Proceedings of the
American Control Conference
Baltimore, Maryland • June 1994

Interpolation and Numerical Differentiation for Observer Design

S. Diop*, J. W. Grizzle†, P. E. Moraal† and A. Stefanopoulou†

Abstract

This paper explores interpolation and numerical differentiation as a basis for constructing a new approach to the design of nonlinear observers.

1 Introduction.

Observer design is a fundamental problem in system theory and control practice. The specific problem addressed herein concerns the situation where one seeks to "estimate" the states of a continuous-time model from observations collected at discrete instants in time. To date, this problem has been addressed only in the context of sampled data systems [4, 5], where the procedure has been to first compute a discrete-time model from the continuous-time model, and then to design an observer on the basis of the discretized system model. Of course, the discretization process sometimes can be quite non-trivial, from a numerical or computational point of view; also, it is usually quite important to assume that the observed data has been collected at regular intervals of time, and thus dealing with problems where event driven sampling is necessary can be difficult.

This paper arises from the observation that on the one hand, the numerical analysis literature contains a substantial body of results on numerical differentiation, while on the other hand, one of the very basic problems in system theory turns out to be that of obtaining the time derivatives of a continuous system variable, which is usually known only through the differential equations of the system, and time samples of this variable. We present some initial results on applying numerical differentiation to the observer design problem; this is done mainly by working out some specific examples, though the presentation is done in such a way as to suggest a number of formal "convergence results" which will be investigated in a future publication.

To see how numerical differentiation and observer design may tie together, consider the following simple observer design problem. Let a system be given by

$$\left\{\begin{array}{rcl} \dot{x}_1 &=& x_2 \\ \dot{x}_2 &=& -ux_1^3 \\ y &=& x_1 \end{array}\right. \tag{1.1}$$

and let the question be to build an observer for (1.1); that is, to reconstruct the states of the system on the basis of the available measurements. It is clear that the variable x_1 need not be computed since it is directly given by the measurement: $x_1 = y$. From the equations of (1.1), we see that $x_2 = \dot{y}$. If, as is usually the case, we have no access to $\dot{y}$ then we need to *compute* $\dot{y}$ from the measurement y. This is where numerical differentiation comes into play: we explore the feasibility of *numerically differentiating* y in order to obtain x_2. If y is known only through its time samples, then numerical differentiation yields an approximation of $\dot{y}$ at the current sampling time t_k by some number $\hat{\dot{y}}(t_k)$. An algorithm for the computation of $\hat{\dot{y}}(t_k)$ from the samples of y, and *estimates of the error bounds* should thus be worked out [1]. Usually, not only $\dot{y}$, but, also, finitely many higher derivatives of y should be approximated at time t_k in order to be able to compute all of the desired state components.

To our knowledge, the work most closely related to that presented here is [6].

2 An observability property.

Given a system described by the state equations

$$\left\{\begin{array}{lll} \dot{x}_i &= g_i(x,u) & (1 \le i \le n) \\ y_j &= h_j(x,u) & (1 \le j \le p) \end{array}\right. \tag{2.1}$$

we will adopt the following specific observability property (see [4, 5]): system (2.1) is **observable** if there is an integer N such that the map $H\left(u, \dot{u}, \cdots, u^{(N-1)}, \cdot\right)$ of the

*Chargé de Recherche CNRS - LAGEP, UCB Lyon I - 43 Bd. 11 Nov. 1918, 69622 Villeurbanne France; work supported in part by a grant from NSF/CNRS and in part by the Control Systems Laboratory of the Dept. Elect. Eng. and Comp. Sci., U. of Michigan.

†Control Systems Laboratory, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122; work supported in part by the National Science Foundation under contract NSF ECS-92-13551; matching funds to this grant were provided by FORD MO. CO.

