15th Triennial World Congress, Barcelona, Spain

IDENTIFICATION OF DYNAMIC ERRORS-IN-VARIABLES MODEL USING A FREQUENCY DOMAIN FRISCH SCHEME [1]

Torsten Söderström†, Kaushik Mahata† and Umberto Soverini‡

† *Department of Systems and Control, Information Technology, Uppsala University. P O Box 337, SE-751 05 Uppsala, Sweden.*
Email: ts@syscon.uu.se, km@syscon.uu.se
‡ *Dipartment of Eletronics, Information Science and Systems, University of Bologna, Viale Risorgimento 2, IT–40136 Bologna, Italy.*
Email: usoverini@deis.unibo.it

Abstract: In this paper we propose a parametric and a non-parametric identification algorithm for dynamic errors-in-variables model. We show that the two-dimensional process composed of the input-output data admits a finite order ARMA representation. The non-parametric method uses the ARMA structure to compute a consistent estimate of the joint spectrum of the input and the output. A Frisch scheme is then employed to extract an estimate of the joint spectrum of noise free input-output data from the ARMA spectrum, which is used to estimate the transfer function. The parametric method exploits the ARMA structure to give consistent estimates of the system parameters. The performances of the algorithms are illustrated using the results obtained from a numerical simulation study. ©IFAC 2002

Keywords: System Identification, Errors-in-variables, Spectrum estimation, Frisch scheme, ARMA model.

1. INTRODUCTION

Identification of linear dynamic systems from noise–corrupted input and output measurements is often referred as a dynamic "errors–in–variables" problem. Several authors have considered this problem with different approaches. For example, we can distinguish between time–domain (Söderström and Stoica, 1989), and frequency-domain methods (Pintelon and Schoukens, 2001), approaches exploiting deterministic signals (e.g. periodic inputs (Forsell *et al.*, 1999)) or stochastic processes (Söderström, 1981). When only second–order statistics are exploited it is a well–known result that, in general, the identification of errors–in–variables models cannot admit a single solution (Anderson and Deistler, 1984).

In this paper we shall show that the two-dimensional process composed of the input-output data admits an ARMA representation of finite order, which we use to construct a parametric and a non-parametric way of identifying the SISO errors-in-variables model. The non-parametric method combines the ARMA spectrum with the ideas of the Frisch scheme (Beghelli *et al.*, 1990). A similar attempt can be found

[1] This work was supported by Swedish Research Council for Engineering Sciences under contract 98-654 and the Ministry for University and Scientific and Technologic Research, Rome.

in (Beghelli *et al.*, 1997), where a periodogram estimate of the joint input-output spectrum is used. The algorithm presented in this work outperforms the method introduced in (Beghelli *et al.*, 1997). The parametric method, which exploits the ARMA structure directly, will be shown to yield the best result. For a detailed mathematical description of the classes of ***observationally equivalent*** systems in the framework of this work, we refer the reader to (Scherrer and Deistler, 1998).

2. PROBLEM STATEMENT

Consider the following system with noise-free input $u_0(t)$ and the undisturbed output $y_0(t)$, linked through the linear difference equation

$$A(q^{-1})\, y_0(t) = B(q^{-1})\, u_0(t), \tag{1}$$

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

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_n q^{-n} \\ B(q^{-1}) &= b_1 q^{-1} + \cdots + b_n q^{-n} \end{aligned} \ . \tag{2}$$

and q^{-1} is the backward shift operator, i.e. $q^{-1} x(t) = x(t-1)$. It is not restrictive to assume that the polynomials $A(q^{-1})$, $B(q^{-1})$ have equal degree n, which represents the order of the system. It is further assumed that $A(z)$ has all zeros outside the unit circle and has no factor in common with $B(z)$.

