Fitting ordinary differential equations to chaotic data

E. Baake*
Institut für Mathematik, Universität Augsburg, Universitätstrasse 8, 8900 Augsburg, Germany

M. Baake
Institut für Theoretische Physik, Universität Tübingen, Auf der Morgenstelle 14, 7400 Tübingen, Germany

H. G. Bock
Institut für Mathematik, Universität Augsburg, Universitätstrasse 8, 8900 Augsburg, Germany

K. M. Briggs
Department of Mathematics, University of Melbourne, Parkville 3052, Australia
(Received 26 August 1991; revised manuscript received 23 December 1991)

We address the problem of estimating parameters in systems of ordinary differential equations which give rise to chaotic time series. We claim that the problem is naturally tackled by boundary-value-problem methods. The power of this approach is demonstrated by various examples with ideal as well as noisy data. In particular, Lyapunov exponents can be computed accurately from time series much shorter than those required by previous methods.

PACS number(s): 05.45.+b, 02.60.+y, 02.70.+d, 42.65.−k

I. INTRODUCTION

Experimental data produced by nonlinear dynamical systems in the chaotic regime are most often evaluated by various methods of time-series analysis [1]. Typical applications are the dynamics of laser emission [2], of glycolysis [3], and of electroencephalogram waves [4]. The aims of the analysis include the reconstruction of state-space representations [5, 7], the estimation of Lyapunov exponents (e.g., [6]), of Hausdorff and correlation dimensions of attractors [1, 7], and the analysis of power spectra [1]. These methods describe some characteristic features of the system but, in general, they do not give physical insight into the mechanism. If one can, however, derive a differential-equation model from more fundamental principles, one is in a potentially more powerful situation because the local representation can now be replaced by a global one. This situation arises, for example, with the dynamics of laser emission in a specific case (see, e.g., [2]), for which the Haken-Lorenz system [8] is derived as the effective equation with as yet undetermined parameters. Please note that local properties or Lyapunov exponents are not sufficient to fix those parameters. In order to judge the quality of such a model, the statistical properties of its numerical solution are compared with those of the original time series.

It would be desirable to fit the free parameters of the ordinary differential equation (ODE) to the observed dynamics directly, but this is often considered impossible because of the difficulties in predicting the behaviour of a chaotic system [9]: Numerical error propagation means that it is certainly not promising to apply a so-called single-shooting approach, i.e., to couple numerical algorithms for initial-value problems and least-squares problems. This is one reason why the analysis in terms of local discrete systems has been favored and highly developed (see, e.g., [10–12], and references above). Nevertheless, this is not fully satisfactory in our opinion, as one may be discarding important information concerning the assumed underlying mechanism.

However, recovering the parameters has been successful in discrete systems by means of embedding and local polynomial fitting [6], see also [12]. From the relationship of discrete and continuous dynamical systems via the Poincaré map, the immediate treatment of ODE's might also be expected to be possible. One method has recently been proposed [13] which is similar to the discrete case in the sense that it works its way piecewise along the time series. Since, so far, this procedure seems to work only for simple examples, we pursue a different strategy here by applying the boundary-value-problem approach for parameter estimation in ODE's and differential algebraic equations [14–17]. Our approach is powerful enough to identify, e.g., the Lorenz system from the most general class of three-dimensional quadratic systems and to estimate its parameters accurately, even with noisy data.

Let us now formulate the problem we are concerned with more precisely. We start from the following three key assumptions.

(i) The dimension of the ODE system is either known or reliably guessed. (ii) The model or the model class to describe the data is specified. (iii) The sampling rate of the data is sufficiently high with respect to all relevant frequencies occurring.

In this article, we do not intend to answer the question of what happens if one of these assumptions is violated, a situation where one expects new assumptions and additional techniques to be necessary. Since this article essentially describes the straightforward though important application of the algorithm mentioned above to the difficult case of chaotic time series, we will be brief in describing the method (Secs. II and III) and present a couple of examples in Sec. IV rather than reproducing the theory from [16].

