

ACCURATE QUANTIFICATION OF VARIANCE ERROR

Brett Ninness ^{*,1} Håkan Hjalmarsson ^{**}

^{*} Dept. of Elec. & Comp. Eng, Uni. Newcastle, Australia.

email:brett@ee.newcastle.edu.au, FAX: +61 49 21 69 93

^{**} Dept. Sensors, Signals and Systems (Automatic Control), The Royal Institute of Technology, S-100 44 Stockholm, Sweden,

email:hakan.hjalmarsson@s3.e.kth.se, FAX: +46 8 790 7329

Abstract: This paper accurately quantifies the way in which noise induced estimation errors are dependent on model structure, underlying system frequency response, measurement noise and input excitation. This exposes several new principles. In particular, it is shown here that when employing Output-Error model structures in a prediction-error framework, then the ensuing estimate variability in the frequency domain depends on the underlying system pole positions. As well, it is also established that the variability is affected by the choice of model structure, in that it is twice as much when system poles are estimated as when they are a-priori known and fixed, even though the model order is the same in both cases. These results are unexpected according to pre-existing theory.

Keywords: Output Error Identification, Estimation Algorithms, Estimation Theory, Identification Algorithms, Frequency Domains, Modelling Errors.

1. INTRODUCTION

When identifying a system model on the basis of observed data, it is essential to quantify the likely error in that estimated model. Typically, this consists of two components. The first, a so-called “bias error”, is the result of the model structure being less complex than the system being estimated. The second, called “variance error”, is caused by corruption of the input-output data measurements.

When the latter can be modelled as an additive stochastic process, and the underlying system is linear, then it is arguable that the total error in any identified model that passes a validation test is dominated by variance error (Ljung and Guo 1997).

In this common case, the quantification of estimation error then becomes a question of assessing variance error. In relation to this, if the widely used prediction-error method with a quadratic criterion is employed, then a seminal result is that noise-induced error, as measured by the variability of the ensuing frequency response estimate $G(e^{j\omega}, \hat{\theta}_N^n)$, may be approximated as (L.Ljung 1985, L.Ljung and Z.D.Yuan 1985, Ljung 1999)

$$\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\} \approx \frac{m}{N} \frac{\sigma^2}{\Phi_u(\omega)}. \quad (1)$$

Here σ^2 and Φ_u are, respectively, the (white) measurement noise and (possibly coloured) input spectral densities, and $\hat{\theta}_N^n$ is the prediction error estimate based on N

observed data points of a vector $\theta^n \in \mathbf{R}^n$ that parameterises a model structure $G(q, \theta^n)$ for which (essentially) $m = \dim \theta^n / (2^d)$ where d is the number of denominator polynomials to be estimated in the model structure. In what follows in this paper, the Output-Error case of $m = n/2$ will be exclusively considered.

Apart from its simplicity, a key factor underlying the importance and popularity of the approximation (1) is that, according to its derivation (L.Ljung 1985, L.Ljung and Z.D.Yuan 1985, Ljung 1999), it applies for a very wide class of so-called ‘shift invariant’ model structures. For example, all the well known FIR, ARX, ARMAX, Output-Error and Box-Jenkins structures are shift invariant (L.Ljung 1985). As well, as shown in (Ljung 1999), it also applies when non-parametric (spectral based) estimation methods (Brillinger 1981, Ljung 1999) are employed provided that the m term in (1) is replaced by one dependent on the number of data points (and the windowing function) used.

Therefore, the only influence that the chosen model structure has on the right hand side of (1) is in terms of its order, and because of this the belief that $\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\}$ is invariant to the particular choice of m 'th order model structure has become a fundamental tenet of system identification.

Furthermore, it is also held as axiomatic that $\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\}$ does not depend on the underlying true frequency response, again on account of the right hand side of (1) being independent of that quantity; see, for example, the work (Forsell and Ljung 1999, Gevers *et al.* 2001, Forsell and Ljung 2000, Zhu 1998).

In relation to these conclusions, a series of recent contributions (Wahlberg 1991, Wahlberg 1994, P.M.J. Van den

¹ This work was supported by the Australian Research Council and the Centre for Integrated Dynamics and Control (CIDAC). It was completed while the first author was visiting the Department of Sensors, Signals and Systems (Automatic Control), The Royal Institute of Technology, Stockholm, Sweden.