[1] An extensive review of the numerical differentiation and interpolation literature is given in [3].

state space into some space of output values, defined by

$$H\left(u, \dot{u}, \cdots, u^{(N-1)}, x\right) = \begin{pmatrix} y \\ \dot{y} \\ \vdots \\ y^{(N-1)} \end{pmatrix}, \qquad (2.2)$$

is *injective* for any fixed *universal input*. It is easiest at this stage to assume that there is actually no input in (2.1), or that all inputs are universal in the sense that the above map is injective for any input.

For an observable system (2.1), we then may write

$$x = L\left(u, \dot{u}, \cdots, u^{(N-1)}, y, \dot{y}, \cdots, y^{(N-1)}\right).$$

The existence of the map L is guaranteed by our definition of observability, but an explicit expression for L may not be easy to obtain so that numerical equation solving methods, such as Newton's algorithm, may be required in order to obtain L. When (2.1) is a rational system in the sense that the g_i's and the h_j's are rational functions of their arguments, then the type of observability we just defined corresponds to *rational observability* of [2, 1], and then the explicit expression of L may be constructively derived through the use of certain differential algebraic techniques.

For observable systems, the observer design problem may thus be seen as a problem of numerical differentiation: once estimates of the derivatives of y and u can be determined from the available measurements, then x can be determined from H or L.

3 Illustrative examples.

In order to better understand some of the issues involved in the use of numerical differentiation and interpolation as a basis for an observer design method, three academic examples have been studied. The first example involves the estimation of a time function and a few of its derivatives from noisy samples collected at discrete instants of time. The second example builds upon the first one by showing that if the signal is generated by a known, continuous-time linear model, then the model's dynamics can be incorporated into the estimation process with advantage. This will bring us into contact with more classical observer theory. The third example is intended to illustrate that the techniques sketched in the first two examples are also applicable to nonlinear system models, which is of course, the main point.

In each case, polynomials are used as the interpolating functions, total least squares at the interpolating points is the metric to be minimized, and the discrete data will be gathered at a uniform rate. These simplifications are made so that the discussion can be focused on other issues.

It is emphasized that no theorems are formulated and thus no proofs of validity are offered for any of the "algorithms" proposed herein. This paper is meant to be an exploration of a set of ideas. Formalization of these ideas into a rigorous design methodology will be pursued in a later publication.

3.1 Sum of two sinusoids.

The purpose of this example is to illustrate the use of numerical differentiation, alone and in combination with standard system theoretic notions, to recover or estimate several derivatives of a signal in noise. The key parameters of interest are: N, the order of the interpolating polynomial; Δt, the time interval between data points; $W * \Delta t$, the length in time of the moving window used for data interpolation (note that $W + 1$ is the number of data points in the window); K, the data node or "knot" within the moving window where the derivative is to be estimated (counted from the left, with the first node in the window numbered zero); and σ^2, the "intensity" of the noise process corrupting the measurements.

Let an analog signal be given by the sum of two sinusoids of frequencies 1 Hz. and 5 Hz., respectively:

$$y(t) = sin(2\pi t) + sin(10\pi t). \qquad (3.1)$$

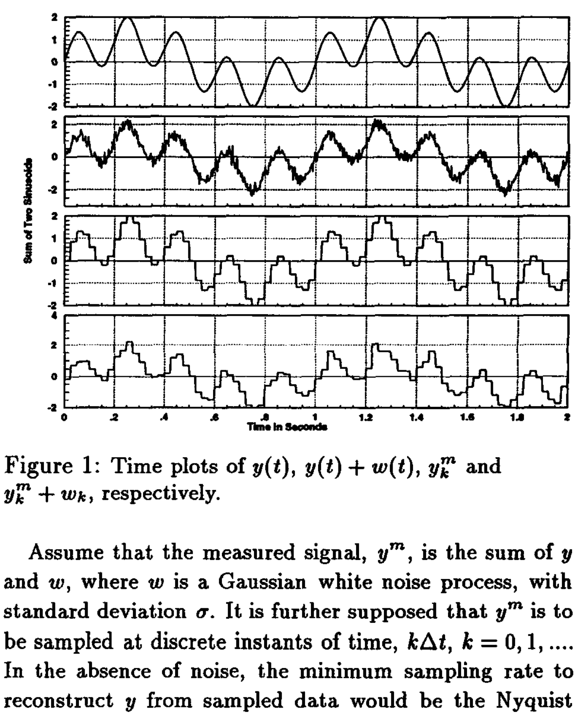

Figure 1: Time plots of $y(t)$, $y(t) + w(t)$, y_k^m and $y_k^m + w_k$, respectively.

