15th Triennial World Congress, Barcelona, Spain

SOME PECULIARITIES OF IDENTIFICATION IN THE PRESENCE OF MODEL ERRORS

R. Pintelon and J. Schoukens

Vrije Universiteit Brussel, Department ELEC, Pleinlaan 2, 1050 Brussel, Belgium

Abstract: Modelling errors are often the limiting factor in identification problems. Therefore, it is important to qualify their impact on the estimated plant model parameters $\hat{\theta}$. This paper qualifies the influence of model errors and disturbing noise level on (i) the asymptotic value θ_* (estimate for an infinite amount of data) of $\hat{\theta}$, and (ii) the asymptotic (amount of data going to infinity) covariance matrix $\mathrm{Cov}(\hat{\theta})$ of $\hat{\theta}$. The theory is elaborated on a time domain and a frequency domain estimator. *Copyright © 2002 IFAC*

Keywords**:** system identification, modelling errors, asymptotic properties, time-domain method, frequency-domain method.

1. INTRODUCTION

Although real life systems are mostly nonlinear and distributed, they are often approximated by a linear lumped model (Ljung, 1999; Pintelon and Schoukens, 2001a). One of the reasons for this is the success of linear lumped modelling in practical applications (e.g. control, forecasting, physical interpretation, measurement …) combined with the difficulty of nonlinear (distributed) modelling. Hence, if the observation time is sufficiently long, modelling errors are often the limiting factor in system identification problems. Therefore, it is important to study their impact on the estimated plant model parameters.

In Ljung (1999) and Pintelon and Schoukens (2001a) the asymptotic (amount of data going to infinity) properties of the estimated plant model parameters $\hat{\theta}$ have been studied in the presence of modelling errors, and general closed form expressions for $\mathrm{Cov}(\hat{\theta})$ are available. Since these expressions are not tractable in the presence of model errors, numerical methods for calculating $\mathrm{Cov}(\hat{\theta})$ have been derived in Hjalmarsson and Ljung (1992) and Tjärnström and Ljung (1999). These numerical techniques, however, give no qualitative insight in the influence of model errors and disturbing noise on $\mathrm{Cov}(\hat{\theta})$.

The contribution of this paper is to give a qualitative study of the influence of model errors and noise level on (i) the asymptotic value θ_* (estimate for an infinite amount of data) of $\hat{\theta}$ (see Section 3), and (ii) the asymptotic (amount of data going to infinity) covariance matrix $\mathrm{Cov}(\hat{\theta})$ of $\hat{\theta}$ (see Section 4). Starting from the asymptotic properties of $\hat{\theta}$ shown in Ljung (1999) and Pintelon and Schoukens (2001a), and without any random behaviour assumption about the model errors, some surprising qualitative conclusions are drawn concerning the dependency of θ_* and $\mathrm{Cov}(\hat{\theta})$ on the noise level. Although the analysis is carried out on two particular estimators (the prediction error method and the sample maximum likelihood method, see Section 2), the same reasoning is applicable to a general class of identification methods. The theoretical results are illustrated by simulation examples in Section 5.

This work is supported by the Belgium National Fund for Scientific Research; the Belgian Government as a part of the Belgian Programme on Interuniversity Poles of Attraction (IUAP4/2) initiated by the Belgian State, Prime Minister's Office, Science Policy Programming; and the Flemish Community (GOA-IMMI).

Due to space limitations the proofs of the theorems are not included in this paper. The reader is referred to Pintelon and Schoukens (2001b) for the complete version.

2. THE ESTIMATORS AND THE SYSTYEM

The theory is elaborated for the prediction error method (see Ljung, 1999) and the sample maximum likelihood method (see Pintelon and Schoukens, 2001a). Both methods identify the plant model parameters θ from input-output observations Z of a system (see Fig. 1). For the open loop experiment (see Fig. 1, solid line) the plant may be nonlinear, so that the plant model errors are due to unmodelled dynamics and/or nonlinear distortions. For the closed loop experiment (see Fig. 1, solid and dashed lines) we explicitly assume that the plant is linear so that in this case the plant model errors are due to unmodelled dynamics only.

