PARAMETER IDENTIFICATION AND STOCHASTIC CONTROL

Girma Mullisa
Department of Electrical Engineering
Addis Ababa University

ABSTRACT

This paper is presented in two parts. PART I deals with the identification of the parameters of discrete systems described by difference equations, using a tailored form of the Kalman filter. PART II describes the methodology of stochastic controller design based on the identified parameters found in PART I to control the original noise-corrupted system. The approach taken is that of optimal prediction based on the solution of a linear Diophantine equation.

INTRODUCTION

Real systems are generally quite complex, not only because they may need high order equations to describe them but also because they may show nonlinear behavior in some range of their operation, and additionally they are often corrupted by noise. The engineer may in many cases do well to consider a mathematical model in place of the real system. For operation about an operating point, linearized equations with unknown parameters of arbitrary order may be introduced, to be identified in such a way that the squared-error between the outputs of the real system and the approximating model is minimized. If the model is found to be satisfactory, then further processing like controller design can be attempted, using the model parameters, to control the real system.

PART I: PARAMETER IDENTIFICATION

Let a real system be described by the difference equation

$$\begin{aligned} y(n) + a_1 y(n-1) + \cdots + a_p y(n-p) \\ = b_o u(n-d) + b_1 u(n-d-1) \\ + \cdots + b_m u(n-d-m) \\ + c_o \omega(n) + c_1 \omega(n-1) \\ + \cdots + c_r \omega(n-r) \end{aligned} \quad (1)$$

in which $a_o = 1$ and the set $a_i = 1 ,..., p$; b_i , $i = 1$,..., m; c_i, $i = 1$,..., r and the control delay d are all unknown. u_i is the control variable and ω_i is the input noise. Equation (1) may be written, using the delay operator z^{-1} (i.e $z^{-1} y(t) = y(t-\Delta)$), as

$$y(z^{-1}) = \frac{z^{-d} B(z^{-1})}{A(z^{-1})} u(z^{-1}) + \frac{C(z^{-1})}{A(z^{-1})} \omega(z^{-1}) \quad (2)$$

where

$$\begin{cases} A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_p z^{-p} \\ B(z^{-1}) = b_o + b_1 z^{-1} + \cdots + b_m z^{-m} \\ C(z^{-1}) = c_o + c_1 z^{-1} + \cdots + c_r z^{-r} \end{cases} \quad (3)$$

This open-loop model can be represented as in Fig. 1:

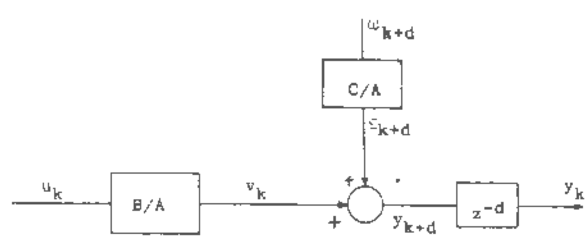

Figure 1 Open-loop model of noise-corrupted system

Obviously the noise may actually appear in the input signal, or it may be generated either internally or in the output section. Since the system is assumed to be linear, the principle of superposition enables us to model the noise as an additive one at the output.

There is an advantage in using this model for parameter estimation as one can obtain the estimates $\hat{A}, \hat{B}$ with $\omega = 0$ and $\hat{A}, \hat{C}$ with $u = 0$. In fact an average value can be obtained for vector $\hat{A}$

During parameter estimation the inputs u_k, ω_{k+d} are made to be software generated random signals to assure that the input remains *persistently exciting*, that is to say that it will always enable us to extract new information from the output as new output data become available.

LEAST SQUARES SOLUTION

Let $c = 0$ in Equation (1), and assume the system has been running for some time. This is to avoid the transient period after starting. A sequence of outputs can be written for a p^{th} order model (for $d = 0$) as

$$\begin{cases} y(N) = -\hat{a}_1 y(N-1) - \hat{a}_2 y(N-2) - \cdots - \hat{a}_p y(N-p) \\ \qquad + \hat{b}_o u(N) + \hat{b}_1 u(N-1) + \cdots + \hat{b}_m u(N-m) \\ y(N+1) = -\hat{a}_1 y(N) - \hat{a}_2 y(N-1) - \cdots - \hat{a}_p y(N-p+1) \\ \qquad + \hat{b}_o u(N+1) + \hat{b}_1 u(N) + \cdots + \hat{b}_m u(N-m+1) \\ \vdots \\ y(N+p) = -\hat{a}_1 y(N+p) - \hat{a}_2 y(N+p-1) - \cdots - \hat{a}_p y(N) \\ \qquad + \hat{b}_o u(N+p) + \hat{b}_1 u(N+p-1) + \cdots + \hat{b}_m u(N-m+p) \end{cases} \tag{4}$$

or, in matrix form

$$\begin{bmatrix} -y(N-1) & -y(N-2) \cdots -y(N-p) & u(N)\, u(N-1) \cdots u(N-m) \\ -y(N) & -y(N-1) \cdots -y(N-p+1) & u(N+1)\, u(N) \cdots u(N-m+1) \\ \vdots & \vdots \qquad \vdots & \vdots \\ -y(N+p) & -y(N+p-1) \cdots -y(N) & u(N+p)\, u(N+p-1)\, u(N+p-m) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_p \\ \hat{b}_o \\ \hat{b}_1 \\ \vdots \\ \hat{b}_m \end{bmatrix} = \begin{bmatrix} y(N) \\ y(N+1) \\ \vdots \\ y(N+p) \end{bmatrix} \tag{5}$$

This can be written in the still simpler form $X\hat{\theta} = y$ for which the well-known least-squares method gives the so-called normal solution set

$$\hat{\theta} = \left[XX^T\right]^{-1} Xy \tag{6}$$

A direct solution of (6) is generally not attempted as the matrix inversion of the large data matrix can be prohibitive. However, a simplification based on the Matrix Inversion Lemma [1] leads to the iterative form

$$\hat{\theta}_{m+1} = \hat{\theta}_m + K_m\left(y_{m+1} - X_{m+1}^T \hat{\theta}_m\right) \tag{7}$$

where

$$K_m = \frac{P_m X_{m+1}}{1 + X_{m+1}^T P_m X_{m+1}} \tag{8}$$

= Kalman gain

$$P_{m+1} = P_m - K_m X_{m+1}^T P_m \tag{9}$$

with

$$P_m = \left(X_m X_m^T\right)^{-1} \tag{10}$$

In the above presentation matrix inversion has been avoided by processing, not a bunch of data, but one output point at a time. To clarify this problem further let the first data vector for a first order model be

$$X_1^T = \left[-y(N-1) \quad u(N)\right]$$

so that

$$P_1^{-1} = X_1 X_1^T = \begin{bmatrix} y^2(N-1) & -u(N)\,y(N-1) \\ -y(N-1)\,u(N) & u^2(N) \end{bmatrix} \tag{11}$$

This matrix is singular and cannot be inverted.

It is only at this starting point that a problem of this type is met as there is no other point at which matrix inversion is required. In this paper the problem is avoided by introducing the following approximation:

$$P_1^{-1} = \left(X_1 X_1^T\right) \approx \begin{bmatrix} y^2(N-1) & 0 \\ 0 & u^2(N) \end{bmatrix} \tag{12}$$

This step immediately leads to the initial estimated parameters

$$\hat{\theta}_1 = \begin{bmatrix} \dfrac{1}{y^2(N-1)} & 0 \\ 0 & \dfrac{1}{u^2(N)} \end{bmatrix} \begin{bmatrix} -y(N-1) \\ u(N) \end{bmatrix} y(N) = \begin{bmatrix} -y(N)/y(N-1) \\ y(N)/u(N) \end{bmatrix} \tag{13}$$

For a 2^{nd} order system one sets the initial data vector

$$X_1^T = \left[-y(N-1) \quad -y(N-2) \quad u(N) \quad u(N-1)\right]$$

with

$$P_1^{-1} = X_1 X_1^T = \begin{bmatrix} y^2(N-1) & y(N-1)\,y(N-2) & & -u(n-1)\,y(N-1) \\ y(N-2)\,y(N-1) & y^2(N-2) & \cdots & u(N-1)\,y(N-2) \\ -u(N)\,y(N-1) & -u(N)\,y(N-2) & \cdots & u(N-1)\,u(N) \\ -u(N-1)\,y(N-1) & -u(N-1)\,y(N-2) & \cdot & u^2(N-1) \end{bmatrix}$$

$$\approx \begin{bmatrix} y^2(N-1) & 0 & 0 & 0 \\ 0 & y^2(N-2) & 0 & 0 \\ 0 & 0 & u^2(N) & 0 \\ 0 & 0 & 0 & u^2(N-1) \end{bmatrix}$$

giving

$$\hat{\theta}_1 = \begin{bmatrix} \dfrac{1}{y^2(N-1)} & 0 & 0 & 0 \\ 0 & \dfrac{1}{y^2(N-2)} & 0 & 0 \\ 0 & 0 & \dfrac{1}{u^2(N)} & 0 \\ 0 & 0 & 0 & \dfrac{1}{u^2(N-1)} \end{bmatrix} \begin{bmatrix} -y(N-1) \\ -y(N-2) \\ u(N) \\ u(N-1) \end{bmatrix} y(N)$$

The initially diagonal covariance matrix P becomes non-diagonal as further data become available, as per equations (8) and (9).

