Technical Research Report

Systems Research Center

Supported by the
National Science Foundation
Engineering Research Center
Program (NSFD CD 8803012),
the University of Maryland,
Harvard University,
and Industry

Conditions for the Equivalence of ARMAX and ARX Systems

by G.A. McGraw, C.L. Gustafson, and J.T. Gillis

TR 91-67

Conditions for the Equivalence of ARMAX and ARX Systems *

G. A. McGraw †   C. L. Gustafson †   J. T. Gillis † ‡

Abstract

It is shown that an autoregressive moving average with exogenous input (ARMAX) system can be represented as an autoregressive with exogenous input (ARX) model if and only if the transfer function from the noise port to the output port has no transmission zeros. A construction using the matrix fractional description of the system is used to prove this result. This construction shows that, by proper addition of sensor measurements and extending the order of the ARX model, accurate parameter estimates of systems driven by unmeasured disturbances can be obtained.

I. Introduction

A problem frequently encountered in identification of dynamic systems is that, in addition to excitation by known inputs, the system may also be driven by unmeasured disturbances. For linear systems in the stationary limit, such systems can be represented by an autoregressive moving average with exogenous input (ARMAX) model:

$$y_t = -\sum_{i=1}^{n} A_i y_{t-i} + \sum_{i=0}^{m} B_i u_{t-i} + \sum_{i=0}^{p} C_i e_{t-i}, \quad t = \ldots, -1, 0, 1, \ldots. \tag{1}$$

Here y_t is the measured $n_y \times 1$ output vector, u_t is the $n_u \times 1$ exogenous (known) input vector, and e_t is the $n_e \times 1$ disturbance vector— which is assumed to be standard white Gaussian noise. Analogously, an autoregressive with exogenous input (ARX) model is defined by taking $C_i = 0$ for $i \geq 1$ in Eq. (1). For single input-single output (SISO) systems, it is well known that identification of the coefficients A_i and B_i of ARMAX systems using ARX model-based least squares methods leads to biased parameter estimates. We show here that a multivariable ARMAX system can be represented *exactly* by an ARX model whenever the multivariable transfer function from the

*Supported by Aerospace Sponsored Research and the Aerospace Corporation Advanced Study Grant Programs.

†The Aerospace Corp., M4/974, P.O. Box 92957, Los Angeles, CA, 90009-2957.

‡Currently Visiting Research Associate, Systems Research Center, University of Maryland, College Park.

noise input to the measured output has no transmission zeros. This result shows that by the proper addition of sensor outputs, ARX models can accurately model a larger class of systems than previously thought, thus extending the utility of linear least square parameter estimation algorithms.

Least square identification with ARMAX models is a nonlinear problem because the product of two unknown quantities, namely the MA coefficients, C_i, and the noise sample path, e_t, enter into the cost function. In contrast, the least square identification using ARX models is linear, since the unknown parameters, A_i and B_i, are multiplied by known signals, y_t and u_t. In many system identification applications, such as adaptive control, it is desirable to use ARX models to avoid convergence and computational difficulties associated with ARMAX or state space models.

In the SISO case, the MA part is easily seen to be equivalent to the presence of zeros in the transfer function from e_t to y_t. The results we develop here generalize this fact to the multi-input multi-output (MIMO) case. Individual component transfer functions in MIMO systems may have zeros, but as we show, it is the presence of multivariable transmission zeros in the MA part which determine whether an ARX model will exist for system (1). We also present a computational example which illustrates the utility of this result in the identification of multivariable systems.

II. Definitions and Preliminary Results

Our main result utilizes a construction based on the theory of matrix fraction descriptions. We shall briefly review some relevant definitions (see Kailath [2] for more details), and then discuss some preliminary results. We assume in this section that $N(z) \in \mathcal{R}(z)^{n\times m}$ and $D(z) \in \mathcal{R}(z)^{n\times n}$, where $\mathcal{R}(z)^{n\times m}$, denotes the ring of $n \times m$ polynomial matrices in z.

Definition 1. The pair $(N(z), D(z))$ is a *left coprime factorization* of the $n \times m$ transfer function $H(z)$ if and only if $H(z) = D(z)^{-1}N(z)$, and there exist $U(z)$ and $V(z)$ such that $D(z)V(z) + N(z)U(z) = I_n$, where $U(z)$ and $V(z)$ are $m \times n$ and $n \times n$ polynomial matrices, respectively.