II. CLASS OF PROBLEM

The problem of parameter fitting in ODE's can be formulated in a rather general but simple way [18] as follows. Determine a vector of parameters $\boldsymbol{p} \in \mathbb{R}^k$ and a trajectory $\boldsymbol{y} : [t_a, t_b] \to \mathbb{R}^n$ such that the ODE

$$\dot{\boldsymbol{y}} = \boldsymbol{f}(t, \boldsymbol{y}, \boldsymbol{p}) \tag{1}$$

is fulfilled, and a least-squares functional is minimized:

$$\mathcal{L}(\boldsymbol{y}, \boldsymbol{p}) \equiv \|\boldsymbol{r}(\boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_\mu), \boldsymbol{p})\|_2^2 \overset{!}{=} \text{minimum}. \tag{2}$$

Here, $\overset{!}{=}$ indicates that equality is the aim of the numerical treatment, and $\boldsymbol{r} \in \mathbb{R}^\ell$ is the vector of least-squares conditions, which depends on the ODE components at specified instants t_i, and on the parameters. In the most common case, the objective functional reduces to

$$\mathcal{L}(\boldsymbol{y}, \boldsymbol{p}) = \sum_{i=1}^{\mu} \sum_{j=1}^{\nu} \frac{1}{\sigma_{ij}^2} [\eta_{ij} - g_j(t_i, \boldsymbol{y}(t_i), \boldsymbol{p})]^2 , \tag{3}$$

where g is a function relating the ODE components to the measured quantities, η_{ij} is the observed value of g_j at instant t_i, and σ_{ij} its standard deviation.

If the ODE (1) is nonlinear and if we consider a chaotic regime, the numerical solution of the corresponding initial-value problems should be spoiled by error propagation due to positive Lyapunov exponents. Consequently, the parameter estimation problem (1) and (2) is expected to be ill-posed when treated in a "forward" manner.

From the "inverse" point of view, however, things look more promising. To see this, let us consider the initial-value problem

$$\dot{\boldsymbol{y}}(t) = \boldsymbol{f}(t, \boldsymbol{y}, \boldsymbol{p}), \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0, \tag{4}$$

and the derivatives of its solution $\boldsymbol{y}(t; t_0, \boldsymbol{y}_0, \boldsymbol{p})$ with respect to parameters and initial values (compare [19] where the notion of "sensitive dependence on parameters" is coined and discussed):

$$\begin{aligned} \boldsymbol{G}(t, t_0, \boldsymbol{y}_0, \boldsymbol{p}) &\equiv \frac{\partial}{\partial \boldsymbol{y}_0} \boldsymbol{y}(t; t_0, \boldsymbol{y}_0, \boldsymbol{p}), \\ \boldsymbol{H}(t, t_0, \boldsymbol{y}_0, \boldsymbol{p}) &\equiv \frac{\partial}{\partial \boldsymbol{p}} \boldsymbol{y}(t; t_0, \boldsymbol{y}_0, \boldsymbol{p}). \end{aligned} \tag{5}$$

It then follows from the variational differential equations

$$\begin{aligned} \dot{\boldsymbol{G}}(t, t_0, \boldsymbol{y}_0, \boldsymbol{p}) &= \boldsymbol{f}_{\boldsymbol{y}}(t, \boldsymbol{y}, \boldsymbol{p}) \cdot \boldsymbol{G}(t, t_0, \boldsymbol{y}_0, \boldsymbol{p}), \\ \boldsymbol{G}(t_0, t_0, \boldsymbol{y}_0, \boldsymbol{p}) &= \mathbb{1}, \\ \dot{\boldsymbol{H}}(t, t_0, \boldsymbol{y}_0, \boldsymbol{p}) &= \boldsymbol{f}_{\boldsymbol{y}}(t, \boldsymbol{y}, \boldsymbol{p}) \cdot \boldsymbol{H}(t, t_0, \boldsymbol{y}_0, \boldsymbol{p}) \\ &\quad + \boldsymbol{f}_{\boldsymbol{p}}(t, \boldsymbol{y}, \boldsymbol{p}), \\ \boldsymbol{H}(t_0, t_0, \boldsymbol{y}_0, \boldsymbol{p}) &= \boldsymbol{0}, \end{aligned} \tag{6}$$