We assume that the observations are corrupted by zero mean, mutually independent, white measurement noise sequences $\tilde{u}(t)$ and $\tilde{y}(t)$, at the input and output, respectively. Therefore, the available signals are of the form

$$\begin{aligned} u(t) &= u_0(t) + \tilde{u}(t) \\ y(t) &= y_0(t) + \tilde{y}(t) \end{aligned} \ . \tag{3}$$

The variances of $\tilde{u}(t)$ and $\tilde{y}(t)$ are λ_u and λ_y respectively. The true input $u_0(t)$ is an ARMA process, i.e.

$$u_0(t) = \frac{C(q^{-1})}{D(q^{-1})} e(t), \tag{4}$$

where $e(t)$ is a zero mean white noise with variance λ_e and

$$\begin{aligned} C(q^{-1}) &= 1 + c_1 q^{-1} + \cdots + c_m q^{-m} \\ D(q^{-1}) &= 1 + d_1 q^{-1} + \cdots + d_m q^{-m} \end{aligned} \ . \tag{5}$$

$u_0(t)$ is independent of both $\tilde{u}(t)$ and $\tilde{y}(t)$. This is a mild assumption since a stationary process can be accurately approximated as an ARMA process by properly adjusting m, $\{c_i\}_{i=1}^{m}$ and $\{d_i\}_{i=1}^{m}$. The problem is to determine the transfer function

$$G(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})} \tag{6}$$

and the variances λ_u and λ_y of the noise sequences, given the available measurements $\{u(t)\}_{t=1}^{N}$ and $\{y(t)\}_{t=1}^{N}$.

3. MODELING OF THE DATA

Let us define the two dimensional process

$$\mathbf{z}(t) = \begin{bmatrix} y(t) & u(t) \end{bmatrix}^T = \mathbf{z}_0(t) + \tilde{\mathbf{z}}(t), \tag{7}$$

where $\mathbf{z}_0(t)$ and $\tilde{\mathbf{z}}(t)$ denote the respective noise-free and noisy part of $\mathbf{z}(t)$ given by

$$\mathbf{z}_0(t) = \begin{bmatrix} y_0(t) & u_0(t) \end{bmatrix}^T,$$
$$\tilde{\mathbf{z}}(t) = \begin{bmatrix} \tilde{y}(t) & \tilde{u}(t) \end{bmatrix}^T.$$

It is straightforward to derive that the spectrum $\mathbf{\Phi}_{\mathbf{z}}(\omega)$ of $\mathbf{z}(t)$ at frequency ω is given by

$$\begin{aligned} &\mathbf{\Phi}_{\mathbf{z}}(\omega) = \mathbf{\Phi}_{\mathbf{z}_0}(\omega) + \mathbf{\Phi}_{\tilde{\mathbf{z}}}(\omega) \\ &= \begin{bmatrix} G(e^{-i\omega}) \\ 1 \end{bmatrix} \begin{bmatrix} G(e^{i\omega}) & 1 \end{bmatrix} \mathbf{\Phi}_{u_0}(\omega) + \begin{bmatrix} \lambda_y & 0 \\ 0 & \lambda_u \end{bmatrix} \end{aligned} \tag{8}$$

The idea here is to extend the frequency domain Frisch scheme in (Beghelli *et al.*, 1997) using a black-box *parametric* modeling of the spectral density $\mathbf{\Phi}_z$. If the spectral density matrix $\mathbf{\Phi}_z$ is known, and an appropriate diagonal matrix is subtracted, then one would get a rank 1 positive semidefinite matrix, corresponding to the first term of (8), for all frequencies ω. In case the decomposition as in (8) can be carried out, the first term would easily lead to estimates of the transfer function $G(e^{i\omega})$ and the true input spectrum $\mathbf{\Phi}_{u_0}$.