The polynomial matrices $N(z)$ and $D(z)$ are generalizations of the numerator and denominator polynomials, respectively, of a single-input single-output (SISO) transfer function. Consistently, we can form a definition for a type of zeros of a MIMO transfer function, commonly called transmission zeros.

Definition 2. A complex number, $z_o \in \mathbf{C}$, is a (finite) *transmission zero* of the $n \times m$ transfer function $H(z)$ if and only if $\text{rank}(N(z_o)) < r$ where the pair $(N(z), D(z))$ is a left coprime factorization of $H(z)$ and $r = \text{rank}(N(z))$ over $\mathcal{R}(z)^{n\times m}$.

We shall formally define transmission zeros of a polynomial matrix $N(z)$, by setting $D(z) = I_n$ in Definition 2. We will also find the following concept useful.

Definition 3. A $n \times n$ polynomial matrix $L(z)$ is said to be *unimodular* if and only if $L^{-1}(z)$ exists and is itself polynomial.

With the above definitions, we have the following standard result.

Lemma 1. [Column Hermite Form] Let $P(z) \in \mathcal{R}(z)^{n\times m}$, with rank r in $\mathcal{R}(z)^{n\times m}$ and $n \geq m$. Then $P(z)$ can be reduced to the following quasi-triangular form by elementary row operations (i.e., by multiplication on the left by a unimodular matrix):

$$P(z) \longrightarrow \left[\begin{array}{ccccc|c} \alpha_{11}(z) & \cdots & \alpha_{1j}(z) & \cdots & \alpha_{1r}(z) & \times \\ & \ddots & \vdots & & & \\ & & \alpha_{jj}(z) & & \vdots & \times \\ & 0 & & \ddots & & \\ & & & & \alpha_{rr}(z) & \times \\ \hline & & & 0_{(n-r)\times m} & & \end{array}\right] \tag{2}$$

where:

1. $\alpha_{jj}(z)$ is monic and $\deg \alpha_{ij}(z) < \alpha_{jj}(z)$ for $i < j$;
2. If $\alpha_{jj}(z) = 1$, then $\alpha_{ij}(z) = 0$, $i < j$;
3. $\times$ indicates that nothing can be said.

Proof: See Kailath [2, Thm 6.3-2]. ∎

We will be interested in applying the Column Hermite Form to the case where $P(z)$ has no transmission zeros:

Lemma 2. Let $P(z) \in \mathcal{R}(z)^{n\times m}$, have rank m and have no transmission zeros. Then the Column Hermite Form of $P(z)$ is

$$P(z) \longrightarrow \begin{bmatrix} I_m \\ 0_{(n-m)\times m} \end{bmatrix}$$

Proof: Denote the Hermite Form of $P(z)$ as $P_H(z)$. Suppose the j^{th} column of $P_H(z)$ is $[\alpha_{1j}(z), \ldots, \alpha_{jj}(z), 0, \ldots, 0]^T$, as in Eq. (2), and let z_o be a root of $\alpha_{jj}(z)$. Then $P_H(z)$ drops rank at z_o, implying that z_o is a transmission zero of $P(z)$, by Definition 2. This is a contradiction; therefore $\alpha_{jj} = 1$, $\alpha_{ij} = 0$, $i < j$. ∎

III. Main Result

In this section we develop necessary and sufficient conditions for the existence of an ARX model for an ARMAX system.

Let z be the unit delay operator, i.e., $z^k x(t) = x(t-k)$. Then the ARMAX model, Eq. (1), can be rewritten as

$$A(z)y_t = B(z)u_t + C(z)e_t, \tag{3}$$

$$A(z) = \sum_{i=0}^{n} A_i z^i, \tag{4}$$

$$B(z) = \sum_{i=0}^{m} B_i z^i, \tag{5}$$

$$C(z) = \sum_{i=0}^{p} C_i z^i. \tag{6}$$

The $n_y \times n_y$ matrix A_0, is assumed to be full rank (usually $A_0 = I_{n_y}$), and without loss of generality $C_0 \neq 0$. The ARMAX model may then be written as a left matrix fractional description:

$$y_t = A(z)^{-1} \begin{bmatrix} B(z) & C(z) \end{bmatrix} \begin{bmatrix} u_t \\ e_t \end{bmatrix}. \tag{7}$$

Note that when the model is an ARX model, the polynomial matrix $C(z)$ becomes $C(z) = C_0$, a constant matrix.