that

$$\boldsymbol{H}(t, t_0, \boldsymbol{y}_0, \boldsymbol{p}) = \int_{t_0}^{t} \boldsymbol{G}(t, s, \boldsymbol{y}_0, \boldsymbol{p}) \cdot \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{p}}(s, \boldsymbol{y}(s; t_0, \boldsymbol{y}_0, \boldsymbol{p}), \boldsymbol{p}) ds. \tag{7}$$

This implies that, at least as long as the time scale of measurement is sufficiently small as compared to the time scale of oscillations in $\boldsymbol{G}$ or $\boldsymbol{f}_p$, the derivatives with respect to parameters may profit from the derivatives with respect to initial values, some of which are large in chaotic systems. Put differently, an observed trajectory may be expected to contain a large amount of information about the parameters. So parameter estimation in chaotic systems may indeed be a well-posed problem, if there is a way to cope with error propagation.

III. NUMERICAL METHOD

The error-propagation problem can adequately be tackled by the boundary-value-problem methods for parameter estimation in ODE [14–16]. The multiple-shooting algorithm PARFIT as the most versatile member of this class of methods is described, e.g., in [14], and a comprehensive treatment can be found in [16]. Its basic idea is to regard the ODE in Eqs. (1) and (2) — independently of its specific nature — not as an initial-value problem but as a multipoint boundary-value problem, and to treat this in discretized form as a nonlinear constraint in the optimization process. To this end, the measuring interval $[t_a, t_b]$ is covered by a suitable grid of multiple-shooting nodes τ_j, such that $\tau_1 < \tau_2 < \cdots < \tau_m$, $[t_a, t_b] \subset [\tau_1, \tau_m]$, and $m - 1$ initial-value problems are considered:

$$\dot{\boldsymbol{y}} = \boldsymbol{f}(t, \boldsymbol{y}, \boldsymbol{p}), \quad \boldsymbol{y}(\tau_j) = \boldsymbol{s}_j, \tag{8}$$

where the additional variables $\boldsymbol{s}_j$ represent the state variables at the nodes; initial guesses for them are suitably chosen to match with the observations, and to exploit any further knowledge about the process, but ingenious guesses are in general not needed. Equation (8) means that the problem is integrated piecewise, starting at the $\boldsymbol{s}_j$, which yields, with the initial guesses for the $\boldsymbol{s}_j$ and $\boldsymbol{p}$, a discontinuous initial trajectory (see Fig. 1). The task then consists in determining the $\boldsymbol{s}_j$ and $\boldsymbol{p}$ such that the objective function is minimized and the final trajectory is made continuous. The basic problem (1)

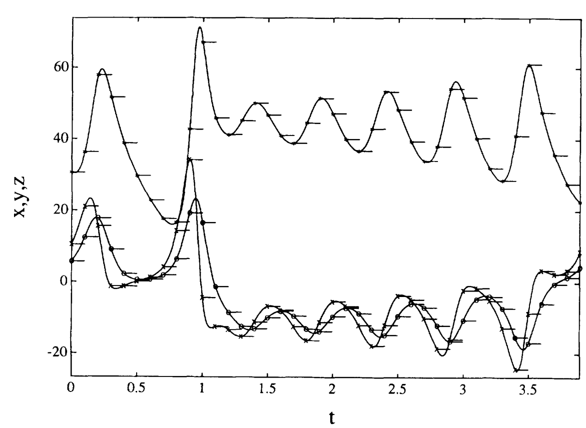

FIG. 1. Operation of multiple-shooting algorithm PARFIT on Lorenz data (test L4 of Table II). o,x,*: x, y, z data, respectively. Discontinuous lines: initial trajectory. Continuous lines: final trajectory.

