ON THE OPTIMAL ESTIMATION OF ERRORS IN VARIABLES MODELS FOR ROBUST CONTROL

Juan C. Agüero [*] **Graham C. Goodwin** [*]
Mario E. Salgado [**]

[*] *The School of Electrical Eng. and Computer Science*
The University of Newcastle, Australia
[**] *Department of Electronic Engineering*
Universidad Técnica Federico Santa María
Valparaíso, Chile

Abstract: There exists a substantial literature dealing with the problem of errors-in-variables identification. It is known, for example, that there is an equivalence class of models that give compatible descriptions of the input-output data. In the current paper, we impose a mild restriction so as to avoid certain singular possibilities. This leads to a parameterization of the equivalence class of models via a single real parameter. We then use this result to show that there exists a model which is optimal in the sense that minimizes the maximal weighted infinity norm of the error between the chosen model and all members of the equivalence class. This model is unique and is independent of the weighting function used in the infinity norm. It is thus the natural choice to be used in applications such as robust control. The result is also compared with more conventional estimates provided by prediction error methods.

Keywords: system identification, errors in variables, robust control, discrete-time systems, sampled-data systems.

1. INTRODUCTION.

The topic of errors-in-variables identification has attracted ongoing interest for almost a century. Work on the non-dynamic case goes back to Gini (1921) and Frisch (1934). These results have also been embellished by Madansky (1959), Moran (1971), Kalman (1981), Anderson *et al.* (1996), among others. The dynamic case has also attracted substantial interest in the engineering and statistics literature-see for example (Söderström, 1981; Green and Anderson, 1986; Tugnait, 1992; Nowak, 1992; Pintelon and Schoukens, 2001).

One conclusion from the above work is that there exists an equivalence class of models which are indistinguishable in term of their second-order input-output behavior. For example, Anderson (1985) has given a succinct description of the equivalence class and has shown that it is finitely parameterized in terms of $(N + 1)$ real variables, provided it is known that the *plant* has N zeros outside the unit circle. The case of white measurement noise has been studied in Stoica and Nehorai (1987) and in Chen (2003). An overview of errors-in-variables estimation is given in Söderström *et al.* (2002).

Our result builds on the earlier work of Anderson (1985). Indeed, it could be fairly said that the

[1] Email Addresses: juan.aguero@newcastle.edu.au, graham.goodwin@newcastle.edu.au, msb@elo.utfsm.cl

result is implicit in this earlier paper. However we see advantages in making the result explicit since it has important implications in the context of robust control design. In particular, we show that, by introducing a very mild restriction, we are able to parameterize the equivalence class of input-output compatible models by a single real parameter. We then utilize this result to show that there exists a model which minimizes the maximal weighted infinity norm of the error between the chosen model and all members of the equivalence class. This model is independent of the weighting function used in the infinity norm. It is thus, a *natural choice*. For example, it would give the best worst case performance in the context of robust control. We also compare the result with estimates provided by prediction error methods.

The layout of the reminder of the paper is as follows: in section 2, we describe the errors-in-variables estimation problem. In section 3, we discuss our main result. In section 4, we make some comments regarding robust control issues, and finally we sketch some conclusions in section 5.

2. MODEL SETUP

Consider the setup shown in Figure 1

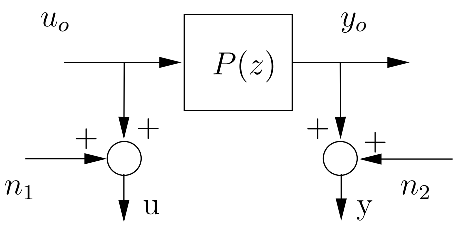

Fig. 1. Errors-in-variables model

We assume that $\{u_o(t)\}$, $\{n_1(t)\}$ and $\{n_2(t)\}$, are independent stationary stochastic processes described by:

$$u_o(t) = L_o(z)\bar{u}_o(t) \quad (1)$$
$$n_1(t) = L_1(z)\bar{n}_1(t) \quad (2)$$
$$n_2(t) = L_2(z)\bar{n}_2(t) \quad (3)$$

with $L_o(z)$, $L_1(z)$ and $L_2(z)$ are **normalized** spectral factors, and where the variances of $\bar{u}_o(t)$, $\bar{n}_1(t)$ and $\bar{n}_2(t)$ are σ_o^2, σ_1^2 and σ_2^2 respectively.