Assume that the measured signal, y^m, is the sum of y and w, where w is a Gaussian white noise process, with standard deviation σ. It is further supposed that y^m is to be sampled at discrete instants of time, $k\Delta t$, $k = 0, 1, \ldots$. In the absence of noise, the minimum sampling rate to reconstruct y from sampled data would be the Nyquist rate, 10 Hz.; with noise, at least four times this rate is warranted (an anti-aliasing filter is not being used though it would be standard practice to do so). Discrete samples of $y + w$ will be collected therefore at 40 Hz.:

$$\begin{aligned} y^m(t) &:= y(t) + w(t) \\ y_k^m &:= y^m(k\Delta t). \end{aligned} \qquad (3.2)$$

The time signals discussed are depicted in Figure 1. The noise intensity was selected to provide a (numerically computed) signal to noise ratio of 20/1, and corresponded to $\sigma = 0.22$. This should provide a feel for the sample rate and noise intensity being used in all later simulations.

Let the interpolating polynomial for the window of data $\{y^m_{k-W}, \ldots, y^m_k\}$ be denoted by

$$\hat{y}_k(t) = a_0 + a_1(t - t_{k-W}) + \ldots + a_N(t - t_{k-W})^N, \quad (3.3)$$

where $t_k := k\Delta t$. The coefficients $\{a_0, a_1, \ldots, a_N\}$ are determined from the least squares solution of

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & \Delta t & \cdots & (\Delta t)^N \\ \vdots & \vdots & \vdots & \vdots \\ 1 & W\Delta t & \cdots & (W\Delta t)^N \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} y^m_{k-W} \\ y^m_{k-W+1} \\ \vdots \\ y^m_k \end{bmatrix}, \quad (3.4)$$

with respect to the Euclidean norm. The estimates of the derivatives of y at time $\bar{t}$ are determined by

$$\widehat{\frac{d^j}{dt^j} y_k}(\bar{t}) := \frac{d^j}{dt^j}\hat{y}_k(t)_{|t=\bar{t}} \; ; \quad (3.5)$$

for simplicity of notation, this is written as $\frac{d^j}{dt^j}\hat{y}_k(\bar{t})$.

A number of numerical experiments or simulations were conducted to ascertain the effects of the parameters N, Δt, W, K and σ on the ability to estimate the derivatives of y. The general trends of these experiments are summarized here; quantifying these observations analytically would be a worthy goal:

- The order of the interpolating polynomial should be selected as low as possible in order to average out the noise in the signal. On the other hand, for the estimation of the 3rd derivative of a signal, for example, at least a 3rd order polynomial is required.
- The window length $W\Delta t$ must be chosen large enough to capture the variations of the signal so that the interpolating polymomial can reproduce its derivatives. However, larger $W\Delta t$ require higher order polynomials to be chosen.
- Estimates of the derivatives made at the end-points of the window $W\Delta t$ are considerably less accurate than those made in the interior. However, selecting $K < W$ introduces a time delay of $(W-K)\Delta t$ in the estimates of the signal and its derivatives.
- For fixed N and $W\Delta t$, decreasing the time between samples, Δt, or increasing the number of points in the window, W, allows a higher noise intensity σ^2 to be tolerated.
- For a fixed Δt, though increasing W can produce more accurate results, it must be considered that it requires more on line compuation and a larger delay must be introduced (i.e., $(W-K)\Delta t$ is increased) in order to retain the accuracy of the estimates.