The prediction error (PE) approach starts from a general output error stochastic framework where the input $u_0(t)$ in Fig. 1 is observed without errors ($m_u = 0$):

$$\begin{aligned} y(t) &= y_0(t) + n_y(t) \\ u(t) &= u_0(t) \end{aligned}, \quad t = 0, 1, \ldots, N-1 \tag{1}$$

with N the number of time domain samples; $n_y = n_p + m_y$ in open loop; and $n_y = n_p$ in closed loop ($m_y = 0$). $u_0(t)$ is an arbitrary, piecewice constant excitation; $y_0(t)$ is the noiseless output signal; $n_y(t)$ is the output error modelled as $n_y(t) = H_0(q)e(t)$, with q the backward shift operator, $H_0(z^{-1})$ a monic rational form in z^{-1}, and $e(t)$ a zero mean white noise source ($E\{e(t_1)e(t_2)\} = \sigma^2\delta(t_1 - t_2)$). The prediction error estimate $\hat{\theta}_{\mathrm{PE}}$ minimizes

$$\begin{aligned} V_{\mathrm{PE}}(\theta, Z) &= \frac{1}{2}\sum_{t=0}^{N-1} \varepsilon^2(t, \theta) \\ &= \frac{1}{2}\sum_{t=0}^{N-1} \left|H^{-1}(q, \theta)(y(t) - G(q, \theta)u(t))\right|^2 \end{aligned} \tag{2}$$

w.r.t. θ. Z contains the input/output observations

$$Z = \left[y(0)\ u(0)\ y(1)\ u(1)\ \ldots\ y(N-1)\ u(N-1)\right]^T \tag{3}$$

$H(z^{-1}, \theta)$ is the parametric noise model (monic rational form in z^{-1}), and $G(z^{-1}, \theta)$ the discrete-time plant model (rational form in z^{-1}). Both models may have common parameters (see Ljung, 1999). Note that the output error $n_y(t)$ in (1) is correlated with the input $u_0(t)$ for a feedback experiment.

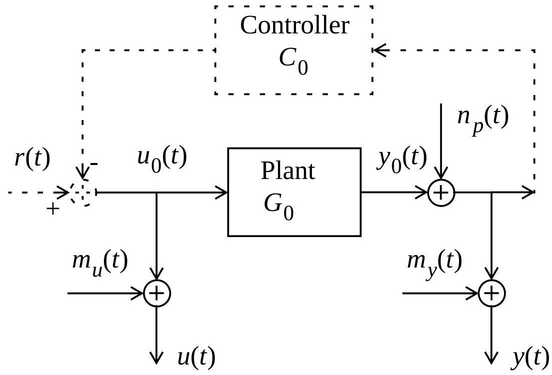

Fig. 1. Identification of a plant in open loop (solid line) or in closed loop (solid and dashed lines). $r(t)$ is the reference signal; $u_0(t)$, $y_0(t)$ the input-output signals; $m_u(t)$, $m_y(t)$ the input-output measurement errors; and $n_p(t)$ the process noise.

The sample maximum likelihood (SML) method starts from an errors-in-variables stochastic framework formulated in the frequency domain where $r(t)$ in Fig. 1 is a periodic signal:

$$\begin{aligned} Y(k) &= Y_0(k) + N_Y(k) \\ U(k) &= U_0(k) + N_U(k) \end{aligned}, \quad k = 1, 2, \ldots, F \tag{4}$$

with F the number of frequencies; and

$$\begin{aligned} &\text{open loop:} \quad \begin{cases} N_Y = N_P + M_Y \\ N_U = M_U \end{cases} \\ &\text{closed loop:} \quad \begin{cases} N_Y = N_P/(1 + G_0C_0) + M_Y \\ N_U = -N_PC_0/(1 + G_0C_0) + M_U \end{cases} \end{aligned} \tag{5}$$