We can now use the results of the previous section to prove our main result.

Theorem 3. Consider a dynamical system S, specified by an $n_y \times (n_u + n_e)$ transfer function matrix, as in Eq. (7). Let the pair $(A(z), [\ B(z) \quad C(z)\ ])$ be left coprime, and satisfy $\text{rank}(C(z)) = n_e$ over $\mathcal{R}(z)^{n_y \times n_e}$. Then S has an ARX representation if and only if $C(z)$ has no finite transmission zeros.

Proof: We will first show that if $C(z)$ has no finite transmission zeros, then S has an ARX representation.

By Lemma 2, there exists an $n_y \times n_y$ unimodular polynomial matrix, $L(z)$, such that

$$L(z)C(z) = \begin{bmatrix} I_{n_e} \\ 0 \end{bmatrix} \tag{8}$$

We can then construct another left coprime factorization of the system (7)

$$y_t = \tilde{A}^{-1}(z)[\ \tilde{B}(z) \quad \tilde{C}(z)\ ] \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{9}$$

where

$$\tilde{A}(z) = L(z)A(z) \tag{10}$$

$$\tilde{B}(z) = L(z)B(z) \tag{11}$$

$$\tilde{C}(z) = L(z)C(z). \tag{12}$$

Note that $\tilde{A}(z)$ is necessarily a polynomial matrix, since $L(z)$ and $A(z)$ are polynomial matrices. From, Eqs. (12) and (8) we see that the MA polynomial, $C(z)$, has been reduced to a single, constant matrix, thus the system S has an ARX representation.

Conversely, suppose the system S has an ARX representation, as in (3), with $C(z) \equiv C_0$. Then clearly $C(z)$ has no transmission zeros. ∎

Remarks:

1. We have assumed no special structure of the ARMAX and equivalent ARX systems: each scalar transfer function in (7) is assumed to be full order. The condition that $(A(z), [B(z)\, C(z)])$ is left coprime is equivalent to the identifiability of the original ARMAX system [3].

2. When an equivalent ARX model exists, the AR polynomial matrix, $\tilde{A}(z)$, will have an order between n and $n+p$. This phenomena was observed in practice [1] and its explanation led to Theorem 3.

3. The theorem shows that identification of ARMAX systems with MIMO ARX models can be successful if there are more outputs than disturbance inputs ($n_y > n_e$) because this will usually eliminate transmission zeros in the noise transfer function.

4. Sensor measurements will always be corrupted by additive noise sources which are independent of other disturbances acting through the system dynamics; this typically precludes the existence of an exact ARX model for the system. For example, a system which is subject to a scalar white input disturbance, w_t, and $n_y \times 1$ additive white sensor noise, v_t,

$$y_t = A(z)^{-1}[B(z)u_t + C(z)w_t] + v_t,$$

has an ARMAX model of the form,

$$A(z)y_t = B(z)u_t + [\ C(z) \quad A(z)\ ] \begin{bmatrix} w_t \\ v_t \end{bmatrix}.$$

In this case the noise transfer function will have transmission zeros at the poles of the system. However, if the sensor noise is small compared to the input disturbance, then the conditions for Theorem 3 can approximately be met and good identification results may still be obtained. Examples of this behavior are discussed in [4].

IV. Example

We now present a computational example to illustrate the utility of Theorem 3. We first demonstrate the transformation of a multivariable ARMAX system without transmission zeros in its MA polynomial to an equivalent ARX model. We then show how multiple output ARX identification gives greatly improved transfer function estimates as compared to single output ARX identification.

Consider a first order, two-output, single-input ARMAX system as in Eq. (3), with

$$\begin{aligned} A(z) &= \begin{bmatrix} 1-0.7z & -0.2z \\ -0.2z & 1-0.7z \end{bmatrix} \\ B(z) &= \begin{bmatrix} 1.5z \\ -0.5z \end{bmatrix} \\ C(z) &= \begin{bmatrix} 0.25+0.5z \\ 0.25-0.5z \end{bmatrix}. \end{aligned}$$

This system has no transmission zeros in the multivariable transfer function from e_t to $y_t = [y_{1t},\ y_{2t}]^T$, although there are zeros in the component transfer functions from e_t to y_{1t} and y_{2t}. Thus, there exists an exact ARX model for the two-output system, but not for either of the two scalar sub-systems.