As a starting point, it is required to estimate $\mathbf{\Phi}_z$. First, to show that $\mathbf{z}(t)$ admits an ARMA representation, let us introduce

$$\begin{aligned} \mathbf{\Delta}(q^{-1}) &= I + \Delta_1 q^{-1} + \cdots + \Delta_K q^{-K} \\ &= \begin{bmatrix} A(q^{-1}) & -B(q^{-1}) \\ 0 & D(q^{-1}) \end{bmatrix}. \end{aligned} \tag{9}$$

It is straightforward to see that

$$\mathbf{\Delta}(q^{-1})\mathbf{z}(t) = \begin{bmatrix} A(q^{-1})\tilde{y}(t) - B(q^{-1})\tilde{u}(t) \\ C(q^{-1})e(t) + D(q^{-1})\tilde{u}(t) \end{bmatrix} \tag{10}$$

is clearly a moving average process in the sense that its covariance function vanishes after $\max(m, n)$ lags. Therefore, $\mathbf{z}(t)$ admits the ARMA representation

$$\mathbf{\Delta}(q^{-1})\mathbf{z}(t) = \mathbf{\Gamma}(q^{-1})\epsilon(t), \tag{11}$$

where $\mathbf{\Gamma}(q^{-1})$ is a 2×2 matrix polynomial of order $K = \max(m, n)$ with the identity matrix as the leading coefficient, i.e.

$$\mathbf{\Gamma}(q^{-1}) = I + \Gamma_1 q^{-1} + \cdots + \Gamma_K q^{-K}. \tag{12}$$

Further, in (11) $\epsilon(t)$ is the two dimensional innovation of $\mathbf{z}(t)$, a zero mean white noise with covariance matrix $\mathbf{\Lambda}$.

4. ESTIMATION METHODS

4.1 *Estimating the two-dimensional ARMA model*

In order to estimate the ARMA model in (11) we adopt a well-known procedure previously used, for example by (Mayne and Firoozan, 1982). The idea is to first estimate the innovation process $\{\epsilon(t)\}_{t=1}^{N}$ and then use it as a second input, thus using a linear regression model to estimate $\mathbf{\Delta}(q^{-1})$ and $\mathbf{\Gamma}(q^{-1})$.

In order to estimate the innovation $\epsilon(t)$, we fit an AR model of large order L to the data $\mathbf{z}(t)$, i.e.

$$[I + A_1 q^{-1} + \cdots + A_L q^{-L}]\,\mathbf{z}(t) = \epsilon(t). \quad (13)$$

The model order L should be sufficiently large, in order to ensure a good approximation of the ARMA model by the AR model. It is straightforward to show that the least squares estimate of $\epsilon(t)$ and $\mathbf{\Lambda}$ is given by

$$\hat{\epsilon}(t) = \mathbf{z}(t) - \left[\sum_{t=1}^{N} \mathbf{z}(t) Z_1^T(t)\right]\left[\sum_{t=1}^{N} Z_1(t) Z_1^T(t)\right]^{-1} Z_1(t) \quad (14)$$

$$\hat{\mathbf{\Lambda}} = \frac{1}{N}\sum_{t=1}^{N} \hat{\epsilon}(t)\hat{\epsilon}^T(t), \quad (15)$$

where $Z_1(t) = \left[-\mathbf{z}^T(t-1) \ldots -\mathbf{z}^T(t-L)\right]^T$. Next, we shall consider $\hat{\epsilon}(t)$ as an additional input to estimate the ARMA model in (11). Note that the order of ARMA model $K = \max(n, m)$ is known. Let us introduce

$$\Theta_2 = [\,\Delta_1 \ldots \Delta_K \; \Gamma_1 \ldots \Gamma_K\,],$$
$$Z_2(t) = [\,-\mathbf{z}^T(t-1) \ldots -\mathbf{z}^T(t-K) \quad \hat{\epsilon}^T(t-1) \ldots \hat{\epsilon}^T(t-K)\,]^T.$$

Then it is straightforward to show that the least squares estimate of Θ_2 is given by

$$\widehat{\Theta}_2 = \left[\sum_{t=1}^{N} \mathbf{z}(t) Z_2^T(t)\right]\left[\sum_{t=1}^{N} Z_2(t) Z_2^T(t)\right]^{-1}. \quad (16)$$

