Radial-Basis Models for Feedback Systems With Fading Memory

David M. Walker, Nicholas B. Tufillaro, and Paul Gross

***Abstract*—We discuss how to build nonlinear input-output models of low-dimensional deterministic systems for both static and dynamic (feedback) systems with "fading memory." To build the dynamic models a new form of radial-basis functions is introduced which, in the absence of an input, have the property that they converge to a constant solution. The utility of these models is illustrated by building accurate and stable models for electronic circuits with dynamic (memory) effects.**

***Index Terms*—Embedding, nonlinear, system identification.**

I. Introduction

This paper describes a dynamical-systems approach to nonlinear system identification [1]. In particular, we examine the question, is it possible to build data-driven, stable, "free-running" models of low-dimensional dynamic systems subject to stochastic drives? The nonautonomous systems we consider can typically be divided into two parts, an "internal" deterministic dynamics, and an "external" (possibly stochastic) drive term. The "low dimensionality" mentioned above refers only to the internal dynamics. Due to the data requirements in higher-dimensional spaces, the methods are practically limited to systems for which the asymptotic solutions of the internal dynamics can be modeled with only a few degrees of freedom (typically less than seven in our applications). Models of this type commonly arise in electronic circuit applications [2]. The models are called "free running" when they have feedback terms (also known as autoaggresive terms) and the inputs to the model are the (time-dependent) drive terms and a single set of (seed) initial conditions. Free-running models often lead to unstable solutions, and thus are only of practical use for short-term prediction [3], [4]. However, the systems we want to consider often have the property that the input signals far in the past have almost no effect on the present state—the so called "fading memory" assumption [5]. As described by Boyd and Chua, fading memory is closely related to the fact that the internal dynamics of the system can have a unique asympototic state [6]. Therefore, in this paper we explicitly build this property into our models in an attempt to create stable, free-running dynamic models. Stable free-running models have practical uses in applications involving simulation where long term prediction is desirable.

The problem of nonlinear systems identification is large with many unresolved issues. Generally, the problem can be divided into an number of sub-problems such as excitation design, model structure selection, modeling fitting and model verification [7]. We will briefly describe how each of these issues is handled in the case studied, however, our main aim is to illustrate how to build qualitative *a priori* information into the identifiation procedure. In this respect, our paper is similar to the recent paper of Aguirre *et al.* [8]

Manuscript received August 16, 2000; revised February 2, 2001 and March 27, 2001. This work was supported in part by the National Science Foundation under Grant NSF #9724707. This paper was recommended by Associate Editor G. Chen.

D. M. Walker is with the Center for Applied Dynamics and Optimization, The University of Western Australia, Nedlands, Perth, WA 6907, Australia.

N. B. Tufillaro and P. Gross are with Agilent Technologies, Palo Alto, CA 94303-0867 USA.

Publisher Item Identifier S 1057-7122(01)07712-1.

We begin with a brief overview of a dynamical-systems approach to input-output modeling. Other methods that have been developed for nonlinear system identification include Volterra Series, neural nets, and cluster weighted models to name a few [9].

A dynamical-systems approach to "black box" or "behavioral modeling" based on the Takens Embedding Theorem was first suggested by Casdagli [10]. The use of delay variables in the structure of these dyanamical models is similar to that originally studied by Leontaritis and Billings [11], and is common in linear time-series analysis and system identification [9].

This approach to nonlinear system identification is sometimes called "dynamic reconstruction theory [12]" and begins with a state-space representation

$$\dot{x} = f(x(t), u(t)) \quad y(t) = h(x(t)) \tag{1}$$

or their numerical version of difference equations

$$x_{n+1} = f(x_n, u_n). \tag{2}$$

In these equations, f, x, and u are typically vectors and $u(t)$ is the input, drive, or stimulus, $x(t)$ is the state, and $h(t)$ is a measurement function. Attempts to build data-driven state-space models appear hard on at least two counts. First, without any specific form for a model, the relevant dynamical variables x appear to be unknown and second, even if we know what variables are needed to be included, they still may not be accessible to experimental measurements. These issues, essentially the nonlinear order and observability of the model, as well as model selection and calibration are discussed below. Another essential issue in building good models is excitation or experiment design. In this paper, we describe the use of band-limited pseudorandom noise in constructing black-box models. Lastly, a simple metric for model validation is considered.