The coefficient vector $\hat{C}$ connected with the noise input is found in exactly the same way. Furthermore, at the end of each iteration the approximate output

$$\begin{aligned} y_{app}(n) = & -\hat{a}_1 y(n-1) - \hat{a}_2 y(n-2) \\ & - \cdots - \hat{a}_q y(n-q) \\ & + \hat{b}_o u(n) + \cdots + \hat{b}_r u(n-r) \\ & + \hat{c}_o \omega(n) + \cdots + \hat{c}_v \omega(n-v) \end{aligned} \quad (14)$$

may be calculated with appropriate initial conditions and the squared-error $[y(n) - y_{app}(n)]^2$ computed to be used as a stopping criterion.

PART II: STOCHASTIC CONTROL

Two main approaches to modern controller design have developed over the years. One approach starts off with a state space model and uses the optimal Kalman filter for state (and/or parameter) estimation and goes on to compute the optimal control signal, making use of the *separation principle* to separate the estimation task from the control one.

The second approach starts from a polynomial representation of the system, makes a d-step-ahead optimal prediction of the output by solving a linear Diophantine equation and goes on to find the necessary optimal control sequence.

The first approach may often need more extensive computation as it deals with matrices. It is also more thorough in the depth of its treatment. The second approach only requires relatively simple operations on polynomials, although certain aspects of these operations are not as convenient as one would wish, in particular for machine computation. In this paper we attempt to present the main features of the second approach which we then apply to some of the parameter identification examples treated in PART I.

OPTIMAL PREDICTION

As already discussed in PART I, a discrete linear system can be modeled by the polynomial

$$A(z^{-1}) y_k = z^{-d} B(z^{-1}) u_k + C(z^{-1}) \omega_k \quad (15)$$

where y_k is the output sequence, u the control sequence, and ω_k a zero-mean white process noise with variance q. d is the delay in control, i.e. u_k affects the output values at times $k+d$ and later.

The polynomials $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$ have the general forms

$$\begin{cases} A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_n z^{-n} \\ B(z^{-1}) = b_o + b_1 z^{-1} + \cdots + b_m z^{-m}, b_o \neq 0 \\ C(z^{-1}) = c_o + c_1 z^{-1} + \cdots + c_n z^{-n} \end{cases} \quad (16)$$

where $C(z^{-1})$ is assumed to be stable, i.e. the zeros of the noise transfer function lie on or inside the unit circle. The reason for this will be clear later on in the discussion.

Now, a d-step-ahead prediction of the output y_{k+d} in terms of the outputs at k, k-1, ... can be made by first making the following division

$$\frac{C(z^{-1})}{A(z^{-1})} = F(z^{-1}) + z^{-d} \frac{G(z^{-1})}{A(z^{-1})} \quad (17)$$

where division continues until z^{-d} can be factored out of the remainder [2,3]. This gives the scalar Diophantine equation

$$C(z^{-1}) = A(z^{-1}) F(z^{-1}) + z^{-d} G(z^{-1}) \quad (18)$$

for which $F(z^{-1})$ and $G(z^{-1})$ are the solutions with the general forms

$$\begin{cases} F(z^{-1}) = f_o + f_1 z^{-1} + \cdots + f_{d-1} z^{-(d-1)} \\ G(z^{-1}) = g_o + g_1 z^{-1} + \cdots + g_{n-1} z^{-(n-1)} \end{cases} \quad (19)$$

Next, multiplication of Equation (15) by $F(z^{-1})$ and use of Equation (18) leads to

$$\begin{aligned} y_{k+d} = & \frac{G(z^{-1})}{C(z^{-1})} y_k + \frac{F(z^{-1}) B(z^{-1})}{C(z^{-1})} u_k \\ & + F(z^{-1}) \omega_{k+d} \end{aligned} \quad (20)$$

It is now easy to see why the zeros of $C(z^{-1})$ must not lie outside the unit circle.

Let $\hat{y}_{k+d/d}$ be the optimal prediction, found by minimizing the mean-squared error

$$J_k = E\left[\left(y_{k+d} - \hat{y}_{k+d/d}\right)^2\right] \tag{21}$$

giving the minimum-variance predictor

$$\hat{y}_{d+d/d} = \frac{G(z^{-1})}{C(z^{-1})}\, y_k + \frac{B(z^{-1})\,F(z^{-1})}{C(z^{-1})}\, u_k \tag{22}$$

with the prediction error $\bar{y}_{k+d/d} = F\omega_{k+d}$, and the minimum mean-square error

$$J_{\min} = q\left(f_o^2 + f_1^2 + \dots + f_{d-1}^2\right) \tag{23}$$

A more general formulation of the prediction problem is to use the quadratic performance index [4]

$$J_k = \left(P\, y_{k+d} - Q\, s_k\right)^2 + (R u_k)^2 \tag{24}$$

which puts weights on the output y_{k+d}, the reference signal s_k, and the control signal u_k. P, Q, R can have the general forms

$$\begin{cases} P(z^{-1}) = 1 + p_1 z^{-1} + \dots + p_{n_p} z^{-n_p} \\ Q(z^{-1}) = q_o + q_1 z^{-1} + \dots + q_{n_o} z^{-n_o} \\ R(z^{-1}) = r_o + r_1 z^{-1} + \dots + r_{n_R} z^{-n_R} \end{cases} \tag{25}$$

where n_p is the degree of $P(z^{-1})$, etc. One can introduce the desired emphasis into the performance index by the proper choice of the coefficients p_i, q_i, r_i. For instance, $P(z^{-1}) = 1$, $Q = R = 0$ corresponds to the simple minimum variance controller. $P, R \neq 0, Q = 0$ corresponds to the regulator problem, whereas $P, R, Q \neq 0$ deals with the tracking problem. Consider the term Py_{k+d} in the performance index:

$$\begin{aligned} P(z^{-1})y_{k+d} &= (1 + p_1 z^{-1} + p_2 z^{-2} \\ &\quad + \dots + p_{n_p} z^{-n_p})\, y_{k+d} \\ &= y_{k+d} + p_1 y_{k+d-1} + \dots + p_{n_p} y_{k+d-n_p} \end{aligned}$$

All values of y for which $d > n_p$ will have to be predicted in terms of $y_k, y_{k-1}, \dots$, so that for $j \leq d$ we have

$$y_{k+j} = \hat{y}_{k+j/k} + \bar{y}_{k+j/k} \tag{26}$$

in which, from the orthogonality principle, $\bar{y}_{k+j/k}$ and $\hat{y}_{k+j/k}$ are orthogonal to each other. Thus, as before, the optimal predictor for $y_{k+j/k}$ is found by making a j-step long division of C/A with the result

$$C(z^{-1}) = A(z^{-1})\,F_j(z^{-1}) + z^{-j}\,G_j(z^{-1}) \tag{27}$$

where F_j, G_j are found for all $d-n_p \leq j \leq d$. In practice one can find F_1, G_1 for $d = 1$, F_2, G_2 for $d = 2$, etc. in a single long-division.

The optimal j-step-ahead predictor is of the form of Equation (22):

$$\hat{y}_{k+j/k} = \frac{G_j}{C}\, y_k + \frac{BF_j}{C}\, u_{k-d+j}\,, \quad 0 < j \leq d \tag{28}$$

It is interesting to note that $\hat{y}_{k+j/k}$ depends, not just on u_k, but on also earlier values of the input.