Hof *et al.* 1995, Ninness *et al.* 1999b) has established a variance error quantification that is an extension of (1) and which is applicable to certain model structures which have poles or zeros fixed according to prior knowledge. For example, if an FIR structure is generalised so that its fixed poles $\{\xi_0, \dots, \xi_{m-1}\}$ are not necessarily all at the origin, then (Wahlberg 1991, Wahlberg 1994, P.M.J. Van den Hof *et al.* 1995, Ninness *et al.* 1999b) has shown that in the interests of maximally accurate approximation, the quantification (1) should be modified to become

$$\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\} \approx \frac{1}{N} \frac{\sigma^2}{\Phi_u(\omega)} \sum_{k=0}^{m-1} \frac{1 - |\xi_k|^2}{|e^{j\omega} - \xi_k|^2}. \quad (2)$$

Note that (2) reverts to (1) for the FIR case of $\xi_k = 0$. Furthermore, in (Ninness *et al.* 1999b) it has been shown that for ARX model structures with fixed noise model zeros, again not necessarily at points $\{\xi_0, \dots, \xi_{m-1}\}$ which are at the origin, then again the expression (2) rather than (1) should be used in the interests of providing the most accurate approximation of $\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\}$.

For both these generalised FIR and ARX cases where (2) is preferable, when actually computing $\hat{\theta}_N^n$, the process of incorporating the fixed poles or zeros may be achieved by first pre-filtering the input data with an all-pole filter $F(q)$, and then using a conventional FIR or ARX structure (Ninness *et al.* 1999b).

The previous work (Wahlberg 1991, Wahlberg 1994, P.M.J. Van den Hof *et al.* 1995, Ninness *et al.* 1999b) has therefore illustrated that the effect of pre-filtering on $\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\}$ cannot be accommodated by simply making the substitutions $\Phi_u \mapsto |F|^2 \Phi_u$, $\sigma^2 \mapsto |F|^2 \sigma^2$. Instead the filter pole locations $\{\xi_k\}$ must be directly accounted for via (2). This establishes that the quantification of $\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\}$ cannot be expressed in a manner that is invariant to the nature (roughly speaking, smoothness) of the input and noise spectral densities.

The contribution of this paper is to extend this result and in fact establish a more fundamental principle. Namely, despite widely held belief, $\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\}$ is *not* invariant to either the model structure, or the underlying true frequency response.

These new ideas are established by examining the estimation of Output–Error model structures in which case when $\{\xi_0, \dots, \xi_{n-1}\}$ are the estimated poles of $G(q, \hat{\theta}_N^n)$, then the main result of this paper is to establish that the expression

$$\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\} \approx \frac{2}{N} \frac{\sigma^2}{\Phi_u(\omega)} \sum_{k=0}^{m-1} \frac{1 - |\xi_k|^2}{|e^{j\omega} - \xi_k|^2} \quad (3)$$

is a significantly more accurate quantification than the widely held one (1).

As an example of the consequences of this expression, comparing (3) with (2) indicates that the variability $\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\}$ associated with estimating a model with fixed known poles $\{\xi_0, \dots, \xi_{m-1}\}$, and hence only estimating a numerator, is only one half the variability associated with a model where the poles are estimated. Since, roughly speaking, twice as much information is being estimated, this result makes intuitive sense. However, it is

completely at odds with pre-existing thought derived from (1) which (since the number of denominators polynomials $d = 1$ in both cases) would indicate that $\text{Var}\{G(e^{j\omega}, \hat{\theta}_N^n)\}$ is invariant to whether poles are estimated or not!

2. PROBLEM FORMULATION

The problem setting considered here is one in which a model structure is used to describe the relationship between an observed input data record $\{u_t\}$ and output data record $\{y_t\}$ as

$$y_t = G(q, \theta^n)u_t + e_t. \quad (4)$$

Here $\{e_t\}$ is a zero-mean white noise sequence that satisfies $\mathbf{E}\{e_t^2\} = \sigma^2$, $\mathbf{E}\{|e_t|^{4+\epsilon}\} < \infty$ for some $\epsilon > 0$ and $G(q, \theta^n)$ is a transfer function, rational in the forward shift operator q , and parameterised by a vector $\theta^n \in \mathbf{R}^n$.