and (2) is thus reformulated as a large constrained least squares problem in the augmented vector of variables $\boldsymbol{x}^T = (\boldsymbol{s}_1^T, \ldots, \boldsymbol{s}_m^T, \boldsymbol{p}^T)$, where $\boldsymbol{R}$ corresponds to $\boldsymbol{r}$ in Eq. (2):

$$\|\boldsymbol{R}(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m, \boldsymbol{p})\|_2^2 \stackrel{!}{=} \text{minimum}, \tag{9}$$

$$\boldsymbol{y}(\tau_{j+1}; \tau_j, \boldsymbol{s}_j, \boldsymbol{p}) - \boldsymbol{s}_{j+1} \stackrel{!}{=} 0, \qquad j = 1, \ldots, m-1. \tag{10}$$

Note that this formulation is similar to that one given in [12] for the treatment of numerical shadowing in discrete systems.

A crucial property of this approach in the context of chaotic dynamics is that, with a sufficiently fine multiple shooting grid, error propagation on subintervals can be kept within arbitrary bounds [16].

Equations (9) and (10) represent a constrained nonlinear least-squares problem of the general form

$$\|\boldsymbol{u}_1(\boldsymbol{x})\|_2^2 \stackrel{!}{=} \text{minimum}, \tag{11}$$

$$\boldsymbol{u}_2(\boldsymbol{x}) \stackrel{!}{=} 0, \tag{12}$$

which is solved with a generalized Gauss–Newton method as described and analyzed in [16]: Starting from an initial guess $\boldsymbol{x}^{(0)}$, the vector of variables $\boldsymbol{x}$ is iterated via

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \lambda^{(k)} \Delta \boldsymbol{x}^{(k)}, \tag{13}$$

where the $\lambda^{(k)} \in [0,1]$ are damping factors which are determined via so-called level functions to ensure global convergence, and $\Delta \boldsymbol{x}^{(k)}$ solves the linearized problem:

$$\left\| \boldsymbol{u}_1(\boldsymbol{x}^{(k)}) + \frac{\partial \boldsymbol{u}_1}{\partial \boldsymbol{x}}(\boldsymbol{x}^{(k)}) \Delta \boldsymbol{x}^{(k)} \right\|_2^2 \stackrel{!}{=} \text{minimum}, \tag{14}$$

$$\boldsymbol{u}_2(\boldsymbol{x}^{(k)}) + \frac{\partial \boldsymbol{u}_2}{\partial \boldsymbol{x}}(\boldsymbol{x}^{(k)}) \Delta \boldsymbol{x}^{(k)} \stackrel{!}{=} 0. \tag{15}$$

The integrations required for the nonstiff examples treated here are carried out with an integrator (based on the Bulirsch–Stoer extrapolation method) which also performs the differentiation via so-called internal numerical differentiation as introduced in [14].

Although this multiple-shooting approach has many more variables than a single-shooting approach, the linearized problem exhibits a special structure which, if carefully exploited, means that the integration and linearization effort is essentially not increased in comparison with single shooting [21].

The algorithm proceeds in precisely the same way when some of the state variables are not measured at the nodes. In this "hidden-variable" case (see examples L2 and R2 below), a very rough initial guess of the $\boldsymbol{s}_j$ is usually sufficient, as long as identifiability is guaranteed, which depends on the mutual coupling of the ODE's and on the actual condition number of the discretized problem.

This method should be more powerful than the method proposed in [13] for at least two reasons: First, no numerical derivatives are needed in the case of hidden variables. Secondly, the number of free variables is drastically reduced by the continuity constraints (10).

It can be shown that the condition of the large constrained least-squares problem is asymptotically that of the presumably well-posed inverse problem, and that it can be solved in a numerically stable way by the generalized Gauss-Newton method. Note that the solution variables can thus be determined with high accuracy, although local integration errors lead to (small) violations of Eq. (10).