(any deviation from the periodic behaviour is considered as noise). $U_0(k)$, $Y_0(k)$ in (4) are the DFT spectra of the PERIODIC PART of the input-output signals $u_0(t)$, $y_0(t)$

$$X(k) = N^{-1/2}\sum_{t=0}^{N-1} x(t)\exp(-j2\pi tk/N) \tag{6}$$

with $x = u, y$ and $X = U, Y$; and $N_U(k)$, $N_Y(k)$ are independent (over the frequency k) zero mean, jointly correlated Gaussian disturbances. Assuming that M independent experiments are available (in practice M consecutive periods of the steady state response to the periodic excitation), the sample maximum likelihood estimate $\hat{\theta}_{\mathrm{SML}}$ minimizes

$$V_{\mathrm{SML}}(\theta, Z) = \sum_{k=1}^{F} \left|\hat{\varepsilon}(\Omega_k, \theta, Z(k))\right|^2 \tag{7}$$

$$= \sum_{k=1}^{F} \frac{|Y(k) - G(\Omega_k, \theta)U(k)|^2}{\hat{\sigma}_Y^2(\Omega_k, \theta)}$$

$$\hat{\sigma}_Y^2(\Omega_k, \theta) = \hat{\sigma}_Y^2(k) + \hat{\sigma}_U^2(k)|G(\Omega_k, \theta)|^2 - 2\mathrm{Re}(\hat{\sigma}_{YU}^2(k)\bar{G}(\Omega_k, \theta))$$

w.r.t. θ (overbar denotes the complex conjugate). Z contains the input/output DFT spectra at the excited frequencies

$$Z = [Z^T(1) \ \ldots \ Z^T(F)]^T, Z^T(k) = [Y(k) \ U(k)] \quad (8)$$

$G(\Omega_k, \theta)$ is the plant model (rational form in Ω); Ω the generalised frequency variable: $\Omega = z^{-1}$ for discrete-time systems, and $\Omega = s$ for continuous-time systems; and $\hat{\sigma}_U^2(k)$, $\hat{\sigma}_Y^2(k)$ and $\hat{\sigma}_{YU}^2(k)$ are the sample covariances (= non-parametric noise model)

$$\begin{aligned}
M\hat{\sigma}_Y^2(k) &= \frac{1}{M-1}\sum_{m=1}^{M} |\hat{N}_Y^{[m]}(k)|^2 \\
M\hat{\sigma}_U^2(k) &= \frac{1}{M-1}\sum_{m=1}^{M} |\hat{N}_U^{[m]}(k)|^2 \\
M\hat{\sigma}_{YU}^2(k) &= \frac{1}{M-1}\sum_{m=1}^{M} \hat{N}_Y^{[m]}(k)\overline{\hat{N}_U^{[m]}}(k) \\
\hat{N}_Y^{[m]}(k) &= Y^{[m]}(k) - Y(k) \\
\hat{N}_U^{[m]}(k) &= U^{[m]}(k) - U(k) \\
Y(k) &= \frac{1}{M}\sum_{m=1}^{M} Y^{[m]}(k) \\
U(k) &= \frac{1}{M}\sum_{m=1}^{M} U^{[m]}(k)
\end{aligned} \quad (9)$$

and where $U^{[m]}(k)$, $Y^{[m]}(k)$ are the input-output DFT spectra of the mth independent experiment (mth signal period). Note that the input-output errors $N_U(k)$, $N_Y(k)$ in (4) are ALWAYS independent of $U_0(k)$, even for a feedback experiment.

3. ASYMPTOTIC VALUE IDENTIFIED PLANT MODEL

The prediction error (PE) and sample maximum likelihood (SML) estimators minimize a quadratic like cost function

$$V_N(\theta, Z) = \frac{1}{2N}\varepsilon^T(\theta, Z)\varepsilon(\theta, Z) \quad (10)$$

where the residual $\varepsilon(\theta, Z) \in \mathbb{R}^N$ is a measure of the difference between the observations Z and the model, and N is a measure for the amount of data ($V_N(\theta, Z) = V_{\mathrm{PE}}(\theta, Z)/N$ for PE, and $V_N(\theta, Z) = V_{\mathrm{SML}}(\theta, Z)/N$ with $N = F$ for SML). The noisy observations Z are related to the true values Z_0 as $Z = Z_0 + N_Z$ with N_Z a zero mean disturbance with covariance matrix $\mathrm{Cov}(N_Z)$

