedure iterated indefinitely we obtain a prediction of the system *attractor,* i.e. its *long term* solution. Long term prediction of the time series *per se* may be quite inaccurate; as a matter of fact it *will* be inaccurate for chaotic time series displaying sensitivity to initial conditions, and therefore exponentially amplifying small errors. Even though the long term solutions (attractors) of the system and the model may differ *point by point,* the ANN-based model can still be considered successful if they lie close to each other in phase space. In this case "success" means that the model attractor accurately approximates certain statistical properties of the system (amplitudes, power spectra, dimension) but more importantly, it can crucially assist understanding the nature of qualitative transitions (bifurcations) as a function of the operating parameters (e.g. Casdagli, 1989). As we will discuss below the *discrete* time form of traditional hidden-layer feedforward ANNs retains short term prediction accuracy and yields attractors close to the "correct" ones in phase space; however, it may prove inaccurate in capturing the detailed structure and the *qualitative* nature of continuous time attractors and their bifurcations. This shortcoming is particularly visible in the case of periodic attractors. We demonstrate this point using time series obtained from our experimental study of the electrodissolution of Cu carried out under potentiostatic conditions in phosphoric acid solutions. Two alternative ANN-based approaches to rectifying this problem are discussed. The first, not applicable when steady state behavior is involved, is based on using traditional ANNs on appropriately constructed (discrete) return maps (Poincaré maps) from the original time series. The second, more systematic approach, is based on constructing ANNs capable of fitting the right-hand side of a continuous dynamical system (a set of Ordinary Differential Equations, ODEs). There are several choices of the actual form of the input to the latter ANN; the particular form we use is the result of an ANN-based "nonlinear principal component" preprocessor of the time series (Kramer, 1991). The second approach is indeed capable of representing transitions involving steady state as well as time-dependent data.

This paper is organized as follows: section II contains a concise description of the experimental setup as well as a presentation of the sequence of time series obtained and our interpretation of the underlying bifurcations. In section III the various approaches we used in processing the data are described along with the structure of the corresponding ANNs, and their predictions are presented and compared.

II EXPERIMENTAL

All time series we processed were obtained from dissolution current measurements during the potentiostatic electrodissolution of a rotating Cu electrode in phosphoric acid. This system is known to show a large variety of complex behavior as the potential is varied (Albahadily *et al.*, 1989, Schell and Albahadily, 1989). A chemically based kinetic model, however, has not yet been derived. The experimental setup consisted of a rotating disc electrode which had a copper rod, 8.26 mm in diameter, imbedded in a 2 cm diameter Teflon cylinder, a platinum sheet with a surface area of 25 cm^2 as counter electrode and a saturated calomel reference electrode (SCE). The rotation speed was maintained at 4350 rpm. The

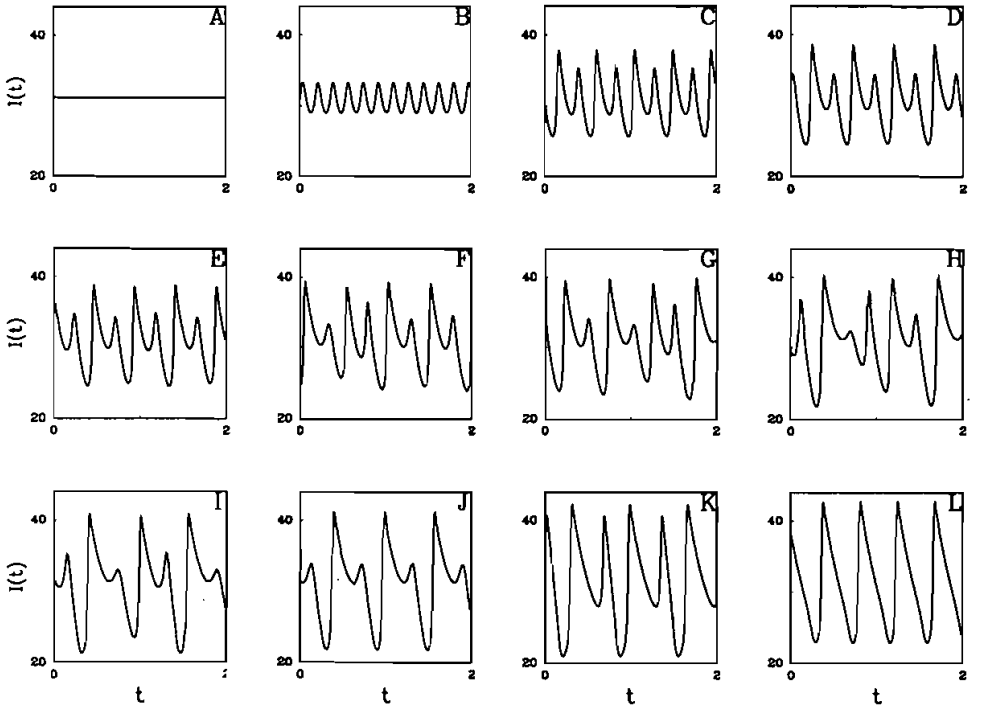

FIGURE 1 Time-series segments (current I [mA] vs. time [s]) from a sequence of Cu electrodissolution experiments. Observe the variations in system behavior as a function of the operating parameter (E, the potential). E [V]; (A) 0.6703, (B) 0.6768, (C) 0.6875, (D) 0.6897, (E) 0.6904, (F) 0.6915, (G) 0.6927, (H) 0.6948, (I) 0.5987, (J) 0.6990, (K) 0.7114, (L) 0.7143.

three electrodes were placed in a three neck 400 ml flask containing 300 ml of 85% phosphoric acid. A water bath was used to maintain the temperature at 11°C. A potentiostat (Princeton Applied Research model 362) was used to regulate the potential of the working disc electrode with respect to the SCE and to monitor the current. Steps in the potential were applied with a PAR model 175 Universal Programmer and the current was digitized by means of a Keithley model 500A measurement and control system and sampled at 500 Hz.

Figure 1 shows (a small part of) typical current measurements versus time which were obtained at different values of the potential. At a potential of 670.3 mV the system exhibits a steady state (Figure 1A). Upon raising it by 6.5 mV the steady state becomes unstable (apparently via a supercritical Hopf bifurcation) and simple oscillatory behavior is observed (Figure 1B). Upon raising the potential further the time series are seen to undergo period doubling, gradually change in shape (Figure 1C-E) and eventually become aperiodic (Figure 1G, 1H). A further increase of the potential restores multipeak periodic oscillations (Figure 1I-K) and finally simple periodic behavior (Figure 1L). At these higher potential values the amplitudes of the oscillations are much larger compared with those of Figure 1B and they appear more relaxation-like in nature.

In Figure 2 the attractors (projected on a two-dimensional phase space)

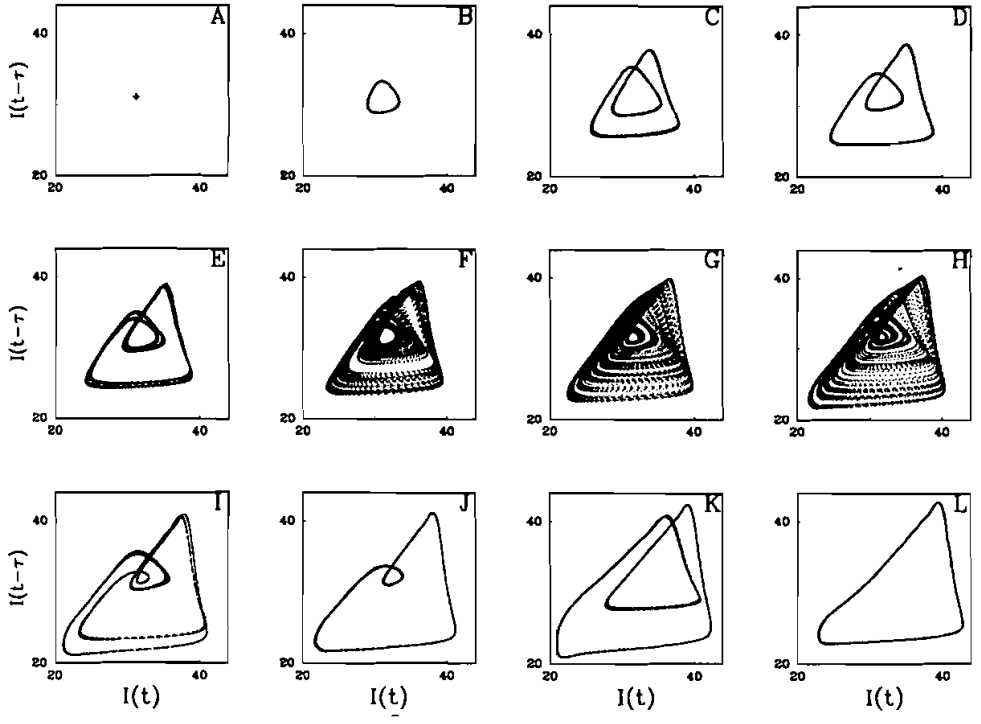

FIGURE 2 2-D reconstruction of the attractors using time delays ($\tau = 0.04$ s). The attractors correspond to the time series shown in Figure 1.

corresponding to the set of time series of Figure 1 are shown. The qualitative changes from a fixed point via a simple periodic, period-2 and -4 limit cycles to an apparently chaotic attractor (most likely due to a period doubling cascade) and the reverse sequence (period doubling) back to a period-1 limit cycle can be seen more clearly. The attractors were reconstructed using the method of delay proposed by Packard *et al.* (1980), and by Takens (1981). Every time series consisted of about 15000 points (or 30s). While in principle any choice of the actual time delay value will give qualitatively similar phase portraits, an extensive literature on selecting the optimal (under certain criteria) delay values exists (e.g. Fraser and Swinney, 1986, Mayer-Kress, 1986, Liebert and Schuster, 1989). Here we have used the "rule of thumb" choice of a delay approximately equal to one fifth of the basic frequency.