The problem of interest is to make statements about the transfer function $P(z)$ from the second order statistics of $\{u(t)\}$ and $\{y(t)\}$.

We introduce the following assumptions:

A.1 L_o, L_1 and L_2 are rational of finite order
A.2 P is strictly stable
A.3 P contains no zero, z_o, satisfying the symmetry conditions

$$P(z_o) = 0: \quad \text{and } P(z_o^{-1}) = 0$$

A.4 P contains no pole or zero ζ, that is either a pole or zero of $L_o(z)$ or $L_o(z^{-1})$, i.e. such that

$$L_o(\zeta) = 0 \textit{ or } L_o(\zeta^{-1}) = 0$$

Remark Note that assumptions (A.3) and (A.4) are relatively mild. They are, of course, restrictive. However, without them, one must resort to the general result of Anderson (1985) and its multivariable generalization (Green and Anderson, 1986). Assumptions A.3 and A.4 rule out various singular cases. These cases lie in a set of measure zero in the set of all systems. This, in itself, is not necessarily important. After all, systems containing a pure integrator also lie in a set of measure zero yet they are, nonetheless, of practical importance. Perhaps it is simply a matter of opinion as to which assumptions one deems *reasonable* and which one does not. In this context, it is interesting to compare assumptions A.3 and A.4 with other assumptions made in the literature. For example, in Stoica and Nehorai (1987) and Nowak (1992) it is shown that unique identifiability is possible, provided certain restrictions are placed on the noise structure and/or the degrees of the AR and MA components of the various signal models. Assumptions A.3 and A.4 appear, at least to the current authors, to be less restrictive than these latter assumptions.

$\nabla\nabla\nabla$

We will consider asymptotic analysis. In this context, most identification procedures can be shown (Ljung, 1999) to be equivalent to analyzing the joint input-output spectrum which is here given by:

$$S(e^{j\omega}) = \begin{bmatrix} S_{11}(e^{j\omega}) & S_{12}(e^{j\omega}) \\ S_{21}(e^{j\omega}) & S_{22}(e^{j\omega}) \end{bmatrix} \quad (4)$$

where

$$S_{11}(e^{j\omega}) = \sigma_o^2|P(e^{j\omega})L_o(e^{j\omega})|^2 + \sigma_2^2|L_2(e^{j\omega})|^2 \quad (5)$$
$$S_{12}(e^{j\omega}) = \sigma_o^2 P(e^{j\omega})|L_o(e^{j\omega})|^2 \quad (6)$$
$$S_{21}(e^{j\omega}) = \sigma_o^2 P(e^{-j\omega})|L_o(e^{j\omega})|^2 \quad (7)$$
$$S_{22}(e^{j\omega}) = \sigma_o^2|L_o(e^{j\omega})|^2 + \sigma_1^2|L_1(e^{j\omega})|^2 \quad (8)$$

3. THE MAIN RESULT

The main result of the current paper is

Theorem 1. Subject to assumptions A.1 to A.4, then the equivalence class of compatible models is given by

$$\mathcal{P} = \left\{ \overline{P}(z) : \overline{P}(z) = \frac{S_{12}(z)}{\overline{\sigma}_o^2 |L_o(z)|^2} \right\} \tag{9}$$

where $\overline{\sigma}_o^2$ is any positive real satisfying

$$\frac{|S_{12}(e^{j\omega})|^2}{S_{11}(e^{j\omega})} \leq \overline{\sigma}_o^2 |L_o(e^{j\omega})|^2 \leq S_{22}(e^{j\omega}) \ \forall \omega \tag{10}$$

and where $L_o(z)$ is uniquely determined by taking those poles and zeros from $S_{12}(z)$ which are symmetrically placed with respect to the unit circle, i.e. were $(z - c)$ and $(1 - zc)$ are factors.