$$\begin{aligned}
\mathrm{PE}: N_Z &= [n_y(0) \ m_u(0) \ \ldots \ n_y(N-1) \ m_u(N-1)]^T \\
\mathrm{SML}: N_Z &= [N_Y(1) \ N_U(1) \ \ldots \ N_Y(F) \ N_U(F)]^T
\end{aligned} \quad (11)$$

with $N_Y(k), N_U(k)$, and $n_y(t)$ defined in Section 2, and $m_u(t)$ the input measurement error (see Fig. 1). The limit $V_*(\theta)$ of the expected value of the cost function (10)

$$V_*(\theta) = \lim_{N \to \infty} V_N(\theta) = \lim_{N \to \infty} E\{V_N(\theta, Z)\} \quad (12)$$

plays a central role in the analysis of the asymptotic ($N \to \infty$) properties of the minimizer $\hat{\theta}$ of (10). Indeed, under some suitable assumptions it can be shown that $\hat{\theta}$ converges for $N \to \infty$ with probability 1 to the minimizer θ_* of (12)

$$(\hat{\theta} = \arg\min_{\theta \in \Theta} V_N(\theta, Z)) \to (\theta_* = \arg\min_{\theta \in \Theta} V_*(\theta)) \quad (13)$$

with Θ a compact set where the cost function and its limit value are "well behaved" (continuous and existing higher order derivatives). For more details see, for example, Ljung (1999) for the prediction error method, and Pintelon and Schoukens (2001a) for the sample maximum likelihood method. Contrary to what has been assumed in (13), $V_*(\theta)$ may have more than one global minimum in case of model errors. An example of this can be found in Kabaila (1983). To handle these cases we restrict the compact set Θ in (13) such that $V_*(\theta)$ has a unique global minimum.

To analyse the dependency of θ_* on the noise level, we replace N_Z by υN_Z and hence $\mathrm{Cov}(N_Z)$ by $\upsilon^2\mathrm{Cov}(N_Z)$ in (12), with υ a real number. If this transforms the θ-dependent part $V_1(\theta)$ of $V_*(\theta)$ to $g(\upsilon)V_1(\theta)$, where $g(\upsilon)$ is independent of θ, then θ_* is independent of the noise level υ, and it makes sense to define θ_* as the noiseless solution (= the estimate $\hat{\theta}$ one would get if $\upsilon \to 0$ and $N \to \infty$). Note, however, that the noiseless solution θ_* defined in this way is only insensitive w.r.t. noise changes where the standard deviation of ALL disturbing noise sources is multiplied with the same factor υ. Hence, it may still depend on the noise colouring and the noise covariance matrix $\mathrm{Cov}(N_Z)$, for example, the ratio of the output variance $\sigma_Y^2(k)$ to the input variance $\sigma_U^2(k)$. If $V_1(\theta)$ is not transformed to $g(\upsilon)V_1(\theta)$ then, in general (but this should be verified for each case), θ_* will depend on the noise level. The results for $\theta_{*\mathrm{PE}}$ and $\theta_{*\mathrm{SML}}$ are given in the following theorem.

Theorem 1 (influence noise level υ on θ_*): In the presence of plant model errors, $G(\Omega, \theta_*) \neq G_0(\Omega)$ with $G_0(\Omega)$ the true plant model, the limit value θ_* of the estimated plant model parameters $\hat{\theta}$ has the following properties

1. the limit value $\theta_{*\mathrm{PE}}$ of the prediction error estimate $\hat{\theta}_{\mathrm{PE}}$ depends on the noise level υ, even in the absence of measurement noise (see Fig. 1, $m_u = 0$ and $m_y = 0$), except for the output error model structure, $H(z^{-1}, \theta) = 1$, identified in
   - 1.a open loop without input measurement noise (see Fig. 1, $m_u = 0$), even if the true noise model is not white, $H_0 \neq 1$,
   - 1.b closed loop without input-output measurement errors (see Fig. 1, $m_u = 0$ and $m_y = 0$), if the true noise model is white, $H_0 = 1$.
2. the limit value $\theta_{*\mathrm{SML}}$ of the sample maximum likelihood estimate $\hat{\theta}_{\mathrm{SML}}$ is independent of the noise level υ.