From the observed bifurcation sequence one can assume that the dynamics underlying the above described data can be completely embedded in a three dimensional phase space. It is possible and often convenient to further reduce the dimension of a system exhibiting oscillatory dynamics by looking at some transverse cross section of its trajectories; this allows description of the dynamics of the system by a map (Poincaré map). Reconstructing the data of Figure 1 in a

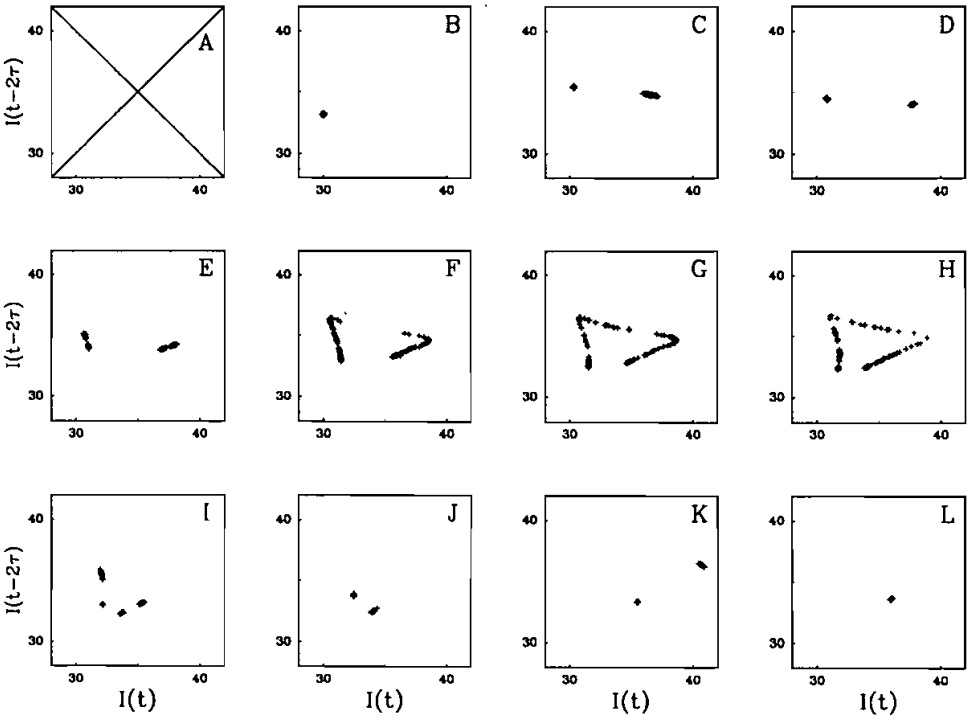

FIGURE 3 Poincaré sections of the 3-D attractors corresponding to the time series shown in Figure 1. The delays used to reconstruct the attractors are 0.04 and 0.08 s; the section is taken at $I(t-0.04\,\mathrm{s}) = 31.5\,\mathrm{mA}$ and only intersections towards increasing values of $I(t-0.04\,\mathrm{s})$ are considered.

3-dimensional phase space and recording the intersection of the trajectory with a plane $I(t-0.04\,\mathrm{s})=31.5\,\mathrm{mA}$ towards increasing values of current results in the Poincaré section shown in Figure 3. The limit cycles appear as fixed or periodic points on the section whereas the aperiodic dynamics lie on what appears to be a curve (but which is known, however, to be very intricately folded with a Cantor-set-like structure).

III. TIME SERIES PROCESSING

III.1 Delay Map Approach

For a deterministic system such as the set of ODEs

$$\dot{\vec{Y}}=\mathbf{f}(\vec{Y};\vec{X}),\qquad \vec{Y}\in\mathcal{R}^n,\qquad \vec{X}\in\mathcal{R}^p,\qquad \mathbf{f}:\mathcal{R}^n\times\mathcal{R}^p\mapsto\mathcal{R}^n$$

the state $\vec{Y}(t)$ at all future times $t>t_0$ depends only on the current state $\vec{Y}(t_0)$ and on the operating parameters $\vec{X}$. The geometrical concept of attractor reconstruction using time delays (Packard *et al.*, 1980 and Takens, 1981) indicates that the future state of the system can also be obtained as a function of measurements of a single state variable augmented with an appropriate number of delayed measurements. This idea also fits nicely in the context of time series processing (linear or nonlinear) and in particular in the context of "traditional" feedforward multilayer artificial neural networks. A number of time delayed measurements is fed to the network and additional input neurons are reserved for operating parameter values at which the measurements were obtained. The ANN output constitutes a prediction of the value of the state variable at a future measurement time. Sigmoidal activation functions are often used for the nonlinear neurons, and training is carried out using either steepest descent backpropagation or the conjugate gradient method. A large number of case studies using variants of this approach can be found in the chemical engineering literature (e.g. Hudson *et al.*, 1990, McAvoy *et al.*, 1989, Naidu *et al.*, 1990, Ydstie, 1990, Bhat and McAvoy, 1989, Hoskins and Hillemblau, 1988, Kramer and Leonard, 1990) along with theoretical studies on the necessary or optimal configuration (number of layers, number of neurons per layer, etc.) (e.g. Cybenko, 1989, Hecht-Nielsen, 1989, Hornick *et al.*, 1988, Irie and Miyake, 1988, Karnin, 1990, Lapedes and Farber 1987a,b, Sanger, 1989). When the training converges, the network constitutes a model dynamical system which can be used for short- and long term (attractor) prediction as well as prediction of parameter dependence if several parameter values have been included in the training.

We used such a standard configuration (Lapedes and Farber, 1987a,b) to process the time series from the electrodissolution experiment contained in Figure 2. As discussed above, the nature of the attractors in Figure 2 suggests $\mathcal{R}^3(I(t), I(t-\tau)$ and $I(t-2\tau))$ as a plausible—and minimal—first attempt at an embedding space. In particular, the ANN used consisted of four layers: two hidden layers with 10 neurons each, a 4 neuron input layer and a single neuron output layer (Figure 4). As there is currently no rigorous way to determine an optimum

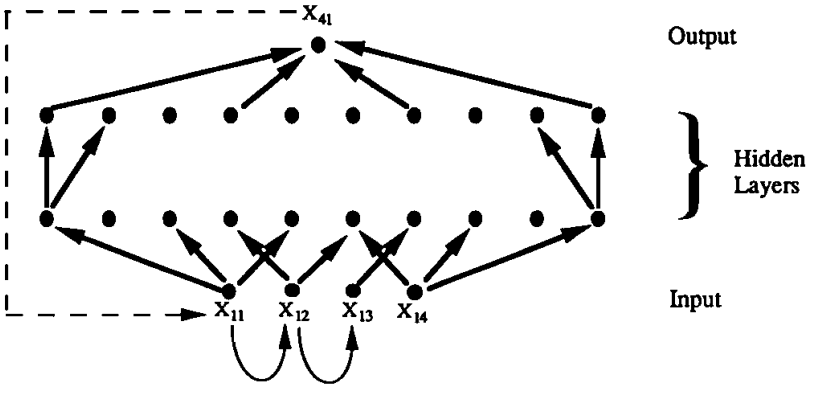

FIGURE 4 ANN architecture used in the delay map case. X_{11}, X_{12} and X_{13} are the inputs to the ANN which represent a state of the system (i.e. $I(t)$, $I(t-\tau)$, $I(t-2\tau)$) and X_{14} is the value of the operating parameter (E). The output of the ANN X_{41} is the prediction (i.e. $I(t+\tau)$). To iterate the ANN, the previous output value is fed as the input X_{11} and the old inputs shifted by one (i.e. now X_{12} is the old X_{11} and X_{13} becomes the old X_{12}). Only a few connections of the fully interconnected neurons are depicted.

number of neurons, the number of neurons in the hidden layers is somewhat arbitrary. Our choice here reflects a compromise between the computational effort required to train the ANN and an estimate of the minimum number of neurons needed to capture the underlying dynamics. This compromise is tested during the process of validation of our results: in this particular case an ANN with 12 neurons per hidden layer was found to yield qualitative and quantitative results similar to the 10 neuron per hidden layer ANN.

The input and output neurons of the ANN used are linear, while the neurons in the hidden layers are nonlinear with activation function $g(X) = \frac{1}{2}(1 + \tanh(X))$. Since this activation function gives values between 0 and 1, the network inputs and target values were accordingly normalized. The time delay (τ) was taken to be 0.04s (the same as used in Figure 2), and the fourth input in the input layer was the value of the potential in volts. The output is the prediction $\hat{I}(t+\tau)$ of $I(t+\tau)$. The training set consisted of a total of 1000 vectors taken from five distinct time series at five different potential values in the interval [0.678–0.699 volts] corresponding to steady, period one, period two, period four and chaotic dynamics. The number of vectors from each type of behavior was not identical; they were apportioned (again somewhat arbitrarily) according to the complexity in phase space of the behavior for every parameter value. While only 20 vectors amply represent the steady state (a single point in phase space), 100 vectors were taken from a period one time series (a closed loop), 200 from a period two time series (a double loop), 300 from a period four time series (a quadruple loop), and 380 vectors from a chaotic attractor lying close to these oscillations in phase space. Note that each data set is obtained at a different value of the potential, which is also used in the training.

