

Nonlinear System Identification in a Noisy Environment using Wavelet based SDP Models

Nguyen-Vu Truong Liuping Wang

*School of Electrical and Computer Engineering, RMIT University,
Melbourne, Australia
e-mail: s3112475@ student.rmit.edu.au, liuping.wang@rmit.edu.au*

Abstract: This paper presents an approach to the identification of nonlinear system in noisy environment using a wavelet based *State Dependent Parameter* (SDP) model to characterize the system's nonlinear dynamics. The obtained model is in the form of a set of linear regressive output/input terms (state) multiplied by the respective SDPs, which are compactly parameterized by wavelet basis functions. In this approach, a modified Instrumental Variable (IV) algorithm is used to solve to the inconsistency problem of linear least squares parameter's estimation in the presence of noise. A simulation example is provided to illustrate the proposed approach.

1. INTRODUCTION

Our recent publications (i.e. Truong *et al.* 2006, 2007a, 2007b) have presented approaches to the identification of nonlinear systems using wavelet based *State Dependent Parameter* (SDP) models. This model structure expresses the nonlinear system as a set of the linear regressive output/input terms (states) multiplied by associated *State Dependent Parameters* to characterize the nonlinearities. These state dependencies, in the first step, are non-parametrically estimated using a SDP algorithm based on recursive fixed interval smoothing (i.e. Young 2001, Young *et al.* 2001, etc). The shapes of the SDP relationships (as defined by the plots of the parameters against the state variables) indicate and visualize the nature of the most significant nonlinearities within the dynamic SDP model. They are then, in the second step, parameterized in a compact manner via wavelet series expansion by employing appropriate types of wavelet basis functions that are selected corresponding to the features of the SDP relationships. This formulates the wavelet based SDP model (WSDP).

In a noisy environment, the parameter's estimates obtained by a standard linear Least Squares (LS) algorithm are biased away from their true values. It is because the regressor matrix contains the process output terms which are correlated with noise. In this situation, to obtain the consistency in the parameter's estimates, other estimation solutions are necessary. One of the simplest approaches to this problem is to use *Instrumental Variable* (IV) methods since they do not require *a priori* knowledge about the additive noise statistical properties, and have been proven to be so effective in the linear model estimation context (i.e. Young 1970, Stoica *et al.* 1983, Young 1984, Söderström *et al.* 1989)

The main idea of a basic IV method is to replace the regressor vector (associated to the noisy process output terms) with the instrumental variable. This variable is chosen in such a way that it correlates with the regression variables

but uncorrelated with the noise variable to represent the noise-free output. However, since the process noise-free output can not be measured in practice, traditionally, the instrument is generated by filtering the process input through an "*auxiliary model*"¹ which is typically obtained by the LS method.

In a nonlinear system identification context as considered in present study, as the regressor matrix is constructed from nonlinear functions of the past sample values of the output data and the input data, the instrument can not be easily obtained by just using the process input. In this paper, to counteract the bias problem in a WSDP setting, a modified Instrumental Variable (MIV) procedure is employed. In this approach, a predicted output is used to be an instrument to replace the noisy output in the regressor matrix (Kalafatis *et al.* 1997). This procedure is implemented iteratively to gradually remove the noise from the predicted output, thus removes the bias from the parameter estimates.

The paper's structure is organized as follows. Background information about WSDP models is provided in Section 2. The modified instrumental variable procedure is described in Section 3. An illustrative example is provided in Section 4 to demonstrate the merit of the proposed approach. Finally, Section 5 concludes the papers.

2. BACKGROUND

It is assumed that a nonlinear system can be represented by following *State Dependent Parameter* (SDP) model:

¹ In an iterative IV method, this "*auxiliary model*" can be later iteratively adapted (Young 1970).

$$y(k) = \sum_{i=1}^{n_y} f_i\{y(k-i)\}y(k-i) + \sum_{i=0}^{n_u} g_i\{u(k-i)\}u(k-i) + e(k) \quad (1)$$

where, $u(k)$ and $y(k)$ are, respectively, the sampled input-output sequences; and n_u, n_y refer to the maximum number of lagged inputs and outputs. The error term $e(k)$ refers to the noise variable, assumed initially to be a zero mean, white noise process that is uncorrelated with the input $u(k)$ and its past values. Here, the parameters $f_i\{\cdot\}$ and $g_i\{\cdot\}$ are dependent on the non-minimal state variables defined by the input and output variables and their past sampled values. In the open literature (i.e. Young 2001, Young *et al.* 2001, etc) they are regarded as State Dependent Parameters (SDP) to carry the system's nonlinearities.