An optimal control law may be developed for the general tracking problem, using the performance index in (24) where y_{k+d} is replaced by $\hat{y}_{k+d/k}$, thus having a deterministic cost function for which $\frac{\partial J_k}{\partial u_k} = 0$ leads to the equation

$$P\hat{y}_{k+d/k} + \frac{r_o}{b_o} R\, u(k) - Q s_k = 0 \tag{29}$$

where

$$\frac{\partial \hat{y}_{k+d/k}}{\partial u_k} = \frac{B(o)\, F_d(o)}{C(o)} \equiv b_o \tag{30}$$

For $j = d, d+1, \cdots, d-n_p$, $\hat{y}_{k+j/k}$ has the forms

$$\hat{y}_{k+d/k} = \frac{G_d}{C} y_k + \frac{BF_d}{C} u_k$$

$$\hat{y}_{k+d-1/k} = \frac{G_{d-1}}{C} y_k + \frac{BF_{d-1}}{C} u_{k-1}$$

$$\vdots$$

$$\hat{y}_{k+d-n_p/k} = \frac{G_{d-n_p}}{C} y_k + \frac{BF_{d-n_p}}{C} u_{k-n_p}$$

so that Equation (29) becomes

$$\left(\frac{G_d}{C} y_k + \frac{BF_d}{C} u_k\right) + p_1\left(\frac{G_{d-1}}{C} y_k + \frac{BF_{d-1}}{C} u_{k-1}\right)$$

$$+ \cdots + p_{n_p}\left(\frac{G_{d-n_p}}{C} y_k + \frac{BF_{d-n_p}}{C} u_{k-n_p}\right)$$

$$+ \frac{r_o}{b_o} R u_k - Q s_k = 0$$

This can be written as

$$T_2(z^{-1}) u_k = T_1(z^{-1}) y_k + C(z^{-1}) Q(z^{-1}) s_k \qquad (31)$$

where

$$\begin{cases} T_1(z^{-1}) = \sum\limits_{j=o}^{n_p} p_i G_{d-j} \\ T_2(z^{-1}) = \sum\limits_{j=o}^{n_p} p_j BF_{d-j} z^{-j} + \dfrac{r_o}{b_o} C(z^{-1}) R(z^{-1}) \end{cases} \qquad (32)$$

Notice that the control in Equation (31) immediately gives a closed-loop controller (Fig. 2) built around the open-loop system already shown in Fig. 1:

The stability of the above closed-loop system depends on the characteristic equation

$$\frac{B(z^{-1})}{A(z^{-1})} z^{-d} \cdot \frac{T_1(z^{-1})}{T_2(z^{-1})} + 1 = 0$$

i.e

$$A(z^{-1})\, T_2(z^{-1}) + B(z^{-1})\, T_1(z^{-1})\, z^{-d} = 0 \qquad (33)$$

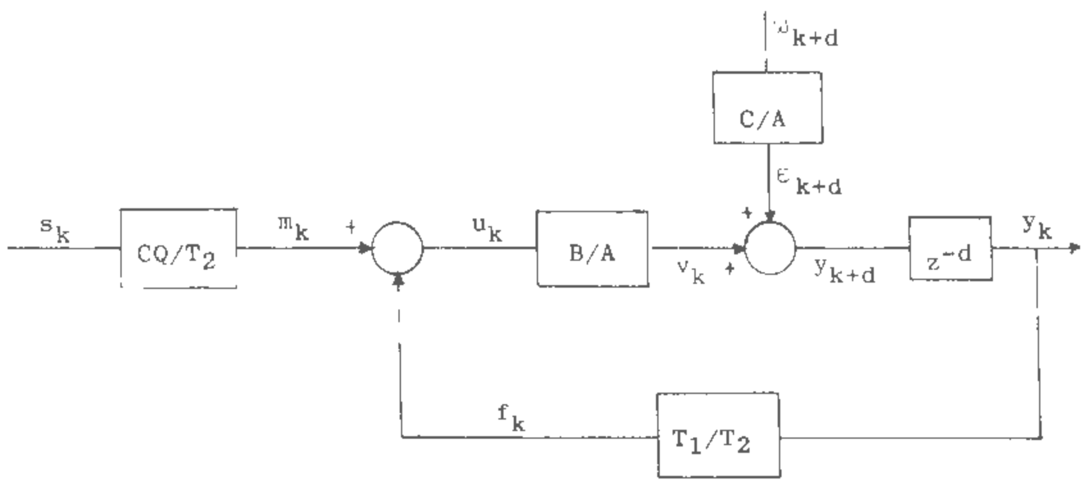

Figure 2 Closed-loop stochastic control system

Thus, whenever a model with parameters $\hat{A}, \hat{B}, \hat{C}$ and $\hat{d}$ is determined using an identification algorithm, the stability of the system should be checked from $\hat{A}(z^{-1})\, T_2(z^{-1}) + \hat{B}(z^{-1})\, T_1(z^{-1})\, z^{-d} = 0$ where $T_1(z^{-1})$ and $T_2(z^{-1})$ are found from (32) using the identified parameters.

The output sequences are computed for the following two cases using the corresponding discrete equations:

i) Open-loop system

$$
\begin{aligned}
&A(z^{-1})v_k = B(z^{-1})s_k \\
&A(z^{-1})\epsilon_{k+d} = C(z^{-1})\,\omega_{k+d} \\
&y_{k+d} \quad v_k + \epsilon_{k+d} \\
&y_k = z^{-d} y_{k+d}
\end{aligned}
\qquad (34)
$$

ii) Controlled closed-loop system

The equations here consist of the set in (34) with s_k replaced by u_k, plus the following additional ones:

$$
\begin{aligned}
&T_2(z^{-1})m_k \quad C(z^{-1})\ Q(z^{-1})\,s_k \\
&T_2(z^{-1})f_k \quad T_1(z^{-1})\,y_k \\
&u_k \quad m_k \quad f_k
\end{aligned}
\qquad (35)
$$

Remark: In the example systems studied in this paper, the output y_k is generated from the original equation using the parameters $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$, and delay d. However, the controller is designed using the estimated parameters $\hat{A}(z^{-1})$, $\hat{B}(z^{-1})$, $\hat{C}(z^{-1})$, and delay $\hat{d}$

Computational results:

For *PART I : Identification*

Example 1

Let the original system be described by

$$
\begin{aligned}
y(n) \quad 0.5y(n \quad 1) + 0.25y(n \quad 3) \\
u(n \quad 1) + 0.5u(n \quad 2) + 0.3u(n \quad 3) \\
+ \omega(n) + 0.25\omega(n \quad 1)
\end{aligned}
$$

This is a system with poles at -0.5, $0.5 \pm j\,0.5$, i.e. a stable system with poles well within the unit circle, and unit delay in the control. Models of various orders with different combinations of u, ω, and delay d can be tried.

Estimated parameters

First model:

a) 1st order $(\hat{a}_1, \hat{b}_0, \hat{c}_o)$

$\hat{A}, \hat{B}: \hat{a}_1 \quad 0.66$;

$\hat{b}_o \quad 0.06$ (approximation poor)

$\hat{A}, \hat{C}: \hat{a}_1 \quad 0.61$;

$\hat{c}_o \quad 1.03$ (approximation poor)

$\hat{A}_{av}: \hat{a}_1 \quad 0.635$

b) 1st order $(\hat{a}_1, \hat{b}_0, \hat{b}_1, \hat{c}_o, \hat{c}_1)$

$\hat{A}, \hat{B}: \hat{a}_1 \quad 0.66$;

$\hat{b}_o - 0.02$, $\hat{b}_1 \quad 1.04$ (approximation poor)

$\hat{A}, \hat{C}: \hat{a}_1 \quad 0.40$,

$\hat{c}_o \quad 1.01$, $\hat{c}_1 \quad 0.39$ (approximation better)

$\hat{A}_{av}: \hat{a}_1 - 0.53$

Second model:

a) 2nd order $(\hat{a}_1, \hat{a}_2, \hat{b}_0, \hat{b}_1, \hat{c}_o)$

$\hat{A}, \hat{B}: \hat{a}_1 - 0.98$, $\hat{a}_2 \quad 0.49$,

$\hat{b}_o \quad 0$, $\hat{b}_1 = 0.99$ (approximation acceptable)

$\hat{A}, \hat{C}: \hat{a}_1 \quad 0.83$, $\hat{a}_2 \quad 0.36$, $\hat{c}_o \quad 1.0$

$\hat{A}_{av}: \hat{a}_1 = 0.905$, $\hat{a}_2 \quad 0.425$

b) 2nd order $(\hat{a}_1, \hat{a}_2, \hat{b}_0, \hat{b}_1, \hat{c}_o, \hat{c}_1)$

$\hat{A}, \hat{B}: \hat{a}_1 \quad 0.98$, $\hat{a}_2 \quad 0.49$,

$\hat{b}_o \quad 0$, $\hat{b}_1 = 0.99$

$\hat{A}, \hat{C}: \hat{a}_1 \quad 1.04$, $\hat{a}_2 \quad 0.49$,

$\hat{c}_o \quad 1.0$, $\hat{c}_1 \quad 0.26$

$\hat{A}_{av}: \hat{a}_1 \quad 1.01$, $\hat{a}_2 \quad 0.49$

Third model:

3^{rd} order $(\hat{a}_1\ ,\ \hat{a}_2\ ,\ \hat{a}_3\ ,\ \hat{b}_0\ ,\ \hat{b}_1\ ,\ \hat{c}_o\ ,\ \hat{c}_1)$

$$\hat{A}\ ,\ \hat{B}\ :\ \hat{a}_1\ =\ -0.97\ ,\ \hat{a}_2\ =\ 0.47\ ,\ \hat{a}_3\ =\ 0.02\ ,$$

$$\hat{b}_o\ =\ 0\ ,\ \hat{b}_1\ =\ 0.99$$

$$\hat{A}\ ,\ \hat{C}\ :\ \hat{a}_1\ =\ -0.54\ ,\ \hat{a}_2\ =\ 0.05\ ,\ \hat{a}_3\ -\ 0.23\ ,$$

$$\hat{c}_o\ =\ 1.0\ ,\ \hat{c}_1\ =\ 0.23$$

Example 2

Let the original system be

$$y(n)\ -\ 1.905\ y(n-1)\ +\ 0.905\ y(n\ \ 2)$$

$$=\ u(n\ \ 1)\ +\ 0.9\ u(n-2)$$

$$+\ \omega(n)\ +\ 0.25\,\omega(n-1)$$

This system has a pole on the unit circle and another at 0.905. Thus, it is a marginally stable system.