The training procedure was considered successfully converged when both the mean square prediction error as well as its rate of decrease fall below preset

bounds; for the particular data studied here this typically involved. $O(10^1)$ complete conjugate gradient cycles (each cycle consisting of $\approx(170)$ batch iterations, for a total of several thousand iterations). A few (2 or 3) complete conjugate gradient cycles after "levelling off" of the error decrease rate were performed in each case to exclude stopping at an inflection point. The final decision on convergence involved comparison with the results of an ANN with more neurons (e.g. 12 neurons per hidden layer as described above). Upon completion of the training procedure, the network is a mapping N': $\mathcal{R}^3 \times \mathcal{R} \mapsto \mathcal{R}$, where $N'([I(t-2\tau),\ I(t-\tau),\ I(t)];\ E) = \hat{I}(t+\tau)$. For a more convenient understanding of the long term predictions of the network we will study the corresponding map N: $\mathcal{R}^3 \times \mathcal{R} \mapsto \mathcal{R}^3$, where $N([I(t-2\tau),\ I(t-\tau),\ I(t)];\ E = [I(t-\tau),\ I(t),\ \hat{I}(t+\tau)]$. One-step ahead prediction for a fixed value of the potential E is obtained by applying this map once, the two-step ahead prediction by the convolution $N \circ N$ (iterating the map twice), and the long term (the attractor of map N) prediction results from iterating this map indefinitely.

Figure 5 shows selected comparisons of the original attractors (Figure 5-I) with the short term predictions (Figure 5-II; this was generated by a single iteration of the map for each one of the vectors—points plotted in the respective attractor—in Figure 5-I) and the long term prediction (Figure 5-III; this was obtained after indefinitely iterating *any* single input vector from Figure 5-I). From these Figures it is evident that although the short term prediction is acceptable (the training set in Figure 5-I is faithfully reproduced after a single iteration of the map in Figure 5-II), the long term behavior (the attractor of the map N) does not resemble the training set. That the training set lies on the long term attractor of the physical system implicitly follows from the fact that these are actual—and thus stable—experimental observations.

A casual inspection of Figure 5 would seem to indicate that there is very little, if any, relation between the attractor of the original system and the infinite time prediction of the network (the attractor of the map induced by the network). While the original system attractors (Figure 2) are mostly "nice" smooth closed curves, network attractors may appear "wrinkly" (Figure 5-IIIA), may consist of many disconnected pieces (Figure 5-IIIB), may be apparently fractal (Figure 5-IIIC, see also Figure 6C) or may actually consist of a finite number of (periodic) points (Figures 6A and 6B). It would seem that the long term network predictions are hopelessly inaccurate and useless; this is actually not the case. Figure 6 shows that even though the map attractors may appear strikingly different from the real ones in *phase space*, the predicted time series (obtained by drawing straight lines between successive discrete time predictions, see Figures 6A′, 6B′ and 6C′) appear qualitatively very similar to the experimental ones (compared with Figure 1). The fact that the long term time series predictions appear qualitatively very reasonable and quite accurate, while the network attractors in phase space appear completely erroneous, is a direct consequence of the discrete-time nature of our approximation to what really is a smooth continuous-time signal. As a partial explanation of this statement (which is based on mathematical considerations) consider the case when the oscillation period happens to be a rational multiple of the delay time; it should be obvious that for a constant delay and a smoothly

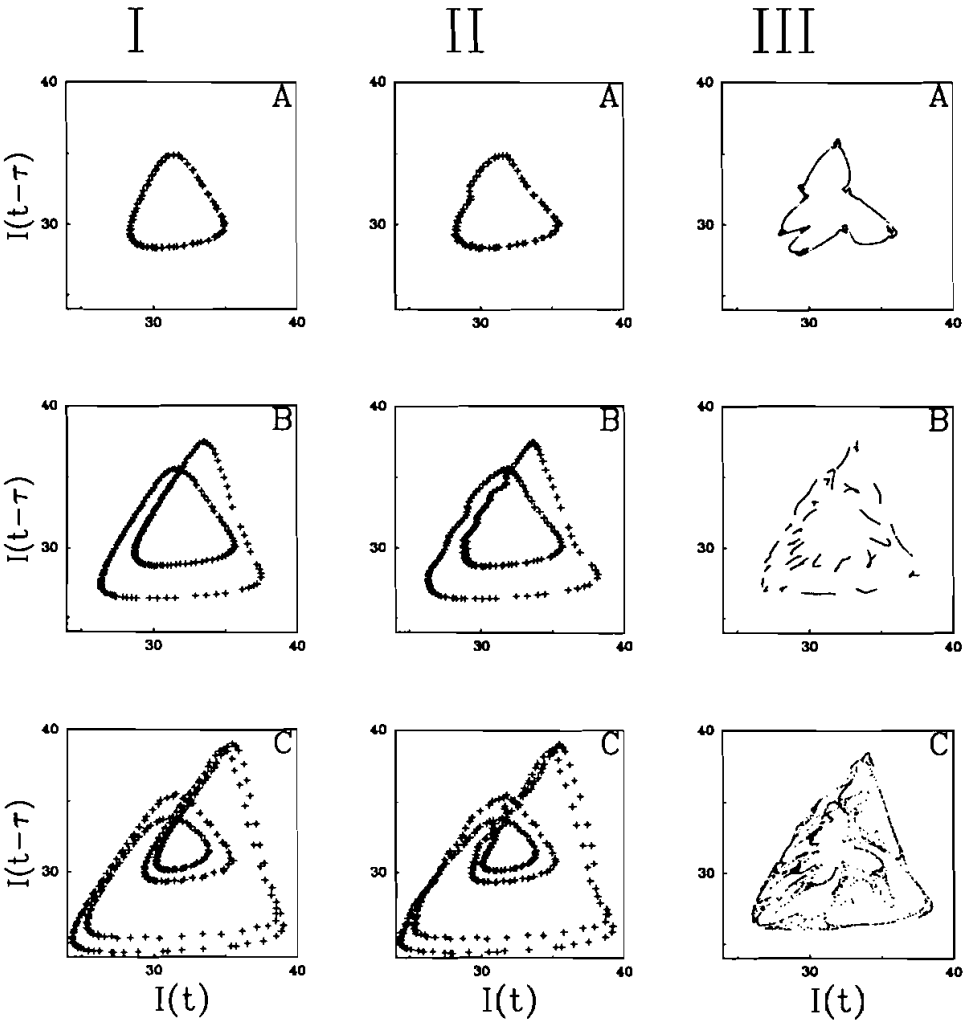

FIGURE 5 Some results of the delay ANN-based map. Figure 5-I shows three experimental attractors for E [V]: (A) 0.6874, (B) 0.6937, (C) 0.6979. Figure 5-II shows the corresponding short term (one-step ahead) predictions of the delay map net. Figure 5-III shows the long term behavior (attractors) the net predicts when using any single input vector and iterating it indefinitely. Clearly, whereas the short term prediction gives satisfactory results, the predicted attractors are somewhat "pathological" in shape and appear wrinkly (IIIA), consist of many disconnected pieces (IIIB) or may even be reminiscent of fractals (IIIC).

changing period with the operating parameter E, this will occur quite often. In such a case even the original smooth curve will appear as a discrete number of points in the delay reconstructed phase space. Of course, changing the voltage infinitesimally should destroy this "resonance" (rational numbers are only countable on the real line). Using discrete maps, however, as the underlying dynamical system tends to enhance such "locking" or "resonance" phenomena: a

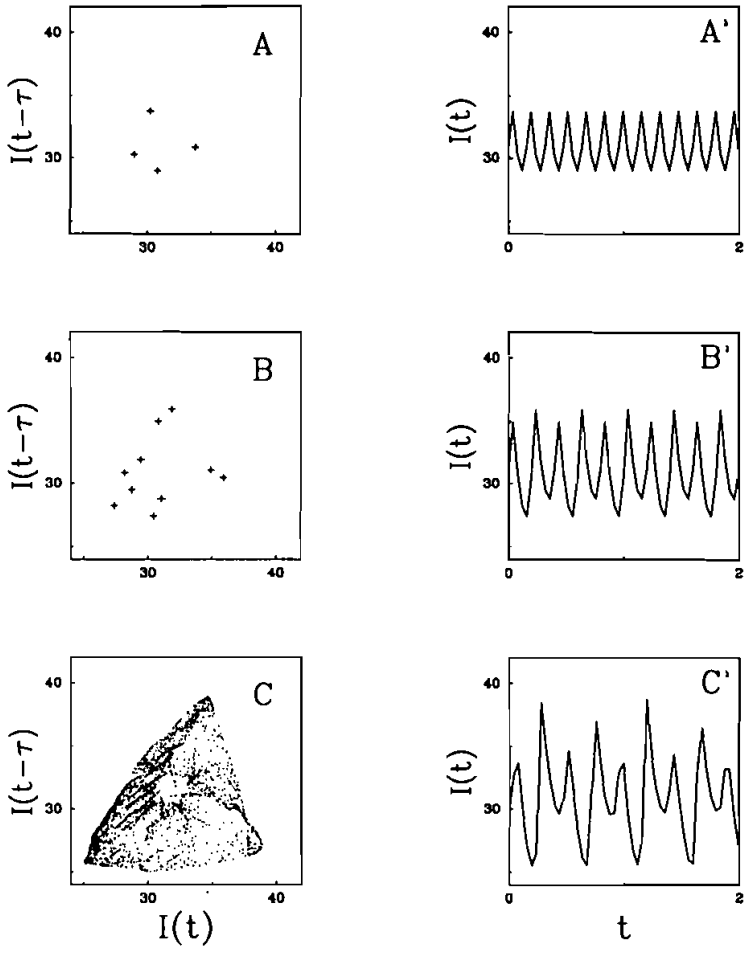

FIGURE 6 Attractors (A, B, C) and corresponding time series segments (A', B', C') predicted by the delay map ANN for E [V]: (A) 0.681 (period-1), (B) 0.691 (period-2) (C) 0.6964 (period-4?). In A and B the attractors are phase locked; this can not occur for limit cycles, but is natural in a one parameter family of maps. Note that though the predicted attractors differ qualitatively from the experimentally observed ones (Figure 2), the predicted time series (obtained by interpolating straight lines between successive discrete time predictions) appear very similar to the experimental ones (Figure 1).

