IDENTIFICATION OF LINEAR SYSTEMS WITH ERRORS IN VARIABLES USING SEPARABLE NONLINEAR LEAST-SQUARES

Mats Ekman

Division of Systems and Control,
Department of Information Technology,
Uppsala University,
PO Box 337, SE-751 05, Uppsala, Sweden
E-mail: mats.ekman@it.uu.se

Abstract: It is well-known that the least-squares identification method generally gives biased parameter estimates when the observed input-output data are corrupted with noise. If the noise acting on both the input and output is white, and if the noise variances are known, or if estimates of the noise variances are available, then the principle of biased-compensated least-squares (CLS) can readily be used to obtain consistent estimates. In this paper an extended version of the CLS (ECLS) method based on an overdetermined linear system of equations is investigated. By considering the ECLS problem as a separable nonlinear LS problem, it is also shown that the noise variance parameters can be obtained from solving a variable projection minimization problem.

Keywords: Parameter estimation, Least-squares method, Noisy input-output systems, Separable nonlinear least squares

1. INTRODUCTION

Least-squares (LS) methods play a fundamental role in identification of linear systems. However, it is well known that the classical LS approach generally gives biased parameter estimates. For systems where the disturbance acts only as a white noise, the bias can be estimated and the parameter estimates can be compensated (compensated-bias least-squares, CLS) to obtain consistent estimates. For the case of white output noise, the CLS estimate can be computed by solving an eigenvalue problem, see for example (Levin, 1964), (Aoki and Yue, 1970) and (Stoica and Söderström, 1982).

Identification of systems with noise-corrupted input and output measurements (errors-in-variables) has received increasing attention recently. A comprehensive description of perspectives on errors-in-variables estimation for dynamic systems can be found in (Söderström *et al.*, 2002). Different approaches for identification of errors-in-variables models can for example be be found in (Fernando and Nicholson, 1985), (Söderström, 1981), (Söderström and Mahata, 2002) and (Mahata and Söderström, 2002). In (Beghelli *et al.*, 1990) it is suggested to use the Frisch scheme for estimating the noise variance and model parameters. Several recursive algorithms for identification of errors-in-variables models (the BELS method) based on CLS techniques, have been proposed in (Zheng and Feng, 1989), (Zheng and Feng, 1992) and (Zheng, 1999).

In this paper an off-line method based on CLS solution of an overdetermined system of equations and separable nonlinear LS, see e.g. (Hannan, 1971) and (Golub and Peryra, 1973), for estimating the parameters in dynamic errors-in-variables models is presented. The method is developed from the fact that the original nonlinear CLS optimization problem is separable. It is observed that the model parameters and the noise variance parameters in the resulting nonlinear loss function can be estimated separately. Thus, instead of using for example the Frisch scheme, see e.g. (Beghelli *et al.*, 1990) and (Soverini and Söderström, 2000), estimates of the noise variances are obtained from solving a variable projection problem. Once the noise variances are available, asymptotically unbiased estimates of the model parameters can be obtained using a CLS technique. The CLS method used in this paper will hereafter be called the extended CLS (ECLS) method.

2. EXTENDED CLS ESTIMATES

The following discrete-time, stochastic SISO system will be considered

$$x(t) = \frac{B(q^{-1})}{A(q^{-1})} u_0(t), \tag{1}$$

where $A(q^{-1})$ and $B(q^{-1})$ are polynomials of the type

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a} \\ B(q^{-1}) &= b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b}. \end{aligned} \tag{2}$$

It is presumed that $A(z)$, z being an arbitrary complex variable replacing q^{-1}, has all zeros outside the unit circle and that $A(z)$ and $B(z)$ have no common factors. Since we are investigating dynamic errors-in-variables models, we assume that the observations are corrupted by additive measurement noise. Thus, the available signals are of the form

$$\begin{aligned} y(t) &= x(t) + e(t) \\ u(t) &= u_0(t) + v(t), \end{aligned} \tag{3}$$

where $y(t)$ is the measured output and $u(t)$ is the measured input. The sequences $e(t)$ and $v(t)$ are mutually uncorrelated zero-mean white noises with unknown variances σ_y^2 and σ_u^2, respectively. It is assumed that the noise sequences are uncorrelated with the noise-free signals $x(t)$ and $u_0(t)$. Moreover, it is assumed that the order of the model is known and that $u_0(t)$ is persistently exciting of a sufficiently high order.