From $\widehat{\Theta}_2$ and $\hat{\mathbf{\Lambda}}$ we estimate $\mathbf{\Phi}_{\mathbf{z}}(\omega)$ as

$$\hat{\mathbf{\Phi}}_{\mathbf{z}}(\omega) = \hat{\mathbf{\Delta}}^{-1}(\omega)\hat{\mathbf{\Gamma}}(\omega)\hat{\mathbf{\Lambda}}\hat{\mathbf{\Gamma}}^T(-\omega)\hat{\mathbf{\Delta}}^{-T}(-\omega), \quad (17)$$

where we have used the notation $X(\omega)$ for $X(e^{-i\omega})$ for convenience.

4.2 *Non-parametric identification*

In non-parametric identification, we determine the estimates of the noise variances λ_u, λ_y and $G(e^{-i\omega})$ for each frequency. From (8), since $\mathbf{\Phi}_{\mathbf{z}_0}(\omega)$ is singular, equating the determinant of $\mathbf{\Phi}_{\mathbf{z}_0}(\omega)$ to zero it is straightforward to derive

$$\det[\mathbf{\Phi}_{\mathbf{z}}(\omega)] = [\,\Phi_{yy}(\omega) \;\; \Phi_{uu}(\omega) \;\; -1\,]\,\Theta_3, \quad (18)$$
$$\text{where} \quad \Theta_3 = [\,\lambda_u \;\; \lambda_y \;\; \lambda_y\lambda_u\,]^T \quad (19)$$

is a vector with *three* unknown parameters. We can derive an overdetermined system of linear equations in Θ_3 by repeating this relation for a large number of frequencies. The LS solution of this system gives an estimate of Θ_3. The structure of Θ_3 can be used to derive a more sophisticated way to obtain $\widehat{\Theta}_3$, comprising of a nonlinear search in one parameter.

Once the estimates $\hat{\lambda}_u$ and $\hat{\lambda}_y$ are obtained, we have the estimate of $\mathbf{\Phi}_{\mathbf{z}_0}(\omega)$ as

$$\hat{\mathbf{\Phi}}_{\mathbf{z}_0}(\omega) = \hat{\mathbf{\Phi}}_{\mathbf{z}}(\omega) - \hat{\mathbf{\Phi}}_{\tilde{\mathbf{z}}}(\omega), \qquad \forall\omega. \quad (20)$$

A simple way to estimate $G(e^{-i\omega})$ from $\hat{\mathbf{\Phi}}_{\mathbf{z}_0}(\omega)$ is

$$\hat{G}(e^{-i\omega}) = \hat{\mathbf{\Phi}}_{z_0}^{(12)} / \hat{\mathbf{\Phi}}_{z_0}^{(22)}. \quad (21)$$

A more sophisticated estimate may be derived by using also the remaining elements of $\hat{\mathbf{\Phi}}_{\mathbf{z}_0}(\omega)$, which would lead us to a two dimensional search, treating the real and the imaginary parts of $G(e^{-i\omega})$ as two unknowns.

4.3 *Parametric identification*

In this section we shall discuss two different parametric identification methods based on the results discussed in the previous sections.

One way to obtain a parametric estimate of the system transfer function is to utilize the non-parametric transfer function estimate (21). Let us introduce the parameter vector

$$\theta = [\,a_1 \ldots a_n \; b_1 \ldots b_n\,]^T.$$

From (2) and (6), we have

$$H(\omega, \theta) = \frac{b_1 e^{-i\omega} + \cdots + b_n e^{-in\omega}}{1 + a_1 e^{-i\omega} + \cdots + a_n e^{-in\omega}}. \quad (22)$$

One way to estimate θ is to minimize the weighted spectral distance between $H(\omega,\theta)$ and $\hat{G}(e^{-i\omega})$ given by (21), i.e.

$$\hat{\theta} = \arg\min_{\theta} \sum_{k=1}^{M} W_k |H(\omega_k, \theta) - \hat{G}(e^{-i\omega_k})|^2, \quad (23)$$

where $\{\omega_k\}_{k=1}^{M} \subset [0, \pi]$ are user defined discrete frequency points, and W_k is a user chosen weight corresponding to frequency ω_k.