"locked" attractor consisting of a discrete number of points (as in Figures 6A and 6B) now persists over entire closed intervals of voltage values.

Even when the attractor of a discrete map looks like a perfectly smooth closed curve, it is inherently different from a continuous-time simple periodic oscillation: a trajectory on the continuous-time oscillation will visit *every single point* on the limit cycle within one period T. On the other hand, a trajectory of the discrete map will visit only a finite number of points in a time interval of length T; only when the discrete map is iterated indefinitely will the points on the trajectory start to "fill up" the closed curve. This is not a limit cycle any more; it is an *invariant*

circle of the map. The dependence of invariant circles of maps on parameters can be much more complicated than that of limit cycles (see any textbook on dynamical systems for a discussion, e.g. Guckenheimer and Holmes, 1983). The "pathological" attractor shapes (lockings, wrinkles, fractals) shown in Figures 5 and 6 simply cannot occur for limit cycles; nevertheless, they are quite natural in a one-parameter family of maps. We do not pretend in this very short qualitative exposition to explain why this is the case; there is a large body of literature and extensive ongoing mathematical research on the bifurcations of invariant circles of maps. We simply conclude this discussion with the following statement: using discrete-time ANN based maps to process time series from continuous-time physical processes does not provide a satisfactory description of the qualitative changes (transitions, bifurcations) of the long term solutions of the system. This approach actually introduces a tremendous number of spurious bifurcations due entirely to the discrete-time nature of the map. The data presented in Figures 5 and 6 are illustrative of these "spurious" transitions. Figure 6 also illustrates that this pathology is "invisible" over short intervals in the time domain. Since it is precisely such short time intervals that are important for ANN applications in real-time control, this long term pathology is not important in that context. Only when we are interested in characterizing the dynamics of our experimental system and its bifurcations as operating parameters vary do these considerations become relevant, and that is the underlying motivation of our study.

III.2 Poincaré Map Approach

Since the discrepancy arises because of using maps to fit ODEs, it would naturally not be present if the data to be fitted were discrete *in nature* (as opposed to continuous in nature and discrete because of our method of sampling). As discussed above, a Poincaré section can be used to reduce the continuous time trajectories to a discrete map (first return map), and as $\mathcal{R}^3$ seems to be a plausible embedding space, this results in a 2-dimensional Poincaré section. An alternative method of reducing the continuous time series to a discrete map is to construct next maximum maps by plotting the value of the subsequent maximum versus the present maximum. Such a 1-dimensional map should be capable of capturing the observed period doubling bifurcations. Both approaches (next maximum maps and 2-dimensional Poincaré sections) were used with similar success; the results of the latter approach are presented here. We can use a traditional ANN architecture to fit a discrete Poincaré map. This procedure requires a preprocessing of the data, i.e. putting them in the form of the Poincaré map. First, an appropriate hypersurface is selected; any hypersurface transversally intersecting the continuous trajectories can be used. In our case, with trajectories in $\mathcal{R}^3$, we simply chose the plane $I(t-\tau) = constant$ (see Figure 3). Since the continuous time data were discretely sampled, a simple (linear) interpolation scheme was used to approximate the intersections of the continuous trajectories with the plane. This obviously results in a considerable reduction in the number of available training vectors, as well as in a different configuration of the input and output layers of the ANN (see Table I for a concise description of the ANN

TABLE I

Number of neurons in each layer of the different neural networks used

Neural network type	Input layer	Output layer	Hidden layer	Bottleneck layer
Poincaré map	3	2	10	—
Delay map	4	1	10	—
NLPC extraction	21	20	10	3
ODE neural network	4	3	6	—

architectures used in our work). The input layer now consists of two neurons for the system state ($I(t)$ and $I(t-2\tau)$, since the constant $I(t-\tau)$ is omitted) and again one neuron for the operating parameter. More importantly, the output layer now has two neurons; these predict the point on the plane ($\hat{I}(t)$ and $\hat{I}(t-2\tau)$) at which the continuous time trajectory will next intersect the plane (the next time $I(t-\tau)$ acquires the given constant value). Both outputs now have to be fed back into the input in order to iterate the net. Typical training information for the two networks used in the Poincaré map approach is as follows: training set $\approx$ (500) points; test set (at potential values within the training interval but not used in the training set) $\approx$ (350) points; conjugate gradient batch iterations $\approx$ (5000); error upon convergence for data normalized between 0 and 1; 5% (training set) and 6% (test set). Because of the linear interpolation involved in obtaining the Poincaré map the error level in training these two ANNs was slightly higher than the $\approx$ (3%) error level in the rest of this paper.

Using this approach we have been able to successfully characterize the dependence of our system on the voltage E and the underlying bifurcations; the results are summarized in Figures 7 and 8. Figure 7a shows a comparison of the long term prediction (attractor) of the Poincaré map based ANN with those of the same Poincaré map of the experimental system. Figures 7a A–F show experimental attractors, while Figures 7a A′–F′ show the corresponding infinite time predictions of the trained net. All the spurious behavior we encountered above has disappeared: the attractors are now both qualitatively and quantitatively captured. What is more important is that the nature of the transitions between them is accurately represented by the net. For example, a period doubling has apparently occurred at some voltage value between Figures 7a A and 7a B. The network does indeed predict that for the intermediate value $E = 0.68956$ V the fixed point of the map becomes unstable with an eigenvalue of the linearization exiting the unit circle in the complex plane through -1, and the bifurcation is predicted to be supercritical (i.e. small amplitude period two solution bifurcates towards increasing values of E).

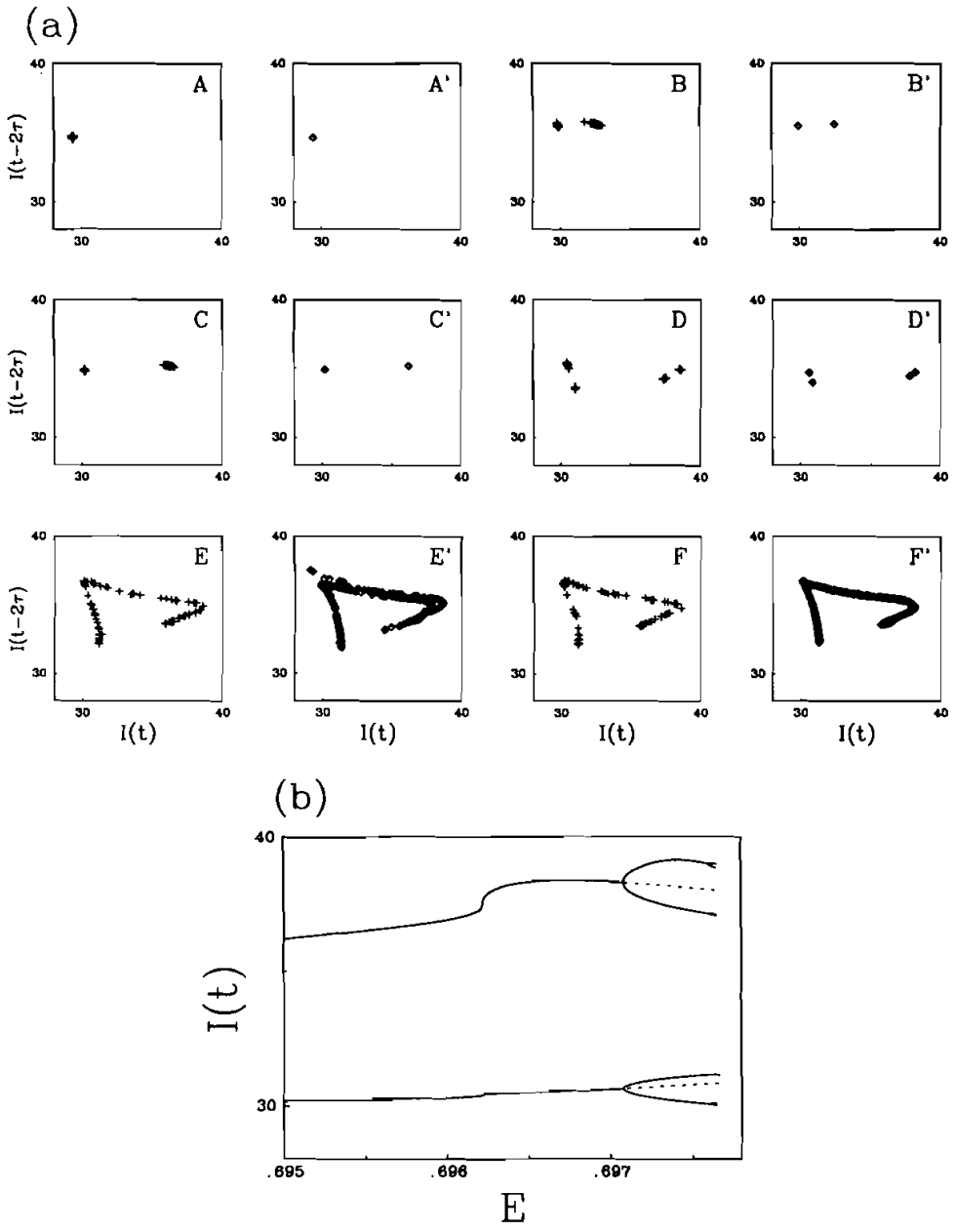

FIGURE 7 Results of the Poincaré map based ANN, trained on data obtained at lower values of E encompassing the initial cascade of period doublings. (a) Comparison of experimental (A–F) and corresponding predicted (A′–F′) attractors for E [V]: (A) 0.6874, (B) 0.6937, (C) 0.6950, (E) 0.6986, (F) 0.6990. Figure (D) shows a period-4 at $E = 0.6979$ V. Because the network predicts the transition to a period-4 at somewhat lower values of E, (D′) shows a predicted period-4 attractor at $E = 0.69754$ V. (b) Segment of the bifurcation diagram predicted by the net. The diagram was computed using AUTO. The first period doubling occurs beyond the range of the graph at $E = 0.68956$ V, the next two period doublings are located at $E = 0.69705$ and $E = 0.69755$ V.