At this point, the nonlinear system identification problem is equivalent to a 2-stage-identification procedure: (1) SDPs' non-parametric estimation, (2) Nonlinear system model's structure determination and parameter's estimation.

2.1 SDP's Nonparametric Estimation

The estimate of the state dependency is based on considering the changes of the SDPs in a transformed space, where the data are re-ordered in a non-temporal manner (normally the simple ascending-order). In this transformed space, the SDPs are assumed to vary in a stochastic manner, according to a specified member of the *Generalised Random Walk* (GRW) family, such as the random walk or the integrated random walk. They are then estimated in a recursive approach that exploits the *Fixed Interval Smoothing* (FIS) algorithm, where the hyper-parameters associated with the stochastic models for the parameter variations are estimated using *Prediction Error Decomposition* (PED), as well as associated *Maximum likelihood* (ML) algorithm (see Young 1993, 2001, Young *et al.* 2001 for more details). The final results of this process are in the form of non-parametric relationships (graphs) between the SDP estimates and the states on which they are dependent. The features of these nonparametric relationships serve as the basis for the selection of wavelet functions as well as the associated scaling parameters for the subsequent nonlinear model's structure selection process.

2.2 Model's Structure Determination and Parameter's Estimation

Using wavelet series expansion, a SDP relationship f can be parameterized with respect to the state variable x as below:

$$f(x) = \sum_{i=i_{\min}}^{i_{\max}} \sum_{k \in L_i} a_{i,k} \Psi(2^{-i}x - k) + \xi(x) \quad (2)$$

where $\Psi(x)$ is any compactly supported mother wavelet function (e.g. a Mexican hat wavelet, Morlet wavelet, and so on), whose supporting range falls within (s_1, s_2) .

Here, i_{\min}, i_{\max} , respectively, refer to the minimum and maximum (finest and coarsest) scales (resolutions) to be employed for approximation. L_i (determined as in (3) and (4)) is the translation library at scale i .

$$L_i = \{k \in (2^{-i}x_{\min} - s_2, 2^{-i}x_{\max} - s_1), k \in \mathbb{Z}\} \quad (3)$$

$$L_{i_{\max}} \subset L_{i_{\max}-1} \subset \dots \subset L_{i_{\min}}. \quad (4)$$

The realization of $f(x)$ as in (2) is often overparameterized since the wavelet function library as derived in (3) and (4) consists all the possible combination of the parameters. In order to obtain a compact parameterization for the SDP relationship under study, a PRESS statistic based term selection algorithm is implemented (Truong *et al.* 2006, 2007a, 2007b). This procedure uses the incremental value of PRESS² ($\Delta PRESS$) as criterion to detect the significance of each terms within the model in which the maximum $\Delta PRESS$ signifies the most significant term, while its minimum reflects the least significant term. Based on this, the algorithm initializes with the initial subset being the most significant term. It then starts to grow to include the subsequent significant terms in a forward regression manner, until a specified performance is achieved.

By doing so, the State Dependent Parameters $f_i(x)$ and $g_i(x)$ can be compactly parameterized in the following forms

$$f_i(x) = \sum_{j=0}^{n_{f_i}} a_{f_i,j} l_{f_i,j}(x) \\ g_i(x) = \sum_{j=0}^{n_{g_i}} a_{g_i,j} l_{g_i,j}(x) \quad (5)$$

where, $L_{f_i} = \{l_{f_i,0}, \dots, l_{f_i,n_{f_i}}\}$, $L_{g_i} = \{l_{g_i,0}, \dots, l_{g_i,n_{g_i}}\}$ are, respectively, the optimized sets of wavelet functions used for parameterization of $f_i(x)$ and $g_i(x)$. Substituting (5) into (1) yields

$$y(k) = \sum_{i=1}^{n_y} \sum_{j=0}^{n_{f_i}} [a_{f_i,j} l_{f_i,j}\{y(k-i)\}]y(k-i) + \sum_{i=0}^{n_u} \sum_{j=0}^{n_{g_i}} [a_{g_i,j} l_{g_i,j}\{u(k-i)\}]u(k-i) + e(k) \quad (6)$$

Equation (6) is regarded as a wavelet based SDP model (WSDP).