Parametric estimates of $A(q^{-1})$, $B(q^{-1})$ and $D(q^{-1})$ can also be obtained from $\hat{\mathbf{\Delta}}(q^{-1})$ using (9). Let us introduce the process $\mathbf{v}(t)$ and its correlation sequence R_τ as

$$\mathbf{v}(t) = \mathbf{\Delta}(q^{-1})\mathbf{z}(t) = \mathbf{\Gamma}(q^{-1})\epsilon(t), \quad (24)$$

$$R_\tau = E\mathbf{v}(t)\mathbf{v}^T(t-\tau) = \sum_{i=\tau}^{K} \Gamma_i \mathbf{\Lambda} \Gamma_{i-\tau}^T, \quad (25)$$

where $\Gamma_0 = I$. Note that we used (24) in deriving the last equality. The previously obtained estimates of $\mathbf{\Lambda}$ and $\{\Gamma_i\}_{i=1}^{K}$ can be substituted in (25) to compute $\hat{R}_\tau$. To explore how $C(q^{-1})$, λ_u, λ_y and λ_e are determined let us examine R_τ elementwise. Combining (10), (24) and (25) it is straightforward to derive

$$R_\tau^{(11)} = \lambda_y \sum_{i=0}^{n-\tau} a_i a_{i+\tau} + \lambda_u \sum_{i=1}^{n-\tau} b_i b_{i+\tau}, 0 \le \tau \le n \quad (26)$$

$$R_\tau^{(12)} = -\lambda_u \sum_{i=0}^{n-\tau} d_i b_{i+\tau}, \quad 0 \le \tau \le n; \quad (27)$$

$$R_\tau^{(21)} = -\lambda_u \sum_{i=1}^{m-\tau} b_i d_{i+\tau}, \quad 0 \le \tau \le m-1; \quad (28)$$

where $b_0 = 0$ and $a_0 = d_0 = 1$. Since the estimates of $A(q^{-1})$, $B(q^{-1})$ and $D(q^{-1})$ are already known, (26), (27) and (28) can be used for different values of $\tau \le K$ to get an overdetermined system of equations in λ_y and λ_u. This system can be solved to obtain the estimates of the unknowns. Now we can use the remaining element of R_τ to estimate $C(q^{-1})$. From (10), (24) and (25), it is straightforward to derive that

$$\lambda_e C(q^{-1})C(q) = \sum_{\tau=-m}^{m} R_\tau^{(22)} q^{-\tau} - \lambda_u D(q^{-1})D(q) \quad (29)$$

The polynomial at the right hand side of (29) can be factorised to obtain $\hat{C}(q^{-1})$. Of course as $\hat{R}_\tau^{(22)}$, $\hat{\lambda}_u$ deviate somewhat from the true values, it may happen that the right hand side cannot be exactly factorised. This problem is encountered precisely when the polynomial is not positive definite.

5. NUMERICAL ILLUSTRATION

To illustrate numerically the identification methods introduced in the previous section, we shall consider the system transfer function as follows

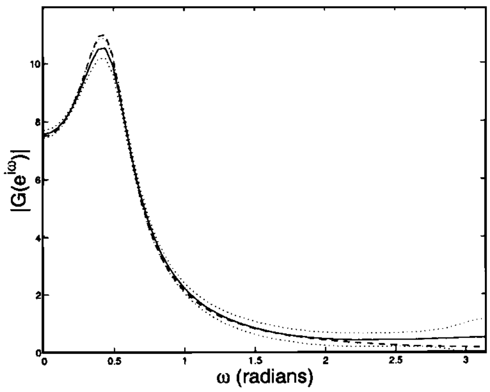
$$A(q^{-1}) = 1 - 1.5q^{-1} + 0.7q^{-2}, \quad (30)$$
$$B(q^{-1}) = 1.0q^{-1} + 0.5q^{-2}, \quad (31)$$
$$C(q^{-1}) = 1 + 0.2q^{-1}, \quad (32)$$
$$D(q^{-1}) = 1 - 0.8q^{-1}. \quad (33)$$