II. Dynamic Reconstruction Theory, Basis Functions, and Excitations

The key idea of dynamic reconstruction is to embed the measured input-output variables in a higher dimensional space built not just with $u(t)$, $y(t)$, but also transforms of u, y, for example their numerical derivatives.

Due to a theorem of Takens (with an extension to the driven case by Stark [13]) these embedded models can be faithful to the dynamics of the original system. In particular, deterministic prediction is possible from an embedded model which will mimic the actual dynamics. Thus, embedding opens the way toward a general solution of extracting black box models for the observable dynamics of nonlinear systems directly from input-output time-series data. It can solve the fundamental existence problem for a class of nonlinear system-identification problems, however, the gulf between these theoretical results and practical implementation is wide.

Practically, the components' behavior is described by embedding both the inputs and outputs in the form

$$y(t) = G[y(t-\tau), y(t-2\tau), \ldots, y(t-l\tau) \\ u(t), u(t-\tau), \ldots, u(t-(k-1)\tau)] \tag{3}$$

where G is fitted to the data using nonlinear modeling methods such as global polynomials, neural nets, or radial-basis functions [9]. The form of the equations shows a "lag" embedding with a time delay τ, input-lag dimension k, and output-lag dimension l, though in practice we find that better quality models can often be built using other embeddings such as linear transforms, integral and differential transforms and wavelets to help bring out the salient dynamical features in the data.

Given the above model form, the problem now reduces to a number of technical issues such as: 1) determination of the dimensions (k and l); 2) determination of lags (τ) or other forms of embeddings and embedding parameters; and 3) determination of model class G and fitting the model parameters, model validation and design of excitation signals (where possible) for a given model/signal class.

It might be helpful to point out that this relation between a continuous dynamical system and an embedded model built from time-delayed input-output signals can be made explicit in the case of linear systems. The details for an algorithm are described in a book by Franklin [14] which shows how to go from the linear system and it's matrix representation to a model based only on delayed variables. Unfortunately, no explicit constructive proof exists for nonlinear systems.

In the nonlinear case, we can attempt to estimate the embedding parameters directly from the data. For embeddings built from a time delay lag τ

$$y(t+1) = G[y(t-\tau), \ldots, y(t-l\tau), u(t), \ldots, u(t-(k-1)\tau)] \quad (4)$$

we can use an extension of the algorithm for the theory of embedded autonomous systems known as "false nearest neighbors [1]". Basically, we find the smallest "k" and "l" by creating a statistic that checks if vectors close in a delay space are also close in a delay space of greater dimension. If they are not, then we have false neighbors and G is not single valued. This diagnostic is independent of G . Examples of using these and related algorithms in circuits are presented reference [15].

For models built from time delays we need to estimate τ. Again, we make use of a diagnostic from autonomous systems theory. We use either the mutual information or the first zero of the autocorrelation function in determining τ [1].

Once we have a suitable embedding, we next turn to function approximation of G. The black-box models we reconstruct are radial-basis function models. We have experience using such basis functions with some success [16] but other basis functions can be used [9]. A radial-basis function which predicts $y(t+1)$ using the reconstructed state $z(t)$ can be expressed as

$$y(t+1) = \beta + \alpha \cdot z(t) + \sum_{i=1}^{M} \omega_i \phi(\|c_i - z(t)\|) \quad (5)$$

where β, α and ω are constant parameters to be estimated. The c_i are referred to as centers and determining their location and number are the main difficulties in reconstructing the radial-basis models. We use the methods developed by Judd and Mees [17] to help solve this problem. These methods attempt to find a subset of centers from a candidate set which best describes the data. Our candidate set will be taken from the reconstructed data points. The function ϕ is the basis function and a common choice is to use a fixed-width Gaussian function.