First model: a) 1^{st} order

$$y(n)\ +\ \hat{a}_1\ y(n-1)$$

$$=\ \hat{b}_o\ u(n-1)\ +\ \hat{c}_o\ \omega(n)$$

Estimated parameters:

$$\hat{A}\,,\hat{B}\ :\ \hat{a}_1 = -1.0\ ;\ \hat{b}_o\ -\ -0.11\ \text{(unreliable)}$$

$$\hat{A}\,,\hat{C}\ :\ \hat{a}_1 = -1.01\ ;\ \hat{c}_o - 1.12$$

b) Another first order model

$$y(n) + \hat{a}_1\ y(n-1) - \hat{b}_o\ u(n-1) + \hat{b}_1\ u(n-2)$$

$$+\,\hat{c}_o\ \omega(n) + \hat{c}_1\ \omega(n-1)$$

Estimated parameters:

$$\hat{A}\,,\hat{B} : \hat{a}_1\quad 1.01\ ,\ \hat{b}_o\quad 0.13\ ,$$

$$\hat{b}_1\ \ 0.99\ \text{(unreliable)}$$

$$\hat{A}\,,\hat{C}^{*} : \hat{a}_1\quad 1.01\ ,\ \hat{c}_o\ \ 1.03\,,\hat{c}_1\ \ 1.24$$

Remark: This is not a suitable model as $\hat{C}$ has a zero outside the unit circle.

Second model: a) 2^{nd} order

$$y(n)\ +\ \hat{a}_1\ y(n-1)\ +\ \hat{a}_2\ y(n-2)$$

$$=\ \hat{b}_0\ u(n-1)\ +\ \hat{b}_1\ u(n-2)\ +\ \hat{c}_o\ \omega(n)$$

Estimated parameters:

$$\hat{A}\,,\hat{B}\ :\ \hat{a}_1 = -1.95\ ,\hat{a}_2 - 0.95\ ,\hat{b}_o - -0.03$$

$$\hat{A}\,,\hat{C}\ :\ \hat{a}_1 = -1.90\ ,\hat{a}_2 - 0.90\ ,\hat{c}_o - 1.07$$

Average values:

$$\hat{a}_1 -\ \ 1.925\ ,\hat{a}_2\ \ 0.925\ ,$$

$$\hat{b}_o = -0.03\ ,\hat{c}_o - 1.07$$

b) Another second order model

$$y(n)\ +\ \hat{a}_1\ y(n\ \ 1)\ +\ \hat{a}_2\ y(n\ \ 2)$$

$$-\ \hat{b}_o\ u(n\ \ 1)\ +\ \hat{b}_1\ u(n\ \ 2)$$

$$+\ \hat{c}_o\ \omega(n)\ +\ \hat{c}_1\ \omega(n\ \ 1)$$

Estimated parameters:

$$\hat{A}\ ,\ \hat{B}\ :\ \hat{a}_1\quad 1.95\ ,\ \hat{a}_2\quad 0.95\ ,$$

$$\hat{b}_o\quad 0.10\ ,\ \hat{b}_1\quad 1.04$$

$$\hat{A}\ ,\ \hat{C}\ :\ \hat{a}_1\quad 1.86\ ,\ \hat{a}_2\quad 0.86$$

$$\hat{c}_o\quad 1.05\ ,\ \hat{c}_1\quad 0.35$$

$$\hat{A}_{\text{average}}\ \ \hat{a}_1\quad 1.905\ ,\ \hat{a}_2\quad 0.905$$

Third model: a) 3^{rd} order

$$y(n)\ +\ \hat{a}_1\ y(n\ \ 1)\ +\ \hat{a}_2\ y(n\ \ 2)\ +\ \hat{a}_3\ y(n\ \ 3)$$

$$-\ \hat{b}_0\ u(n\ \ 1)\ +\ \hat{b}_1\ u(n\ \ 2)\ +\ \hat{c}_o\ \omega(n)$$

Estimated parameters

$$\hat{A}\ ,\ \hat{B}\ :\ \hat{a}_1\ -\ -2.37\ ,\ \hat{a}_2\quad 1.84\ ,$$

$$\hat{a}_3\ -\quad 0.44\ ,\ \hat{b}_o\ -\ 0.03$$

$$\hat{A}\ ,\ \hat{C}\ :\ \hat{a}_1\ -\ -1.86\ ,\ \hat{a}_2\ =\ 0.82\ ,$$

$$\hat{a}_3\ =\ 0.05\ ,\quad \hat{c}_o\ -\ 1.07$$

Remark: values of $\hat{A}$ are widely different.

b) Another third order model

In this case one gets

$\hat{A}$, $\hat{B}$: $\hat{a}_1 = -2.37$, $\hat{a}_2 = 1.81$, $\hat{a}_3 = -0.44$,

$\hat{b}_o = -0.04$, $\hat{b}_1 = 1.04$

$\hat{A}$, $\hat{C}$: $\hat{a}_1 = -1.26$, $\hat{a}_2 = -0.34$, $\hat{a}_3 = 0.59$,

$\hat{c}_o = 1.04$, $\hat{c}_1 = 0.96$

Remark: values of $\hat{A}$ are widely different.

The graphs of the exact and approximate responses using some of the models in Examples 1 and 2 are shown in Fig. 3, in which the horizontal axis shows the number of samples.

For PART II : Stochastic Control

EXAMPLE 1

Using 1st order model (*b*):

$$\hat{A}(z^{-1}) = 1 \;\; 0.53z^{-1}$$

$$\hat{B}(z^{-1}) = -0.02 + 1.04z^{-1}$$

$$\hat{C}(z^{-1}) = 1.01 + 0.39z^{-1}$$

For $d = 1$, $F_1 = 1.01$; $G_1 = 0.925$

For $d = 2$, $F_2 = 1.01 + 0.925\, z^{-1}$; $G_2 = 0.49$

For $d = 1$, $T_1 = 0.925$

$$b_o = \frac{B(o)\,F_d(o)}{C(o)} = \frac{-0.02 \times 1.01}{1.01} = -0.02$$

$$T_2 = \hat{B}F_1 + \frac{r_o}{b_o}\hat{C}R = \hat{B}F_1 + \frac{r_o^2}{b_o}\hat{C}$$

$$= (-0.02 + 1.04z^{-1})(1.01) \;\; 50\,r_o^2(1.01 + 0.39z^{-1})$$

$$= 1.01(-0.02 - 50r_o^2) + (1.05 - 19.5r_o^2)z^{-1}$$

For $r_o = 1$, $T_2 = -50.5 - 18.45\,z^{-1}$

A check on stability shows that this is a stable system.

Then

$$\frac{m_k}{s_k} = \frac{C(z^{-1})}{T_2(z^{-1})} = \frac{1.01 + 0.39z^{-1}}{-50.5 \;\; 18.45z^{-1}}$$

$$= \frac{0.02 - 0.0077z^{-1}}{1 + 0.365z^{-1}}$$

i.e.

$$m_k = -0.025\,s_k \;\; 0.0077s_{k-1} - 0.365\,m_{k-1}$$

$$\frac{f_k}{y_k} = \frac{T_1(z^{-1})}{T_2(z^{-1})} = \frac{0.925}{-50.5 \;\; 18.45z^{-1}}$$

$$= \frac{0.0183}{1 + 0.365z^{-1}}$$

i.e.

$$f_k = \;\; 0.0183\,y_k \;\; 0.365f_{k-1}$$

Similarly, for $r = 0.5$, one gets the equations

$$T_2 = -12.645 - 3.825z^{-1}$$

with

$$m_k = -0.08\,s_k \;\; 0.031\,s_{k-1} \;\; 0.302\,m_{k-1}$$

$$f_k = -0.073y_k - 0.302f_{k-1}$$

Using 2nd order model (*b*):

$$\hat{A}(z^{-1}) = 1 - 1.01z^{-1} + 0.49z^{-2}$$

$$\hat{B}(z^{-1}) = 0.99 \quad (\text{delay } d = 2)$$

$$\hat{C}(z^{-1}) = 1.0 - 0.26z^{-1}$$

Then,

For $d = 1$, $F_1 = 1$; $G_1 = 0.75 - 0.49\,z^{-1}$
For $d = 2$, $F_2 = 1 + 0.75\,z^{-1}$; $G_2 = 0.268 - 0.368\,z^{-1}$