By letting

² The difference between the overparameterized (original) model's PRESS value and the one calculated by excluding a term from the original model.

$$\begin{aligned}
 \theta_{fi} &= [a_{fi,0}, \dots, a_{fi,n_{fi}}]^T \\
 \theta_{gi} &= [a_{gi,0}, \dots, a_{gi,n_{gi}}]^T \\
 L_{k,fi} &= y(k-i)[l_{fi,0}\{y(k-i)\}, \dots, l_{fi,n_{fi}}\{y(k-i)\}] \\
 L_{k,gi} &= u(k-i)[l_{gi,0}\{u(k-i)\}, \dots, l_{gi,n_{gi}}\{u(k-i)\}] \\
 \theta &= [\theta_{f1}, \dots, \theta_{fn_y}, \theta_{g0}, \dots, \theta_{gn_u}]^T \\
 L_k &= [L_{kf1}, \dots, L_{kfn_y}, L_{kg0}, \dots, L_{kgn_u}]^T \\
 L &= [L_0, \dots, L_{N-1}] \\
 E &= [e(0), \dots, e(N-1)]^T \\
 Y &= [y(0), \dots, y(N-1)]^T
 \end{aligned} \tag{7}$$

Equation (6) is now written in the following matrix form:

$$Y = L^T \theta + E \tag{8}$$

which is a standard least squares formulation.

Let us define the cost function as below

$$J = [Y - L^T \theta]^T [Y - L^T \theta] \tag{9}$$

and as usual, the estimate $\hat{\theta}$ of the parameter vector θ to minimize the cost function J is obtained in the usual least squares manner, i.e.

$$\hat{\theta} = [LL^T]^{-1} [LY] \tag{10}$$

3. MAIN RESULTS

3.1 Sensitivity of the Least Squares Solution to Noise

In the presence of noise, the parameter's estimate as obtained in (10) will be biased since the regressor matrix contains the process output terms which are correlated with noise. However, this bias is dependent upon the signal to noise ratio (SNR). If SNR is high, the bias in the parameter's estimate can be significantly reduced, and vice versa. To demonstrate that, let us consider the following example.

Consider a nonlinear system described by the following equation:

$$\begin{aligned}
 y(k) &= [-0.5\Psi_{1,3}(x) + 0.45\Psi_{0,1}(x)]_{y(k-1)} y(k-1) \\
 &+ [0.25\Psi_{2,1}(x) - 0.65\Psi_{1,0}(x)]_{u(k)} u(k)
 \end{aligned} \tag{11}$$

where,

$$\Psi_{i,k}(x) = \Psi(2^{-i}x - k), \text{ and } \Psi(x) = (1 - x^2)e^{-0.5x^2}$$

With $u(k) = \sin(\frac{k}{50})$ and the initial condition to be $y(0) = 0$, (11) is simulated to generate 1000 data points (Figure 1).

To simulate the output measurement noise, a zero mean, white noise sequence is added to the noise-free output signal. In this case, the output signal is redefined as below:

$$y(k) = \bar{y}(k) + \vartheta(k) \quad \vartheta(k) = N(0, \sigma^2) \tag{12}$$

in which $\bar{y}(k)$ denotes the noise-free signal.

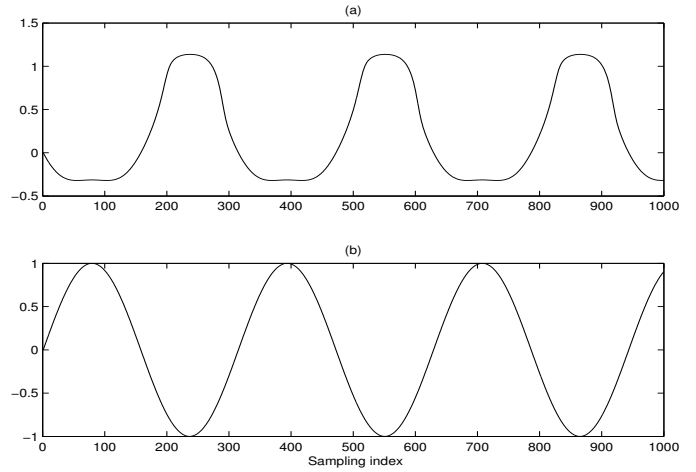


Fig. 1. Benchmark example: (a) noise-free output (b) input

For this example, to investigate various noise levels added to the output signal, σ is selected to be several values (0.0057, 0.0284 and 0.0398) to respectively add 1%, 5% and 7% noise (by standard deviation) to the noise-free signal $\bar{y}(k)$.