Proof: see Pintelon and Schoukens (2001b). G

Recall that the plant is by assumption linear for the closed loop experiment and possibly nonlinear for the open loop experiment. From Theorem 1 it follows that sample maximum likelihood estimate $\hat{\theta}_{\mathrm{SML}}$ converges to the solution of the noiseless problem which is obtained by decreasing the noise level to zero while maintaining the noise colouring (linear/nonlinear plants), or by increasing the excitation level to infinity while maintaining the colouring of its power spectrum (linear plants only). This is not true for the prediction error estimate $\hat{\theta}_{\mathrm{PE}}$: if the identification experiment is repeated with a different noise and/or excitation level then $\hat{\theta}_{\mathrm{PE}}$ converges to other limit values $\theta_{*\mathrm{PE}}$. The reason for the different behaviour of $\hat{\theta}_{\mathrm{PE}}$ and $\hat{\theta}_{\mathrm{SML}}$ is that the sample (co-)variances (= non-parametric noise model) used in (7) are estimated independently of the plant model, while the estimated parametric noise model in (2) strongly depends on the plant model errors. It can be concluded that estimating (in time or frequency domain) with non-parametric noise models, is less sensitive to the experimental conditions (noise level, excitation level) than with parametric noise models.

4. UNCERTAINTY IDENTIFIED PLANT MODEL

Define $\tilde{\theta}$ as the minimizer of the expected value of the cost function (10)

$$\tilde{\theta} = \arg\min_{\theta \in \Theta} V_N(\theta) \text{ with } V_N(\theta) = E\{V_N(\theta, Z)\} \quad (14)$$

where we restrict the compact set Θ such that $V_N(\theta)$ has a unique global minimum (see also the discussion after eq. (13)). Under some suitable assumptions it can be shown that $\sqrt{N}(\hat{\theta} - \tilde{\theta})$ is asymptotically ($N \to \infty$) normally distributed with zero mean and covariance matrix P_N

$$\begin{aligned} &\sqrt{N}(\hat{\theta} - \tilde{\theta}) \in AsN(0, P_N) \\ &P_N = V_N''^{-1}(\tilde{\theta}) Q_N(\tilde{\theta}) V_N''^{-1}(\tilde{\theta}) \\ &Q_N(\tilde{\theta}) = N E\{V_N'^T(\tilde{\theta}, Z) V_N'(\tilde{\theta}, Z)\} \end{aligned} \quad (15)$$

with ' the derivative w.r.t. θ (Pintelon and Schoukens, 2001a). Note that the expectations in (15) are taken w.r.t. the disturbing noise AND the excitation signal. This should be kept in mind when studying the uncertainty of $\hat{\theta}$ in the presence of modelling errors. Using $\theta_* = \tilde{\theta} + O(N^{-1})$ (see Pintelon and Schoukens, 2001a, Theorems 7.21 and 8.3), and taking the limit of (15) for $N \to \infty$, gives

$$\begin{aligned} &\sqrt{N}(\hat{\theta} - \theta_*) \in AsN(0, P) \\ &P = V_*''^{-1}(\theta_*) Q V_*''^{-1}(\theta_*) \\ &Q = \lim_{N \to \infty} N E\{V_N'^T(\theta_*, Z) V_N'(\theta_*, Z)\} \end{aligned} \quad (16)$$

which is the classical result given in Ljung (1999). In the following theorem (15) is analysed for "small" noise errors N_Z and "small" plant model errors $\Delta G_* = G_0(\Omega) - G(\Omega, \theta_*)$. To get tractable qualitative expressions for P_N and P, we replace in (15) N_Z, $\mathrm{Cov}(N_Z)$, and ΔG_* by respectively υN_Z, $\upsilon^2 \mathrm{Cov}(N_Z)$, and $\mu \Delta G_*$, where $\upsilon \to 0$ and $0 < \mu \ll 1$ are "small" real numbers.