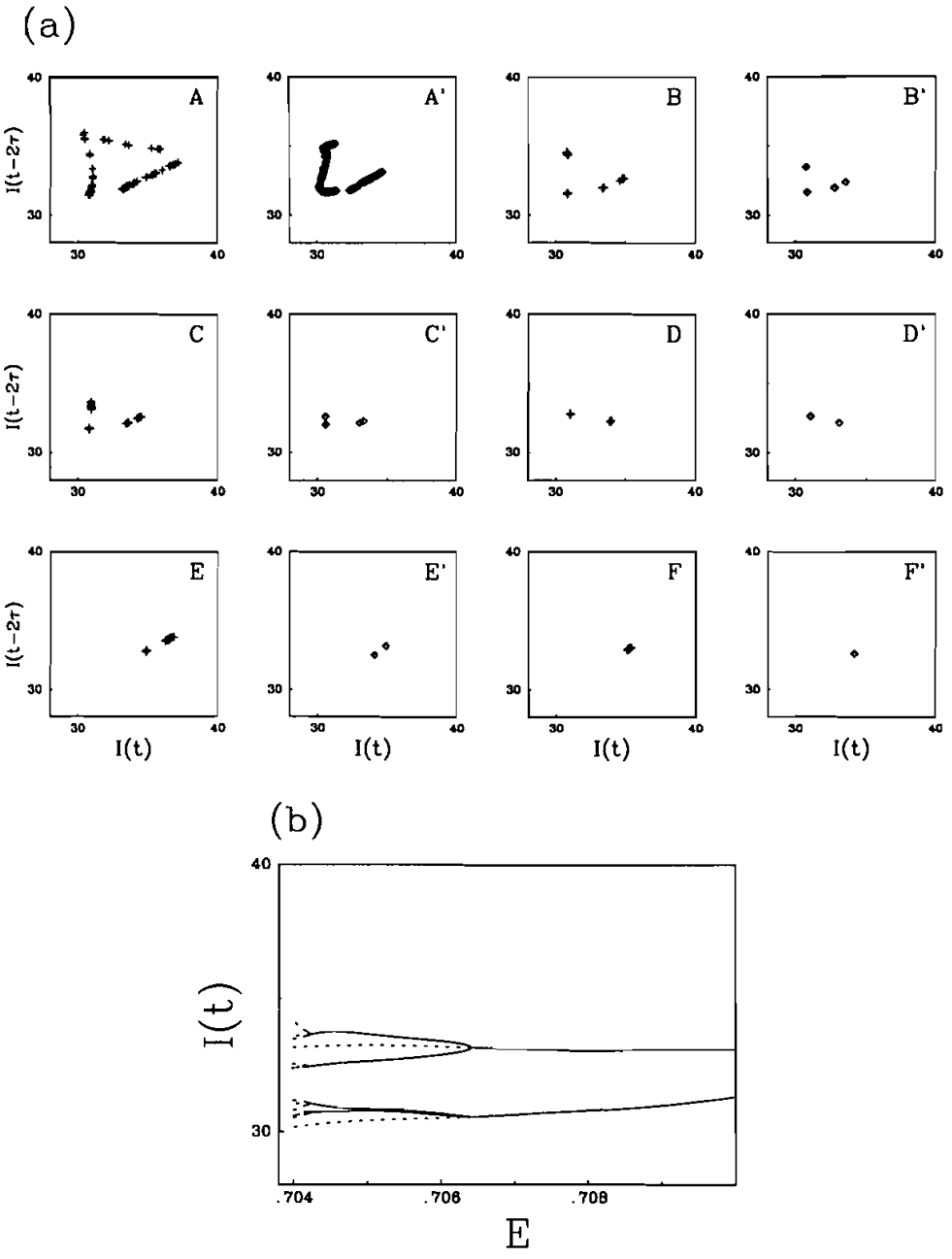

FIGURE 8 Results of the Poincaré map based ANN. Here the training was performed on data obtained at higher values of E encompassing the reverse cascade of period doublings. (a) Comparison of experimental (A–F) and corresponding predicted (A'–F') attractors for E [V]: (A) 0.7040, (B) 0.7055, (C) 0.7063, (D) 0.7092, (E) 0.7220, (F) 0.7236. Notice that the shape of the only fully developed chaotic attractor used in the training (A) is somewhat inaccurately captured; this is probably due to the lack of sufficient data in this regime. (b) Analogous to 7b. The first three period doublings occur at $E = 0.72255$ (off-scale), $E = 0.70642$ and $E = 0.70423$ V, respectively.

Having the closed form for the (approximate) Poincaré map allows us to use a wealth of powerful analytical and computational tools from dynamical systems theory. For example, we can use numerical continuation/bifurcation algorithms to systematically analyze the behavior as a function of E and the relevant bifurcations. This offers a considerable advantage over simple simulations forward in time, especially close to bifurcation values of the parameter. At such values marginal stability of the solutions makes long integrations necessary, and often obscures the nature of the solutions. Figure 7b shows a segment of the bifurcation diagram predicted by the Poincaré map based ANN. This bifurcation diagram was computed using the software package AUTO developed by E. Doedel (Doedel, 1981, Doedel and Kernévez, 1986). The period-2 to period-4 period doubling is predicted to occur at $E = 0.69705$ V. The period-4 to period-8 period doubling is barely visible in the graph at $E = 0.69755$ V; the first period doubling, not shown in Figure 7b, is predicted at $E = 0.68956$ V. Between the extreme training values of the voltage $E = 0.6874$ and $E = 0.699$ V that were used in this training set, the map does predict the sequence of period doublings from period-1 to apparently chaotic behavior and does indeed reasonably capture the structure of the apparently chaotic attractor in phase space. It is important, however, to note that while the qualitative sequence of bifurcations is correctly reproduced, the exact predicted bifurcation values of the voltage (dynamics at intermediate values of E not contained in the training set, which the network has to interpolate) are not necessarily quantitatively accurate. While for example the network predicts the period-4 to period-8 doubling at $E = 0.69755$ V, we know that the original system has this bifurcation at slightly higher values of E. This quantitative discrepancy does not affect the qualitative sequence of predicted bifurcations. Furthermore, these predictions can be used to design further experiments to accurately pinpoint the exact bifurcation values. To the extent that the experimental data can sufficiently resolve the dynamics, incorporating this new data in the training set will naturally improve the accuracy of the predicted bifurcation.

Figure 8 shows the results of another Poincaré map based set of predictions. The training set in this case consisted of time series obtained at the higher range of values of E (between 0.704 and 0.724 V) when a reverse cascade of period doublings from chaos to period-1 oscillations is observed. Figure 8a shows comparisons of experimental and predicted attractors (analogous to Figure 7a), while Figures 8b shows a portion of the corresponding bifurcation diagram computed using AUTO (analogous to Figure 7b). Two issues should be stressed again. The first is that the qualitative sequence of predicted bifurcations appears to be correct, even if their exact location is not perfectly accurate. The "correct" bifurcation sequence is considered here to be the simplest rational explanation of qualitative transitions between successive phase portraits. The second is that one-step ahead prediction is essentially perfect through the entire parameter range (and thus not shown). But in this case even the infinite time behavior is well captured over the parameter interval.

While this Poincaré map based approach does bypass the spurious bifurcation behavior we discussed above, it does have certain drawbacks. Predictions are

made for the particular Poincaré map used in training. While we can have a very good idea of where the trajectory will next intersect our chosen plane, we do not know how long it will take, nor what it will look like during this time interval. More importantly, unless a steady state happens to lie on the chosen plane, it will not be seen at all; this procedure is only appropriate for solutions that are periodic in time or "worse" (i.e. more complicated). As a result, bifurcations involving steady states, from simple turning points and Hopf bifurcations to global bifurcations (e.g. Shil'nikov loops) will not be captured using this approach. In order to meet this latter problem directly and still avoid the spurious bifurcations we discussed above, one must attempt to fit a continuous time system with a model of similar nature.