Proof The key observation is that assumptions A.3 and A.4 allow $L_o(z)$ to be uniquely determined from $S_{12}(z)$. Then equation (6) immediately leads to (9).

Now we also require that

$$\overline{\sigma}_o^2 |L_o(e^{j\omega})|^2 + \sigma_1^2 |L_1(e^{j\omega})|^2 = S_{22}(e^{j\omega})$$

Thus

$$S_{22}(e^{j\omega}) - \overline{\sigma}_o^2 |L_o(e^{j\omega})|^2 \quad \text{must be nonzero } \forall \omega$$

Also from (4)-(8)

$$\frac{|S_{12}(e^{j\omega})|^2}{\overline{\sigma}_o^2 |L_o(e^{j\omega})|^2} + \sigma_2^2 |L_2(e^{j\omega})|^2 = S_{11}(e^{j\omega})$$

Thus we also require $\overline{\sigma}_o^2$ to satisfy

$$\overline{\sigma}_o^2 |L_o(e^{j\omega})|^2 \geq \frac{|S_{12}(e^{j\omega})|^2}{S_{11}(e^{j\omega})}$$

Thus the constraints on $\overline{\sigma}_o$ are that

$$\frac{|S_{12}(e^{j\omega})|^2}{S_{11}(e^{j\omega})} \leq \overline{\sigma}_o^2 |L_o(e^{j\omega})|^2 \leq S_{22}(e^{j\omega})$$

□□□

An immediate consequence of the above result is:

Corollary 1. Consider the equivalence class $\mathcal{P}$. Then there exists a unique model $\overline{P}^{\infty}(z) \in \mathcal{P}$ which satisfies

$$\overline{P}^{\infty}(z) = \arg \min_{\overline{P} \in \mathcal{P}} \max_{P \in \mathcal{P}} \left\{ |W(z)(\overline{P}(z) - P(z))|_{\infty} \right\} \tag{11}$$

$$\overline{P}^{\infty}(e^{jw}) = \frac{1}{2}[\overline{P}_{max} + \overline{P}_{min}] \tag{12}$$

$$\overline{P}_{max} = \frac{S_{11}}{S_{12}^*} \tag{13}$$

$$\overline{P}_{min} = \frac{S_{12}}{S_{22}} \tag{14}$$

for any weighting function $W(z)$.

Proof We note from Theorem 1 that at any frequency ω, every $\overline{P}(z) \in \mathcal{P}$ satisfies

$$\overline{P}(e^{j\omega}) = \frac{S_{12}(e^{j\omega})}{\overline{\sigma}_o^2 |L_o(e^{j\omega})|^2}$$

where $\overline{\sigma}_o^2 > 0$. Since all frequency responses lie on a straight line passing through the origin but not including the origin, then the solution to (11) is clearly (see Figure 2)

$$\overline{P}^{\infty}(z) = \frac{S_{12}(z)}{[\overline{\sigma}_o^{\infty}]^2 |L_o(z)|^2}$$

where

$$\frac{1}{[\overline{\sigma}_o^{\infty}]^2} = \frac{1}{2} \left[\frac{1}{\overline{\sigma}_{min}^2} + \frac{1}{\overline{\sigma}_{max}^2} \right]$$

and where $\overline{\sigma}_{min}$ and $\overline{\sigma}_{max}$ are the minimum and maximal values of $\overline{\sigma}$ such that (10) is satisfied.