The parameter's estimates are shown in Table 1 in comparison to the true values. To quantify the bias, the following measure (Normalized Squared Error-SE) is introduced:

$$SE = \frac{\|\hat{\theta} - \theta_0\|}{\theta_0} \tag{13}$$

in which θ_0 denotes the true parameters. As shown on

	True values	LS estimates		
		σ		
		0.0057	0.0284	0.0398
$a_{1,3,f1}$	-0.5	-0.4915	-0.3939	-0.3313
$a_{0,1,f1}$	0.45	0.4506	0.4580	0.4598
$a_{2,1,g0}$	0.25	0.2518	0.2795	0.3050
$a_{1,0,g0}$	-0.65	-0.6516	-0.6681	-0.6809
SE		0.92%	11.56%	18.63%

Table 1. Sensitivity of the Least Squares Estimates to Noise

Table 1, the bias increases as SNR decreases. When $\sigma = 0.0057$ (1% of noise added by standard deviation), the bias is insignificant. The Least Squares (LS) estimates are very close to their true values. When the noise level increases, the LS estimates are significantly biased away from the true values.

3.2 Modified Instrumental Variable Parameter Estimation

Assuming that the true system can be represented as:

$$y(k) = \bar{L}_k^T \theta_0 + \vartheta(k) \tag{14}$$

where θ_0 represents the true estimate of θ , and \bar{L}_k^T represents the noise free regressor vector which is formulated from the noise free output $\bar{y}(k)$, and the noise free input $u(k)$ as described in (7), i.e.

$$\bar{L}_k = \begin{bmatrix} y(k-1)[l_{f1,0}\{y(k-1)\}, \dots \\ \dots, l_{fi,n_{f1}}\{y(k-1)\}]^T \\ \vdots \\ y(k-n_y)[l_{fn_y,0}\{y(k-n_y)\}, \dots \\ \dots, l_{fn_y,n_{fn_y}}\{y(k-n_y)\}]^T \\ u(k)[l_{g0,0}\{u(k)\}, \dots \\ \dots, l_{g0,n_{g0}}\{u(k)\}]^T \\ \vdots \\ u(k-n_u)[l_{gnu,0}\{u(k-n_u)\}, \dots \\ \dots, l_{gnu,n_{gnu}}\{u(k-n_u)\}]^T \end{bmatrix} \quad (15)$$

As

$$\hat{\theta} = [LL^T]^{-1}[LY] = [LL^T]^{-1}[L(\bar{L}^T\theta_0 + V)] \quad (16)$$

Thus, we can express the difference between $\hat{\theta}$ and θ_0 as below

$$\begin{aligned} \hat{\theta} - \theta_0 &= [LL^T]^{-1}[L(\bar{L}^T\theta_0 + V)] - \theta_0 \\ &= \{[LL^T]^{-1}L\bar{L}^T - 1\}\theta_0 + [LL^T]^{-1}LV \end{aligned} \quad (17)$$

in which,

$$V = [\vartheta(0), \dots, \vartheta(N-1)]^T \quad (18)$$

As some elements in L are terms such as $y(k-1)[l_{f1,0}\{y(k-1)\}, \dots, l_{fi,n_{f1}}\{y(k-1)\}, \dots, y(k-i)[l_{fi,0}\{y(k-i)\}, \dots, l_{fi,n_{fi}}\{y(k-i)\}, \dots]$, which contains the noise terms, i.e. $\vartheta(k-1), \dots, \vartheta(k-i), \dots$, the least squares estimate of θ will be biased. It is because LV does not tend to zero even if $\vartheta(k)$ is a zero mean white noise sequence.