The unimodular matrix, $L(z)$, which transforms the MA polynomial matrix, $C(z)$, to a constant matrix may be constructed by the following sequence of elementary row operations:

$$\begin{aligned} \begin{bmatrix} 0.25+0.5z \\ 0.25-0.5z \end{bmatrix} &\xrightarrow{L_1} \begin{bmatrix} 0.5 \\ 0.25-0.5z \end{bmatrix} \\ &\xrightarrow{L_2} \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix} \\ &\xrightarrow{L_3} \begin{bmatrix} 0.25 \\ 0.25 \end{bmatrix} \end{aligned}$$

where

$$L_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$
\begin{aligned}
L_2 &= \begin{bmatrix} 1 & 0 \\ z & 1 \end{bmatrix} \\
L_3 &= \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.
\end{aligned}
$$

When implementing recursive least square algorithms, it is typically assumed that the AR polynomial in Eqn. (4) has its leading coefficient given by $A_0 = I_{n_y}$. We have chosen L_3 to normalize $L(z)$, resulting in the unimodular polynomial matrix

$$
\begin{aligned}
L(z) &= L_3 L_2 L_1 \\
&= \begin{bmatrix} 1-z & -z \\ z & 1+z \end{bmatrix}.
\end{aligned}
$$

Performing the transformations as in Eqs. (10)–(12) yields the ARX model

$$
\tilde{A}(z) = \begin{bmatrix} 1 - 1.7z + 0.9z^2 & -1.2z + 0.9z^2 \\ 0.8z - 0.9z^2 & 1 + 0.3z - 0.9z^2 \end{bmatrix} \tag{13}
$$

$$
\tilde{B}(z) = \begin{bmatrix} 1.5z - z^2 \\ -0.5z + z^2 \end{bmatrix} \tag{14}
$$

$$
\tilde{C}(z) = \begin{bmatrix} 0.25 \\ 0.25 \end{bmatrix}. \tag{15}
$$

Note that the order of the equivalent ARX model has increased to two.

Next we demonstrate how the existence of an equivalent ARX model greatly affects the quality of identification results. The above ARMAX system was simulated with the exogenous input, u_t, and the noise, e_t, both consisting of unity Gaussian white noise. A total of 1500 data points were generated. The results from three sets of identification experiments using recursive least squares are discussed here:

1. Separate identification of y_{1t} and y_{2t}, using fourth-order scalar ARX models. The transfer function estimates are shown in Figure 1. The estimate for y_2 is quite poor; the estimate for y_1 is better because of the higher signal-to-noise ratio.

2. Joint identification of y_t using a first-order matrix ARX model. Figure 2 shows the transfer function estimates; the fit is improved over case 1, but the y_2/u transfer estimate has a 2–3 dB error over most of the frequency range.

3. Joint identification using a second-order matrix ARX model. The transfer function estimates are shown in Figure 3; the fit is seen to be excellent across the entire the frequency range.

	True System	Estimates
poles	0.9, 0.5	0.9043, 0.4576, -0.1537, 0.1960
zeros y_1/u	0.7667	0.7625, -0.1543, 0.1953
zeros y_2/u	1.3	1.2910, -0.1548, 0.1950

Table 1: Estimated poles and zeros for second-order matrix ARX model from example.

These results show that considerable improvement over scalar ARX model transfer function estimates can be achieved by the addition of a sensor output and the proper extension of the matrix ARX model order. The results of cases 1 and 2 indicate that neither the additional output measurement nor the increase in model order individually permit accurate identification.

We now examine the poles and zeros of the ARX parameters from case 3. To conform with the standard convention that the stability region for discrete time systems is the unit disk of the complex plane, we shall express the estimated model in terms of the forward shift operator, $q = z^{-1}$. Thus, $\hat{A}(q) = q^2\hat{A}(z)$ and $\hat{B}(q) = q^2\hat{B}(z)$, where " ˆ " indicates estimated quantities. The MIMO transfer function estimate $\hat{G}(q) = \hat{A}(q)^{-1}\hat{B}(q)$ consists of two, strictly proper, fourth order transfer functions. The poles and zeros of the true system and those from the case 3 ARX estimates are shown in Table 1.