Theorem 2 (asymptotic expression P): For small noise levels, $N_Z \to \upsilon N_Z$ with $\upsilon \to 0$, and small plant model errors, $\Delta G_* \to \mu \Delta G_*$ with $0 < \mu \ll 1$, the asymptotic covariance matrix P in (16) of the estimated plant model parameters $\hat{\theta}$ is given by

$$P_{\mathrm{PE}} = \begin{cases} O(\mu^2) & \text{if } (\Delta G_* \neq 0) \text{ and (random input)} \\ O(\upsilon^2) & \text{if } (\Delta G_* = 0) \text{ or (deterministic input)} \end{cases} \quad (17)$$

for the prediction error method (2), and

$$P_{\mathrm{SML}} = \begin{cases} O(\mu^2) & \text{if } \Delta G_* \neq 0 \\ O(\upsilon^2) & \text{if } \Delta G_* = 0 \end{cases} \quad (18)$$

for the sample maximum likelihood method (7).
Proof: see Pintelon and Schoukens (2001b). G

The result $P = O(\mu^2)$ for $\Delta G_* \neq 0$ and $\upsilon \to 0$ can be understood as follows. In the presence of plant model errors, $\Delta G_* \neq 0$, the estimate $\hat{\theta}$ depends on the periodogram $|U_0(k)|^2$ of the input signal (see Ljung, 1999 and Pintelon and Schoukens, 2001a). For finite N and random input signals, $|U_0(k)|^2$ is a random variable. Hence, for small noise levels ($\upsilon \to 0$), the variability of $\hat{\theta}$ will be mainly due to the variability of $|U_0(k)|^2$ over different realisations of the input signal.

The different behaviour of $\hat{\theta}_{\mathrm{PE}}$ and $\hat{\theta}_{\mathrm{SML}}$ for deterministic inputs and plant model errors can be explained intuitively by the different stochastic behaviour of the parametric noise model and the sample (co-)variances. Indeed, for deterministic inputs and plant model errors, the uncertainty of the estimated parametric noise model $H(z^{-1}, \hat{\theta}_{\mathrm{PE}})$ tends to zero as $O(\upsilon)$ (for small noise levels the residuals $\varepsilon(t, \theta)$ in (2) are dominated by the plant model errors), while the uncertainty of the scaled sample (co-)variances $\hat{\sigma}_Y^2(k)/\upsilon^2$, $\hat{\sigma}_U^2(k)/\upsilon^2$ and $\hat{\sigma}_{YU}^2(k)/\upsilon^2$ is independent of the noise level υ. The technical reason for the different behaviour of $\hat{\theta}_{\mathrm{PE}}$ and $\hat{\theta}_{\mathrm{SML}}$ is that in eq. (15)

$$V_{\mathrm{PE}}'(\tilde{\theta}_{\mathrm{PE}}, Z_0) = E\{V_{\mathrm{PE}}'(\tilde{\theta}_{\mathrm{PE}}, Z_0)\} = 0 \qquad (19)$$

for deterministic inputs, while due to the uncertainty of the sample (co-)variances

$$V_{\mathrm{SML}}'(\tilde{\theta}_{\mathrm{SML}}, Z_0) \neq E\{V_{\mathrm{SML}}'(\tilde{\theta}_{\mathrm{SML}}, Z_0)\} = 0 \qquad (20)$$

even for deterministic inputs (see Pintelon and Schoukens 2001b).

Theorem 2 has also an impact on the ongoing discussion (see, for example, Tjärnström and Ljung, 2000) whether a low order approximation of a complex system should be obtained either via direct identification of the low order model, or via a two step procedure consisting of the estimation of a validated complex model (removing in some optimal way the disturbing noise) followed by a deterministic model reduction step (for example, in weighted L_2 or L_∞ sense). Indeed, from Theorem 2 it follows that asymptotically ($\upsilon \to 0$) the uncertainty of the two step procedure tends to zero ($P = O(\upsilon^2)$ for the first step and the second step is deterministic), while that of the direct approach is bounded below by $P = O(\mu^2)$ (except for the prediction error method with deterministic input where $P = O(\upsilon^2)$). Hence, the two step procedure should be preferred over the direct approach if the model errors are larger than the noise errors, except for the prediction error method with deterministic input where Theorem 2 cannot make any distinction.