The following notations are introduced

$$\begin{aligned} \phi(t) &= [Y_t, U_t]^T \\ Y_t &= [-y(t-1), \ldots, -y(t-n_a)] \\ U_t &= [u(t-1), \ldots, u(t-n_b)] \\ \hat{\theta} &= [\hat{a}_1, \ldots, \hat{a}_{n_a}, \hat{a}_1, \ldots, \hat{b}_{n_b}]^T, \end{aligned}$$

where $\hat{\theta}$ is the vector of estimated model parameters. It is obvious that the true system is given by

$$y(t) = \phi(t)^T \theta_0 + w(t), \tag{4}$$

where $w(t)$ is a stochastic disturbance term and θ_0 is the vector of 'true' parameters. Assume that $z(t)$ is a vector with dimension $(n_z|1)$, where $n_z \geq n_a + n_b + 2$. Now, let us consider the following overdetermined system of equations

$$\frac{1}{N}\sum_{t=1}^{N} z(t)w(t) = \frac{1}{N}\sum_{t=1}^{N} z(t)(y(t) - \phi(t)^T\theta), \tag{5}$$

where the unknown model parameter vector, θ, is required to satisfy (5). Choosing the entries of the vector $z(t)$ as signals uncorrelated with the disturbance $w(t)$ will give rise to the well-known extended instrumental variable (IV) estimates of θ_0, see e.g. (Söderström and Stoica, 1989). However, since we are interested of estimating the noise variances as well as the model parameters, we will choose at least some of the entries in $z(t)$ correlated with $w(t)$.

Using the result from e.g. (James *et al.*, 1972), it follows that the extended CLS estimate for the overdetermined system of equations (5) is given by

$$\hat{\theta}_{ECLS} = (R_{z\phi} - S(\sigma))^{\dagger}(R_{zy} - \xi\sigma), \tag{6}$$

where $(R_{z\phi} - S(\sigma))^{\dagger}$ is the pseudo-inverse of $(R_{z\phi} - S(\sigma))$. It is assumed that $(R_{z\phi} - S(\sigma))$ has full column rank, which in general is a mild assumption. $S(\sigma)$ is a $(n_z|n_a + n_b)$ matrix function of the noise variances $\sigma = [\sigma_y^2, \sigma_u^2]^T$ and ξ is a $(n_z|2)$ matrix. Further, we have

$$R_{z\phi} = \frac{1}{N}\sum_{t=1}^{N} z(t)\phi(t)^T, \; R_{zy} = \frac{1}{N}\sum_{t=1}^{N} z(t)y(t). \tag{7}$$

The choice of the entries in the vector $z(t)$ will determine the structure of the matrices $S(\sigma)$ and ξ in (6). Choosing $z(t) = \phi(t)$ will give rise to the basic CLS method, where $\xi = 0$ and the inverse can be used instead of the pseudo-inverse in (6).

The elements of the augmented $z(t)$ vector for the extended CLS can be chosen in many ways. One special choice, which will be investigated here, is to let the entries of $z(t)$ be given by

$$z(t) = [y(t), Y_t, Y_p, u(t), U_t, U_p]^T, \tag{8}$$

where

$$\begin{aligned} Y_p &= [-y(t-n_a-1), \ldots, -y(t-n_a-p)] \\ U_p &= [u(t-n_b-1), \ldots, u(t-n_b-p)]. \end{aligned}$$

Then, we have $n_z = n_a + n_b + 2p + 2$, and it is now possible to write the structure of the noise matrices as

$$S(\sigma) = \begin{bmatrix} 0 & 0 \\ \sigma_y^2 I_{n_a} & 0 \\ 0 & 0 \\ 0 & \sigma_u^2 I_{n_b} \\ 0 & 0 \end{bmatrix} \begin{array}{l} \} \; 1 \\ \} \; n_a \\ \} \; p+1, \\ \} \; n_b \\ \} \; p \end{array} \qquad \xi = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{array}{l} \\ \} n_z - 1 \end{array}, \tag{9}$$

where I_{n_a} and I_{n_b} are $n_a \times n_a$ and $n_b \times n_b$ unity matrices, respectively. The structures above follow from e.g. (James *et al.*, 1972), and from the fact that Y_p and U_p are uncorrelated with $w(t)$.