In this case, the relationship (4) is commonly known as an ‘Output–Error’ model structure, and the mean-square optimal one-step ahead predictor $\hat{y}_t(\theta^n)$ based on this model structure is (Ljung 1999)

$$\hat{y}_t(\theta^n) = G(q, \theta^n)u_t$$

with associated prediction error

$$\varepsilon_t(\theta^n) \triangleq y_t - \hat{y}_t(\theta^n) = y_t - G(q, \theta^n)u_t. \quad (5)$$

Using this, a quadratic estimation criterion may be defined as

$$V_N(\theta^n) = \frac{1}{2N} \sum_{t=1}^N \varepsilon_t^2(\theta^n)$$

and then used to construct the prediction error estimate $\hat{\theta}_N^n$ of θ^n as

$$\hat{\theta}_N^n \triangleq \arg \min_{\theta^n \in \mathbf{R}^n} V_N(\theta^n). \quad (6)$$

As has been established in (Ljung 1999), under certain mild assumptions on the nature of the input $\{u_t\}$ (which will be discussed in detail later), the estimate $\hat{\theta}_N^n$ converges with increasing N according to

$$\lim_{N \rightarrow \infty} \hat{\theta}_N^n = \theta_\circ^n \triangleq \arg \min_{\theta^n \in \mathbf{R}^n} \lim_{N \rightarrow \infty} \mathbf{E}\{V_N(\theta^n)\} \quad (7)$$

with probability one. As well, it also holds that as N increases, the estimate $\hat{\theta}_N^n$ converges in law to a Normally distributed random variable with mean value θ_\circ^n according to (L.Ljung and P.E.Caines 1979, Caines 1988, Ljung 1999)

$$\sqrt{N}(\hat{\theta}_N^n - \theta_\circ^n) \xrightarrow{\mathcal{D}} \mathcal{N}(0, P_n), \quad \text{as } N \rightarrow \infty. \quad (8)$$

The $n \times n$ ‘covariance matrix’ P_n in (8) is defined in terms of two other matrices R_n and Q_n as

$$P_n \triangleq R_n^{-1} Q_n R_n^{-1} \quad (9)$$

which themselves are specified as

$$R_n \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N [\mathbf{E} \{ \psi_t(\theta_\circ^n) \psi_t^T(\theta_\circ^n) \} - \mathbf{E} \left\{ \varepsilon_t(\theta_\circ^n) \left(\frac{d\psi_t(\theta_\circ^n)}{d\theta^n} \right)^T \right\}] \quad (10)$$

and

$$Q_n \triangleq \lim_{N \rightarrow \infty} \sum_{t,\ell=1}^N \mathbf{E} \{ \psi_t(\theta_\circ^n) \psi_\ell^T(\theta_\circ^n) \varepsilon_t(\theta_\circ^n) \varepsilon_\ell(\theta_\circ^n) \}. \quad (11)$$

The quantity $\psi_t(\theta^n)$ in the preceding expressions is the prediction error gradient given by

$$\psi_t(\theta^n) \triangleq \frac{d\hat{y}_t(\theta^n)}{d\theta^n} = \frac{dG(q, \theta^n)}{d\theta^n} u_t. \quad (12)$$

While an asymptotic distributional result like (8) is very satisfying theoretically, for practical applications it is rather less appealing, mainly due to the (just presented) intricate definition of P_n via Q_n , R_n and $\psi_t(\theta^n)$.

In response to this, the seminal work (Hannan and Nicholls 1977, L.Ljung 1985, L.Ljung and Z.D.Yuan 1985, Ljung 1999) has proposed a solution by investigating how (8) manifests itself in the variability of the frequency response $G(e^{j\omega}, \hat{\theta}_N^n)$; the result being approximations such as (1).