Lastly, depending on the application, we build our models from training sets using smooth, band-limited, aperiodic excitations. For the examples shown, we use an excitation signal based on (an ISO95) CDMA (code division multiple access) specification which are of the form

$$u(t) = A\cos(2\pi f_c t) \sum_{i=1}^{48} b_i(t) p_i \quad (6)$$

where the b_i are random 48-dimensional vectors taking the values -1 or $+1$ [18]. The p_i are coefficients of a low pass filter. A is a constant amplitude and f_c is a carrier frequency. The chip rate is one and the bandwidth is roughly 5 kHz. Although the models are very sensitive to the center frequency and bandwidth of the training signal, they appear to be less sensitive to the exact form of the random excitation. The excitation signal was chosen to excite several harmonics.

In the absence of any external excitation, many of the devices we hope to model, converge to a unique (usually constant) solution. We would like our models to have this property. We have developed the following basis function to use in our models

$$\phi(\|\mathbf{c} - z\|) = e^{-(1/2v^2)\|c-x\|^2} \times \left[e^{-(1/2w^2)\|d\|^2} - e^{-(1/2w^2)\|d-u\|^2}\right] \quad (7)$$

where

- x part of z reconstructed using the outputs;
- u parts of z reconstructed from the inputs;
- c "output" centers with the same dimension of x;
- d "input" centers with the same dimension as u;
- v, w fixed-widths of "output" and "input".

We set their values to be the standard deviations of the output- and input-data, respectively. We note that, by design, when $u = 0$ the contribution of the basis function is zero leaving only the autoregressive portion of the model in (5). Since the autoregressive portion will be stable, the models prediction will converge to the type of response we are looking for when there is no stimulus. The subset-selection method based on a minimum-description-length criterium is used to determine the centers d for the input space and c for the output space and is described in detail by Judd and Mees [17]. Roughly, a set of centers is chosen from the data points and the goodness of fit is determined from both a least squares error term plus a penalty term for the size of the model. Quadratic programming techniques are used to grow or shrink the basis set so as to optimize the least squares mean error. Alternatively, one could also try to fit the models by orthogonal least squares, or stepwise regression [9].

It can be difficult to reconstruct a dynamic black-box model which simulates well, that is, one whose outputs are stable and at least qualitatively the same as the actual device, if not exactly quantitatively accurate. We will demonstrate that by using our proposed-basis function of (7) accurate dynamic models under simulation can be reconstructed.

In addition to reconstructing feedback models we are also interested in constructing static models, in part to understand when feedback models are necessary (there exists a large literature on how to model the "static" nonlinearities of a system, see, for example, reference [7]). Static models are functions which directly map the input to the output, that is, no past simulated outputs are "fed back" into the model. Static models are not expected to perform well when the device under study exhibits strong memory effects. In this case, models with knowledge of the internal state of the device should be expected to perform better. The use of past outputs provides an approximation to the internal state. For the purposes of comparison we reconstruct static radial-basis models as well. A static model is of the form

$$y(t) = H[u(t), u(t-\tau), \ldots, u(t-(l-1)\tau] \quad (8)$$

and in this example we find the form

$$y(t+1) = H[u(t+1), u(t)] \quad (9)$$

produced the best static models. Thus, our aim in this paper is twofold. We want to show how a spread-spectrum excitation design is suitable building black-box models and to introduce the basis function of (7) as a good basis function to use in feedback models of electronic devices.