Remark 1. Assuming that we have a consistent estimate of σ, it is possible to conclude, from previous consistency result concerning basic CLS and IV methods, see for example (James *et al.*, 1972), (Stoica and Söderström, 1982) and (Söderström and Stoica, 1989), that the ECLS estimate in (6) is consistent.

3. SEPARABLE NONLINEAR LEAST-SQUARES

Considering the ECLS estimate in (6), it can be seen that in order to obtain the estimate $\hat{\theta}_{ECLS}$ we need an estimate of the noise variances σ. One way to estimate σ from the observations is to solve (5) in a LS sense. Formulating the loss function as

$$f(\theta,\sigma) = ||R_{zy} - \xi\sigma - (R_{z\phi} - S(\sigma))\theta||^2, \qquad (10)$$

an estimate of σ is given as the minimal point

$$\hat{\sigma} = arg\min_{\sigma}\Big[\min_{\theta} f(\theta,\sigma)\Big]. \qquad (11)$$

However, the optimization problem (11) is separable and can be solved for σ and θ separately. If the loss function (10) is minimized analytically with respect to θ, then, for a given σ the minimum is achieved by the ECLS estimate in (6). Substituting (6) in (10), we have the loss function

$$f_1(\sigma) = ||R_{zy} - \xi\sigma - (R_{z\phi} - S(\sigma))(R_{z\phi} - S(\sigma))^{\dagger}(R_{zy} - \xi\sigma)||^2, \qquad (12)$$

and the optimization problem in (11) reduces to

$$\hat{\sigma} = arg\min_{\sigma} f_1(\sigma). \qquad (13)$$

The optimization problem in (13) is referred to as a variable projection problem. Once a minimizing $\hat{\sigma}$ is found from (13), then $\hat{\theta}_{ECLS}$ can be obtained from (6) by replacing σ by $\hat{\sigma}$ in (6).

One interesting observation is that the loss function in (10) can be rewritten as

$$f(\theta,\sigma) = ||R_{zy} - R_{z\phi}\theta + \big((S_1\theta V_1 + S_2\theta V_2) - \xi\big)\sigma||^2, \qquad (14)$$

where S_1 and S_2 are $(n_z|n_a + n_b)$ matrices defined by

$$S_1 = \begin{bmatrix} 0 & 0 \\ I_{n_a} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},\ S_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & I_{n_b} \\ 0 & 0 \end{bmatrix}, \qquad (15)$$

and

$$V_1 = [\,1 \quad 0\,],\ V_2 = [\,0 \quad 1\,]. \qquad (16)$$

Thus, for a given θ the loss function (14) can be analytically minimized with respect to σ and the minimum is then achieved at

$$\hat{\sigma} = \big((S_1\theta V_1 + S_2\theta V_2) - \xi\big)^{\dagger}(R_{z\phi}\theta - R_{zy}). \qquad (17)$$

Substituting (17) in (14), we have the loss function

$$f_2(\theta) = ||R_{zy} - R_{z\phi}\theta - \big((S_1\theta V_1 + S_2\theta V_2) - \xi\big) \times \big((S_1\theta V_1 + S_2\theta V_2) - \xi\big)^{\dagger}(R_{zy} - R_{z\phi}\theta)||^2, \qquad (18)$$

and the solution to the variable projection problem can instead be written as

$$\hat{\theta}_{ECLS} = arg\min_{\theta} f_2(\theta). \qquad (19)$$

Following the previous procedure, the noise variance estimates are obtained in the second step by replacing θ by $\hat{\theta}_{ECLS}$ in (17).

Remark 2. Note that the ECLS method should more correctly be referred to as the estimates given by (6), assuming that the noise variances are known. However, in this section the ECLS method will also include solving the variable projection problem (13). We will also refer the solving of the variable projection problem (19) and the corresponding equation (17) as the ECLS method.