The path towards achieving this involves noting that according to a first order Taylor expansion, the relationship between frequency domain and parameter space estimation errors is given as

$$G(e^{j\omega}, \hat{\theta}_N^n) - G(e^{j\omega}, \theta_\circ^n) = \left[\frac{dG(q, \theta^n)}{d\theta^n} \Big|_{\theta^n = \theta_\circ^n} \right]^T \times (\hat{\theta}_N^n - \theta_\circ^n) + o(\|\hat{\theta}_N^n - \theta_\circ^n\|). \quad (13)$$

Therefore, a consequence of (8) is that as $N \rightarrow \infty$

$$\sqrt{N} \left[G(e^{j\omega}, \hat{\theta}_N^n) - G(e^{j\omega}, \theta_\circ^n) \right] \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Delta_n(\omega)) \quad (14)$$

where

$$\Delta_n(\omega) \triangleq \left[\frac{dG(e^{j\omega}, \theta^n)}{d\theta^n} \Big|_{\theta^n = \theta_\circ^n} \right]^T P_n \times \left[\frac{dG(e^{-j\omega}, \theta^n)}{d\theta^n} \Big|_{\theta^n = \theta_\circ^n} \right]. \quad (15)$$

The main contribution of this paper is to rigorously establish that this quantity $\Delta_n(\omega)$ may be accurately approximated (or in some cases exactly quantified) by the simple expression (3), and furthermore that the pre-existing quantification (1) can be unreliable.

3. MAIN RESULT

It is also important to emphasise that a crucial aspect of this paper is the recognition of the need to carefully

consider the relationship between the model order m for which a variance error quantification is required and any underlying ‘true’ system order. Indeed, given the usual complexity of real-world dynamics, any assumption the existence of a true model order could be quite inappropriate.

In relation to this issue, the work here takes the perspective that, while on the one hand it is reasonable to assume that undermodelling-induced error decreases with increasing model order m , it is also reasonable to assume that the model order of interest has not surpassed an underlying true order, and hence does not imply pole-zero cancellations in the (asymptotic in N) estimated system.

This last premise is considered to be a realistic way of avoiding the supposition of a true model order, while still considering that some sort of model validation procedure, that checks for the appropriateness of an Output–Error structure (eg. in terms of residual whiteness), and at the very least checks for pole-zero cancellation, is part of an overall estimation and error-quantification process.

With these preliminary comments in mind, the following Theorem and Corollary provide formal statements of the main technical results of this paper.

Theorem 3.1. Suppose that $\hat{\theta}_N^n$ is calculated via (6) using the Output-Error model structure (4) so that $m = n/2$ and that the resultant asymptotic in N estimate $G(z, \theta_\circ^n)$ defined via (7) has poles $\{\xi_0, \dots, \xi_{m-1}\}$ all lying in a closed subset of the open unit disk \mathbf{D} . Then with

$$K_m(\omega) \triangleq \sum_{k=0}^{m-1} \frac{1 - |\xi_k|^2}{|e^{j\omega} - \xi_k|^2} \quad (16)$$

and in the limit as $N \rightarrow \infty$

$$\sqrt{N} \begin{bmatrix} K_m^{-1/2}(\omega) & 0 \\ 0 & K_m^{-1/2}(\lambda) \end{bmatrix} \begin{bmatrix} G(e^{j\omega}, \hat{\theta}_N^n) - G(e^{j\omega}, \theta_\circ^n) \\ G(e^{j\lambda}, \hat{\theta}_N^n) - G(e^{j\lambda}, \theta_\circ^n) \end{bmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_m(\omega, \lambda))$$

where provided

- (1) $\{u_t\}$ is quasi-stationary with $\Phi_u(\omega) \in \text{Lip}(\alpha)$ for some $\alpha > 0$;
- (2) $G(z, \theta_\circ^n)$ contains no pole-zero cancellation for any model order $m = n/2$;
- (3) Certain bounds quantifying how much $\{\varepsilon_t(\theta_\circ^n)\}$ differs from white noise are satisfied (Ninness and Hjalmarsson 2001a),

then for $\omega \neq \lambda$

$$\lim_{m \rightarrow \infty} \Sigma_m(\omega, \lambda) = 2\sigma^2 \begin{bmatrix} \Phi_u^{-1}(\omega) & 0 \\ 0 & \Phi_u^{-1}(\lambda) \end{bmatrix}. \quad (17)$$