Now, for $d = 2$, $r_o = 1$, $T_2 = 2 + 0.48\,z^{-1}$, giving a stable system having the dynamic controller equations (in part)

$$m_k = 0.5\,s_k - 0.13s_{k-1} - 0.42m_{k-1}$$

$$f_k = 0.134y_k - 0.184y_{k-1} - 0.24f_{k-1}$$

EXAMPLE 2

Using second model (*b*): but with change in $\hat{B}(z^{-1})$.

$$\hat{A}(z^{-1}) = 1 - 1.905z^{-1} + 0.905z^{-2}$$

$$\hat{B}(z^{-1}) = 1.04 \quad \text{(but with } d = 2\text{)}$$

$$\hat{C}(z^{-1}) = 1.05 + 0.35z^{-1}$$

$$Q(z^{-1}) = 0.5 + z^{-1} \quad \text{(arbitrary choice)}$$

d=2 : $F_2 = 1.05 + 2.35\, z^{-1}$; $G_2 = 3.527 - 2.127\, z^{-1}$

$$b_o = \frac{B(o)\,F_d(o)}{C(o)} = \frac{1.04 \times 1.05}{1.05} = 1.04$$

$$T_1 = G_2 = 3.527 - 2.127z^{-1}$$

$$T_2 = BF_2 + \frac{r_o^2}{b_o} C(z^{-1})$$

$$= 1.04(1.05 + 2.35z^{-1})$$

$$+ 0.962\, r_o^2 (1.05 + 0.35z^{-1})$$

$$= 1.092 + 2.444z^{-1} + 1.01r_o^2$$

$$+ 0.337r_o^2 z^{-1}$$

Let $r_o = 4$: $T_2 = 17.25 + 7.836\, z^{-1}$

Then

$$\frac{m_k}{s_k} = \frac{CQ}{T_2} = \frac{(1.05 + 0.35z^{-1})\ (0.5 + z^{-1})}{17.25 + 7.836z^{-1}}$$

$$= \frac{0.03 + 0.071z^{-1} + 0.02z^{-2}}{1 + 0.454z^{-1}}$$

$$m_k = 0.03s_k + 0.071s_{k-1} + 0.02s_{k-2} - 0.454m_{k-1}$$

$$\frac{f_k}{y_k} = \frac{T_1}{T_2} = \frac{3.527 - 2.127z^{-1}}{17.25 + 7.836z^{-1}}$$

$$= \frac{0.204 - 0.123z^{-1}}{1 + 0.454z^{-1}}$$

$$f_k = 0.204y_k - 0.123y_{k-1} - 0.454f_{k-1}$$

Open loop and controlled system responses are shown in Fig. 4, in which the horizontal axis is the number of samples.

CONCLUSION

The parameter identification algorithm works quite well even on systems that are unstable. The stochastic controller works well on stable systems if the identified vector $\hat{C}$ lies inside the unit circle. However, the controller fails to work if the original system is unstable. This means that an unstable system must first be stabilized before stochastic control is attempted.

The identification problem and the control problem are treated separately in the paper. An obvious improvement is the machine computation of the solution to the Diophantine equation so that identification and control can be done in one. This is necessary if, for instance, adaptive control is to be attempted.

REFERENCES

[1] Girma Mullisa "Introduction to Optimization" a teaching material, Addis Ababa University, 1979.

[2] Clarke, D.W., : "Introduction to Self-tuning Controllers", in "Self-tuning and Adaptive Control", Harris, C.J., and S.A. Billings, editors, Peter Peregrinus, 1985.

[3] Goodwin, F.C., and K.S. sin: "Adaptive Filtering, Prediction, and Control", Prentice-Hall, 1984.

[4] Lewis, F.L : "Optimal Estimation", Wiley, 1986.

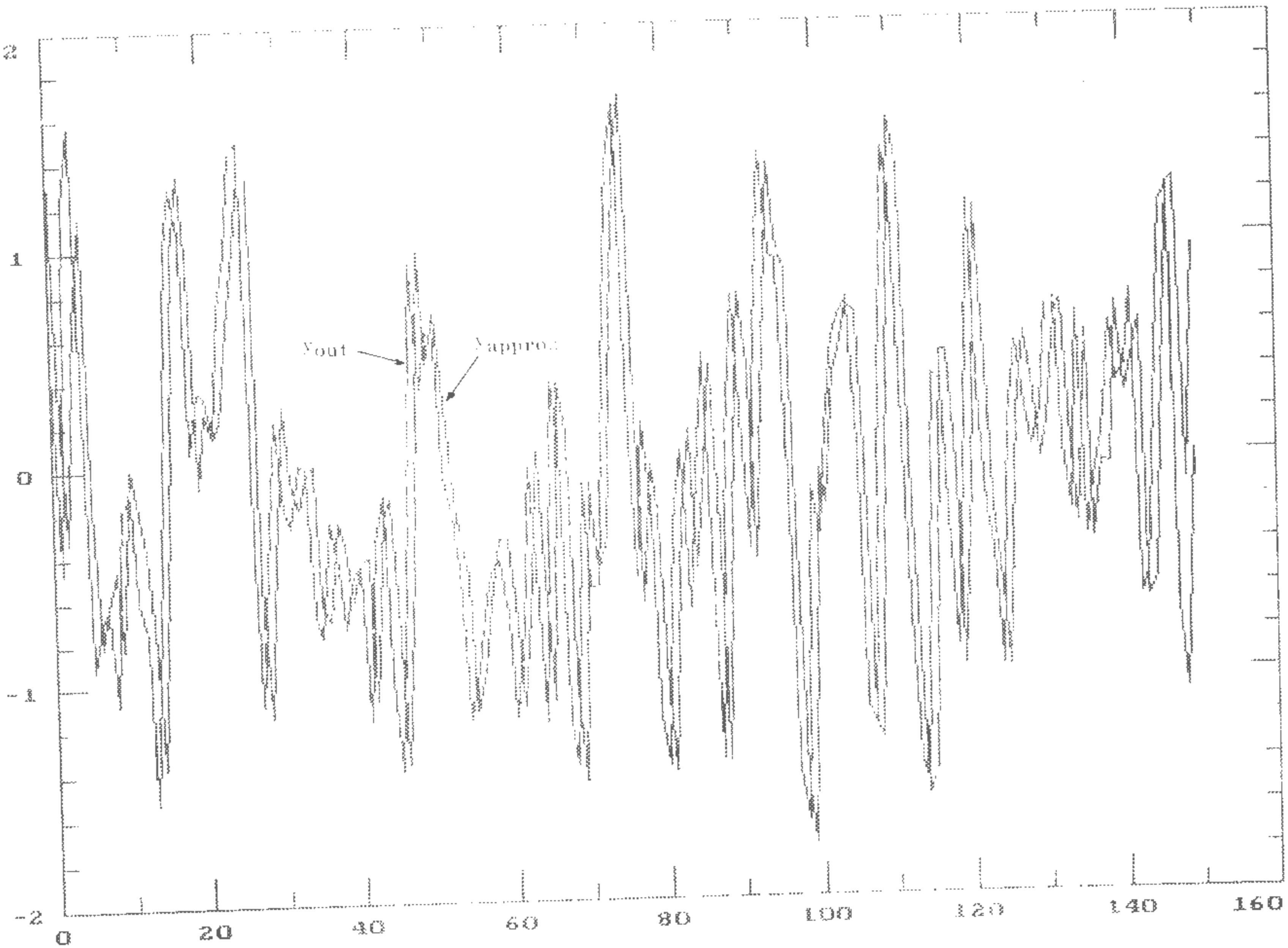

Figure 3-1 Responses of actual (y_{out}) and identified (y_{approx}) systems: Example 1 first model