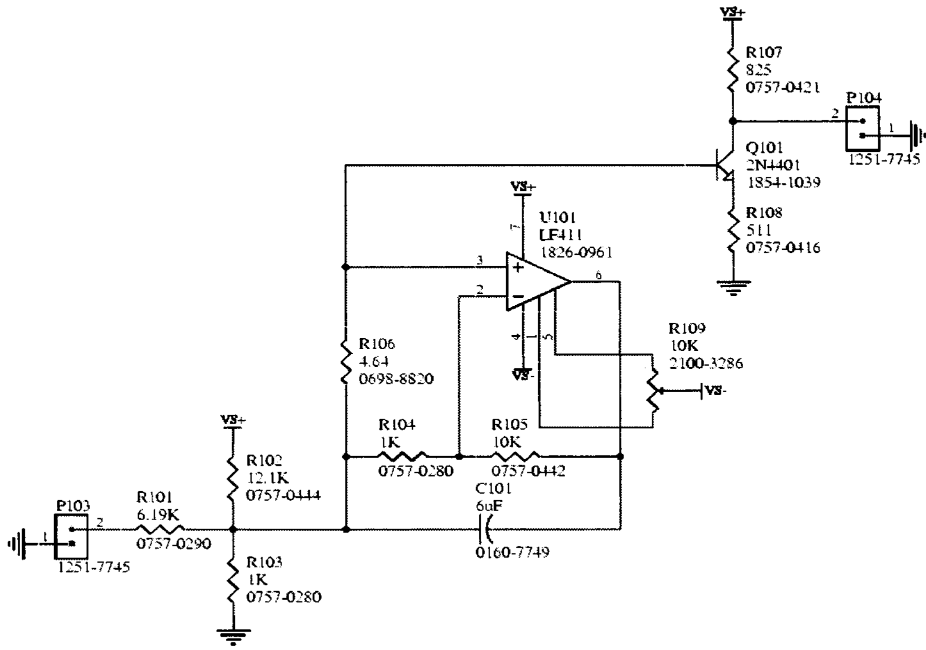

Fig. 1. A high-frequency "analog" transistor circuit.

III. Data Sets and Measurements

We developed black-box models for a number of simple electronic components (e.g., transistors) and simple circuits (e.g., amplifiers). An example of a simple circuit we built a model for is shown in Fig. 1.

This particular circuit is meant to be a transistor 'analog' of a high-frequency microwave transistor [19]. That is, the circuit attempts to capture some of the dynamic effects that should be present in microwave transistor but which are difficult to measure in the time-domain due to the high frequency of its typical operation. Circuit 'self-analogs' of this type were built in the 1960s to study the dynamics of microwave circuits. This particular circuit acts as an amplifier and also tries to mimic certain (memory dependent) charge-storage effects which should be active in the microwave amplifier in the gigahertz regime and in our 'analog' circuit in a frequency range of around 0.5 kHz. Models are also built from numerical models and data from simulations.

For the experimental data, voltage excitations are supplied and their amplified response are measured using a nonlinear circuit measurement system. The measurement system combines nonlinear-circuit-device models, arbitrary wave-form generation cards, analog-to-digital cards, and a numerical software package all developed in and controlled by, Matlab [20]. We can thus generate time-domain input-output (stimuli-response) data for nonlinear circuit devices either in measurements or simulations. The particular example we use in this study is the high-frequency amplifier analog in Fig. 1 with the resonance frequency set to 650 Hz. The drive signals we use as stimuli are the aforementioned CDMA type signal as well as periodic drive signals for some additional tests. The data sets are labeled by the type of signal and the carrier frequency. Thus, C_{500} refers to a CDMA signal with carrier frequency 500 Hz. We generate numerous data sets of different types with frequencies at 50-Hz intervals starting at 50 Hz and ending at 1200 Hz. Voltage samples are equally spaced with a sampling frequency usually about $1/64$th of the center frequency of the CDMA carrier. Thus, we are sampling 16 kHz at Nyquist.

This is oversampled because memory constraints are not a consideration. As described below, we typically decimate the data sets and use only a fraction of them in building models. The number of points used to build a given model is usually less than 20 000 points and can be as little as 2000.

IV. Models

In our first modeling attempt, the models we built used an embedding of the following form:

$$y(t+1) = F\left[y(t-1), y(t), u(t), u(t+1)\right]. \tag{10}$$

It is known, however, from autonomous time-series studies that not all data sets are best embedded using a lag of one, yet a dynamic model with, say, an "optimal" lag of five did not simulate well. A possible explanation for this discrepancy is that for models which predict one time-step in the future we need to keep track of "extra dimensions" when the lag is not one. For example, suppose we simulate the following model:

$$y(t+1) = F\left[y(t-5), y(t)\right]. \tag{11}$$

We see that although the model includes two state variables, we must keep track of six values of y, so implicitly the model is six dimensional. Now after five iterations our implicit internal state consists entirely of predicted values all of which have errors compounded. Keeping the unit lags in our model implicitly could slightly decrease these compounded errors.