One question that appears is which of the two loss functions, $f_1(\sigma)$ and $f_2(\theta)$, in (12) and (18) respectively, that should be used as a first minimization step in the identification problem. The question will be partly answered in the next section through simulation studies. However, some concluding remarks about this issue can be given already at this stage. Minimizing $f_1(\sigma)$ and $f_2(\theta)$ is a nonlinear optimization problem, which in general has to be solved numerically. Despite of minimization techniques, it is obvious that if $dim(\theta) > dim(\sigma)$, then the minimizing of the loss function $f_2(\theta)$ will require a larger computational load than minimizing $f_1(\sigma)$. Thus, for systems of high orders it is more preferable to minimize $f_1(\sigma)$ than $f_2(\theta)$. However, there is another reason why it might be easier to solve the, in general non-convex, variable projection problem (13) instead of (19). This is due to the fact that the search space for a global minimum of the loss function $f_1(\sigma)$ can be reduced by considering the result from (Beghelli *et al.*, 1990). By defining new matrices as

$$\begin{aligned} R_{YY} &= \frac{1}{N}\sum_{t=1}^{N}\big([y(t), -Y_t, -Y_p]^T[y(t), -Y_t, -Y_p]\big), \\ R_{UU} &= \frac{1}{N}\sum_{t=1}^{N}\big([u(t), U_t, U_p]^T[u(t), U_t, U_p]\big), \\ R_{YU} &= \frac{1}{N}\sum_{t=1}^{N}\big([y(t), -Y_t, -Y_p]^T[u(t), U_t, U_p]\big), \\ &= R_{UY}^T, \end{aligned} \qquad (20)$$

the following theorem can be stated.

Theorem 1. The maximal admissible value for the output noise variance σ_y^2 is the least eigenvalue of the matrix

$$R_{YY} - R_{UY}^T R_{UU}^{-1} R_{UY} \qquad (21)$$

and, similarly, the maximal admissible value for the input noise variance σ_u^2 is the least eigenvalue of the matrix

$$R_{UU} - R_{UY} R_{YY}^{-1} R_{UY}^T. \qquad (22)$$

Proof. See ((Beghelli *et al.*, 1990)).

Remark 3. It has not been mentioned anything about consistency of the estimated parameters, obtained from solving the separable nonlinear LS optimization problem (11)-(13) or (17)-(19). However, in Remark 1 the case where it is assumed that we have consistent estimate from the variable projection problem (13) is discussed. Interesting results concerning consistency and accuracy for general variable projection problems can be found in (Mahata, 2003).

4. SIMULATION RESULTS

In this section the same standard example as the one described in (Söderström, 1981), and also used in e.g. (Soverini and Söderström, 2000), has been used in the simulations. The system is described by

$$x(t) = \frac{B(q^{-1})}{A(q^{-1})} u_0(t), \tag{23}$$

where

$$\begin{aligned} A(q^{-1}) &= 1 - 1.5q^{-1} + 0.7q^{-2} \\ B(q^{-1}) &= 1.0q^{-1} + 0.5q^{-2}. \end{aligned} \tag{24}$$

The undisturbed input $u_0(t)$ is modeled as

$$u_0(t) = \frac{C(q^{-1})}{D(q^{-1})} r(t), \tag{25}$$

where

$$\begin{aligned} C(q^{-1}) &= 1 \\ D(q^{-1}) &= 1 - 0.9q^{-1}, \end{aligned} \tag{26}$$

and $r(t)$ is white noise with variance $\sigma_r^2 = 1$.

A comparative study, where the above derived method is compared with a standard IV method and the Frisch Scheme approach (FR), has been performed. For the IV method the system was identified using the delayed inputs, $U_{p=4}$, as instruments. Both the ECLS and the CLS method were used in the study. For the ECLS method, $z(t)$ was chosen as in (8) where $p = 2$, see further Remark 2. Moreover, the variable projection problem (13) and (19) were solved by the Nelder-Mead minimization method using the *fminsearch* function in Matlab optimization toolbox. In the standard CLS method, the same noise variance estimates as for the ECLS method were used, i.e. the estimates obtained from the variable projection problem (13). Moreover, as mentioned above, for the CLS method we have $z(t) = \phi(t)$.

4.1 Example 1.

In the first comparative study, Monte Carlo simulations based on one hundred independent realizations with $N = 500$ data points were performed. The measurement noise variances are $\sigma_y^2 = \sigma_u^2 = 1$. For the IV method, the estimates were rejected if $|\theta_0 - \hat{\theta}| > 10$. The initial values, $\sigma = [0 \quad 0]^T$ and $\theta = \theta_{LS}$, were used for the variable projection problems (13) and (19), respectively. θ_{LS} is the standard LS estimate. Tables 1 and 2 show the mean values and the standard deviations of the estimated model parameters. From the tables it can be seen that the IV estimates are not very accurate. The FR and the CLS methods have approximately the same level of accuracy. The ECLS method shows an overall better accuracy than the other methods. Table 3 reports the mean values and the standard deviations of the estimated noise variance parameters for the ECLS and the FR methods. From the table it can be seen that the accuracy is approximately the same for both the output noise variance estimate and the input noise variance estimate.