Figure 2 compares the true and estimated values of y, $\frac{dy}{dt}$, $\frac{d^2y}{dt^2}$, and $\frac{d^3y}{dt^3}$ when $N = 4$, $\Delta t = .025$, $W = 8$ and $K = 5$. Figure 3 displays the normalized errors in the estimates: the errors in the estimates for $(y, \ldots, \frac{d^3}{dt^3}y)$ are normalized by

$$lim_{T\to\infty}\frac{1}{T}\int_0^T |\frac{d^j}{dt^j}y(t)|dt. \quad (3.6)$$

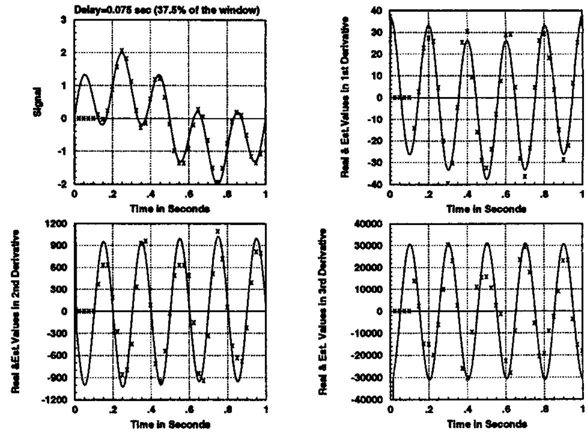

Figure 2: Comparison of the true and estimated (×) values of y, $\frac{dy}{dt}$, $\frac{d^2y}{dt^2}$, and $\frac{d^3y}{dt^3}$, respectively using interpolation/numerical differentiation.

Thus, the error in y is divided by 0.796, $\frac{dy}{dt}$ by 19.3, $\frac{d^2y}{dt^2}$ by 650, and $\frac{d^3y}{dt^3}$ by 22,700. The wide range in the magnitudes of y and its derivatives makes the estimation problem quite challenging.

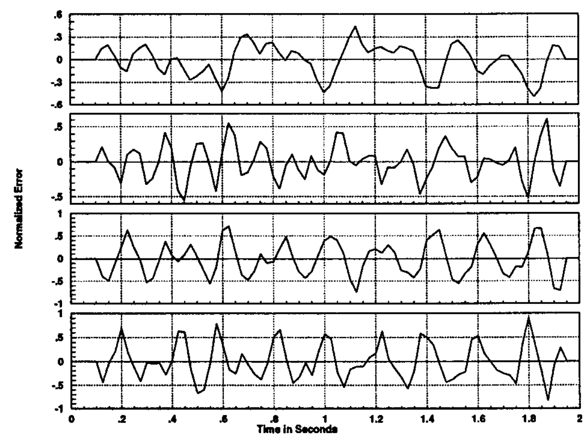

Figure 3: Normalized estimation errors in y, $\frac{dy}{dt}$, $\frac{d^2y}{dt^2}$ and $\frac{d^3y}{dt^3}$, respectively using interpolation/numerical differentiation (eq. (3.4)).

Estimating the signal and its derivatives on the basis of the $W+1$ data points in the moving window corresponds to using an FIR (finite impulse response) filter. By introducing a simple modification to the above, a sort of IIR (infinite impulse response) filter can be achieved. The modification is somewhat analogous to what is done in spline approximations, but to a systems person, it would be called a state.

The idea is to append to (3.4) the estimates of y and its derivatives at node K from the previous window, that is, one appends the relations

$$\frac{d^j}{dt^j}\hat{y}_k(t_{k-1-W+K}) = \frac{d^j}{dt^j}\hat{y}_{k-1}(t_{k-1-W+K}), \quad (3.7)$$

for $j = 0, \dots, J$, for some $0 \le J \le N$.

