Proceedings of the 37th IEEE
Conference on Decision & Control
Tampa, Florida USA • December 1998

IDENTIFICATION IN CLOSED LOOP: ASYMPTOTIC HIGH ORDER VARIANCE FOR RESTRICTED COMPLEXITY MODELS

Håkan Hjalmarsson* and Brett Ninness†
*Dept. of Signals, Sensors and Systems, The Royal Institute of Technology
S-100 44 Stockholm, Sweden, hakan.hjalmarsson@s3.kth.se
†Dept. of Electrical and Computer Engineering
University of Newcastle, Australia, brett@ee.newcastle.edu.au

Abstract

Asymptotic high order variance expressions for identified models have found widespread use in e.g. optimal experiment design, analysis of model accuracy and control. The expressions derived in the 1970's and 1980's by Berk, Ljung and others are valid for a wide range of operating conditions and models such as restricted complexity models identified from open loop data as well as full order models identified from closed loop data. Throughout the 1990's much attention has been devoted to the issue of identification of models based on closed loop data. In order to, at least asymptotically in the model order, be able to analyze the quality of such models a corresponding high order variance theory for restricted complexity models identified from closed loop data is necessary. This paper provides this for models with a fixed noise model. A novel variance expression, valid for Gaussian signals, is derived. Simulations show that this expression is surprisingly accurate also for non-Gaussian signals.

1 Introduction

A result that has found great utility in practical applications of least-squares system identification methods is that the sensitivity to measurement noise of the ensuing frequency response estimate $\hat{G}(e^{j\omega})$ may be quantified as [1] [2]

$$\mathrm{Var}\{\hat{G}(e^{j\omega})\} \approx \frac{n}{N}\frac{\Phi_v(\omega)}{\Phi_u(\omega)} \tag{1}$$

where $\Phi_v(\omega)$ and $\Phi_u(\omega)$ are the measurement noise and input excitation spectral densities (respectively), N is the length of the available data record, and n is the order of the model $G(e^{i\omega})$.

The expression (1) has been instrumental for understanding the design variables in system identification [3], [4], [5] and for optimal experimental design [6], [7], [8]. Extensions of the expression (1) which cover other model structures have been derived. Analysis for Laguerre basis functions is provided in [9]. Orthonormal bases where the poles are repeated cyclically is treated in [10]. Recently, fixed denominator model structures and extensions of ARX structures with fixed noise model zeros have been treated in [11].

One missing piece has been high order variance expressions for biased models in closed loop. This means that there is currently no high order variance expression for methods that employ *fixed noise models* such as output error, finite impulse response, Laguerre, Kautz and fixed denominator models. The contribution of this paper is such a result.

2 Problem Formulation

The problem studied in this paper is one where N point data records of an input sequence $\{u_t\}$ and output sequence $\{y_t\}$ of a linear time invariant system are available. It is assumed that this data is generated as follows

$$y_t = G(q)u_t + H(q)e_t$$

Here $G(q)$ is an unknown transfer function describing, in terms of the forward shift operator q, the system dynamics that are to be identified by means of the observations $\{u_t\}$ and $\{y_t\}$. The output measurements $\{y_t\}$ are corrupted by a zero mean stationary noise sequence $v_t = H(q)e_t$ where $H(q)$ is stable and stably invertible and monic transfer function and $\{e_t\}$ is a zero mean white noise sequence with variance $\mathbf{E}\{e_t^2\} = \sigma^2 < \infty$ and with $\mathbf{E}\{|e_t|^{4+\delta}\} < \infty$ for some $\delta > 0$. The input $\{u_t\}$ is assumed to be quasi-stationary in the sense used by Ljung [7] so that it has an associated spectral density $\Phi_u(\omega) > 0$ which is assumed to be Lipschitz continuous of some order $\epsilon > 0$. The spectral density of the noise process $\{v_t\}$ is denoted as $\Phi_v(\omega) = \sigma^2|H(e^{i\omega})|^2$ and is also assumed to be positive and Lipschitz continuous of order $\epsilon > 0$.