IV. RESULTS AND DISCUSSION

Let us illustrate these ideas by some examples. As model systems we choose the Lorenz system

$$\dot{x} = \sigma(y - x), \qquad \dot{y} = -xz + rx - y, \qquad \dot{z} = xy - bz \tag{16}$$

with initial values $x_0 = 5.765\,40, y_0 = 10.505\,47, z_0 = 30.589\,41$, and parameters $\sigma = 10$, $r = 46$, $b = \frac{8}{3}$; and the Rössler system

$$\dot{x} = -y - z, \qquad \dot{y} = x + ay, \qquad \dot{z} = b + (x - c)z \tag{17}$$

with initial values $x_0 = 1.132\,93, y_0 = -1.749\,53, z_0 = 0.022\,07$, and parameters $a = 0.15$, $b = 0.2$, $c = 10$, and generated "data" by numerical integration of the initial value problem (requested local relative accuracy: 10^{-12}) at intervals of $\Delta t = 0.1$ and $\Delta t = 0.2$ for the Lorenz and the Rössler system, respectively.

Let us briefly comment on the question of what may be the meaning of "data" in this context. No problems arise if, for the system considered, a suitable shadowing property is known to guarantee that the numerical orbit is close to a true one (e.g., [20]). If this is not the case, however, we may rely on a generalization of the boundary

TABLE I. Lorenz system. Comp : components used to determine parameters ("noise" indicates Gaussian random numbers of standard deviation 2 are added to all components). Absolute errors: modulus of deviation of estimated parameters from true values. Confidence intervals: half widths of symmetric intervals, applicable to noisy data only. Data used: number of observations used for each component of column 2.

Test	Comp.	Absolute errors in σ, r, b (95% confidence intervals)	Data used	Lyapunov exponents
L1	x, y, z	$5 \times 10^{-6}, 1 \times 10^{-6}, 8 \times 10^{-7}$	5	$1.24, -0.01, -14.90$
L2	x	$< 1 \times 10^{-6}, 3 \times 10^{-6}, 8 \times 10^{-7}$	15	$1.22, 0.00, -14.88$
L3	x, y, z plus noise	$2 \times 10^{-2}, 1 \times 10^{-1}, 4 \times 10^{-2}$ (0.5, 1.0, 0.1)	40	$1.23, 0.00, -14.92$

TABLE II. Lorenz system. For explanations, see Table I. Here, the number of data used is 40, and the Gaussian noise is of standard deviation 0.0125.

Test	Comp.	Absolute errors in A_i, b_i, c_i, $i = 1, 2, 3$	Lyapunov exponents
L4	x, y, z	$\begin{pmatrix} 9\times10^{-7} & 8\times10^{-7} & 3\times10^{-7} \\ & 2\times10^{-7} & 5\times10^{-8} \\ & & 1\times10^{-7} \end{pmatrix}, \begin{pmatrix} 1\times10^{-5} \\ 3\times10^{-6} \\ 6\times10^{-6} \end{pmatrix}, 6\times10^{-5}$	1.24
		$\begin{pmatrix} 4\times10^{-7} & 4\times10^{-7} & 2\times10^{-6} \\ & 2\times10^{-7} & 1\times10^{-7} \\ & & 2\times10^{-7} \end{pmatrix}, \begin{pmatrix} 1\times10^{-5} \\ 3\times10^{-6} \\ 1\times10^{-5} \end{pmatrix}, 2\times10^{-4}$	0.00
		$\begin{pmatrix} 1\times10^{-6} & 1\times10^{-6} & 3\times10^{-7} \\ & 3\times10^{-7} & 1\times10^{-7} \\ & & 1\times10^{-7} \end{pmatrix}, \begin{pmatrix} 2\times10^{-5} \\ 1\times10^{-6} \\ 5\times10^{-6} \end{pmatrix}, 3\times10^{-5}$	−14.91
L5	x, y, z Plus Noise	$\begin{pmatrix} 1\times10^{-3} & 3\times10^{-3} & 9\times10^{-4} \\ & 1\times10^{-3} & 4\times10^{-4} \\ & & 2\times10^{-4} \end{pmatrix}, \begin{pmatrix} 5\times10^{-2} \\ 3\times10^{-2} \\ 1\times10^{-2} \end{pmatrix}, 2\times10^{-1}$	1.24
		$\begin{pmatrix} 2\times10^{-3} & 2\times10^{-3} & 6\times10^{-3} \\ & 4\times10^{-4} & 5\times10^{-6} \\ & & 1\times10^{-4} \end{pmatrix}, \begin{pmatrix} 3\times10^{-2} \\ 7\times10^{-2} \\ 3\times10^{-3} \end{pmatrix}, 3\times10^{-2}$	0.00
		$\begin{pmatrix} 4\times10^{-4} & 7\times10^{-4} & 1\times10^{-3} \\ & 3\times10^{-4} & 8\times10^{-4} \\ & & 8\times10^{-5} \end{pmatrix}, \begin{pmatrix} 7\times10^{-2} \\ 4\times10^{-2} \\ 9\times10^{-3} \end{pmatrix}, 2\times10^{-1}$	−14.93