The right-hand side of (3.7) must be initialized to start the interpolation process, and thus becomes a state. This state allows derivative information to be passed from one window of data to the next. Weights can be added to trade-off errors in interpolating the measured data points versus errors in interpolating the derivatives in (3.7). Due to space limitations, no simulations employing this modification are presented; a related idea is explored in the next subsection.

3.2 Including a system model in the interpolation process.

The goal of this example is to demonstrate two methods for incorporating a dynamic model of a system into the interpolation/numerical differentiation process, whenever the signal in question is generated by a finite set of ordinary differential equations. This will bring us into more direct contact with observer design because we will tightly connect estimating the derivatives of a system's output with estimating the states of the model.

The signal (3.1) can obviously be generated by an initialized, uncontrolled, linear model with four states. Let

$$\begin{aligned} \frac{dx}{dt} &= Ax \\ y &= Cx \end{aligned} \tag{3.8}$$

be such a model; it will be observable. This means that x can be recovered from estimates of y and its derivatives through

$$\begin{bmatrix} y \\ \frac{d}{dt}y \\ \vdots \\ \frac{d^\eta}{dt^\eta}y \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^\eta \end{bmatrix} x \tag{3.9}$$

for η large enough, which in the particular case of (3.8) is $\eta = 3$.

The first way to include the model in the estimation process is now explained. Suppose that in general the model (3.8) has dimension n, and for simplicity, assume that it has a single output. As before, let

$$\begin{bmatrix} \hat{y}_{k-1}(t_{k-1-W+K}) \\ \frac{d}{dt}\hat{y}_{k-1}(t_{k-1-W+K}) \\ \vdots \\ \frac{d^{n-1}}{dt^{n-1}}\hat{y}_{k-1}(t_{k-1-W+K}) \end{bmatrix} \tag{3.10}$$

be estimates of y and its first $n-1$ derivatives from the previous window. By substituting (3.10) for the right-hand side of (3.9) with $\eta = n-1$, one can determine $\hat{x}_{k-1}(t_{k-1-W+K})$. If the resulting estimate of x is "accurate", then it satisfies (3.9) for any η; if not, choosing $\eta \ge n$ introduces constraints whose errors can be minimized in order for x to be more closely compatible with the dynamics (3.8). Thus, in place of (3.7), one augments (3.4) with

$$\frac{d^j}{dt^j}\hat{y}_k(t_{k-1-W+K}) = CA^j \hat{x}_{k-1}(t_{k-1-W+K}), \tag{3.11}$$

For $j = 0, \dots, \eta$, and proceeds as indicated previously. Note that taking $\eta = J \le n-1$ reduces (3.11) to (3.7).

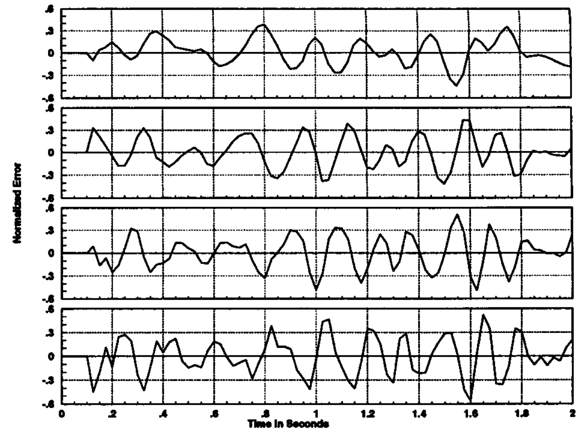

Figure 4: Normalized estimation errors in y, $\frac{dy}{dt}$, $\frac{d^2y}{dt^2}$, and $\frac{d^3y}{dt^3}$, respectively using interpolation/numerical differentiation (3.4) and $\eta = 5$ in (3.11). Note the reduced error with respect to Figure 3.