III.3 Continuous Time Approach

An n-th order ODE, e.g.

$$y''' = f(y, y', y''; x), \qquad y' \equiv \frac{dy}{dt}$$

can be easily transformed into a set of n coupled first order ODEs,

$$y_1' = y_2$$

$$y_2' = y_3$$

$$y_3' = f(y_1, y_2, y_3; x)$$

$$y_1 \equiv y, \qquad y_2 \equiv y', \qquad y_3 \equiv y''$$

with additional dependent variables including up to the $(n-1)$ time derivative of the original dependent variable (x is the operating parameter). Packard *et al.* (1980) actually used these derivatives (instead of time delays) in order to reconstruct attractors in phase space. An obvious approach to fitting an ODE to our continuous time data would be to evaluate time derivatives ($y_2 \equiv y'$, $y_3 \equiv y''$, and $y_3' \equiv y'''$ in our example) from the data by numerical differentiation. One would then train a traditional ANN of the type described in section III.1 (with the inputs in our example being the measured y, and the numerically obtained y' and y'') to approximate the right-hand side of the ODE ($f(y, y', y'')$ in our example, with target value the numerically obtained y'''). In the case of a single time series it will obviously be sufficient to train the network to predict only the right hand side of the n-th ODE. Upon successful completion of the training, the set of n first order ODEs can be integrated to produce predictions at any desired time using numerical integration. The most obvious problem of this approach is the notorious numerical sensitivity of the estimation of numerical derivatives, especially high order ones, from time series, which is aggravated by the presence of noise.

We attempt here to construct a set of ODEs from the data using state measurements only, without direct numerical evaluation of time derivatives from the time series. Both the inputs and the outputs of the net we want to construct will involve only state measurements. The target values (the state of the system at some

future time $t + \tau$) are the result of integrating the (unknown) right-hand side of the ODEs; this integration is performed by the physical system itself. We must therefore attempt to approximate the unknown right-hand side of the ODEs from our experimental knowledge of the *results* of integrating these ODEs (for time τ, with known initial conditions, the state at time t). Let us assume that the experimental data (the result of "physically" integrating the "true" underlying ODE) are practically indistinguishable from the result of numerically integrating this ODE using a simple numerical integration scheme. We will use an ANN which will emulate this numerical integration scheme, and we will train it on the experimental data. In the numerical experiments presented below, we have constructed a network emulating a fourth order Runge–Kutta integrator; such an explicit integrator was satisfactory for our data.

Consider the autonomous ODE

$$\dot{\vec{Y}} = \mathbf{f}(\vec{Y}; \vec{X})$$

$$\vec{Y} \in \mathcal{R}^n, \qquad \vec{X} \in \mathcal{R}^p, \qquad \mathbf{f}\colon \mathcal{R}^n \times \mathcal{R}^p \mapsto \mathcal{R}^n$$

The result $\vec{Y}_{n+1}$ of numerically integrating this equation for fixed values of the operating parameters $\vec{X}$ with initial conditions $\vec{Y}_n$ using a fourth order Runge–Kutta method and a time step of h is given by:

$$\vec{Y}_{n+1} = \vec{Y}_n + \frac{1}{6}(\vec{k}_1 + 2\vec{k}_2 + 2\vec{k}_3 + \vec{k}_4)$$

where:

$$\vec{k}_1 = h\mathbf{f}(\vec{Y}_n; \vec{X})$$

$$\vec{k}_2 = h\mathbf{f}\left(\vec{Y}_n + \frac{\vec{k}_1}{2}; \vec{X}\right)$$

$$\vec{k}_3 = h\mathbf{f}\left(\vec{Y}_n + \frac{\vec{k}_2}{2}; \vec{X}\right)$$

$$\vec{k}_4 = h\mathbf{f}\left(\vec{Y}_n + \vec{k}_3; \vec{X}\right)$$

Figure 9 shows a schematic representation of our implementation of this scheme. The four boxes marked "Neural Net" are really the same box, repeated four times in this schematic representation to clarify its role in the Runge–Kutta method; this box is a "traditional" ANN with 4 input neurons (a three dimensional state and a single operating parameter), and two hidden layers with 6 nonlinear neurons each. The output layer of this box consists of 3 neurons which, upon eventual convergence of the training, will provide our estimates of the three elements of the right-hand side of the ODEs. The remaining connections in the figure are linear, with preset weights dictated by the Runge–Kutta algorithm. The target values used in the training are the states of the system at the next time step. Because of the repeated iterations through the network involved in evaluating $\vec{Y}_{n+1}$, it is not possible to use the chain rule in the simple "layer by

$$\vec{Y}_{n+1} = \vec{Y}_n + \frac{1}{6}(\vec{k}_1 + 2\vec{k}_2 + 2\vec{k}_3 + \vec{k}_4)$$

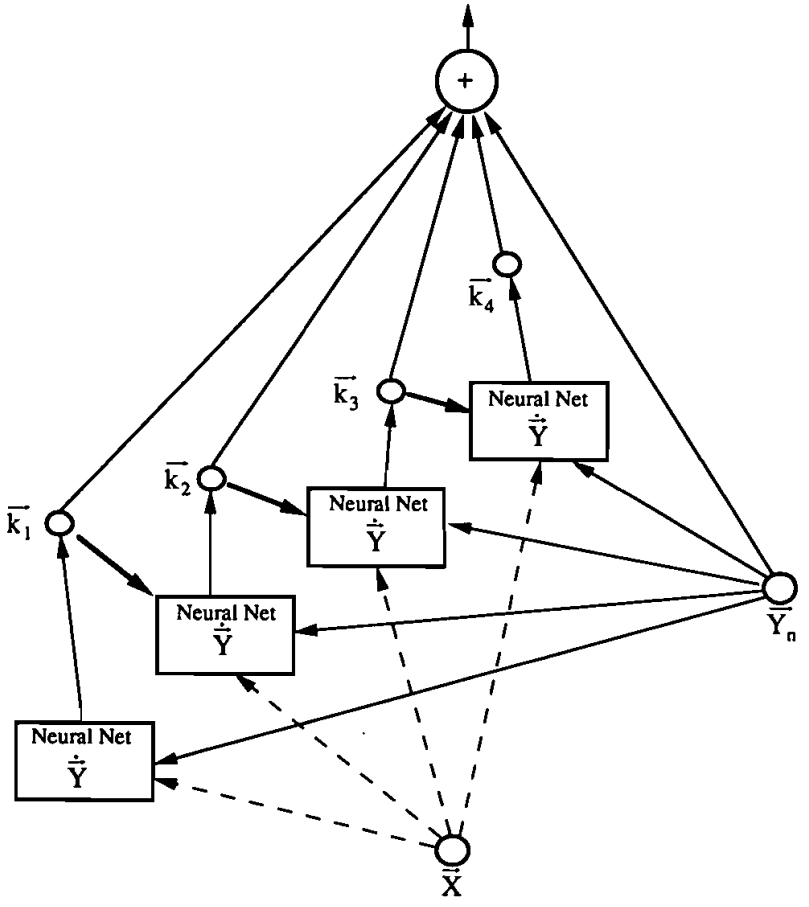

FIGURE 9 Schematic representation of the ANN-based implementation of the 4th order Runge–Kutta method. A single two hidden layer traditional ANN (marked $\dot{\vec{Y}}$ and repeated four times for clarity in the diagram) is trained to evaluate the right-hand side (derivatives) of the ODEs given the current states ($\vec{Y}_n$) and parameters ($\vec{X}$). Its output is recurrently processed through it according to the Runge–Kutta formulae given in the text. The output layer performs the final summation, and predicts the state values ($\vec{Y}_{n+1}$) after one time step h.

layer" fashion of a feedforward net to evaluate partial derivatives. Instead, the expression for $\vec{Y}_{n+1}$ was used to evaluate these partial derivatives with respect to all weights simultaneously.

Before we proceed to the presentation of the results of using this approach on our experimental data, there is one final important issue to be discussed. This concerns the actual input data we chose to use in training the ODE-based net. We chose to preprocess the time series using what Kramer (1991) calls "Nonlinear Principal Component Analysis" (NLPCA). The nature of the experimentally observed solutions and their bifurcations suggest a three dimen-

sional phase space, which means that we want to approximate a set of three coupled ODEs. We already reconstructed the data in such a 3-D space using time delays; an obvious choice will be to try and fit the trajectories in this delay-reconstructed space. The time delayed measurements, however, are not uncorrelated, and it would be desirable to use a set of coordinates in which the data would be as uncorrelated as possible. We chose to extract the three first principal components from windows of our time series consisting of 20 measurements separated by three sampling intervals. The length of this window (0.12 s) is comparable to the interval used in the delay representation; the reason for not using all sixty measurements in this window is a practical issue of computational economy. The sampling rate is so fast that the intermediate points do not add important information; this is supported by the success we have had in eventually reproducing the data. A linear principal component analysis has been used (Bromhead and King, 1986, Albano *et al.*, 1989) to produce better embeddings for attractor reconstruction purposes. The issues of optimal window length, overlap of successive windows, number of principal components retained, and noise rejection capabilities of the procedure are discussed by Broomhead and King (1986). We chose instead to use an ANN based NLPCA procedure, which is in principle capable of detecting nonlinear correlations in the data. Kramer (1991) presents a detailed discussion of the motivation behind such an approach, its properties, its relation to linear principal component analysis (Oja, 1982), and its history in the context of neural network research.