The identification schemes considered here are ones in which the prediction error framework [7, 12] is used. This requires a model structure

$$y_t = G(q, \theta^n)u_t + H(q, \theta^n)e_t \tag{2}$$

parameterised by a vector $\theta^n \in \mathbf{R}^n$ to be employed which is of the form

$$G(q,\theta^n) = \sum_{k=1}^{\infty} g_k(\theta^n)q^{-k}, \quad H(q,\theta^n) = 1+\sum_{k=1}^{\infty} h_k(\theta^n)q^{-k}.$$

This structure implies the following one step ahead predictor

$$\widehat{y}_t(\theta^n) = [1-H^{-1}(q,\theta^n)]y_t+H^{-1}(q,\theta^n)G(q,\theta^n)u_t \quad (3)$$

and associated prediction error

$$\varepsilon_t(\theta^n) \triangleq y_t - \widehat{y}_t(\theta^n) \quad (4)$$

so that if the quadratic criterion

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

is employed, then based on the N point data observation, a least squares estimate $\widehat{\theta}^n_N$ of θ^n may be found as

$$\widehat{\theta}^n_N \triangleq \arg\min_{\theta^n\in\mathbf{R}^n} V_N(\theta^n). \quad (5)$$

The theory pertaining to the properties of such a method is very rich. Germane to this paper are the properties that [7, 12]

$$\widehat{\theta}^n_N \xrightarrow{\text{a.s.}} \theta^n_\circ \quad \text{as } N \to \infty$$

where with $\mathbf{E}\{\cdot\}$ denoting expectation over the underlying probability space that any random variables are defined on

$$\theta^n_\circ \triangleq \arg\min_{\theta^n\in\mathbf{R}^n} \lim_{N\to\infty} \mathbf{E}\left\{V_N(\theta^n)\right\}. \quad (6)$$

As well [7, 12]

$$\sqrt{N}(\widehat{\theta}^n_N - \theta^n_\circ) \xrightarrow{\mathcal{D}} \mathcal{N}(0,P_n), \quad \text{as } N \to \infty$$

where

$$P_n \triangleq R_n^{-1}Q_nR_n^{-1} \quad (7)$$

and with the definition of the prediction error gradient $\psi_t(\theta^n)$ as

$$\psi_t(\theta^n) = \frac{\mathrm{d}\widehat{y}_t(\theta^n)}{\mathrm{d}\theta^n}$$

then

$$R_n \triangleq \lim_{N\to\infty} \mathbf{E}\left\{\left.\frac{\mathrm{d}^2V_N(\theta)}{\mathrm{d}\theta\theta^T}\right|_{\theta=\theta^n_\circ}\right\} \quad (8)$$

and $Q_n \triangleq \lim_{N\to\infty} N\mathbf{E}\left\{\frac{\mathrm{d}V_N(\theta^n_\circ)}{\mathrm{d}\theta^n}\left(\frac{\mathrm{d}V_N(\theta^n_\circ)}{\mathrm{d}\theta^n}\right)^T\right\}$

$$= \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\sum_{\ell=1}^{N}\mathbf{E}\left\{\psi_t(\theta^n_\circ)\psi_\ell(\theta^n_\circ)^T\varepsilon_t(\theta^n_\circ)\varepsilon_\ell(\theta^n_\circ)\right\}. \quad (9)$$

A key contribution of [1, 2] is to recognise that in applications, often the quantification of the parameter space properties of $\widehat{\theta}^n_N$ are of secondary importance to their influence on the associated properties of $G(e^{j\omega},\widehat{\theta}^n_N)$ and $H(e^{j\omega},\widehat{\theta}^n_N)$. For the purposes of analysing the latter, it is expedient to define the composite transfer function

$$\Pi(q,\theta^n) \triangleq [G(q,\theta^n),H(q,\theta^n)] \quad (10)$$

and argue by Taylor expansion that

$$\begin{aligned}\widetilde{\Pi}_N(\omega) &\triangleq \Pi(e^{j\omega},\widehat{\theta}^n_N) - \Pi(e^{j\omega},\theta^n_\circ)\\ &= \left[\frac{\mathrm{d}}{\mathrm{d}\theta^n}\Pi(e^{j\omega},\theta^n_\circ)\right]^T(\widehat{\theta}^n_N - \theta^n_\circ) + o(\|\widehat{\theta}^n_N - \theta^n_\circ\|)\end{aligned}$$

where here $\|\cdot\|$ denotes Euclidean norm so that with the notation $\Pi'(e^{j\omega},\theta^n_\circ) \triangleq \mathrm{d}\Pi(e^{j\omega},\theta^n)/\mathrm{d}\theta^n$

$$\sqrt{N}\,\widetilde{\Pi}_N(\omega) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\Lambda_n(\omega)) \quad (11)$$

where with $\cdot^\star$ denoting 'conjugate transpose'

$$\Lambda_n(\omega) = [\Pi'(e^{j\omega},\theta^n_\circ)]^\star P_n\Pi'(e^{j\omega},\theta^n_\circ) \quad (12)$$

and this suggests the approximation

$$N\mathbf{E}\left\{\widetilde{\Pi}_N(\omega)\widetilde{\Pi}_N^\star(\omega)\right\} \approx \Lambda_n(\omega).$$

Unfortunately, the evaluation of $\Lambda_n(\omega)$ is always too complicated to be useful. A key contribution of [1, 2] is to observe that in contrast to the intractability of $\Lambda_n(\omega)$, the limit

$$\lim_{N,n\to\infty}\frac{N}{n}\mathbf{E}\left\{\widetilde{\Pi}_N(\omega)\widetilde{\Pi}_N^\star(\omega)\right\} = \lim_{n\to\infty}\frac{1}{n}\Lambda_n(\omega) \triangleq \Lambda(\omega)$$

does have a simple and useful form; for example $\Lambda(\omega) = \Phi_\nu(\omega)/\Phi_u(\omega)$ for FIR model structures [2].