The process $e(t)$ is zero mean white noise and its variance equals unity. The number of data points N was 1000 for each Monte-Carlo simulation. A natural choice of the weighting vector W_k in (23) would be

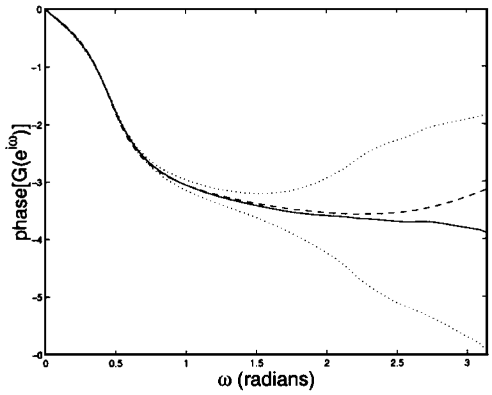
$$W_k = \hat{\mathbf{\Phi}}_{u_0 u_0}(\omega_k). \quad (34)$$

Using this choice during the optimisation in (23), the non-parametric estimate of the transfer function at a particular frequency is given a weight directly proportional to the SNR at that frequency.

In Figure 1 we have compared the mean value and the standard deviation of non-parametric transfer function estimates $\hat{G}(e^{-i\omega})$ from 50 independent Monte-Carlo simulations with the true transfer function $G(e^{-i\omega})$. The SNR at both the input and the output are 15 dB approximately. The order L of the AR model involved in estimation of the innovation $\epsilon(t)$ in (13), has been taken as 10. The order K of the ARMA model in (11) is 2. We have used 128 equally spaced discrete frequency points



(a) Magnitude



(b) Phase

Fig. 1. Non-parametric transfer function estimates at 15 dB input–output SNR and $L = 10$. The mean from Monte Carlo simulations (solid line), the mean $\pm$ standard deviation (dotted line) and the true values (dashed line).

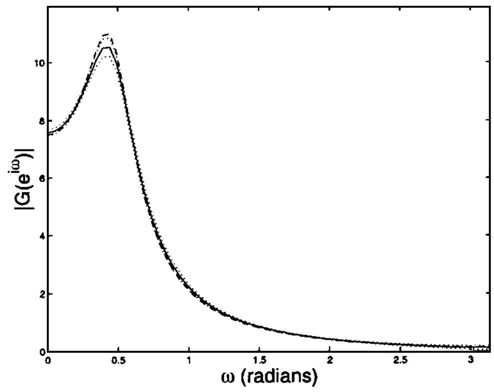
in the interval $[0, \pi]$ to set up the optimisation problem in order to obtain Θ_3 defined in (19). The parametric estimate of the transfer function $G(e^{-i\omega})$ obtained from (23) is shown in Figure 2.

The noise sensitivity of the algorithms can be observed in Figure 3 where we have shown the results of Monte-Carlo simulations at 20 dB input-output SNR. The value of L is set to 10.

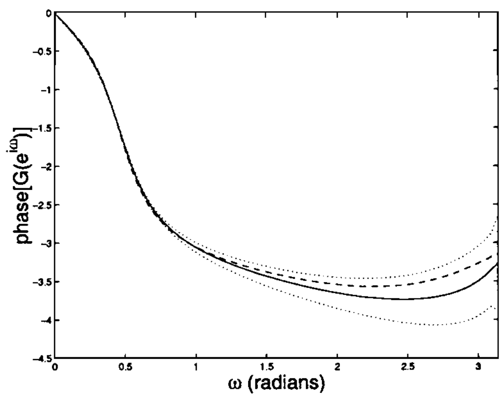
In Figure 4 we have shown the non-parametric transfer function estimates when the polynomials

$$C(q^{-1}) = 1 + 0.1q^{-1}, \quad (35)$$
$$D(q^{-1}) = 1 - 0.2q^{-1} \quad (36)$$

give more flat spectrum of the true input. The system under consideration remains unchanged. The input and output SNR are 20 dB each and L is 10.