The bias as in (17) will be disappeared if we substitute the noisy process output $y(k)$ with the noise free output $\hat{y}(k)$ in the L matrix (constructed as in (7)). However, as this information is unavailable in practice, its prediction $\hat{y}(k)$ is proposed to be used (Kalafatis *et al.* 1997). This procedure is proposed to be implemented iteratively to gradually remove the bias from the estimates as described below:

1. Form the data matrix $L_{Y,U}^T$ by using the process input $u(k)$ and the measured output $y(k)$ as in (7), and obtain the initial estimate for the parameters as:

$$\hat{\theta}_0 = [L_{Y,U}L_{Y,U}^T]^{-1}[L_{Y,U}Y] \quad (19)$$

2. Generate the predicted output \hat{Y}_{LS}^0 as

$$\hat{Y}_{LS}^0 = L_{Y,U}^T\hat{\theta}_0$$

For the ease of representation, let us denote $L_{\hat{Y},U}^T$ be the data matrix that is formulated as in (7) by using $\hat{y}(k)$ and $u(k)$ instead of the actual noise corrupted output $y(k)$, and $u(k)$, i.e.

$$L_{\hat{Y},U} = [\hat{L}_0, \dots, \hat{L}_{N-1}] \quad (20)$$

in which,

$$\hat{L}_k = \begin{bmatrix} \hat{y}(k-1)[l_{f1,0}\{\hat{y}(k-1)\}, \dots \\ \dots, l_{fi,n_{f1}}\{\hat{y}(k-1)\}]^T \\ \vdots \\ \hat{y}(k-n_y)[l_{fn_y,0}\{\hat{y}(k-n_y)\}, \dots \\ \dots, l_{fn_y,n_{fn_y}}\{\hat{y}(k-n_y)\}]^T \\ u(k)[l_{g0,0}\{u(k)\}, \dots \\ \dots, l_{g0,n_{g0}}\{u(k)\}]^T \\ \vdots \\ u(k-n_u)[l_{gnu,0}\{u(k-n_u)\}, \dots \\ \dots, l_{gnu,n_{gnu}}\{u(k-n_u)\}]^T \end{bmatrix} \quad (21)$$

3. Set $\hat{Y} = \hat{Y}_{LS}^0$.

4. Form a new data matrix $L_{\hat{Y},U}^T$ by using $\hat{y}(k)$, and $u(k)$, and compute $\hat{\theta}$ as

$$\hat{\theta} = [L_{\hat{Y},U}L_{\hat{Y},U}^T]^{-1}[L_{\hat{Y},U}Y] \quad (22)$$

5. Generate the predicted output \hat{Y} as

$$\hat{Y} = L_{\hat{Y},U}^T\hat{\theta} \quad (23)$$

6. Repeat Step 4 through 5 until the parameter estimates converge.

Upon the convergence of $L_{\hat{Y},U}$ to \bar{L} , based on (17), it can be shown that $\hat{\theta}$ is unbiased estimate of θ_0 . The proof of convergence is, however, still under study and remains an open question.

This approach is, on one hand, closely related to the Instrumental Variable (IV) method (i.e. Young 1970, Stoica *et al.* 1983, Young 1984, Söderström *et al.* 1989, etc) in the sense that the predicted output $\hat{y}(k)$ is used as an instrument for $y(k)$ in the regressor matrix. On the other hand, this differs from the standard IV approach since the estimate is actually based on the following model:

$$Y = L_{\hat{Y},U}^T\theta + V \quad (24)$$

which leads to the estimator

$$\hat{\theta} = [L_{\hat{Y},U}L_{\hat{Y},U}^T]^{-1}[L_{\hat{Y},U}Y] \quad (25)$$

while a standard IV estimator is as follows

$$\hat{\theta} = [L_{\hat{Y},U}L_{Y,U}^T]^{-1}[L_{\hat{Y},U}Y] \quad (26)$$

As a result, it might be regarded as a pseudo-instrumental variable (PIV) method.

4. SIMULATION EXAMPLE

To demonstrate the merit of the proposed approach, in this section, a simulation example is provided. For simplicity, throughout this section, a form of Mexican hat wavelet function as defined in (27) is used.

$$\Psi(x) = \begin{cases} (1-x^2)e^{-0.5x^2} & \text{if } x \in (-4, 4) \\ 0 & \text{otherwise} \end{cases} \quad (27)$$

Consider a nonlinear system described by the following model:

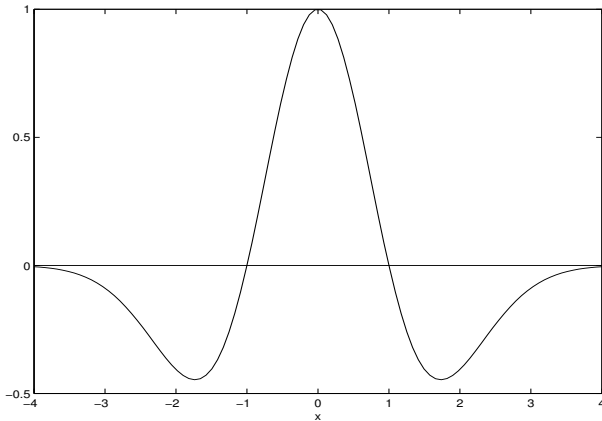


Fig. 2. Mexican hat wavelet function

$$y(k) = \left[0.5 + 0.2e^{-0.5y(k-1)^2} \right] y(k-1) + \frac{y(k-2)^2}{1 + y(k-2)^2} + [u(k)^2 - 0.7u(k) + 0.3] u(k) \quad (28)$$

With $u = \sin(\frac{k}{50})$ and the initial conditions to be $y(0) = 0$, $y(1) = 0.0057$, model (28) is simulated to generate 1000 input/output data points. In this example, in order to evaluate the effect of measurement noise, it is assumed that the output is noisy, in the sense that a zero mean, white noise sequence is added to the noise-free output signal (Figure 3). In this situation the measured output $y(k)$ is redefined as:

$$y(k) = \bar{y}(k) + \vartheta(k) \quad \vartheta(k) = N(0, \sigma^2) \quad (29)$$

where $\sigma^2 = 0.0844$ is selected to add 5% noise (by standard deviation) to $\bar{y}(k)$, which now denotes the noise-free output sequence.

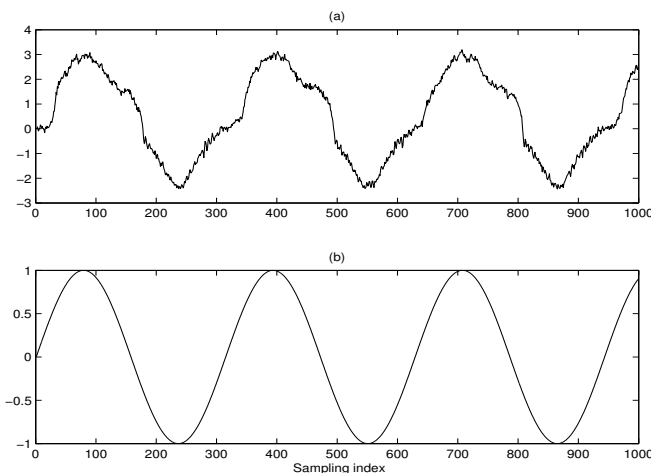


Fig. 3. Simulation data: (a) Noisy output (b) input.

Using a discrete time model form for this system, its SDP model structure is identified as follows:

$$y(k) = f_1\{y(k-1)\}y(k-1) + f_2\{y(k-2)\}y(k-2) + g_0\{u(k)\}u(k) \quad (30)$$

With the finest and coarsest scaling parameters chosen to be 0 and 3, the SDP parameters are identified in the following general parametric forms:

$$\begin{aligned} f_1(x) &= a_{0,0,f1}\Psi_{0,0}(x) + a_{0,1,f1}\Psi_{0,1}(x) \\ &+ a_{0,3,f1}\Psi_{0,3}(x) + a_{0,2,f1}\Psi_{0,2}(x) \\ &+ a_{3,0,f1}\Psi_{3,0}(x) + a_{1,-1,f1}\Psi_{1,-1}(x) \\ &+ a_{1,2,f1}\Psi_{1,2}(x) \\ f_2(x) &= a_{0,2,f2}\Psi_{0,2}(x) + a_{2,1,f2}\Psi_{2,1}(x) \\ g_1(x) &= a_{0,-1,g0}\Psi_{0,-1}(x) + a_{0,1,g0}\Psi_{0,1}(x) \\ &+ a_{3,0,g0}\Psi_{3,0}(x) \end{aligned} \quad (31)$$

where,

$$\Psi_{i,k}(x) = \Psi(2^{-i}x - k), \text{ and } \Psi(x) = (1 - x^2)e^{-0.5x^2}$$