5. SIMULATION RESULTS

As simulation example we take an open loop experiment without input measurement noise (see Fig. 1, $m_u = 0$), a second order plant

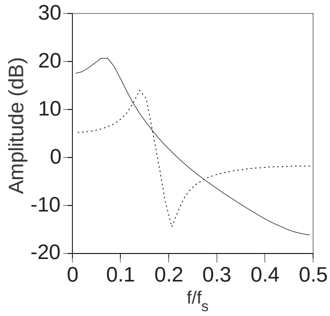

Fig. 2. True plant (solid line) and noise (dashed line) model.

$$G_0(z^{-1}) = \frac{z^{-1} + 0.5z^{-2}}{1 - 1.5z^{-1} + 0.7z^{-2}} \qquad (21)$$

and a second order disturbing noise process $n_p(t) + m_y(t) = H_0(q)e(t)$ with

$$H_0(z^{-1}) = \frac{1 - 0.53549z^{-1} + 0.88021z^{-2}}{1 - 1.11111z^{-1} + 0.85204z^{-2}} \qquad (22)$$

and $e(t) \in N(0, \sigma^2)$ (see Fig. 2). The plant is excited with periodic, normally $N(0, 1)$ distributed noise (period length of 150 samples). $M = 6$ consecutive periods of the steady state response are used to calculate the prediction error (PE) and the sample maximum likelihood (SML) estimates. The PE method (2) uses a Box-Jenkins model structure with a second order noise model, while the SML method (7) calculates the sample variance $\hat{\sigma}_Y^2(k)$ from the $M = 6$ consecutive periods. For each run of the Monte Carlo simulation a new disturbing noise sequence and a new excitation signal is generated.

Figures 3 and 4 show the results of a Monte-Carlo simulation (1000 runs for each disturbing noise level) for respectively a first and a second order plant model

$$G(z^{-1}, \theta) = \frac{b_1 z^{-1}}{1 + a_1 z^{-1}} \qquad \text{(a)}$$
$$G(z^{-1}, \theta) = \frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} \qquad \text{(b)} \qquad (23)$$

These two figures illustrate nicely the influence of model errors on the noise behaviour of the estimates. In the presence of model errors (see Fig. 3) it follows that (i) $G(z^{-1}, \theta_{*\mathrm{SML}})$ is independent of the noise level, while $G(z^{-1}, \theta_{*\mathrm{PE}})$ strongly depends on the noise level; and (ii) the sample standard deviations of $G(z^{-1}, \hat{\theta}_{\mathrm{SML}})$ and $G(z^{-1}, \hat{\theta}_{\mathrm{PE}})$ converge to a non-zero value as the noise level decreases to zero. Both observations are in agreement with Theorems 1 and 2. In the absence of model errors (see Fig. 4) it follows that (i) $G(z^{-1}, \theta_{*\mathrm{SML}})$ and $G(z^{-1}, \theta_{*\mathrm{PE}})$ are independent of the noise level ($\theta_{*\mathrm{SML}} = \theta_{*\mathrm{PE}} = \theta_0$); and (ii) the sample standard deviations of $G(z^{-1}, \hat{\theta}_{\mathrm{SML}})$ and $G(z^{-1}, \hat{\theta}_{\mathrm{PE}})$ converge to zero as the noise level decreases to zero.

6. CONCLUSION

Some qualitative analysis tools for studying the influence of the disturbing noise level and the modelling errors on the asymptotic value and the uncertainty of the estimated plant model parameters have been presented. The theory is illustrated on the prediction error (PE) and the sample maximum likelihood (SML) methods. The following peculiarities result: in the presence of plant model errors (i) the asymptotic value of the PE estimate strongly depends on the noise/excitation level, while that of the SML estimate is independent of the noise level (linear/nonlinear plants) and the excitation level (linear plants only); (ii) for random inputs the uncertainty of the PE and SML estimates does not decrease to zero for vanishing noise levels and is bounded below by the model errors; (iii) for deterministic inputs the uncertainty of the PE decreases to zero for vanishing noise levels while that of the SML estimates remains bounded below by the modelling errors.