□□□

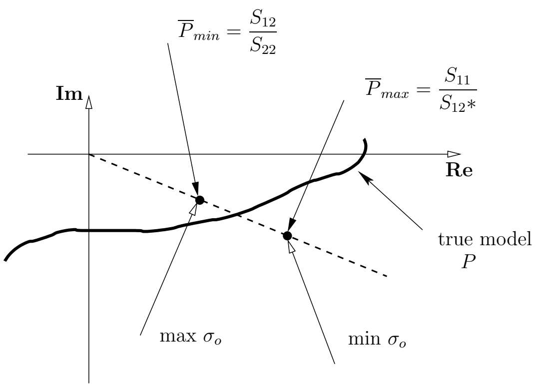

Fig. 2. Nyquist Plot (The dashed line represents all possible estimates)

4. IMPLICATIONS IN ROBUST CONTROL

Since $\overline{P}^{\infty}(z)$ minimizes the maximal weighted infinity norm of the error among the class P, then it is a natural choice in the context of robust control since it will yield the smallest value of $|T_o\overline{P}^{-1}(P-\overline{P})|_\infty$, where T_o is any nominal complementary sensitivity function. This is the best one can hope for since all members of $\mathcal{P}$ are indistinguishable from the given second order data on $\{y(t),u(t)\}$

Notice that the bias in the result is minimized over the class of compatible models. This result should be compared with estimates blindly obtained by other methods (e.g. prediction error methods (PEM) , (Ljung, 1999)). The latter schemes will, in general, yield results which are at the extreme edge of the equivalence class. This is illustrated below.

Theorem 2. Let us assume that the system is estimated by a PEM based on the following hypothesized model:

$$y(t) = Gu(t) + H\epsilon(t) \tag{15}$$

where G and H are independently parameterized.

Then, the asymptotic PEM estimate is given by

$$\hat{G} = \frac{1}{1+\rho}P = \overline{P}_{min}$$

$$\rho = \frac{\Phi_{n_1}}{\Phi_{u_o}} = \frac{|L_1|^2\sigma_1^2}{|L_o|^2\sigma_o^2}$$

where Φ_{n_1} and Φ_{u_o} are the input noise and input spectra respectively.

Proof The prediction error is given by

$$\epsilon(t) = H^{-1}[y(t) - Gu(t)] \tag{16}$$
$$= H^{-1}[\tilde{G}u(t) + x(t)] \tag{17}$$

where $\tilde{G} = P - G$, and $x(t) = n_2(t) - Pn_1(t)$.

Then, the prediction error spectrum is given by

$$\Phi_\epsilon = \frac{1}{|H|^2}\left[\tilde{G}\ 1\right]\Phi_X\begin{bmatrix}\tilde{G}\\ 1\end{bmatrix}^* \tag{18}$$

where

$$\Phi_X = \begin{bmatrix}\Phi_u & \Phi_{ux}\\ \Phi_{xu} & \Phi_x\end{bmatrix} \tag{19}$$

and $\Phi_u = S_{22}(e^{jw})$.

It is possible to write the prediction error spectrum as:

$$\begin{aligned}\Phi_\epsilon =& \frac{\Phi_u}{|H|^2}\left(\tilde{G}+\frac{\Phi_{xu}}{\Phi_u}\right)\left(\tilde{G}+\frac{\Phi_{xu}}{\Phi_u}\right)^*\\ &+\frac{1}{|H|^2}\left[\Phi_x - \frac{|\Phi_{xu}|^2}{\Phi_u}\right]\end{aligned} \tag{20}$$

We also have that, since u_o, n_1, and n_2 are uncorrelated:

$$\Phi_{xu} = -P\Phi_{n_1} = -P|L_1|^2\sigma_1^2 \tag{21}$$

Then, assuming that G, and H are independently parameterized we have that the asymptotic estimate is given by:

$$\hat{G} = P + \frac{\Phi_{xu}}{\Phi_u} = P\left(1 - \frac{\Phi_{n_1}}{\Phi_u}\right) \tag{22}$$

We also have that

$$\frac{\Phi_{n_1}}{\Phi_u} = \frac{\Phi_{n_1}}{\Phi_{u^o}+\Phi_{\tilde{u}}} = \frac{\rho}{1+\rho} \tag{23}$$

Then, noting that $S_{12} = P\Phi_{u_o}$, we obtain the result.