Table 1. Means and standard deviations for coefficient estimates of $A(q^{-1})$.

Method	$a_1 = -1.5$	$a_2 = 0.7$
$ECLS$	-1.502 $\pm$ 0.025	0.700 $\pm$ 0.019
CLS	-1.502 $\pm$ 0.048	0.701 $\pm$ 0.030
FR	-1.509 $\pm$ 0.034	0.705 $\pm$ 0.024
IV	-1.491 $\pm$ 0.602	0.692 $\pm$ 0.391

Table 2. Means and standard deviations for coefficient estimates of $B(q^{-1})$.

Method	$b_1 = 1$	$b_2 = 0.5$
$ECLS$	0.996 $\pm$ 0.155	0.493 $\pm$ 0.191
CLS	0.986 $\pm$ 0.249	0.507 $\pm$ 0.400
FR	1.007 $\pm$ 0.199	0.459 $\pm$ 0.272
IV	0.773 $\pm$ 4.780	0.736 $\pm$ 5.448

Table 3. Means and standard deviations for noise variance estimates.

Method	$\sigma_y^2 = 1$	$\sigma_u^2 = 1$
$ECLS$	0.995 $\pm$ 0.094	0.992 $\pm$ 0.120
FR	0.996 $\pm$ 0.104	0.971 $\pm$ 0.110

Some minor studies concerning the computational load has also been performed. It can be established that both the FR and the IV method are faster than the ECLS method. This is due to to the fact that in the ECLS method a nonlinear optimization problem has to be solved. However, even if there exist faster and more efficient methods than the Nelder-Mead method, it is interesting to compare the computational load required when minimizing the loss functions $f_1(\sigma)$ and $f_2(\theta)$, respectively. Minimizing $f_1(\sigma)$ and then solve the equation (6) will give the same result as minimizing $f_2(\theta)$ and then solve the equation (17), i.e. the same parameter values are obtained. However, simulation studies reveal that solving the variable projection problem (19) and the corresponding equation (17) is approximately 2 times slower than solving (13) and (6). Moreover, in order to find the global minima, simulation studies showed that it is more important to have a good initial guess of the parameters when solving the minimization problem (19) compared to solving (13).

4.2 Example 2.

The robustness of the noise variance estimates for the ECLS and the FR methods has been investigated. Six different condition for the noise variances, i.e. all possible pairs of $\sigma_y^2 = \{1,\ 10,\ 50\}$ and $\sigma_u^2 = \{1,\ 3\}$, have been tested. Again, Monte Carlo simulations based on one hundred independent realizations with $N = 500$ data points were performed. The result is reported in Table 4. The result shows that, the accuracy is approximately the same and for both methods the estimates get worst when the variances of the added noise are high. However, the result indicates that the FR method might have a slightly more robust behavior when the amounts of noise on the data are unbalanced, even if the ECLS method for some noise conditions show better mean values.

Table 4. Means and standard deviations for noise variance estimates.

Variances / Method	$ECLS$	FR
$\sigma_y^2 = 1$	0.995 $\pm$ 0.094	0.996 $\pm$ 0.104
$\sigma_u^2 = 1$	0.992 $\pm$ 0.120	0.971 $\pm$ 0.110
$\sigma_y^2 = 10$	9.714 $\pm$ 1.776	9.726 $\pm$ 0.838
$\sigma_u^2 = 1$	0.936 $\pm$ 0.396	0.860 $\pm$ 0.304
$\sigma_y^2 = 50$	48.230 $\pm$ 7.700	49.53 $\pm$ 5.208
$\sigma_u^2 = 1$	1.449 $\pm$ 0.769	0.907 $\pm$ 0.362
$\sigma_y^2 = 1$	0.978 $\pm$ 0.100	0.975 $\pm$ 0.122
$\sigma_u^2 = 3$	2.987 $\pm$ 0.264	2.932$\pm$ 0.239
$\sigma_y^2 = 10$	9.684 $\pm$ 2.198	10.443 $\pm$ 1.094
$\sigma_u^2 = 3$	3.025 $\pm$ 0.556	1.357$\pm$ 0.476
$\sigma_y^2 = 50$	48.000 $\pm$ 3.405	47.426 $\pm$ 4.733
$\sigma_u^2 = 3$	3.405 $\pm$ 0.965	2.525$\pm$ 0.651