value problem approach to multiple experiments [21], in which the time range is split and the corresponding data are treated as separate time series.

The numerical experiments performed are summarized in Tables I–IV. We choose the multiple-shooting nodes to be identical with the instants of "measurement" and use the simulated "data" as initial guesses for the state variables s_j at the nodes, if these quantities are available. Otherwise (for unobserved variables), the s_j are initialized at zero.

We consider two scenarios.

(1) A model of the underlying dynamics is available, with only individual parameters unspecified (Tables I and III)—an advantageous situation as, e.g., in the laser experiment already mentioned.

In all cases tested, the parameters (and initial values; not shown in the Tables for brevity) are correctly recovered from a small number of data points, not only in the case of complete observation of the Lorenz system (tests L1 and R1), but also from the x component (L2) or from the x and z components (R2) alone. The procedure even copes with a considerable amount of noise added to the data (L3 and R3). To be more specific, the noise is Gaussian distributed with standard deviation 2 for the Lorenz system (which corresponds to roughly 20%, 20%, and 5% of the average modulus of the x, y, and z component, respectively) and with standard deviation 1 for the Rössler system (20%, 20%, and 200% of the average modulus of the x, y, and z component, respectively). In no case do the estimated parameters deviate from their true values by more than 1.5%.

(2) Only the class of model is specified, with all coefficients to be determined. Such a case occurs frequently in practice. Consider, for example, (chemical) mass-action kinetics. Here, only polynomials up to second order occur in the right-hand side of the ODE [22], and a good guess concerning the number of reactants, i.e. the embedding dimension, is often available. In physical examples, harmonic or next-to-harmonic approximations will also produce an ansatz with a polynomial of bounded order on the right-hand side. If a guess for the embedding dimension is not provided, it may be arrived at by state space reconstructions [7], but this is not our concern here.

For our examples, we assume a three-dimensional system with the most general quadratic right-hand side:

$$\dot{y}_i = \frac{1}{2} y^T (A_i + A_i^T) y + b_i^T y + c_i, \qquad i = 1, 2, 3, \qquad (18)$$

TABLE III. Rössler system. For explanations, see Table I. Here, noise is of standard deviation 1.

Test	Comp.	Absolute errors in a, b, c (95% confidence intervals)	Data used	Lyapunov exponents
R1	x, y, z	$6\times10^{-8}, 2\times10^{-6}, 2\times10^{-5}$	200	$0.09, 0.01, -9.88$
R2	x, z	$< 1\times10^{-8}, 4\times10^{-8}, < 1\times10^{-6}$	200	$0.10, 0.03, -9.74$
R3	x, y, z plus noise	$8\times10^{-4}, 2\times10^{-2}, 3\times10^{-2}$ $(9\times10^{-3}, 6\times10^{-2}, 4\times10^{-1})$	200	$0.07, 0.01, -9.81$

TABLE IV. Rössler system. For explanations, see Table I. Here, the number of data used is 200, and noise is of standard deviation 0.005.