The estimates closely match the true system poles and zeros, although the lower frequency pole at 0.5 is not as well estimated. More interesting are the estimates of the extra poles and zeros corresponding to $L(z)$. There is practically exact cancellation of these poles and zeros, thus permitting easy model reduction to a minimal representation of $\hat{G}(q)$. However, our experience is that much longer data records may be required to obtain accurate pole-zero estimates than accurate ARX transfer function estimates.

V. Conclusions

In this paper we have shown that for multi-output systems it is possible to *exactly* model an ARMAX system by an ARX model if the multivariable transfer function from the noise input port to the output port has no transmission zeros. This does not mean that the noise transfer function has no element zeros, or that there are no transmission zeros in the transfer function from the known input to the output. This is an important point, since zero order hold sampling introduces element zeros. Transmission zeros can often be eliminated by the addition of sensors. Dramatic improvements in identification results can be obtained in this manner.

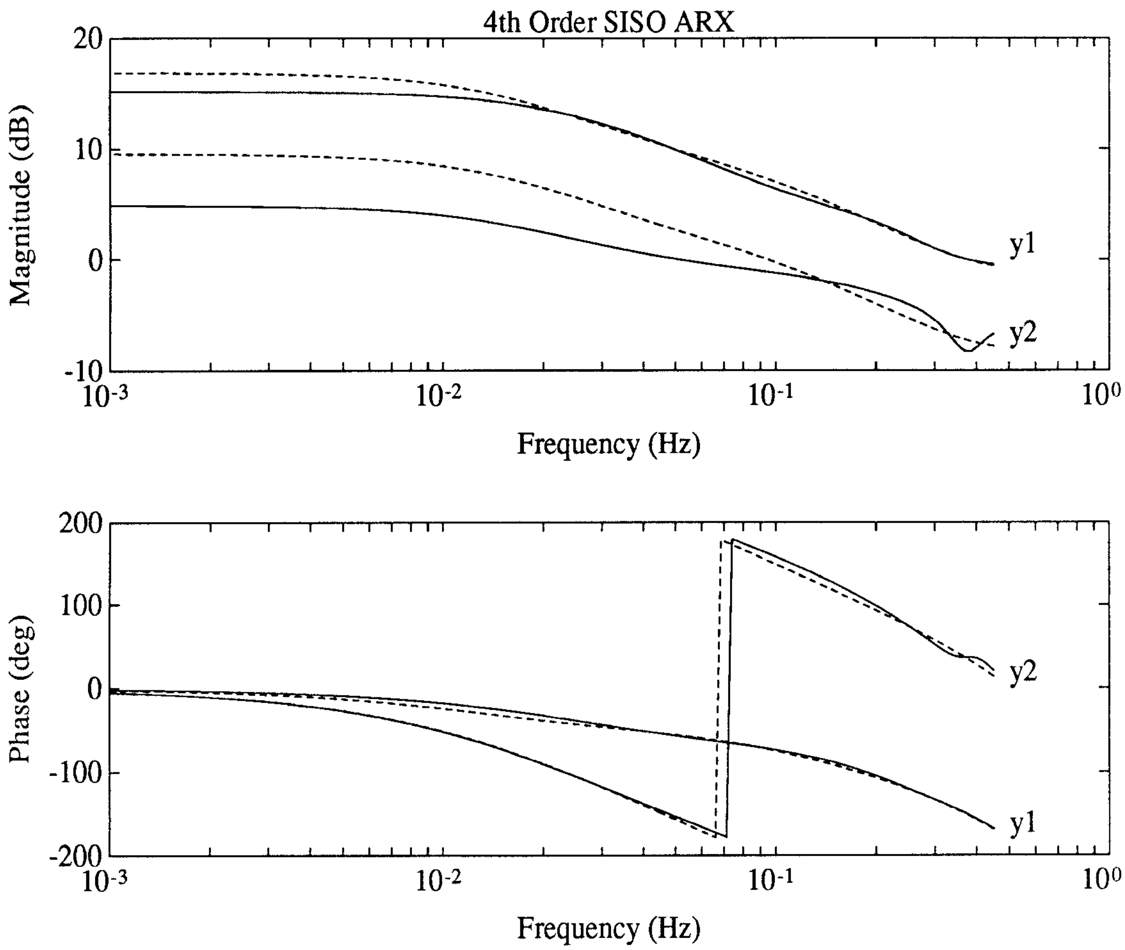

Figure 1: Fourth-order scalar ARX transfer function estimates for example. Solid lines: estimates; Dashed lines: true system.