Proof: See (Ninness and Hjalmarsson 2001a). ■

The implication of this result is that since it asserts that

$$\lim_{m, N \rightarrow \infty} \frac{\sqrt{N}}{K_m(\omega)} \left[G(\hat{\theta}_N^n) - G(\theta_\circ^n) \right] \sim \mathcal{N} \left(0, \frac{2\sigma^2}{\Phi_u(\omega)} \right)$$

then one could expect that the equality should nearly hold for finite m and N so that

$$G(e^{j\omega}, \hat{\theta}_N^n) - G(e^{j\omega}, \theta_0^n) \sim \mathcal{N}\left(0, \frac{2\sigma^2}{N} \frac{K_n(\omega)}{\Phi_u(\omega)}\right)$$

and

$$\mathbf{E} \left\{ |G(e^{j\omega}, \hat{\theta}_N^n) - G(e^{j\omega}, \theta_0^n)|^2 \right\} \approx \frac{2}{N} \frac{\sigma^2}{\Phi_u(\omega)} \sum_{k=0}^{m-1} \frac{1 - |\xi_k|^2}{|e^{j\omega} - \xi_k|^2}. \quad (18)$$

are good approximations.

However, in many applications, it may be very unappealing that quantifications like (18) depend on asymptotic in model order m arguments. In response to this, it is in fact possible to provide expressions that are valid for arbitrarily small m , but (in general) at the expense of more restrictive assumptions; see also (Ninness and Hjalmarsson 2001b).

Corollary 3.1. Under the conditions imposed in Theorem 3.1, together with further assumptions that

- (1) $\Phi_u(\omega) = \kappa$ a constant;
- (2) $\{\varepsilon_t(\theta_0^n)\}$ is a zero mean i.i.d. process of variance σ^2

then

$$\lim_{N \rightarrow \infty} N \mathbf{E} \left\{ |G(e^{j\omega}, \hat{\theta}_N^n) - G(e^{j\omega}, \theta_0^n)|^2 \right\} = \frac{2\sigma^2}{\kappa} K_m(\omega).$$

Proof: See (Ninness and Hjalmarsson 2001a). ■

That is, in the case of white input excitation $\Phi_u = \kappa$, then the accuracy of the quantification

$$\begin{aligned} \mathbf{E} \left\{ |G(e^{j\omega}, \hat{\theta}_N^n) - G(e^{j\omega}, \theta_0^n)|^2 \right\} &\approx \frac{2\sigma^2}{N\kappa} K_m(\omega) \\ &= \frac{2\sigma^2}{N\kappa} \sum_{k=0}^{m-1} \frac{1 - |\xi_k|^2}{|e^{j\omega} - \xi_k|^2} \end{aligned}$$

depends only on the amount N of data. It is therefore applicable for arbitrarily low model order, as will be illustrated in the following simulation section.

4. SIMULATION EXAMPLE

These studies are organised according to the type of system simulated, the colouring of the input spectrum and the amount N of observed data. In particular, the following systems are considered.

System1: Low-Order

$$G(q) = \frac{0.1}{(q - 0.9)},$$

System2: Mid-Order

$$G(q) = \frac{0.06(q - 0.8)(q - 0.9)}{(q - 0.99)(q - 0.7)(q - 0.6)},$$

System3: Low-Order Resonant

$$G(q) = \frac{0.0342q + 0.0330}{(q - 0.95e^{j\pi/12})(q - 0.95e^{-j\pi/12})},$$

Sys	Input Spectrum			
	Coloured		White	
	N = 10000	N = 200	N = 10000	N = 200
1	Fig 1(a)	Fig 1(b)	Fig 2(a)	Fig 2(b)
2	Fig 3(a)	Fig 3(b)	Fig 4(a)	Fig 4(b)
3	Fig 5(a)	Fig 5(b)	Fig 6(a)	Fig 6(b)
4	Fig 7(a)	Fig 7(b)	Fig 8(a)	Fig 8(b)

Table 1. Organisation of Simulation Examples

System4: Mid-Order Resonant

$$G(q) = \frac{0.1176(q + 8.0722)(q + 0.8672)(q + 0.0948)}{(q - 0.75e^{\pm j\pi/3})(q - 0.95e^{\pm j\pi/12})}.$$

For each of these systems, two cases of input spectrum are examined

$$\Phi_u(\omega) = \frac{1}{1.25 - \cos \omega} \quad \text{and} \quad \Phi_u(\omega) = 1$$

and for each of these spectra, both long ($N = 10,000$) and short ($N = 200$) data lengths are employed.

For all these situations, white Gaussian measurement noise of variance $\sigma^2 = 0.0001$ is added, and an Output–Error model of order equal to the true system is fitted over 10000 different input and measurement noise realisations. This allows the computation of the true estimate variability via sample average over these Monte–Carlo simulations, which is then compared to the new expression (3) as well as the pre-existing one (1) in figures 1–8 and according to the organisation given in table 1.