Test	Comp.	Absolute errors in $\boldsymbol{A}_i, \boldsymbol{b}_i, c_i,\ i = 1,2,3$	Lyapunov exponents
		$\begin{pmatrix} 8\times10^{-9} & 4\times10^{-7} & 5\times10^{-6} \\ & 1\times10^{-7} & 3\times10^{-6} \\ & & 2\times10^{-6} \end{pmatrix}, \begin{pmatrix} 2\times10^{-7} \\ 1\times10^{-6} \\ 7\times10^{-6} \end{pmatrix}, 2\times10^{-6}$	0.09
R4	x,y,z	$\begin{pmatrix} 2\times10^{-7} & 9\times10^{-8} & 2\times10^{-7} \\ & 2\times10^{-7} & 8\times10^{-7} \\ & & 3\times10^{-7} \end{pmatrix}, \begin{pmatrix} 1\times10^{-6} \\ 2\times10^{-7} \\ 9\times10^{-6} \end{pmatrix}, 5\times10^{-6}$	0.01
		$\begin{pmatrix} 1\times10^{-7} & 1\times10^{-7} & 2\times10^{-5} \\ & 2\times10^{-8} & 2\times10^{-5} \\ & & 1\times10^{-5} \end{pmatrix}, \begin{pmatrix} 5\times10^{-7} \\ 1\times10^{-7} \\ 3\times10^{-4} \end{pmatrix}, 4\times10^{-6}$	−9.79
		$\begin{pmatrix} 3\times10^{-4} & 7\times10^{-4} & 6\times10^{-3} \\ & 3\times10^{-4} & 4\times10^{-3} \\ & & 9\times10^{-4} \end{pmatrix}, \begin{pmatrix} 6\times10^{-5} \\ 3\times10^{-4} \\ 1\times10^{-2} \end{pmatrix}, 3\times10^{-3}$	0.08
R5	x,y,z Plus Noise	$\begin{pmatrix} 2\times10^{-4} & 3\times10^{-5} & 1\times10^{-3} \\ & 3\times10^{-4} & 7\times10^{-4} \\ & & 2\times10^{-4} \end{pmatrix}, \begin{pmatrix} 1\times10^{-4} \\ 5\times10^{-4} \\ 1\times10^{-3} \end{pmatrix}, 1\times10^{-3}$	0.00
		$\begin{pmatrix} 1\times10^{-4} & 3\times10^{-4} & 4\times10^{-4} \\ & 3\times10^{-5} & 2\times10^{-2} \\ & & 1\times10^{-2} \end{pmatrix}, \begin{pmatrix} 1\times10^{-3} \\ 1\times10^{-4} \\ 2\times10^{-1} \end{pmatrix}, 7\times10^{-3}$	−9.68

with real numbers c_i, real vectors $\boldsymbol{b}_i$, and real upper triangular matrices $\boldsymbol{A}_i$ (for technical reasons, we prefer this to working with symmetric matrices), which amounts to 30 parameters altogether (L4 and R4, Fig. 1). As might be expected, however, this approach is sensitive to noise. Noise of standard deviation 0.0125 (L5) and 0.005 (R5) already leads to non-negligible errors in the parameter estimates. Nevertheless, the parameters together with their confidence intervals (not included in the tables) lead to an unambiguous identification of the correct model within the class considered [23]. It turned out that this approach is much more robust against "relative" noise as often used in numerical examples (e.g., [6], and references therein), but the assumption of noise proportional to the signal amplitude is rather artificial on physical grounds.

As an application, we now discuss Lyapunov-exponent estimates, which have been used in previous attempts to diagnose chaos in experimental data [1, 6, 24]. Actually, these attempts have not been successful when applied to real, noisy data. The present ODE fitting method suggests a new possibility: once the ODE's have been fitted, the Lyapunov exponents can be calculated directly and accurately by the method of Wolf *et al.* [25]. We calculated the exponents from the fitted parameters in the Lorenz and Rössler system (see tables). We believe that the values are sufficiently accurate to allow the diagnosis of sensitive dependence on initial conditions (i.e., the largest exponent is positive). From the Lyapunov exponents, we calculated the Kaplan-Yorke dimension $D_{\rm KY}$ [1] for the Lorenz and the Rössler system and found that $D_{\rm KY} = 2.083 \pm 0.001$ and $D_{\rm KY} = 2.009 \pm 0.002$, respectively, covers all examples of the tables. They are unambiguously noninteger, suggesting the presence of a fractal attractor, which is a powerful conclusion from a small amount of data.