The estimation model is then obtained by substituting (31) into (30). Using the input-output data, the associated parameters are estimated using the proposed modified instrumental variable (MIV) algorithm. After 25 iterations, the MIV estimates' convergence is achieved. The predicted output $\hat{y}(k)$ converges to the noise free output $\bar{y}(k)$ (Figure 4) and the final identified model is found to be: $y(k) =$

$$\begin{aligned} &\left[\begin{array}{l} 0.3943\Psi_{0,0}(x) + 0.6952\Psi_{0,1}(x) \\ + 0.1163\Psi_{0,2}(x) + 0.2874\Psi_{0,3}(x) \\ + 0.4828\Psi_{1,-1}(x) + 0.0400\Psi_{1,2}(x) \\ + 0.4668\Psi_{3,0}(x) \end{array} \right]_{y(k-1)} y(k-1) \\ &+ [0.2900\Psi_{0,2}(x) + 0.5785\Psi_{2,1}(x)]_{y(k-2)} y(k-2) \\ &+ \left[\begin{array}{l} 1.4036\Psi_{0,-1}(x) + 0.9998\Psi_{0,1}(x) \\ + 0.2010\Psi_{3,0}(x) \end{array} \right]_{u(k)} u(k) \end{aligned} \quad (32)$$

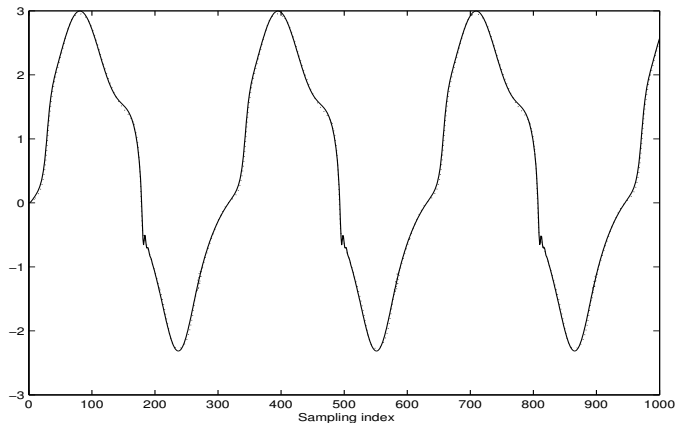


Fig. 4. Comparison between noise-free (solid) and predicted (dot-dot) process output after 25 iterations

The obtained MIV parameter estimates along with the noise-free estimates are given in Table 2, showing that they are very close to each other. Figure 5 compares the identified model's iterative (simulated) output to the noise-free output signal. They, in turn, imply that the identified model (32) excellently characterizes this system.

To further validate the simulation results, 100 independent Monte Carlo simulation tests (1000 data points each) are generated based on (28) and (29). Here, to quantify the

Parameter	LS estimates	Noise-free estimates	MIV estimates
$a_{0,0,f1}$	0.3012	0.3821	0.3943
$a_{0,1,f1}$	0.6173	0.6548	0.6952
$a_{0,2,f1}$	0.2813	0.0871	0.1163
$a_{0,3,f1}$	0.2716	0.2118	0.2874
$a_{1,-1,f1}$	0.2110	0.4788	0.4828
$a_{1,2,f1}$	0.2028	0.0735	0.0400
$a_{3,0,f1}$	0.6194	0.4908	0.4668
$a_{0,2,f2}$	0.0729	0.2784	0.2900
$a_{2,1,f2}$	0.2541	0.5834	0.5785
$a_{0,-1,g0}$	1.3560	1.4108	1.4036
$a_{0,1,g0}$	0.9656	1.0764	0.9998
$a_{3,0,g0}$	0.1624	0.1795	0.2010

Table 2. Noise-free versus MIV estimates

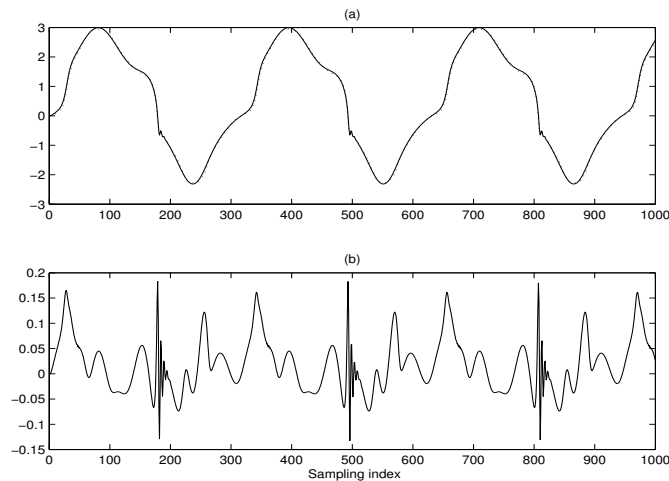


Fig. 5. (a) Comparison between the noise-free output (solid) and model iterative output of (32) (dot-dot), and (b) their difference