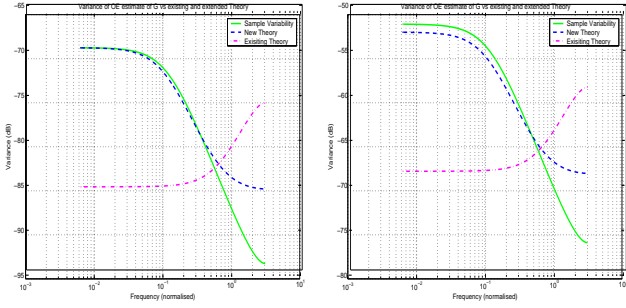
In each of these figures, (the estimate of) the true variability is shown as a solid line, the new variance expression (3) of this paper is shown as a dashed line, and the pre-existing approximation (1) is illustrated via a dash-dot line. The examination of all these examples reveals some important points.

Firstly, the new approximation (3) is clearly quite robust. It provides an informative quantification across the full range of scenarios, even for the case of very low model order $m = 1$ and very low data length $N = 200$ as shown in figure 1(b).

Secondly, as shown in the cases of white input, the new approximation (3) is essentially exact in these cases regardless of model order, save for small errors at very low data lengths. This, of course, is consistent with Corollary 3.1.

Thirdly, as illustrated in the case of resonant systems, even when the true variability has a quite complicated nature, the new approximation (3) is able to provide an informative, and in most cases accurate quantification.

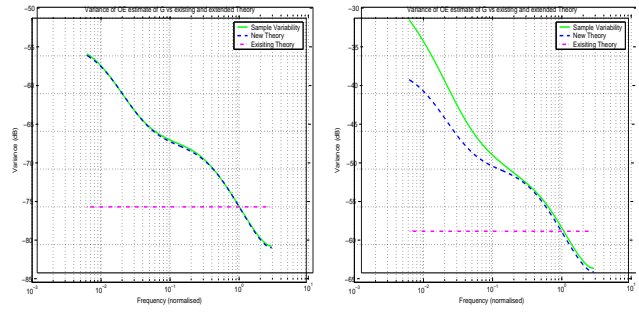
Finally, as suggested by examination of the dash-dot line representing (1) in each of figures 1–8, that pre-existing and widely used quantification can be unreliable, which leads to the suggestion of this paper that in fact it should be replaced by (3).



(a) Low-Order, Coloured Φ_u , $N = 10,000$

(b) Low-Order, Coloured Φ_u , $N = 200$

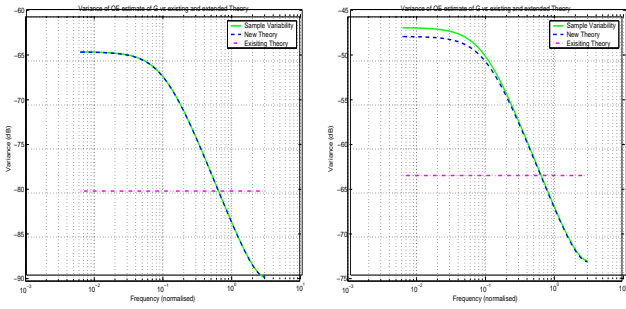
Fig. 1. System 1, very low-order. True variability is solid line, new quantification (3) is the dashed line, and the existing quantification (1) is the dash-dot line.



(a) Mid-Order, White Φ_u , $N = 10,000$

(b) Mid-Order, White Φ_u , $N = 200$

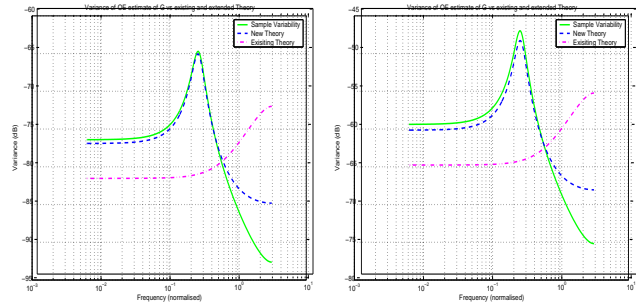
Fig. 4. System 2, mid-order. True variability is solid line, new quantification (3) is the dashed line, and the existing quantification (1) is the dash-dot line.