The suggestion of [1, 2] is to then assume that even for finite N and n it can reasonably be expected that $n^{-1}\Lambda_n(\omega)$ has approximately converged to the simple form $\Lambda(\omega)$ so that

$$\mathbf{E}\left\{\widetilde{\Pi}_N(\omega)\widetilde{\Pi}_N^\star(\omega)\right\} \approx \frac{n}{N}\Lambda(\omega) \quad (13)$$

is a good approximation. The validity of this strategy is argued via numerical example in [1, 2], and indeed it has won widespread acceptance as a tool for analysing the performance of least squares estimation schemes; see for example [13, 14, 15, 8].

3 Fixed Denominator Model Structures

In this contribution we shall use a common form of fixed denominator model structures

$$G(q,\theta^n) \triangleq D_n^{-1}(q)\sum_{k=0}^{n-1}\theta^n_k q^k, \quad H(q,\theta^n) = 1 \quad (14)$$

with $D_n(q) = \prod_{k=0}^{n-1}(q-\xi_k)$ for some user chosen poles $\{\xi_k\} \in \mathbf{D} \triangleq \{z \in \mathbf{C} : |z| < 1\}$ where $\mathbf{C}$ denotes the field of complex numbers. For a more general type of 'fixed denominator' model structures see [11].

Specialised 'orthonormal' versions of this structure have recently attracted significant research attention [9, 16] where it has been suggested that although the choice $\xi_k = 0$ in (14) gives the common FIR structure, it is intuitively more reasonable to choose the poles $\{\xi_k\}$ according to prior knowledge so as to be close to the suspected true poles of $G(q)$.

4 Quadratic Forms of Generalized Toeplitz Matrices

This section provides some results on generalized Toeplitz matrices which are useful when deriving asymptotic variance expressions.

Let the set $\{\mathcal{B}_k(q)\}$ be a particular orthonormal rational basis [17] defined by

$$\mathcal{B}_k(q) \triangleq \frac{\sqrt{1-|\xi_k|^2}}{q-\xi_k}\prod_{m=0}^{k-1}\left(\frac{1-\overline{\xi_m}q}{q-\xi_m}\right). \qquad (15)$$

where the poles ξ_k are restricted to lie strictly inside the unit disc. For any continuous complex-valued function f we define the generalized Toeplitz matrix $M_n(f)$ as

$$M_n(f) \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi}\Gamma_n(\omega)\Gamma_n^\star(\omega)f(\omega)\,\mathrm{d}\omega \qquad (16)$$

where

$$\Gamma_n(z) \triangleq [\mathcal{B}_\circ(z), \mathcal{B}_1(z), \cdots, \mathcal{B}_{n-1}(z)]^T. \qquad (17)$$

Define the 'reproducing kernel' $K_n(z,\mu)$ associated with the space $X_n \triangleq \mathrm{Span}\{\mathcal{B}_0, \mathcal{B}_1, \cdots, \mathcal{B}_{n-1}\}$ as

$$K_n(z,\mu) = \sum_{k=1}^{n-1}\overline{\mathcal{B}_k(\mu)}\mathcal{B}_k(z), \quad z,\mu \in \mathbf{C}. \qquad (18)$$

The following result is shown in in [18].

Theorem 4.1 *Provided $\sum(1-|\xi_k|) = \infty$ then for any possibly complex valued but continuous f*

$$\lim_{n\to\infty}\frac{1}{K_n(\omega,\omega)}\Gamma_n^\star(\omega)M_n(f)\Gamma_n(\lambda) = \begin{cases} f(\omega) & ;\omega = \lambda, \\ 0 & ;\omega \neq \lambda \end{cases} \qquad (19)$$

Two $n \times n$ matrices A_n and B_n are asymptotically equivalent, with notation

$$A_n \sim B_n \quad \text{as } n \to \infty, \qquad (20)$$

if for all $\omega \in [-\pi, \pi]$

$$\lim_{n\to\infty}\frac{\Gamma_n^\star(\omega)[A_n - B_n][A_n - B_n]^\star\Gamma_n(\omega)}{K_n(\omega,\omega)} = 0. \qquad (21)$$

The utility of this idea is that if $A_n \sim B_n$ as $n \to \infty$ then the asymptotic consideration of quadratic forms of A_n and B_n in $\Gamma_n(\omega)$ are identical.