(a) Magnitude (Parametric)



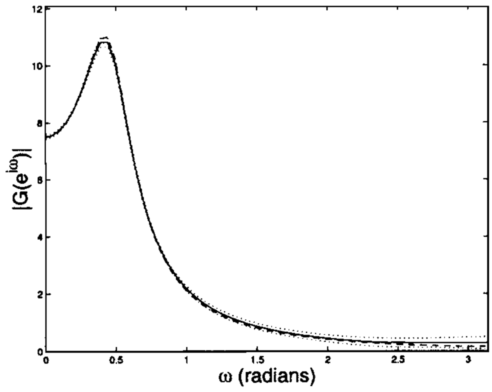
(b) Phase (Parametric)

Fig. 2. Parametric transfer function estimates at 15 dB input–output SNR and $L = 10$. The mean from Monte Carlo simulations (solid line), the mean ± standard deviation (dotted line) and the true values (dashed line).

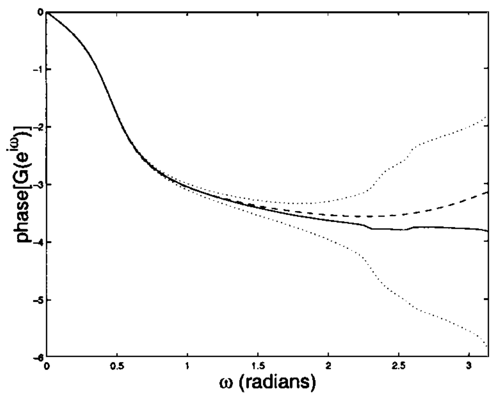
	True Values	Mean Value	Std Deviation	CRB
a_1	−1.5	−1.49	0.007	0.0042
a_2	0.7	0.69	0.005	0.0032
b_1	1.0	0.95	0.177	0.0210
b_2	0.5	0.57	0.151	0.0288
c_1	0.2	0.23	0.071	0.0332
d_1	−0.8	−0.80	0.011	0.0174
λ_u	0.1	0.11	0.042	0.0396
λ_y	1.0	1.0	0.041	0.0070

Table 1. Parametric identification results from 500 Monte Carlo Simulations. Input SNR = 15 dB, Output SNR = 23 dB, $L = 50$ and $N = 2000$.

Next, in Table 1 we have summarised the results from parametric estimation scheme which is directly related to the ARMA model, mentioned in (11) and (25)-(29). The value of L was taken as 50, while the number of data point N is 2000 in each simulation. The SNR at the input is 15 dB and



(a) Magnitude



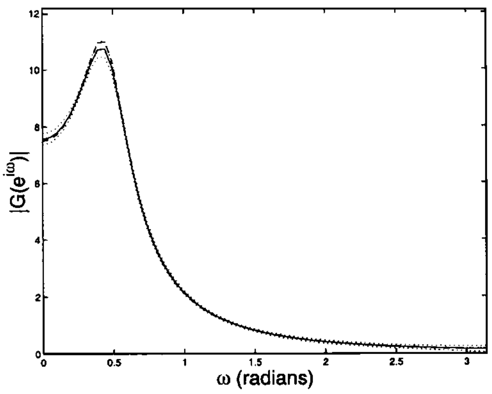
(b) Phase

Fig. 3. Non-parametric transfer function estimates at 20 dB input and output SNR and $L = 10$. The mean from Monte Carlo simulations (solid line), the mean ± standard deviation (dotted line) and the true value (dashed line).