(a) Low-Order, White Φ_u , $N = 10,000$

(b) Low-Order, White Φ_u , $N = 200$

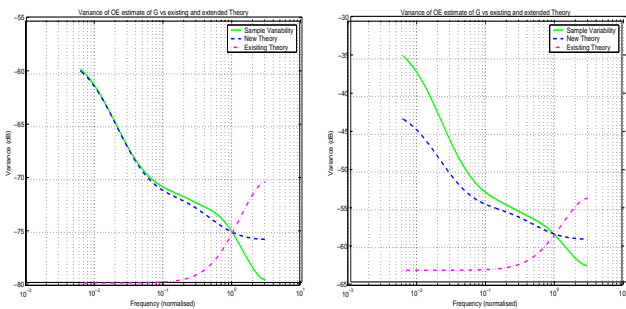
Fig. 2. System 1, very low-order. True variability is solid line, new quantification (3) is the dashed line, and the existing quantification (1) is the dash-dot line.



(a) Low-Order Resonant, Coloured Φ_u , $N = 10,000$

(b) Low-Order Resonant, Coloured Φ_u , $N = 200$

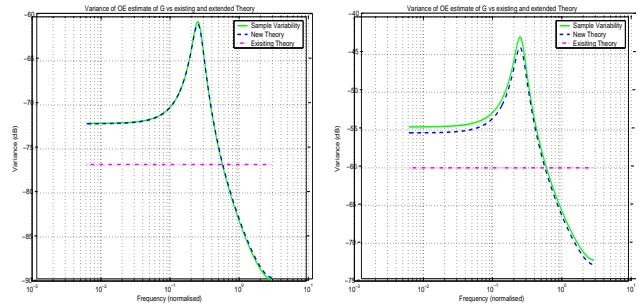
Fig. 5. System 3, low-order resonant. True variability is solid line, new quantification (3) is the dashed line, and the existing quantification (1) is the dash-dot line.



(a) Mid-Order, Coloured Φ_u , $N = 10,000$

(b) Mid-Order, Coloured Φ_u , $N = 200$

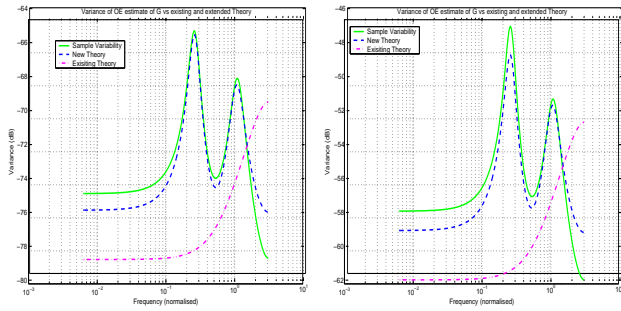
Fig. 3. System 2, mid-order. True variability is solid line, new quantification (3) is the dashed line, and the existing quantification (1) is the dash-dot line.



(a) Low-Order Resonant, White Φ_u , $N = 10,000$

(b) Low-Order Resonant, White Φ_u , $N = 200$

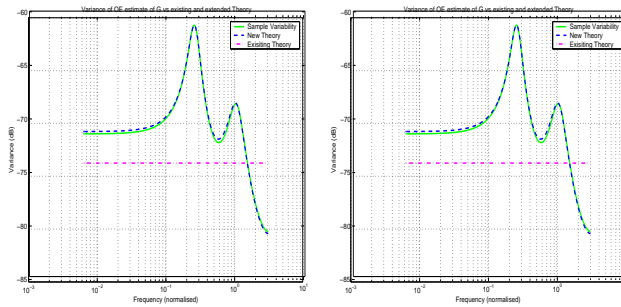
Fig. 6. System 3, low-order resonant. True variability is solid line, new quantification (3) is the dashed line, and the existing quantification (1) is the dash-dot line.



(a) Mid-Order Resonant, Coloured Φ_u , $N = 10,000$

(b) Mid-Order Resonant, Coloured Φ_u , $N = 200$

Fig. 7. System 4, mid-order resonant. True variability is solid line, new quantification (3) is the dashed line, and the existing quantification (1) is the dash-dot line.