V. CONCLUDING REMARKS

We have demonstrated by means of examples that the multiple-shooting method is appropriate to the specific problem of parameter estimation in chaotic ODE's. Important remaining questions include the applicability of the method in dimensions higher than three, the treatment of very long-time series, and the tractability of real data. We hope to tackle these aspects in the near future.

ACKNOWLEDGMENTS

We wish to thank R. Delbourgo and J. Roberts for critically reading the manuscript and J. P. Schlöder for helpful discussions. Financial support from Deutsche Forschungsgemeinschaft is gratefully acknowledged.

* Present address: Max-Planck-Institut für Entwicklungsbiologie, Spemannstrasse 35/IV, 7400 Tübingen, Germany.

[1] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
[2] U. Hübner, N. B. Abraham, and C. O. Weiss, Phys. Rev. A **40**, 6354 (1990).
[3] M. Markus, Biophys. Chem. **22**, 95 (1985).
[4] D. Gallez and A. Babloyantz, Biol. Cybern. **64**, 381 (1991).
[5] F. Takens, in *Dynamical Systems and Turbulence*, edited by D. Rand and L. S. Young (Springer, Berlin, 1981).
[6] K. M. Briggs, Phys. Lett. A **151**, 27 (1990).
[7] T. Sauer, J. A. Yorke, and M. Casdagli, J. Stat. Phys.

65, 579 (1991).
[8] H. Haken, Phys. Lett. A **53**, 77 (1975).
[9] M. Markus, Habilitationsschrift, Universität Dortmund 1988.
[10] E. Kostelich and J. Yorke, Phys. Rev. A **38**, 1649 (1988).
[11] E. Kostelich and J. Yorke, Physica D **41**, 183 (1990).
[12] J. D. Farmer and J. J. Sidorowich, Physica D **47**, 373 (1991).
[13] J. L. Breeden and A. Hübler, Phys. Rev. A **42**, 5817 (1990).
[14] H. G. Bock, in *Modelling of Chemical Reaction Systems*, edited by K. H. Ebert, P. Deuflhard, and W. Jäger, Springer Series in Chemical Physics Vol. 18 (Springer, Heidelberg, 1981).
[15] H. G. Bock, in *Automatic Control in Petrol, Petrochemical and Desalination Industries (IFAC)*, edited by P. Kotobh (Pergamon, Oxford, 1986).
[16] H. G. Bock, Ph.D. thesis, in Bonner Mathematische Schriften, edited by E. Brieskorn *et al.* (internal report, Bonn, 1987), Vol. 183 (in German).
[17] H. G. Bock, E. Eich, and J. P. Schlöder, in *Numerical Treatment of Differential Equations*, edited by B. Strehmel (Teubner, Stuttgart, 1988).
[18] This is the unconstrained case with instants t_i given explicitly. Much more general problems, which include equality and inequality constraints and implicitly determined switching functions, are tractable with the algorithms described in [14].
[19] J. D. Farmer, Phys. Rev. Lett. **85**, 351 (1985).
[20] S. M. Hammel, J. A. Yorke, and C. Grebogi, Bull. AMS **19**, 465 (1988).
[21] J. P. Schlöder, Ph.D. thesis, in Bonner Mathematische Schriften, edited by E. Brieskorn *et al.* (internal report, Bonn, 1988), Vol. 187 (in German).
[22] P. Erdi and J. Toth, *Mathematical Models of Chemical Reactions* (Manchester University Press, Manchester, 1989).
[23] A similar result has been obtained with polynomial fitting of time series generated by discrete maps [6], e.g., the full recovery of the Hénon map from the class of two-dimensional quadratic systems.
[24] X. Zeng, R. Eykolt, and R. A. Pielke, Phys. Rev. Lett. **66**, 3229 (1991).
[25] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Physica D **16**, 285 (1985).