Our experiments with our data sets appear to bear this out. However, for the higher-frequency data sets where we have many samples per carrier cycle using a lag of one is not appropriate. We overcome this by decimating the data to have approximately 12 to 16 points per cycle and then we reconstruct models of the form of (10) with this decimated data. For example, for the data set C_{1200} we have approximately 48 points per cycle. We create four data sets by decimating the original data by four, i.e., we take every fourth point. So, when we reconstruct a model of the form of (10) we are essentially reconstructing a feedback model of the form

$$y(t+4) = F\left[y(t-4), y(t), u(t), u(t+4)\right]. \tag{12}$$

TABLE I
RESULTS OF SIMULATING RECONSTRUCTED FEEDBACK VS STATIC MODELS

Data Set	Decimation	#Parms	RMS/std(y)	SNR	Static #Parms	Static SNR
C_{100}	1	21	0.06	25.06	7	20.76
C_{200}	1	16	0.05	26.22	3	11.77
C_{300}	1	14	0.06	24.75	7	7.58
C_{400}	1	15	0.08	21.64	5	3.24
C_{500}	1	24	0.09	21.13	3	1.71
C_{600}	2	26	0.09	21.08	3	2.15
C_{700}	3	29	0.06	23.33	2	1.85
C_{800}	3	29	0.08	21.77	2	1.43
C_{900}	4	34	0.08	21.94	3	1.80
C_{1000}	4	36	0.1046	19.60	2	1.01
C_{1100}	4	41	0.1077	19.36	2	0.75
C_{1200}	4	33	0.12	18.78	2	0.66