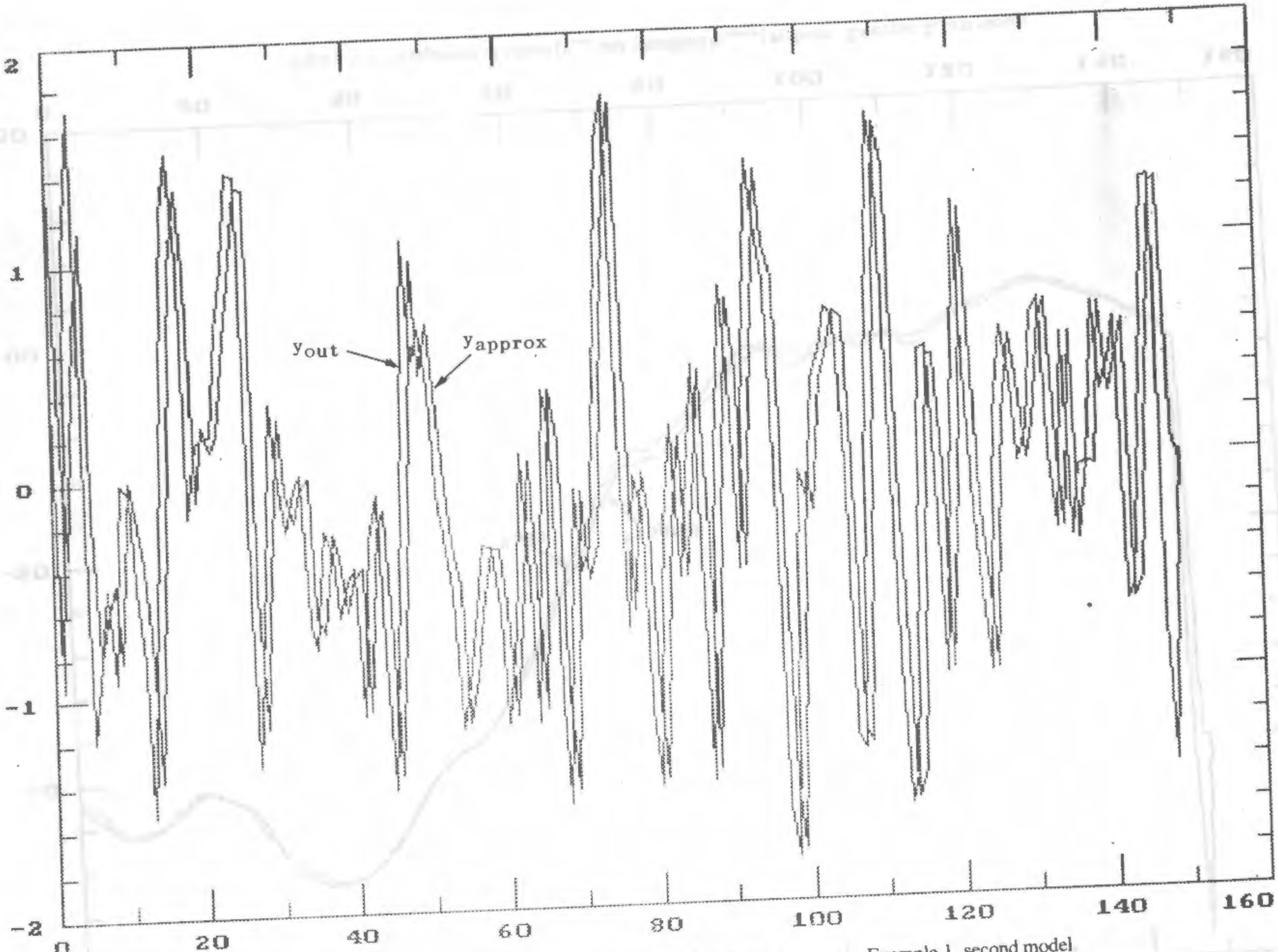

Figure 3-2 Responses of actual (y_{out}) and identified (y_{approx}) systems: Example 1, second model

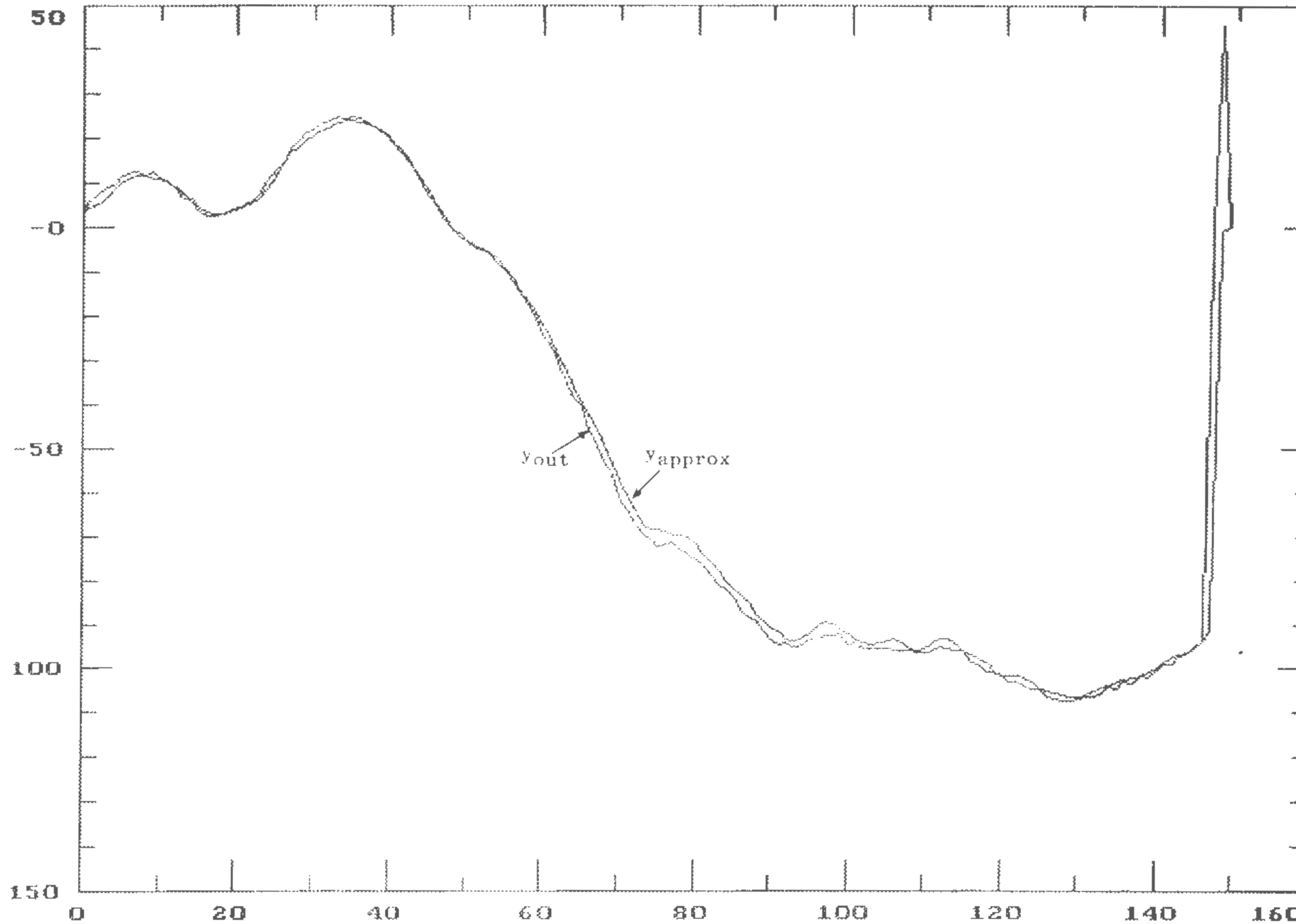

Figure 3-3 Responses of actual (y_{out}) and identified (y_{approx}) systems: Example 2, first model