REFERENCES

Hjalmarsson, H. and L. Ljung (1992). Estimating model variance in the case of undermodeling. *IEEE Trans. Autom. Contr.*, **vol. 37**, no. 7, pp. 1004-1008.

Kabaila, P. (1983). Parameter values of ARMA models minimizing the one-step-ahead prediction error when the true system is not in the model set. *J. Applied Prob.*, **vol. 20**, no. 2, pp. 405-408.

Ljung, L. (1999). *System Identification: Theory for the User.* Prentice-Hall, Upper Saddle River, NJ.

Pintelon, R. and J. Schoukens (2001a). *System Identification: a Frequency Domain Approach.* IEEE Press, Piscataway, NJ.

Pintelon, R. and J. Schoukens (2001b). Some peculiarities of identification in the presence of model errors. *Internal report no RP072001*, ReportRP072001.pdf @ ftp://elecftp.vub.ac.be/ Papers/RikPintelon/.

Tjärnström, F. and L. Ljung (1999). Estimating the variance in case of undermodeling using bootstrap, *Proceedings of the 38th IEEE Conference on Decision and Control*, Phoenix, Arizona (USA), pp. 2394-2399.

Tjärnström, F. and L. Ljung (2000). L_2-Model reduction and variance reduction. *Preprints of the 12th IFAC Symposium on System Identification*, Santa Barbara, California (USA), June 21-23.

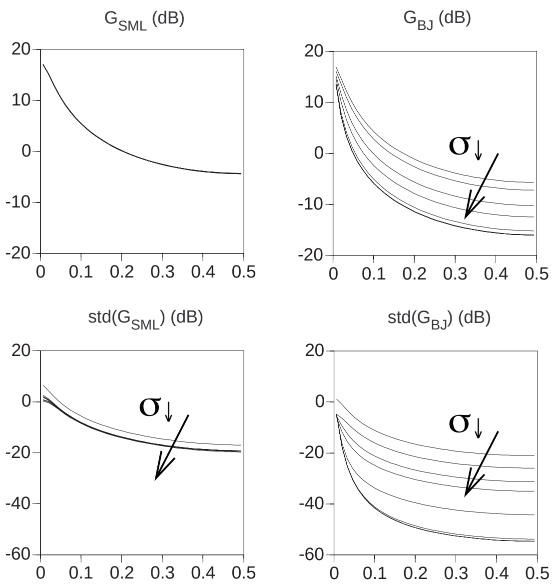

Fig. 3. Simulation results for a first order plant (eq. (20a): model errors). Estimated plant and its sample standard deviation as a function of the normalised frequency for eight different values of the noise level: $\sigma = 5, 2, 0.5, 0.1, 0.01, 0.001$ and 0.0001 (left: sample maximum likelihood, right: Box-Jenkins).

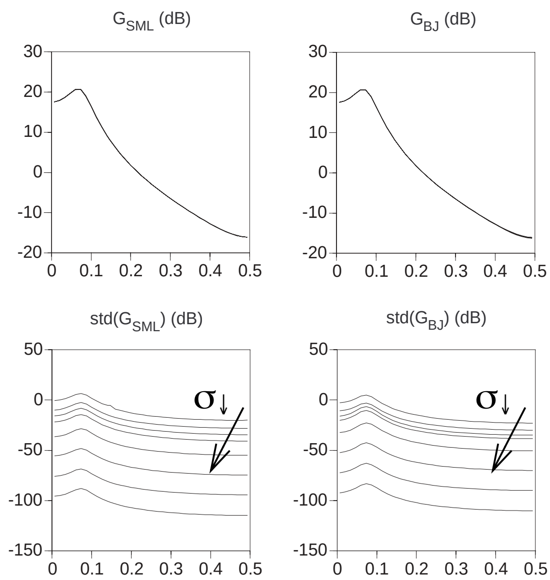

Fig. 4. Simulation results for a second order plant (eq.(20b): no model errors). Estimated plant and its sample standard deviation as a function of the normalised frequency for eight different values of the noise level: $\sigma = 5, 2, 0.5, 0.1, 0.01, 0.001$ and 0.0001 (left: sample maximum likelihood, right: Box-Jenkins).