Figure 4 shows the effect of this modification on the estimation process for $N = 6$, $\Delta t = .025$, $W = 8$, $J = 6$ and $K = 5$. The weight on the constraints was chosen as $Q = \mathrm{diag}\{1.26,\ 5.19\ 10^{-2},\ 1.54\ 10^{-3},\ 4.4\ 10^{-5},\ 1.3\ 10^{-6},\ 4.4\ 10^{-8}\}$. The rationale for selecting the entries of the weight Q is that the magnitude of each successive derivative grows by a approximatley 30 (recall (3.6), or see (3.1)), and thus Q corresponds to equal "relative weighting" on y and its derivatives; this same rationale will be applied to subsequent examples. Including the model interpolation constraints allowed the same window size and delay as in Figure 3 to be used with an increased polynomial order, without "tracking" the noise.

The method just outlined for estimating x from y and its derivatives is akin to a "dead-beat" observer for x, though, as explained earlier, there is smoothing in the interpolation of y, which is "IIR". Some additional smoothing from the model can be obtained, and this constitutes the second method for introducing the model into the interpolation/numerical differentiation process. This technique, which is reminiscent of [4, 5], will be directly applicable to nonlinear systems.

Let A_d denote the discretization of the linear model, (3.8), and let $\hat{x}^+_{k-1}(t_{k-1-W+K})$ be the previous estimate of x. Update x based on the latest estimate of y and its derivatives through a damped Newton method:

$$\begin{aligned} \hat{x}^-_k(t_{k-W+K}) &= A_d \hat{x}^+_{k-1}(t_{k-1-W+K}) \\ \hat{x}^+_k(t_{k-W+K}) &= \hat{x}^-_k(t_{k-W+K}) + \epsilon P^{-1} \cdot \\ \left(\begin{bmatrix} \hat{y}_k(t_{k-W+K}) \\ \frac{d}{dt}\hat{y}_k(t_{k-W+K}) \\ \vdots \\ \frac{d^{n-1}}{dt^{n-1}}\hat{y}_k(t_{k-W+K}) \end{bmatrix} \right. &- \left. P\hat{x}^-_k(t_{k-W+K}) \right), \end{aligned} \tag{3.12}$$

where $P = col(C, CA, \dots, CA^{n-1})$ and $0 < \epsilon < 1$; using $\epsilon = 1$ reduces the above to a "dead-beat" estimator for x. The damped Newton update can be used with (3.4) alone or with any of the additions to (3.4) discussed previously.

Selected simulations of the above estimation methods are now discussed. In what follows, the state in (3.8) was chosen as

$$x = col(y, \frac{d}{dt}y, \frac{d^2}{dt^2}y, \frac{d^3}{dt^3}y).$$

It was properly initialized so that its output is precisely given by (3.1). The same noise sequence used in Figure 3 was added to the output of (3.8).

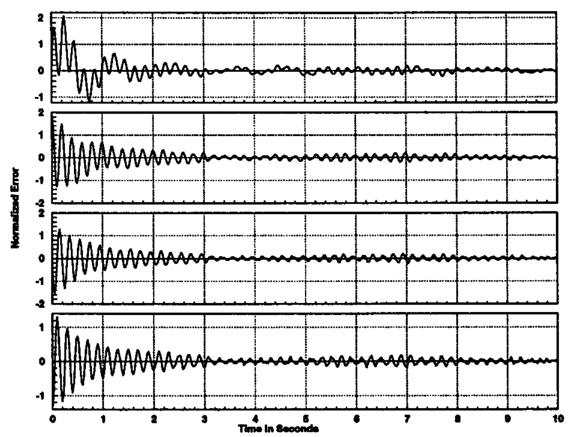

Figure 5: Normalized estimation errors in y, $\frac{dy}{dt}$, $\frac{d^2y}{dt^2}$ and $\frac{d^3y}{dt^3}$, respectively, incorporating "spline method" and Newton update in interpolation/numerical differentiation (equations (3.4), (3.7) and (3.11)).