Parameter	Noise-free estimates	MIV estimates
$a_{0,0,f1}$	0.3821	0.3849±0.0359
$a_{0,1,f1}$	0.6548	0.7081±0.0631
$a_{0,2,f1}$	0.0871	0.0900±0.0677
$a_{0,3,f1}$	0.2118	0.3039±0.0773
$a_{1,-1,f1}$	0.4788	0.4286±0.0864
$a_{1,2,f1}$	0.0735	0.0431±0.0892
$a_{3,0,f1}$	0.4908	0.4766±0.0701
$a_{0,2,f2}$	0.2784	0.2975±0.0604
$a_{2,1,f2}$	0.5834	0.5107±0.1189
$a_{0,-1,g0}$	1.4108	1.4200±0.1460
$a_{0,1,g0}$	1.0764	1.0587±0.1200
$a_{3,0,g0}$	0.1795	0.2379±0.0478
	<i>RE</i>	7.2%
	<i>MSE</i>	2.46%

Table 3. Monte Carlo test performance

simulation results, a number of performance measures (Relative Error-**RE** and Normalized Mean Squared Error-**MSE** of $\hat{\theta}_m$ with respect to the noise-free estimated value θ_0) are introduced to assess the estimation accuracy (shown in Table 3) as follows:

$$RE = \frac{\|m(\hat{\theta}) - \theta_0\|}{\|\theta_0\|} \quad (33)$$

$$MSE = \frac{1}{M} \sum_{m=1}^M \frac{\|\hat{\theta}_m - \theta_0\|}{\|\theta_0\|} \quad (34)$$

in which, $\hat{\theta}_m$ denotes the MIV parameter estimates in the m^{th} test over the total of $M = 100$ independent tests; $m(\hat{\theta}) = \frac{1}{M} \sum_{m=1}^M \hat{\theta}_m$.

5. CONCLUSION

This paper has presented a nonlinear system identification algorithm which uses a wavelet based SDP (WSDP) model to identify and characterize the system's nonlinear dynamics. In a noisy environment, to obtain the consistent estimates of the model's parameters, a modified instrumental variable (MIV) procedure is employed. This procedure uses the predicted value as an instrument to substitute for the noise disturbed process output in the regressor matrix, and gradually removes the bias in the parameter estimates in an iterative manner. The results obtained from the simulation example demonstrate the merits of the proposed approach.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge helpful discussion with Prof. Peter C. Young (University of Lancaster, UK) who visited RMIT in 2005.

REFERENCES

A.D. Kalafatis, L. Wang and W.R. Cluett. Identification of Weiner type nonlinear systems in noisy environment. *International Journal of Control*, 66:923-941, 1997.

P.Stoica, and T. Söderström. Optimal instrumental variable estimation and approximate implementation. *IEEE Trans on Automatic control*, AC-28:757-772,1983.

T. Söderström and P. Stoica. *System Identification*. Prentice Hall, New York,1989.

N.V. Truong, L. Wang and P.C. Young. Nonlinear system modeling based on nonparametric identification and linear wavelet estimation of SDP models. *Proc. 45th IEEE Conf. on Decision and Control*, San Diego, USA, pp. 2523-2528, 2006.

N.V. Truong, L. Wang and P.C. Young. Nonlinear system modeling based on nonparametric identification and linear wavelet estimation of SDP models. *International Journal of Control*, 80:774-788, 2007a.

N.V. Truong, L. Wang and J.M. Huang. Nonlinear modeling of a magnetic bearing using SDP model and linear wavelet parameterization. *Proc. 2007 American Control Conference*, New York, pp. 2254-2259, 2007b.

P.C. Young. An Instrumental variable for real time identification in a noisy process. *Automatica*, 6:271-287, 1970.

P. C. Young. *Recursive Estimation and Time Series Analysis*. Springer-Verlag: Berlin, 1984.

P.C. Young. The identification and estimation of nonlinear stochastic systems. In A.I. Mees (ed), *Nonlinear dynamics and statistics*, pages 127-166,Boston:Birkhauser, 2001.

P.C. Young, P. Mckenna, and J. Bruun. Identification of non-linear sochastic systems by state dependent parameter estimation. *International Journal of Control*, 74: 1837-1857, 2001.