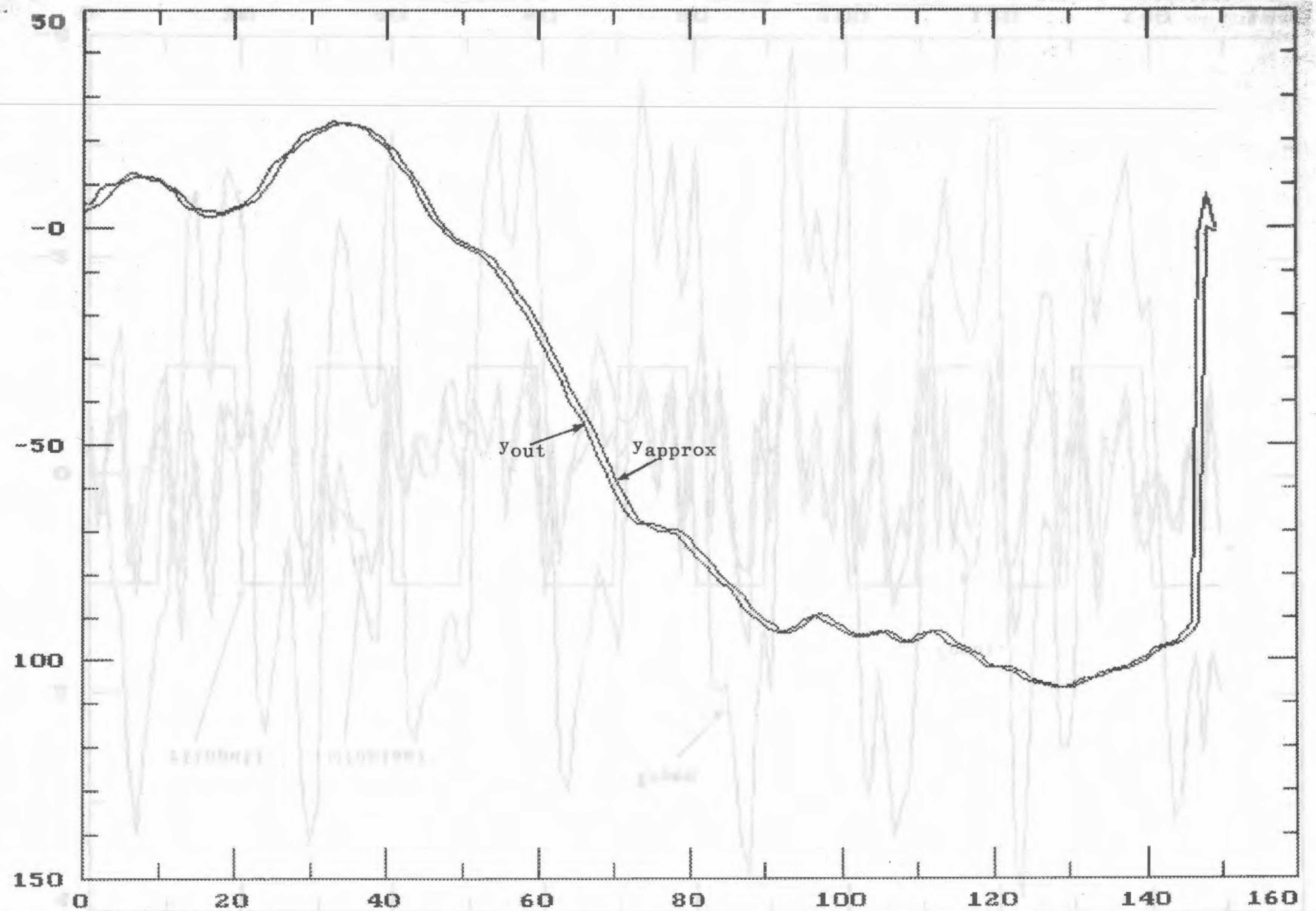

Figure 3-4 Responses of actual (y_{out}) and identified (y_{approx}) systems: Example 2, second model

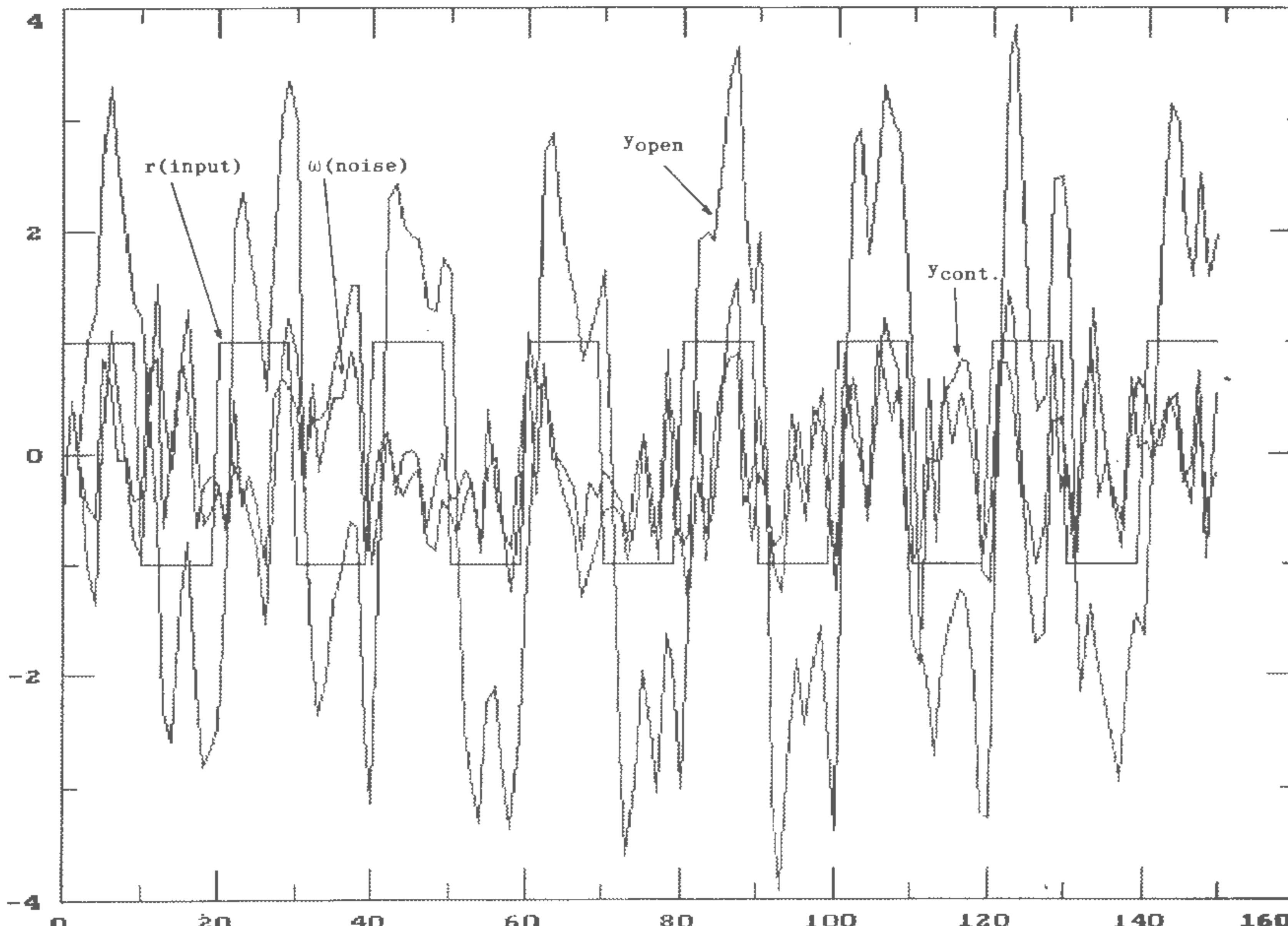

Figure 4-1 Open-loop response (y_{open}) and controlled response (y_{cont}): Example 1, first order model, 100% noise

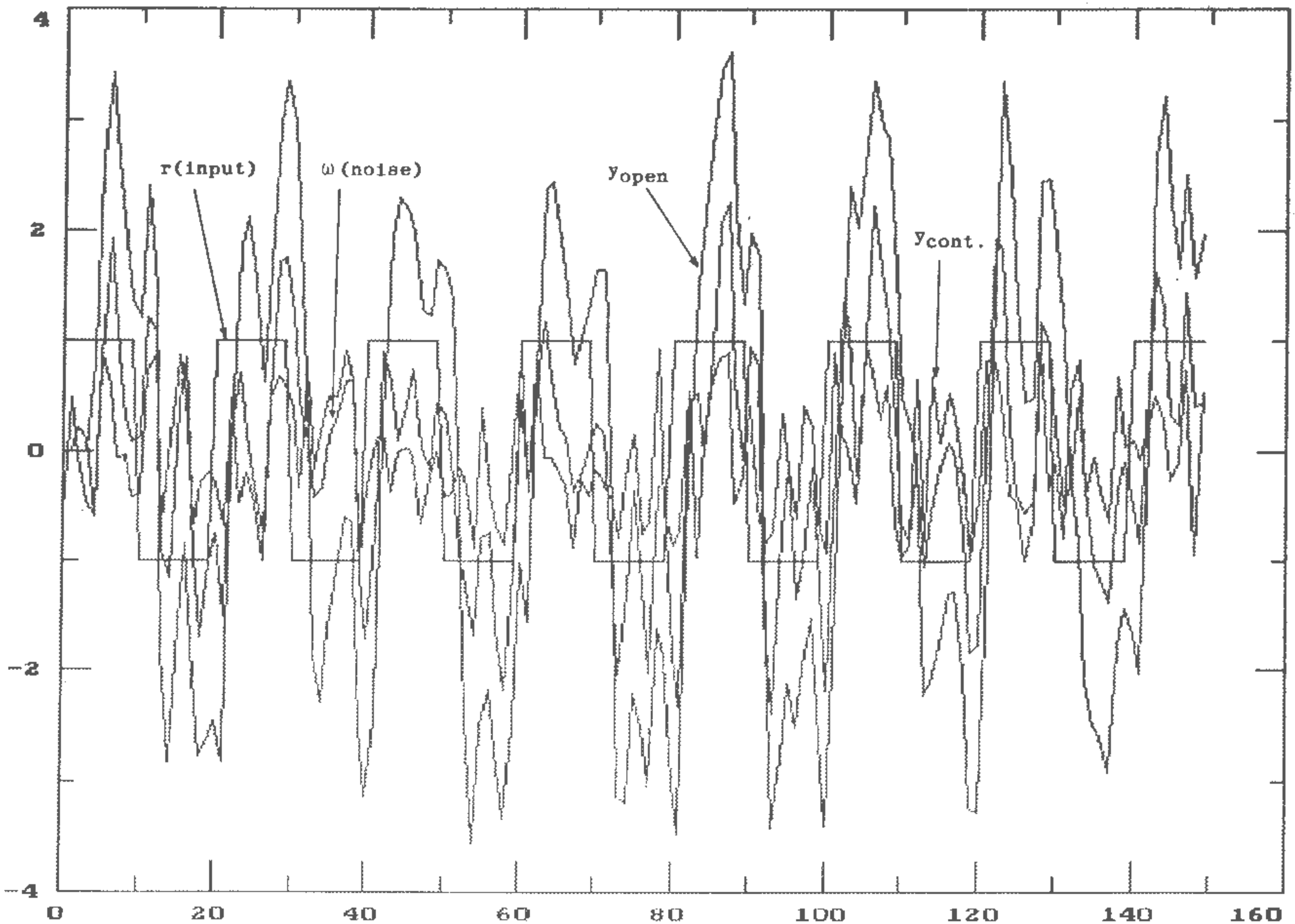

Figure 4-2 Open-loop response (y_{open}) and controlled response ($y_{cont.}$): Example 1, second order model, 100% noise