5. CONCLUSIONS

A new approach for estimating parameters in dynamic errors-in-variables models is investigated in this paper. The method is a version of the CLS method based on an overdetermined system of equations, and is therefore called the extended CLS (ECLS) method. Utilizing the observation that the noise variance parameters and the model parameters can be treated separately, a two step procedure is developed, based on the principle of separable nonlinear LS. Simulation experiments show that the ECLS approach provides a good accuracy of the parameter estimates, and a comparative analysis has indicated that the ECLS method might give an even better accuracy of the model parameter estimates than the FR method. On the other hand, the FR method seems to have a slightly more robust behavior of the noise variance estimates. Moreover, the ECLS approach relies on computationally more demanding iterative minimizing procedures than the FR approach.

Current research within the framework of this paper is directed towards developing a recursive variant of the ECLS method. Other topics for further research could include a theoretical investigation of the accuracy of the parameter estimates when the CLS method is used. A study of the performance of the ECLS method under more general noise conditions could also be a topic for further research.

Acknowledgment The financial support by the MISTRA program Sustainable Urban Water Management is gratefully acknowledged.

REFERENCES

Aoki, M. and P. C. Yue (1970). On a priori error estimate of some identification methods. *IEEE Trans. on Automatic Control* **AC–15**, 541–548.

Beghelli, S., R. P. Guidorzi and U. Soverini (1990). The Frisch scheme in dynamic system identification. *Automatica* **26**, 171–176.

Fernando, K. V. and H. Nicholson (1985). Identification of linear systems with input and output noise: the Koopmans-Levin method. *Proc. IEE, Part D* **132**, 30–36.

Golub, G. H. and V. Peryra (1973). The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal of Numerical Analysis* **10**(2), 413–432.

Hannan, E. J. (1971). Non-linear time series regression. *J. Appl. Prob.* **8**, 767–780.

James, P. N., P. Souter and D. C. Dixon (1972). Suboptimal estimation of the parameters of discrete systems in the presence of correlated noise. *Electron. Lett.* **8**, 411–412.

Levin, M. J. (1964). Estimation of a system pulse transfer function in the presence of noise. *IEEE Trans. on Automatic Control* **AC–9**, 229–235.

Mahata, K. (2003). Estimation using low rank signal models. PhD thesis. Uppsala University. Division of Systems and Control.

Mahata, K. and T. Söderström (2002). Identification of dynamic errors-in-variables model using prefiltered data. In: *15th World Congress of IFAC*. Barcelona, Spain.

Söderström, T. and K. Mahata (2002). On instrumental variable and total least squares approaches for identification of noisy systems. *Int. J. Control* **75**(6), 381–389.

Söderström, T., U. Soverini and K. Mahata (2002). Perspectives on errors-in-variables estimation for dynamic systems. *Signal Processing* **82**(8), 1139–1154.

Söderström, T. (1981). Identification of stochastic linear systems in presence of input noise. *Automatica* **17**, 713–725.

Söderström, T. and P. Stoica (1989). *System Identification*. Prentice Hall International, Hemel Hempstead, UK.

Soverini, U. and T. Söderström (2000). Identification methods of dynamic systems in presence of input noise. In: *Proceedings IFAC SYSID '00*. Santa Barbara, California.

Stoica, P. and T. Söderström (1982). Bias correction in least-squares identification. *Int. J. Control* **35**(3), 449–457.

Zheng, W. X. (1999). On least-squares identification of stochastic linear systems with noisy input-output data. *International Journal of Adaptive Control and Signal Processing* **13**, 131–143.

Zheng, W. X. and C. B. Feng (1989). Unbiased parameter estimation of linear systems in presence of input and output noise. *International Journal of Adaptive Control and Signal Processing* **3**, 231–251.

Zheng, W. X. and C. B. Feng (1992). Identification of a class of dynamic errors-in-variables models. *International Journal of Adaptive Control and Signal Processing* **6**, 431–440.