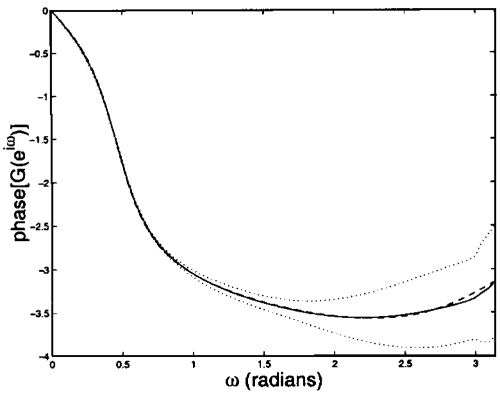
that at output is 23 dB. The standard deviation obtained has been compared with the Cramer-Rao bound (Karlsson *et al.*, 2000).

6. DISCUSSION

In Figures 1 and 3 we see that low data SNR at a particular frequency causes degradation of the transfer function estimate. The phase of the estimated transfer function, in particular, is more sensitive to noise. The estimation of θ in (23) is the only computationally intensive operation in this algorithm, but the corresponding improvement in the estimate of the transfer function is also impressive as shown in Figure 2 compared with Figure 1. It is also interesting to observe the significant improvement in the estimation when true input has a flat spectrum, as in Figure 4, while the SNR of the data remains the same. The choice of the parameter L is important. Too small



(a) Magnitude



(b) Phase

Fig. 4. Non-parametric transfer function estimates at 20 dB input and output SNR with input signal having more flat spectrum and $L = 10$. The mean from Monte Carlo simulations (solid line), the mean $\pm$ standard deviation (dotted line) and the true value (dashed line).

L leads to biased estimate of $A(q^{-1})$, $B(q^{-1})$ and $D(q^{-1})$. On the other hand, a large choice of L causes significantly larger computational burden. One way to chose a reasonably good value of L is to apply whiteness test, see for example (Söderström and Stoica, 1989), on the estimated innovation $\hat{\varepsilon}(t)$.

7. CONCLUSIONS

Two parametric and one non-parametric algorithm based on a frequency domain approach for dynamic errors-in-variables system are proposed. The algorithms provide reasonably good estimates with low computational cost. The parametric method described by (11) and (25)-(29) is computationally economical, because all the operations involved with it are linear. The non-parametric method given by (21) is also fast because of the same reason. The non-paramatric method is sensitive to noise. Hence it is sometimes required to modify the non-parametric estimate using another parametric estimate given by (22). This modification involves non-linear optimisation and is computationally expensive.

The accuracy of the parametric method is dependent upon the choice of L. The optimal choice of L depends upon the the system characteristics. In general, it is required to have a large L, if the joint spectrum of the noisy input and output data is peaky. The non-parametric method is less sensitive to the choice of L.

8. REFERENCES

Anderson, B. D. O. and M. Deistler (1984). Identifiability of dynamic errors-in-variables models. *J. Time Series Analysis* **5**, 1–13.

Beghelli, S., P. Castaldi and U. Soverini (1997). A frequential approach for errors-in-variables models. In: *Proc. European Control Conference, ECC '97*. Brussels, Belgium.

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

Forsell, U., F. Gustafsson and T. McKelvey (1999). Time-domain identification of dynamic errors-in-variables systems using periodic excitation signals. In: *Proc. IFAC 14th World Congress*. Beijing, P.R.China.

Karlsson, E, T Söderström and P Stoica (2000). The Cramér-Rao lower bound for noisy input output systems. *Signal Processing* **80**(11), 2421–2447.

Mayne, D. Q. and F. Firoozan (1982). Linear identification of ARMA processes. *Automatica* **18**, 461–466.

Pintelon, R. and J. Schoukens (2001). *System Identification: A Frequency Domain Approach*. IEEE Press and John Wiley.

Scherrer, W. and M. Deistler (1998). A structure theory for linear dynamic errors-in-variables models. *SIAM Journal on Control and Optimization* **36**(6), 2148–2175.

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.