(a) Mid-Order Resonant, White Φ_u , $N = 10,000$

(b) Mid-Order Resonant, White Φ_u , $N = 200$

Fig. 8. System 4, mid-order resonant. True variability is solid line, new quantification (3) is the dashed line, and the existing quantification (1) is the dash-dot line.

5. CONCLUSIONS

This paper is closely related to one by the same authors (Ninness *et al.* 1999a) where the same estimation problem was considered and the dependence of variance error on system poles was established.

The key advancement made here is to establish the new factor of 2 in the variance quantification (3). There are two main reasons for the significance of this development. First, it resolves the paradox inherent in previous variance quantifications which predicted no benefit in knowing system poles rather than having to estimate them.

Second, the method used to establish the factor of 2, which is developed in detail in (Ninness and Hjalmarsson 2001a), invents new techniques for handling quadratic forms in inverses of Toeplitz matrices via properties of reproducing kernels. Of itself, this is important since it allows for variance quantifications that are not asymptotic in the model order; see Corollary 3.1 and (Ninness and Hjalmarsson 2001b).

Finally, this paper also presents more extended simulation results further validating the need (previously raised

in (Ninness *et al.* 1999a)) for incorporating knowledge of estimated system poles in variance error quantifications.

6. REFERENCES

- Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*. Holden-Day.
- Caines, P. (1988). *Linear Stochastic Systems*. John Wiley and Sons. New York.
- Forsell, U. and L. Ljung (1999). ‘Closed-loop identification revisited’. *Automatica* **53**, 1215–1241.
- Forsell, U. and L. Ljung (2000). ‘Some results on optimal experiment design’. *Automatica* **36**(5), 749–756.
- Gevers, M., Lennart Ljung and Paul Van den Hof (2001). ‘Asymptotic variance expressions for closed loop identification’. *Automatica* **37**, 781–786.
- Hannan, E. and D.F. Nicholls (1977). ‘The estimation of the prediction error variance’. *Journal of the American Statistical Association* **72**(360, part 1), 834–840.
- Ljung, L. (1999). *System Identification: Theory for the User, (2nd edition)*. Prentice-Hall, Inc.. New Jersey.
- Ljung, L. and Lei Guo (1997). ‘The role of model validation for assessing the size of the unmodeled dynamics’. *IEEE Trans. Automatic Control* **AC-42**(9), 1230–1240.
- L.Ljung (1985). ‘Asymptotic variance expressions for identified black-box transfer function models’. *IEEE Trans. on Aut. Cont* **AC-30**(9), 834–844.
- L.Ljung and P.E.Caines (1979). ‘Asymptotic Normality of prediction error estimators for approximate system models’. *Stochastics* **3**, 29–46.
- L.Ljung and Z.D.Yuan (1985). ‘Asymptotic properties of black-box identification of transfer functions’. *IEEE Transactions on Automatic Control* **30**(6), 514–530.
- Ninness, B. and Håkan Hjalmarsson (2001a). Accurate quantification of estimation error: I. Technical report. EE2001050 Department of Electrical & Computer Engineering, University of Newcastle.
- Ninness, B. and Håkan Hjalmarsson (2001b). Exact quantification of variance error. Technical report. EE2001051 Department of Electrical & Computer Engineering, University of Newcastle.
- Ninness, B., Håkan Hjalmarsson and Fredrik Gustafsson (1999a). Asymptotic variance expressions for output error model structures. In ‘Proceedings of the 14th IFAC World Congress, Beijing’.
- Ninness, B., Håkan Hjalmarsson and Fredrik Gustafsson (1999b). ‘The fundamental role of general orthonormal bases in system identification’. *IEEE Transactions on Automatic Control* **44**(7), 1384–1406.
- P.M.J. Van den Hof, P.S.C. Heuberger and J. Bokor (1995). ‘System identification with generalized orthonormal basis functions’. *Automatica* **31**(12), 1821–1834.
- Wahlberg, B. (1991). ‘System identification using Laguerre models’. *IEEE Transactions on Automatic Control* **AC-36**(5), 551–562.
- Wahlberg, B. (1994). ‘System identification using Kautz models’. *IEEE Trans. Auto. Cont* **AC-39**(6), 1276–1282.
- Zhu, Y. (1998). ‘Multivariable process identification for mpc: the asymptotic method and its applications’. *Journal of Process Control* **8**(2), 101–115.