In practice, the true system is at best partially known, therefore the ARMAX-to-ARX transformation cannot be determined. However, if a nominal analytical model is available, then Theorem 3 provides an upper bound on the ARX model order. ARX models of a range of model orders can be identified, and a model order criterion used to determine the best model. This procedure was used in [1].

Transmission zeros in the MA model are affected by sensor placement and noise. In [4], examination of the noise model transmission zeros is used as a tool to analyze these system identification design issues. In practice, the existence of an equivalent MIMO ARX model will only be approximate, due to the presence of sensor noise. However, the location of the transmission zeros in the noise transfer function can indicate whether it may be possible to use linear least squares algorithms. If these zeros are located at frequencies remote from the system dynamics, then an ARX

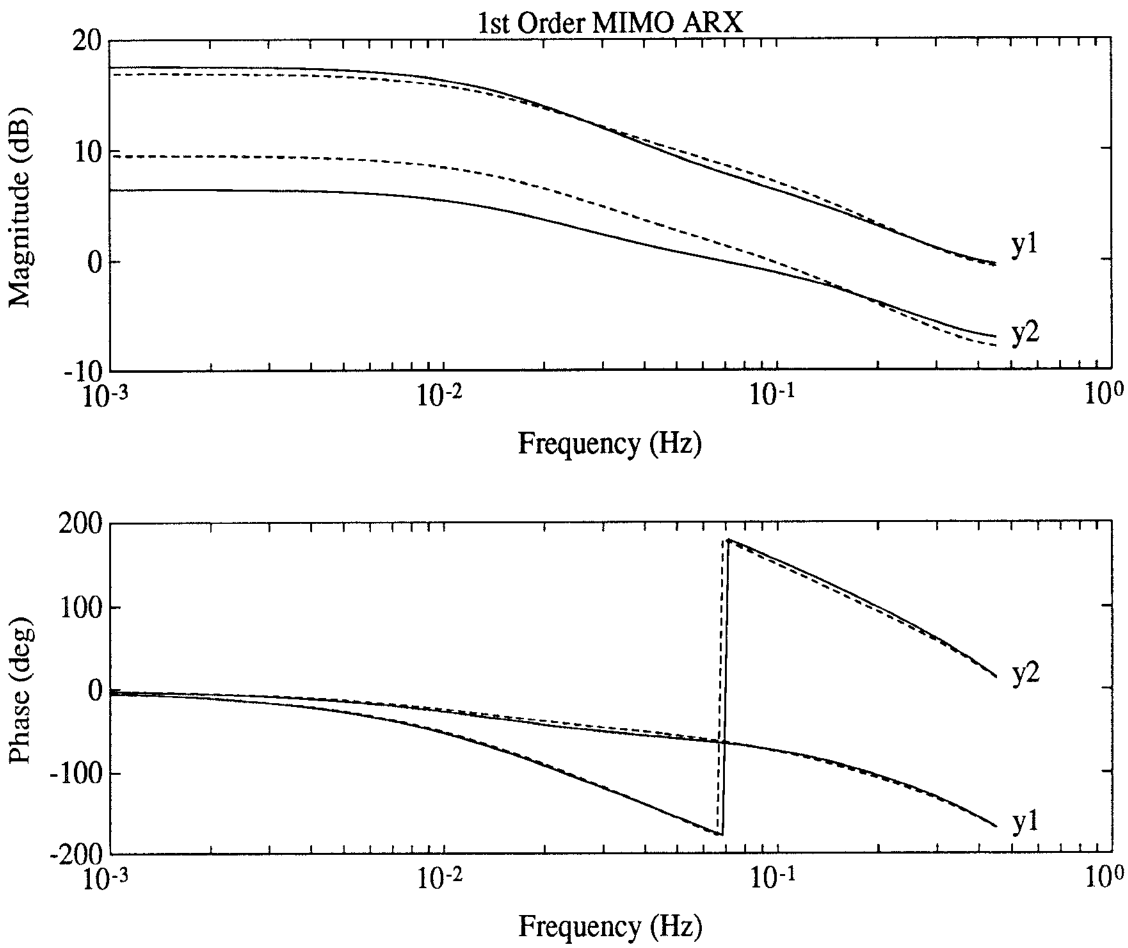

Figure 2: First-order matrix ARX transfer function estimates for example. Solid lines: estimates; Dashed lines: true system.

model may be quite adequate.