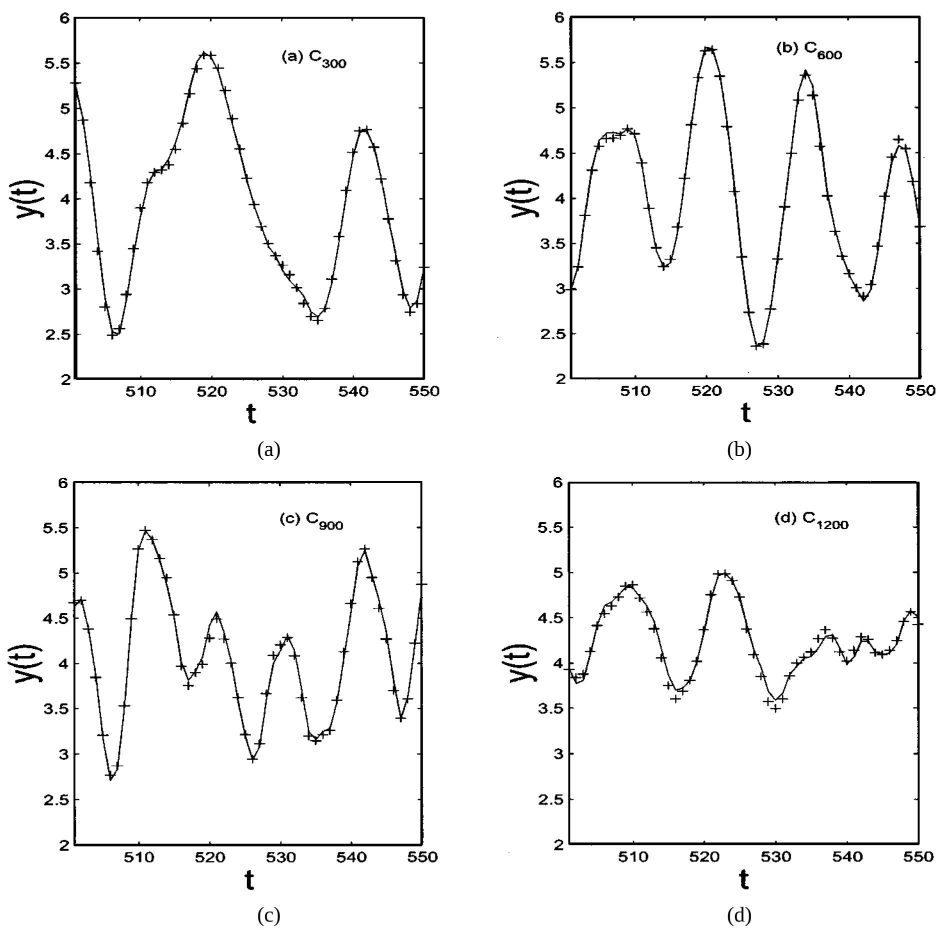

Fig. 2. Sections of feedback simulations for data sets. (a) C_{300}. (b) C_{600}. (c) C_{900}. (d) C_{1200}. The solid lines are the actual device response and the crosses are the simulated predictions.

We will see that by following this procedure, good results can be obtained.

We present the results of modeling and simulating the CDMA data sets with static and dynamic models in Table I. Table I has seven columns. The first column indicates the data set and the second column indicates the decimation used. The third column shows the "size" of our best reconstructed feedback model, i.e., the number of model coefficients. In column four, we give an error measure of the feedback model for out-of-sample drive signals and express this error in terms of signal-to-noise in column five. We calculate the signal-to-noise ratio using

$$\text{SNR} = 20\log_{10}\left(\frac{\text{std(actual)}}{\text{std(errors)}}\right)\text{dB}. \quad (13)$$

We show the analogous results obtained by reconstructing and testing static models on the same data in the remaining columns.

Typical results are presented in Fig. 2(a)–(d) where we show sections of the time series produced by simulating the models compared to the actual measured values of the device. We show the results obtained by simulating the models reconstructed using the C_{300}, C_{600}, C_{900} and C_{1200} data sets. Good agreement is seen in the figures as expected from the numbers given in Table I. These simulations are also superior to the results we obtained using the best static models we could reconstruct. In most of the cases examined, the long-term solutions are not sensitive to the initial seed value and, in the case of periodic drives, they appear to converge to a unique solution.

V. Conclusion

We have shown how to construct stable, free-running, input-output models for a class of electronic devices and circuits having fading memory and (in the absence of a drive signal) converge to a constant solution. The models are built from band-limited, spread-spectrum excitations and such excitations provide a sufficiently rich training set to make accurate predictions of periodic or similar spread spectrum drive signals. Adding additional local and global properties appears to be a promising avenue for research in building stable and accurate black-box models.

Acknowledgment

The authors wish to thank R. Brown for some early discussions about this investigation.