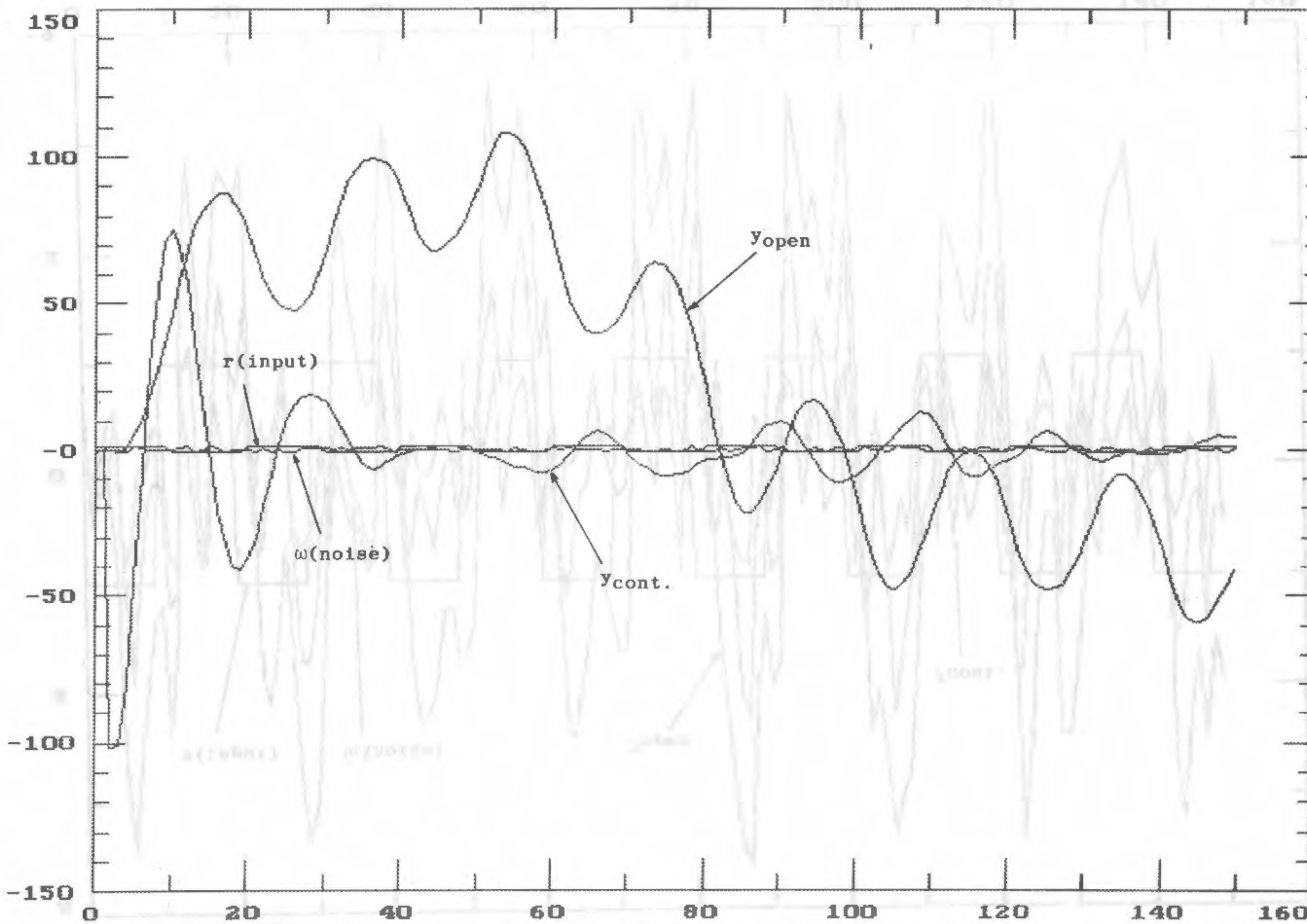

Figure 4-3 Open-loop response (y_{open}) and controlled response (y_{cont}): Example 2, second order model, 100% noise

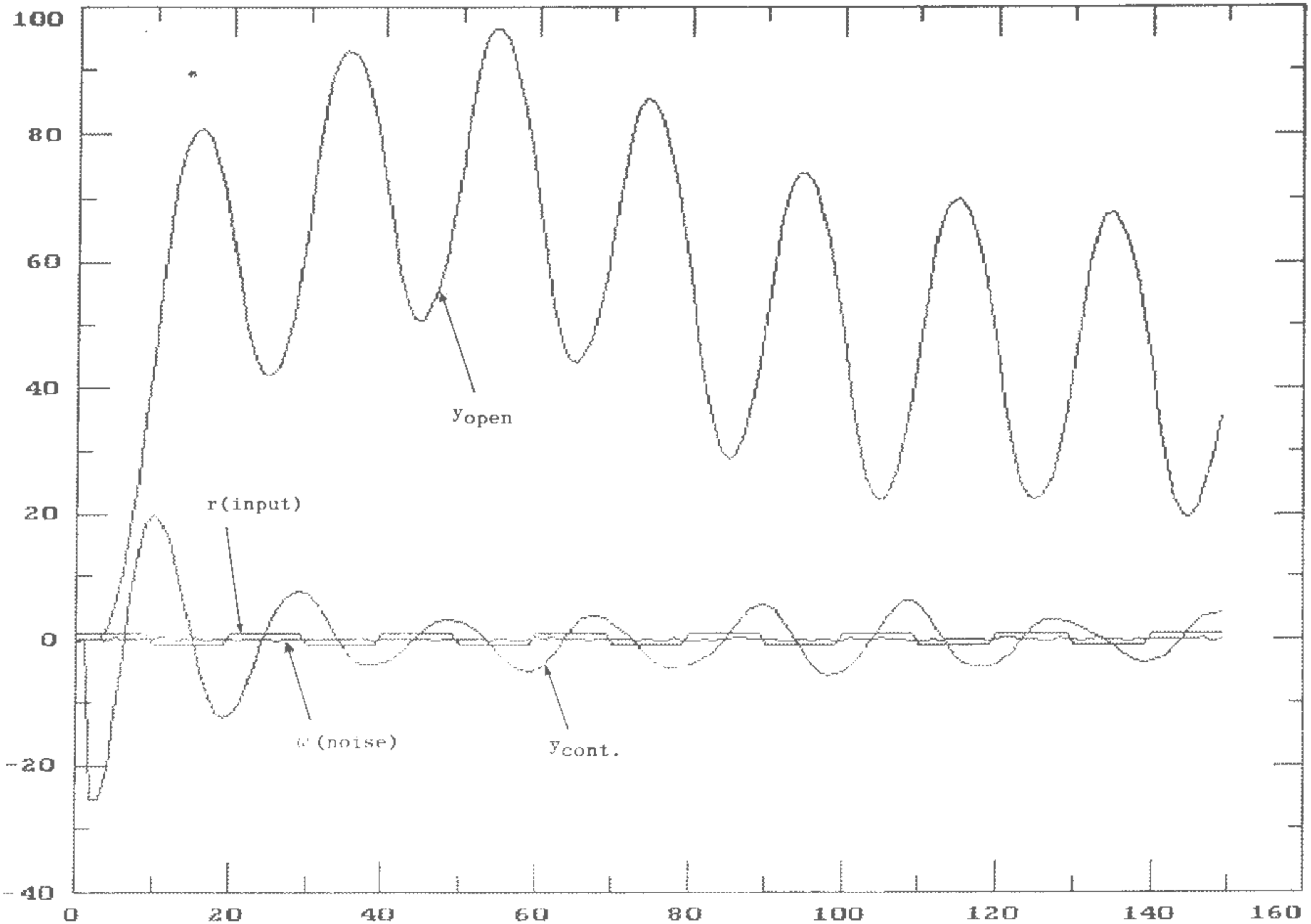

Figure 4-4 Open-loop response (y_{open}) and controlled response (y_{cont}): Example 2, second order model, 25% noise