References

[1] J. T. Gillis, G. N. Smit, and K. Yong. Structural identification by lattice least squares: a preliminary report. In *AAS/AIAA Astrodynamics Specialist Conference*, Vail, CO, August 12-15, 1985.

[2] T. Kailath. *Linear Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1980.

[3] L. Ljung. *System Identification: Theory for the User.* Englewood Cliffs, NJ: Prentice-Hall, 1987.

[4] G. A. McGraw. *Robust Adaptive Control Design Techniques.* PhD thesis, UCLA, 1989.

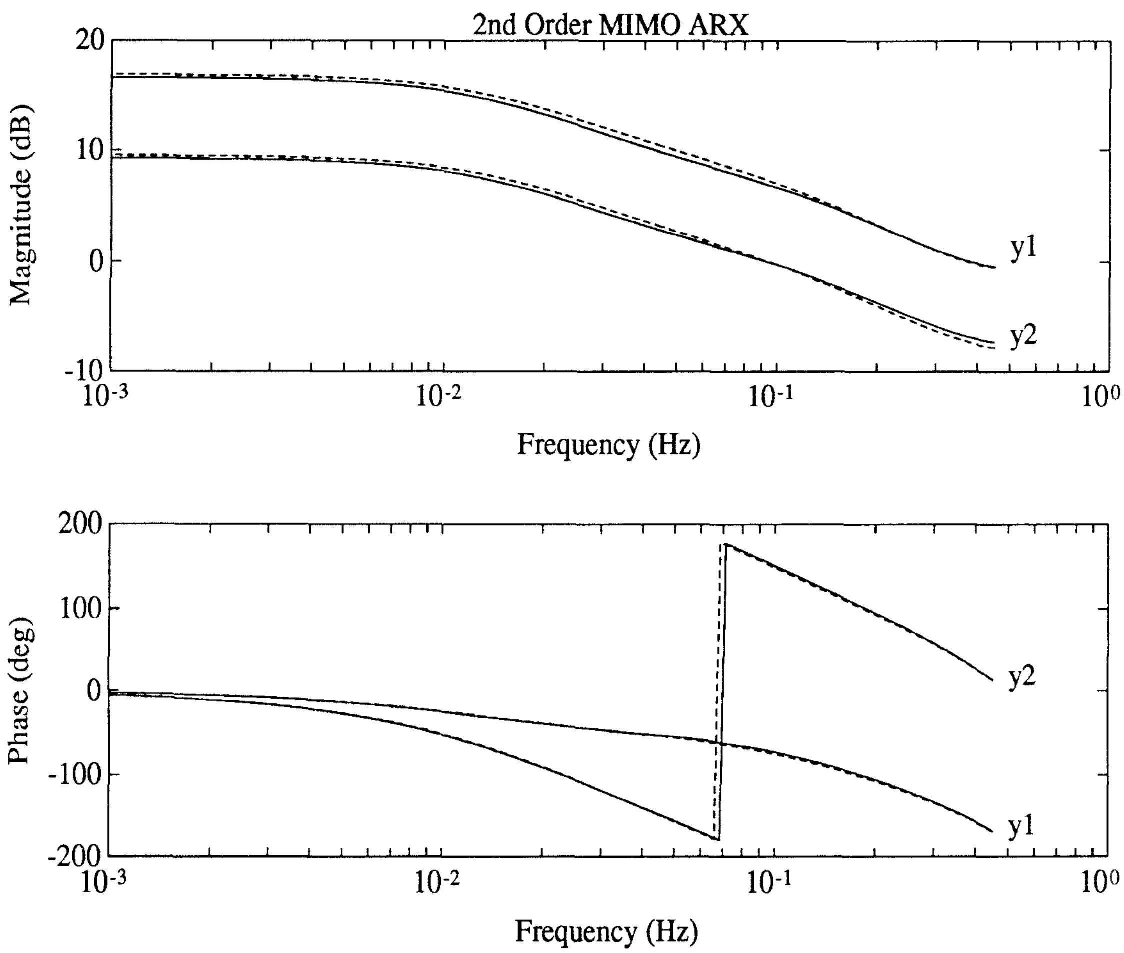

Figure 3: Second-order matrix ARX transfer function estimates for example. Solid lines: estimates; Dashed lines: true system.