The particular NLPC network is schematically represented in Figure 10. The input layer consisted of 20 linear neurons in which the vector of measurements $(I(t), I(t-3\theta), I(t-6\theta), \ldots, I(t-57\theta))$ was fed, along with one additional neuron for the parameter value. $\theta = 0.002$ s is the sampling interval. After this

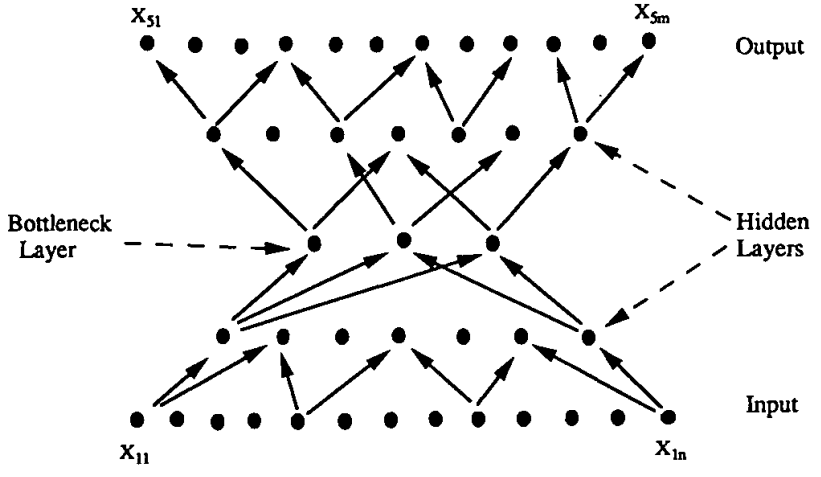

FIGURE 10 Schematic representation of the autoassociative ANN used for data preprocessing. The input consists of a (long) vector of time series measurements, which also constitutes the target values. The outputs of the three neurons of the intermediate bottleneck layer are the processed data ("nonlinear principal components", the objective of the training).

information is processed through a hidden nonlinear layer (consisting here of 10 neurons), the data were passed through a "bottleneck" layer consisting of 3 linear neurons. The output of these neurons constitutes the three "nonlinear principal components" we seek. Since these three values are expected to contain the information in the original time series window, the network is trained to learn the identity mapping ("autoassociation", "self-supervised backpropagation"). The output of the three bottleneck neurons is processed through one more 10-hidden nonlinear neuron layer, and finally through the 20 linear neuron output layer, for which the target values are the original time series measurements. Kramer discusses both a simultaneous and a sequential architecture for determining nonlinear principal components; we used the simultaneous approach. The training of the network was again performed using conjugate gradients on a set of 620 vectors sampled from time series in the interval ($E = 0.678$ to $E = 0.6979$ V) containing the original Hopf bifurcation of the steady state and the first pair of period doublings. The results of this procedure (i.e. the outputs of the three bottleneck neurons) were then used to train the ODE-based network discussed above. The input vectors for this latter training consisted of the three NLPCs (at time t) plus an additional input for the parameter value. The target values of the training were the three NLPCs at time $t + h$ ($h = 3\theta$). We use regularly spaced data here, although the algorithm is capable of using data at irregular time intervals. The training of the ODE-based network took $O(10^1)$ complete conjugate gradient cycles to reach convergence.

Figure 11a shows the short term prediction capabilities of the ODE approach: time series obtained from integration (for time interval 6θ) of the set of ODEs resulting from the training (11a-II) are compared in real space to experimental time series (11a-I). The ODE integration is carried out in NLPC space and the results are transformed back to real space. A projection (in NLPC space) of the long time attractors of the ODE integration is also included in Figure 11a-II', and is compared to the corresponding experimental attractors in the same projection. The attractors are not quantitatively the same, but their shapes are remarkably qualitatively similar. More importantly, Figure 11b shows the bifurcation diagram for the ODEs obtained using AUTO. Not only are the period doublings of the limit cycles satisfactorily captured, but also the original Hopf bifurcation of the steady state to a period-1 limit cycle is contained in the diagram. Furthermore, the ODE is capable of extrapolating further period doublings and chaotic behavior for values beyond the training range. The correct shape of the periodic attractors (not captured by the traditional ANN approach in secton III.1) and the steady state and its bifurcations (not captured by the Poincaré-map based approach in section III.2) show that our ODE approach is indeed capable of reproducing the entire range of dynamical characteristics of the experimental data.

IV. DISCUSSION AND CONCLUSIONS

We have discussed some of the problems traditional discrete time ANNs present when used to approximate the long term dynamics of continuous time systems.

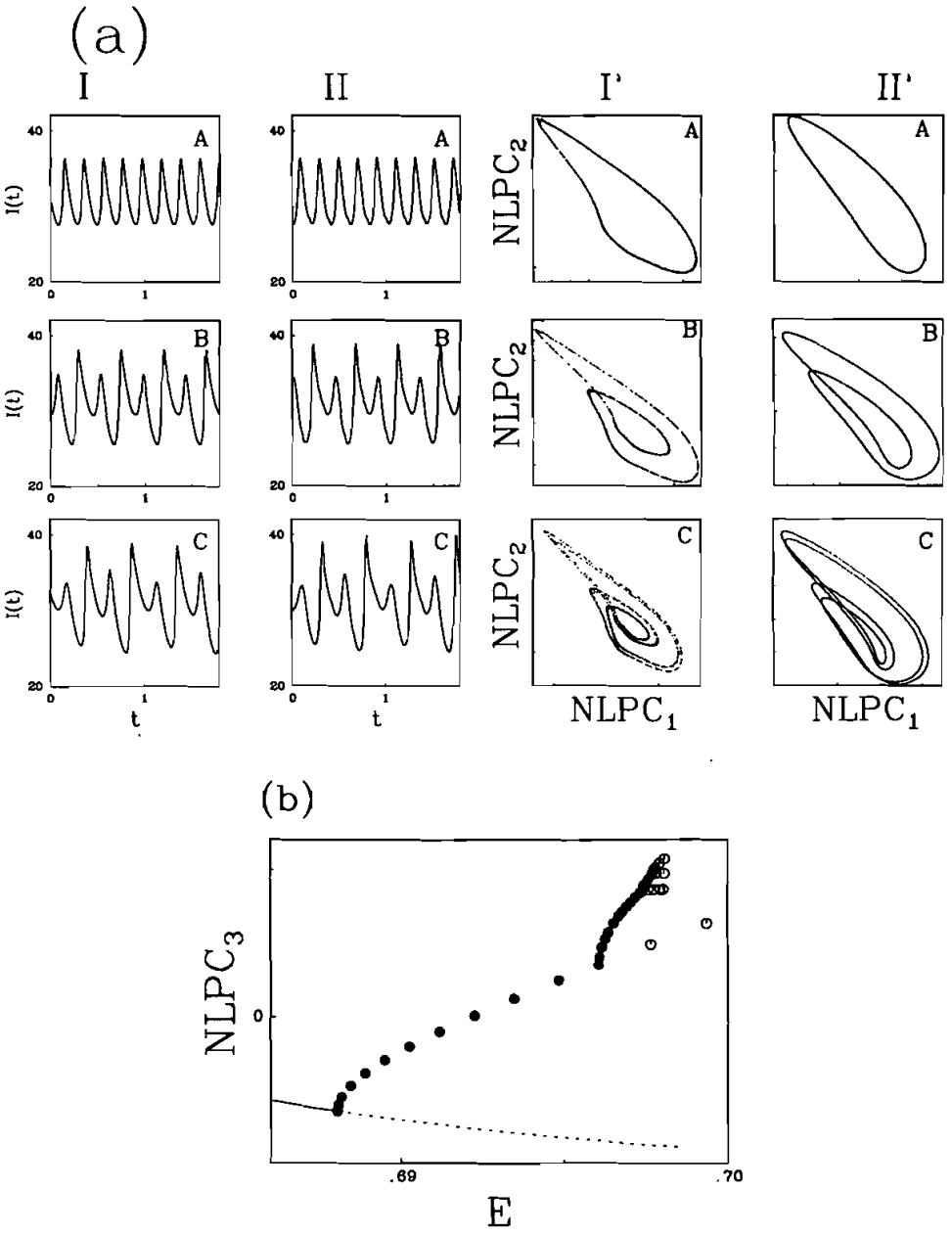

FIGURE 11 (a) Experimental time series (I) and corresponding short term prediction of the ODE-based ANN (II). E [V]: (A) 0.6919, (B) 0.6956, (C) 0.6979. The corresponding experimental attractors projected in NLPC-space are shown in Figure (I'). (II') shows ODE-based long term attractors of similar nature; E [V]: 0.6919 (II'-A), 0.6966 (II'-B) and 0.6976 (II'-C). (b) Partial bifurcation diagram of the ODE-based ANN computed using AUTO. Solid circles denote stable and open circles denote unstable limit cycles. Notice the accurately captured supercritical Hopf bifurcation of the steady state (solid line: stable steady state, dashed line: unstable). Only three of the (infinitely many) predicted period doublings are shown.

We presented an alternative ANN-based method which does not exhibit these shortcomings; its capabilities were illustrated by successfully reproducing a nontrivial range of experimentally observed dynamic transitions in Cu electrodissolution. A number of alternative approaches to characterizing the dynamics of nonlinear systems are currently used for the same purposes. Most approaches employ discrete map techniques; representative examples include local linearization (Farmer and Sidorowich 1987, 1988, Kostelich and Yorke, 1990), rational polynomial approximations (Bayly *et al.*, 1987), ANN-based methods with sigmoidal (Lapedes and Farber, 1987a,b, Casdagli, 1989, Hudson *et al.*, 1990) and alternative (e.g. radial) basis functions (Moody and Darken, 1988, Casdagli, 1989). Attempts have also been made to fit ODEs (e.g. Cremers and Hübler (1987)). Our efforts were directed towards an ANN-based approach which would successfully capture continuous-time dynamics with emphasis on the detection of bifurcations.

Data preprocessing was an important issue in our study: finding the correct embedding space, choosing a good set of coordinates for this space, achieving satisfactory noise-rejection without corrupting the underlying dynamics, all hinge upon successfully performing this task. We chose the simultaneous "nonlinear principal component" analysis here with satisfactory results. We feel it is important to further experiment with the comparative success of using the results of alternative linear and particularly nonlinear approaches to data preprocessing in identifying bifurcations. Theoretical advances in the direction of NLPCA should be helpful in directing such efforts.