Figure 5 shows the normalized estimation error resulting from the application of (3.4) in combination with (3.7) and (3.12). The parameter values were chosen as: $N = 5$, $\Delta t = 0.025$ sec., $W = 6$ and $K = 3$ in (3.4), $\eta = 4$ in (3.11) with $J = 4$ and a weight on (3.7) selected as $Q = \text{diag}\{1.26\ 10^{-1}, 5.19\ 10^{-2}, 1.54\ 10^{-3}, 4.4\ 10^{-5}\}$ and $\epsilon = 0.03$ in (3.12). For comparison purposes, a steady state, discrete-time, Kalman filter was designed with state noise covariance equal to $0.01\ \sigma^2\ Q^{-1}$ and output noise variance equal to σ^2; even though (3.8) does not have any process noise per se, the pair consisting of the matrix A and the state noise convariance matrix had to be made stabilizable for a stable Kalman filter to exist. The magnitudes of the steady state errors were essentially identical to those in Figure 5.

The normalized estimation errors resulting from the application of (3.4) in combination with (3.11) and (3.12) with the parameter values chosen as: $N = 5$, $\Delta t = 0.025$ sec., $W = 6$ and $K = 3$ in (3.4), $\eta = 4$ in (3.11), $J = 5$ and a weight selected as $Q = \text{diag}\{1.26, 5.19\ 10^{-2}, 1.54\ 10^{-3}, 4.4\ 10^{-5}, 1.3\ 10^{-6}\}$ were, again, essentially identical to those in Figure 5.

3.3 Third order, nonlinear dynamical system.

Consider the following system which is a "convex combination" of an unstable linear dynamics and a stable linear dynamics:

$$\begin{aligned} \begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{bmatrix} &= f(x) = \begin{bmatrix} x_2 \\ x_3 \\ f_3(x) \end{bmatrix} \qquad (3.13) \\ y &= h(x) = x_1, \end{aligned}$$

where,

$$\begin{aligned} f_3(x) &= \alpha(x)g_s(x) + (1-\alpha(x))g_u(x), \\ \alpha(x) &= 2\frac{x'x}{1+x'x} \\ g_s(x) &= -54x_1 - 36x_2 - 9x_3 \\ g_u(x) &= 54x_1 - 36x_2 + 9x_3. \end{aligned}$$

A state estimator was constructed using the interpolation/numerical differentiation method of (3.4) using $N = 6$, $W = 8$, $K = 4$, and a nonlinear version of (3.12). The sample rate was set at 4 Hz. As before, a noise process with a 20/1 signal to noise ratio was added to the measurement. Very encouraging results were obtained despite the fact that the system exhibits "complicated behaviour", that is, there appear to be multiple locally attractive limit cycles near the origin. The interested reader is referred to [3].

References

[1] S. Diop. Rational system equivalence, and generalized realization theory. In *Proceedings of the IFAC-Symposium NOLCOS'92*, pages 402–407, Bordeaux, 1992. IFAC.

[2] S. Diop and M. Fliess. On nonlinear observability. In C. Commault, D. Normand-Cyrot, J. M. Dion, L. Dugard, M. Fliess, A. Titli, G. Cohen, A. Benveniste, and I. D. Landau, editors, *Proceedings of the First European Control Conference*, pages 152–157, Paris, 1991. Hermès.

[3] S. Diop, J. W. Grizzle, P. E. Moraal and A. Stefanopoulou, "Interpolation and Numerical Differentiation for Observer Design", University Michigan's College of Engineering Control Group Report Series, Report No. CGR-93-14, September 1993.

[4] J. W. Grizzle and P. E. Moraal. Newton observers and nonlinear discrete-time control. In *Proceedings of the 29th IEEE Conference on Decision and Control*, 1990, pages 760–767.

[5] P. E. Moraal and J. W. Grizzle. Nonlinear discrete-time observers using Newton's and Broyden's methods. In *Proceedings of the Americal Control Conference*, June 1992, pages 3086–3090.

[6] U. Schönwandt, "Approximations to nonlinear observers", *Automatica*, Vol. 9,1973, pages 349–356.