References

[1] H. Kantz and T. Schrieber, *Nonlinear Time Series Analysis*. Cambridge, U.K: Cambridge Univ. Press, 1997.
[2] N. Tufillaro, D. Usikov, L. Barford, D. Walker, and D. Schreurs, "Measurement driven models of nonlinear electronic components," in *Dig. 54th ARFTG Conf.*, Boston, MA, June 2000.
[3] K. Judd and M. Small, "Towards long-term prediction," *Physica D*, vol. 136, no. 1/2, pp. 31–44, 2000.
[4] L. A. Aguirre and S. A. Billings, "Dyanamical effects of overparametrization in nonlinear models," *Physica D*, vol. 80, no. 1,2, pp. 26–40, 1995.
[5] L. Chua, C. Desoer, and E. Kuh, *Linear and Nonlinear Circuits*. San Francisco, CA: McGraw-Hill, 1987.
[6] S. Boyd and L. Chua, "Fading memory and the problem of approximating nonlinear operators with Volterra series," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 1150–1161, Nov. 1985.
[7] R. Harber and H. Unbchauen, "Structure identification of nonlinear dyanamic systems, A survey on input/output approaches," *Automatica*, vol. 26, no. 4, pp. 651–677, 1990.
[8] L. A. Aguirre, P. F. Donoso-Garcia, and R. Santos-Filho, "Use of a priori Information in the Identification of Global Nonlinear Models–A Case Study Using a Buck Converter," *IEEE Trans. Circuits Syst. I*, vol. 47, pp. 1081–1085, July 2000.
[9] N. A. Gershenfeld, *The Nature of Mathematical Modeling*. Cambridge, U.K: Cambridge Univ. Press, 1999.
[10] M. Casdagli, "A dynamical systems approach to modeling input-output systems," in *Nonlinear Modeling and Forecasting, SFI Studies in the Sciences of Complexity, Proc*, M. Casdagli and S. Eubank, Eds. Reading, MA: Addison-Wesley, 1992, vol. XII.
[11] I. J. Leontaritis and S. A. Billings, "Input-Output parametric models for nonlinear systems part I: Deterministic nonlinear systems," *Int. J. Control*, vol. 41, no. 2, pp. 303–328, 1985.
[12] S. Haykin and J. Principe, "Making sense of a complex world," *IEEE Signal Processing Mag.*, pp. 66–81, May 1998.
[13] J. Stark, "Delay Embeddings of Forced Systems: I deterministic forcing," *J. Nonlinear Science*, vol. 9, pp. 255–332, 1999.
[14] G. F. Franklin, *Feedback Control of Dynamic Systems*. Reading, MA: Addison-Wesley, 1994.
[15] D. M. Walker and N. B. Tufillaro, "Phase space reconstruction using input-output time-series data," *Phys. Rev. E*, vol. 60, no. 4, pp. 4008–4013, 1999.
[16] D. M. Walker, R. Brown, and N. B. Tufillaro, "Constructing transportable behavioral models for nonlinear electronic devices," *Phys. Letts A*, vol. 255, no. 4/6, pp. 236–242, 1999.
[17] K. Judd and A. I. Mees, "On selecting models for nonlinear time series," *Physica D*, vol. 82, no. 4, pp. 426–444, 1995.
[18] J. S. Lee and L. E. Miller, *CDMA Systems Engineering Handbook*. Boston, MA: Artech, 1998.
[19] B. T. Murphy, "Diode and transistor self-analogues for circuit analysis," *Bell Syst. Techn. J.*, no. 4, pp. 487–502, 1968.
[20] J. King, *Matlab for Engineers*. San Francisco, CA: Addison-Wesley, 1998.

Generating Chaos in Chua's Circuit via Time-Delay Feedback

Xiao Fan Wang, Guo-Qun Zhong, Kit-Sang Tang, Kim F. Man, and Zhi-Feng Liu

***Abstract*—A time-delay chaotification approach can be applied to the Chua's circuit by adding a small-amplitude time-delay feedback voltage to the circuit. The chaotic dynamics of this newly derived time-delay Chua's circuit is studied by theoretical analysis, verified by computer simulations as well as by circuit experiments.**

***Index Terms*—Chaos, stability, time delay.**

I. Introduction

Chua's circuit is one of the physical systems for which the presence of chaos (in the sense of Shil'nikov) has been established experimentally, confirmed numerically, and proven mathematically. In recent years, Chua's circuit has become a standard model for studying chaos in systems described by finite-dimensional ordinary differential equations [1].

Synchronization of chaotic Chua's circuit with application to secure communication has also been investigated. However, a classic Chua's circuit is a third-order continuous-time autonomous system which can only produce low-dimensional chaos with one positive Lyapunov exponent.

Manuscript received January 4, 2000; revised August 1, 2000 and December 21, 2000. This work was supported in part by the Research Grants Council, Hong Kong, under Competitive Earmarked Research Grant 9040565, and in part by the City University of Hong Kong under Strategic Grant 7000956. This paper was recommended by Associate Editor M. Di Bernardo.

X. F. Wang was with the Department of Automatic Control, Nanjing University of Science & Technology, Nanjing, 210094, China. He is now with the Department of Mechanical Engineering, University of Bristol, Bristol, U.K. (e-mail: X.Wang@bristol.ac.uk).

G.-Q. Zhong, K.-S. Tang, K. F. Man and Z.-F. Liu are with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong.

Publisher Item Identifier S 1057-7122(01)07711-X.