If the underlying physical system is described by a stiff ODE the simple procedure we described here based on an explicit integrating scheme will most probably fail. In our case, however, the agreement between the experimental and the ANN-based attractors (Figure 11) should be an indication that, whatever the "real" ODEs underlying the system are, they were not stiff in the regime experimentally studied. Applying a similar procedure based on an *implicit* integration scheme would obviously make the process much more computationally intensive, since an iterative nonlinear equation solver will be required to evaluate the network output at every single time step. This is drastically different from the simple function evaluations required for an explicit integrator like the one we are using here. The difference is also reflected in its ANN implementation (Figure 9) which is *not* a recurrent network (Pineda, 1987): different information is processed through the same net at every step, and its inputs and outputs are *not* expected to be identical upon convergence. Because the outputs of an implicit integrator scheme are not closed form functions of the inputs and the various parameters, evaluating the derivatives necessary for training such a network would also become a difficult task; the implementation of such an implicit integrator would indeed constitute a recurrent ANN.

A final issue worth mentioning is that all the experimental data presented here were obtained after initial transients die out, and should therefore lie *on* the corresponding attractors. Training the network on the attractor does not provide information regarding the stability of the attractor to finite perturbations (beyond the level of experimental noise). Our networks were capable of reasonably well extrapolating attractor stability. In order to make this extrapolation quantitative

it is important to experimentally capture information about the stability of the attractor; this can be accomplished by applying appropriate perturbations, sampling the resulting transients as they approach the attractor in phase space, and including these data in the training sets.

ACKNOWLEDGEMENTS

This work was supported in part by DARPA/ONR (N00014-91-J-1850), the National Science Foundation (IGK and JLH) and Shell Development Company. The support of the DFG through a Fellowship to KK, the Packard Foundation through a Fellowship to IGK, and the Mexican Ministry of Public Education (Instituto Tecnológico de Celaya) to RRM is also gratefully acknowledged. Jack Hudson thanks George Bankoff for his guidance and inspiration, and for the lively discussions at Northwestern. Fortunately, we have been able to continue the interchange which has remained as interesting and as animated as that we first had 30 years ago.

Note added in Proof: The use of ANNs to obtain derivative information and continuous models is an active research topic. Since this manuscript was first submitted, a number of publications on the subject have appeared. We consider the following particlarly relevant:

Cardaliaguet, P., and Euvrard, G. Approximation of a function and its derivative with a neural network. *Neural Networks,* **5,** 207–220 (1992).

Chu, S.R., and Shoureshi, R. A neural network approach for identification of continuous-time nonlinear dynamic systems. *Proceedings of the* 1991 *American Control Conference,* **1,** 1–5 (1991).

Gallant, A.R., and White, H. On learning the derivaties of an unknown mapping with multilayer feedforward networks. *Neural Networks,* **5,** 129–138 (1992).

REFERENCES

Albano, A.M., Muench, J., Schwartz, C., Mees, A.I., and Rapp, P.E., Singular-value decomposition and the Grassberger-Procaccia algorithm. *Phys. Rev. A,* **38,** 3017–3026 (1988).

Albahadily, F.N., Ringland, J., and Schell, M., Mixed-mode-oscillations in an electrochemical system. I. A Farey sequence which does not occur on a torus. *J. Chem. Phys.* **90,** 813–821 (1989).

Bayly, B., Goldhirsch, I., and Orszag, S., Independent degrees of freedom of dynamical systems. *J. Sci. Comp.,* **2,** 111–121 (1987).

Bhat, N., and McAvoy, T.J., Use of neural nets for dynamic modeling and control of chemical process systems. *Proc.* 1989 *American Control Conf.,* Pittsburgh, PA, June, 1989, 1342–1347.

Broomhead, D.S., and King, G.P., Extracting qualitative dynamics from experimental data. *Physica D,* **20,** 217–236 (1986).

Casdagli, M., Nonlinear prediction of chaotic time series. *Physica D,* **35,** 335–356 (1989).

Cremers, J., and Hübler, A., Construction of differential equations from experimental data. *Z. Naturforsch.,* **42a,** 797–802 (1987).

Cybenko, G., Approximation by superposition of a sigmoidal function. *Math. Control Signals Systems,* **2,** 303–314 (1989).

Doedel, E.J., and Kernévez, J.P., AUTO: A program for continuation and bifurcation problems in ordinary differential equations. Applied Mathematics Report, California Institute of Technology, 226 pages (includes the AUTO 86 User Manual), 1986.

Doedel, E.J., AUTO: A program for the automatic bifurcation analysis of autonomous systems. *Cong. Num.,* **30,** 265–284 (1981).

Farmer, J.D., and Sidorowich, J.J., Exploiting chaos to predict the future and reduce noise. In: *Evolution, Learning and Cognition* (Y.C. Lee, ed.), World Scientific, Singapore, 277–330 (1988).

Farmer, J.D., and Sidorowich, J.J., Predicting Chaotic Time Series. *Phys. Rev. Letters*, **59**, 845–848 (1987).

Fraser, A.M., and Swinney, H.L., Independent coordinates for strange attractors from mutual information. *Phys. Rev. A*, **33**, 1134–1140 (1986).

Guckenheimer, J., and Holmes, P., *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Springer, Heidelberg (1983).

Hecht-Nielsen, R., Theory of the backpropagation neural network. *Proc. Intl. Joint Conf. Neural Networks*, Washington, DC, (1989).

Hornick, K., Stinchcombe, M., and White, H., Multi-layer feedforward networks are universal approximators. *D-008 UCSD*, La Jolla, CA (1988).

Hoskins, J.C., and Himmelblau, D.M., Artificial Neural Network models of knowledge representation in chemical engineering. *Computers chem. Engng.*, **12**, 881–890 (1988).

Hudson, J.L., Kube, M., Adomaitis, R.A., Kevrekidis, I.G., Lapedes, A.S., and Farber, R.F., Nonlinear signal processing and system identification: Applications to time series from electrochemical reactions. *Chem. Eng. Sci.*, **45**, 2075–2081 (1990).

Irie, B., and Miyake, S., Capabilities of three layer perceptrons. *Proc. Intl. Joint Conf. on Neural Networks*, San Diego, **1**, 641–648 (1988).

Karnin, E.D., A simple procedure for pruning back-propagation trained neural networks. *IEEE Trans. Neural Networks*, **1**, 239–242 (1990).

Kostelich, E.J., and Yorke, J.A., Noise reduction: finding the simplest dynamical system consistent with the data. *Physica D*, **41**, 183–196 (1990).

Kramer, M.A., and Leonard, J.A., Diagnosis using backpropagation neural networks-analysis and criticism. *Computers chem. Engng.*, **14**, 1323–1338 (1990).

Kramer, M.A., Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.*, **37**, 233–243 (1991).

Lapedes, A.S., and Farber, R.F., How neural networks work. In: *Neural Information Processing Systems* (D.Z. Anderson, ed.), AIP Press, 422–456 (1987a).

Lapedes, A.S., and Farber, R.F., Nonlinear signal processing using neural networks: prediction and system modeling. *Los Alamos Report LA-UR* 87-2662 (1987b).

Liebert, W., and Schuster, H.G., Proper choice of the time delay for the analysis of chaotic time series. *Phys. Letters A*, **142**, 107–111 (1989).

Mayer-Kress, G. (ed.). *Dimensions and Entropies in Chaotic Systems*, Springer, Heidelberg (1986).

McAvoy, T.J., Wang, N.S., Naidu, S., Bhat, N.V., Gunter, J., and Simmons, M., Interpreting biosensor data via back propagation. *Proc. Intl. Joint Conf. Neural Networks*, Washington, DC, **1**, 227–233 (1989).

Moody, J., and Darken, C., Learning with localized receptive fields. In: *Proceedings of the* 1988 *Connectionist Models Summer School*, Touretzky, Hinton and Sejnowski (eds.), Morgan Kaufmann, Publishers (1988).

Naidu, S.R., Zafiriou, E., and McAvoy, T.J., Use of neural networks for sensor failure detection in a control system. *IEEE Control Systems Magazine*, April 1990, 49–55.

Packard, N.H., Crutchfield, J.P., Farmer, J.D., and Shaw, R.S., Geometry from a time series. *Phys. Rev. Letters*, **45**, 712–716 (1980).

Pineda, F.J., Generalization of back-propagation to recurrent neural networks. *Phys. Rev. Letters*, **59**, 2229–2232 (1987).

Oja, E., A simplified neuron model as a principal component analyzer. *J. Math. Biology*, **15**, 267–273 (1982).

Rummelhart, D., McClelland, J., and the PDP research group. *Parallel distributed processing*, Vol. 1. M.I.T. Press, Cambridge, MA (1986).

Sanger, T.D., Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, **2**, 459–473 (1989).

Schell, M., and Albahadily, F.N., Mixed-mode oscillations in an electrochemical system. II. A periodic-chaotic sequence. *J. Chem. Phys.*, **90**, 822–828 (1989).

Takens, F., Detecting strange attractors in turbulence. In: *Dynamical Systems and Turbulence* (D.A. Rand and Young, L.S., eds.), *Lect. Notes in Math*, Springer, Heidelberg 366–381 (1981).

Usui, S., Nakauchi S., and Nakano, M. Reconstruction of munsell color space by a five-layered neural network. *Proceedings 1990 IEEE INNS IJCNN*, San Diego, CA. **2**, 515–520 (1990).

Ydstie, B.E., Forecasting and control using adaptive connectionist networks. *Computers chem. Engng.*, **14**, 583–599 (1990).