□□□

Remark The above asymptotic result for PEM shows that $\hat{G}$ is a scaled version of P, where the scaling is frequency dependent. Notice that the bias in $\hat{G}$ can be larger or smaller than that achieved by $\overline{P}^\infty$ depending on the value of ρ. On the other hand, Corollary 1 shows that $\overline{P}^\infty$ minimizes the maximal bias among the indistinguishable members of the equivalence class $\mathcal{P}$.

▽▽▽

5. CONCLUSIONS

This paper has given a parameterization of the equivalence class of all compatible models in dynamic errors-in-variables estimation. A mild restriction has been introduced ensuring that this equivalence class is parameterized by a single real variable. Comparisons with the estimates by prediction error methods have been given. Implications of the result in robust control have also been briefly examined.

REFERENCES

Anderson, B. D. O. (1985). Identification of scalar errors-in-variables models with dynamics. *Automatica* **Vol. 21, No 6**, 709–716.

Anderson, B. D. O., M. Deistler and W. Scherrer (1996). Solution set properties for static errors-in-variables problems. *Automatica* **Vol. 32, No 7**, 1031–1035.

Chen, H. F. (2003). Strongly consistent parameter estimate for error-in-variables model. *13th IFAC Symposium on System Identification, Rotterdam, The Netherlands.*

Cheng, C. L. and J. W. Van Ness (1999). *Kendall's library of statistics: Statistical Regression with measurement error.* Vol. 6. Edward Arnold.

Frisch, R. (1934). Statistical confluence analysis by means of complete regression systems. *Pub. No 5, University of Oslo, Economics Institute*.

Fuller, W. A. (1987). *Measurement error models*. Wiley.

Gini, C. (1921). Sull interpolazione di una retta quando i valori della variable independente sono affetti da errori accidentali. *Metron 1, 63*.

Green, M. and B. D. O. Anderson (1986). Identification of multivariable errors in variables models with dynamics. *IEEE Transactions on Automatic Control* **Vol. AC-31, No 5**, 467–471.

Kalman, R. E. (1981). System identification for noisy data. *Proc. Int. Symp. on Dyn. Syst., Gainesville, Florida*.

Ljung, L. (1999). *System Identification: Theory for the user*. 2nd ed.. Prentice Hall.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to errors. *Journal of the american statistical association* **Vol. 54, No 285**, 173–205.

Moran, P. A. P. (1971). Estimating structural and functional relationships. *Journal of multivariable analysis*.

Nowak, E. (1992). Identifiability in multivariable dynamic linear error-in-variables models. *Journal of the american statistical association* **Vol. 87, No 419**, 714–723.

Pintelon, R. and J. Schoukens (2001). *System Identification: A frequency domain approach*. IEEE press.

Söderström, T. (1981). Identification of stochastic linear systems in presence of input noise. *Automatica* **Vol. 17, No 5**, 713–725.

Söderström, T., U. Soverini and K. Mahata (2002). Perspectives on errors-in-variables estimation for dynamic systems. *Signal Processing* **Vol. 82**, 1039–1154.

Stoica, P. and A. Nehorai (1987). On the uniqueness of prediction error models for systems with noisy input-output data. *Automatica* **Vol. 23, No 4**, 541–543.

Tugnait, J. K. (1992). Stochastic system identification with noisy input using cumulant statistics. *IEEE Transactions on Automatic Control* **Vol. 37, No 4**, 476–485.

Van Huffel, S. and J. Vandewalle (1991). *The Total Least Squares Problem: Computational aspects and analysis*. SIAM Frontiers Applied Math. Philadelphia.