In [18] the following group structure of generalized Toeplitz matrices is shown.

Theorem 4.2 *Let f and g be any (possibly complex valued) functions that are invertible and Lipschitz continuous of some order $\epsilon > 0$, and if all the poles $\{\xi_k\}$ are chosen in some closed subset of the open unit disc* $\mathbf{D}$, *then*

$$M_n(f)M_n(g) \sim M_n(fg) \quad \textit{as } n \to \infty, \qquad (22)$$
$$M_n^{-1}(f) \sim M_n(1/f) \quad \textit{as } n \to \infty. \qquad (23)$$

5 Variance Error for Fixed Denominator Model Structure

For fixed denominator structures (14), (12) reduces to

$$\mathbf{E}\left\{|G(e^{j\omega},\hat\theta_N^n) - G(e^{j\omega},\theta_\circ^n)|^2\right\} = \Gamma_n^\star(\omega)P_n\Gamma_n(\omega) \qquad (24)$$

where P_n is defined by (7). In order to obtain a useful asymptotic expression for the right-hand side of (24), the idea is to massage P_n into a form such that Theorem 4.1 can be used.

The key observation is that the analysis of models of the type (14) can be facilitated by a reparameterization in terms of the basis functions (15)

$$G(q,\theta^n) = \sum_{k=0}^{n-1}\theta_k^n\mathcal{B}_k(q), \quad H(q,\theta^n) = 1 \qquad (25)$$

Note that in the fixed denominator case the prediction error gradient ψ_t is given by

$$\psi_t(\theta_\circ^n) = \Gamma_n(q)D_n^{-1}(q)u_t = \Gamma_n(q)u_t'$$

where $u_t' \triangleq D_n^{-1}(q)u_t$. Therefore, $\mathrm{d}\psi_t/\mathrm{d}\theta^n = 0$ so that from (8)

$$R_n = \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\mathbf{E}\left\{\psi_t(\theta_\circ^n)\psi_t^T(\theta_\circ^n)\right\}. \qquad (26)$$

Using (26) and Parseval's Formula R_n is a generalised Toeplitz matrix of the form

$$R_n = M_n(\Phi_u). \qquad (27)$$

For Q_n such a simple expression exists only asymptotically. The classical case when data are collected in open loop or the noise spectrum can be modelled perfectly has been thoroughly treated [1] [2, 9, 10] for various types of model structures. The novelty in this contribution is that we consider the case where *data may be collected in closed loop with a fixed (possibly incorrect) noise model.* For Gaussian signals, Lemma A.1 gives that

$$Q_n \sim M_n(\Phi_u(\omega)\Phi_\varepsilon(\omega,\theta_\circ^n)) \qquad (28)$$

regardless of how the data has been generated and if the noise model is correct or not.

Using (27)–(28) in (7), Theorem 4.2 gives that

$$P_n \sim M_n \left(\frac{\Phi_\varepsilon(\omega, \theta_\circ^n)}{\Phi_u(\omega)} \right) \tag{29}$$

Combining this with Theorem 4.1 gives the following main result of this paper.

Theorem 5.1 *Let both $\{u_t\}$ and $\{e_t\}$ be Gaussian distributed, zero mean stationary stochastic processes.*

With $\widehat{\theta}_N^n$ calculated via (5) using the model structures (14), (25) or any other structure with the same fixed poles $\{\xi_k\}$ all chosen to lie within the open unit disk $\mathbf{D}$ then with $K_n(\omega,\omega) \equiv K_n(e^{j\omega}, e^{j\omega})$ given by (18) and in the limit as $N \to \infty$

$$\sqrt{N} \begin{bmatrix} K_n(\omega,\omega) & 0 \\ 0 & K_n(\lambda,\lambda) \end{bmatrix}^{-1/2} \begin{pmatrix} G(e^{j\omega}, \widehat{\theta}_N^n) - G(e^{j\omega}, \theta_\circ^n) \\ G(e^{j\lambda}, \widehat{\theta}_N^n) - G(e^{j\lambda}, \theta_\circ^n) \end{pmatrix}$$

$$\xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma_n(\omega, \lambda))$$

where, provided all the poles $\{\xi_k\}$ are chosen in a closed subset of $\mathbf{D}$ and $\Phi_u(\omega)$, $\Phi_\nu(\omega)$ are Lipschitz continuous of some order $\epsilon > 0$,

$$\lim_{n\to\infty} \Sigma_n(\omega,\lambda) - \begin{bmatrix} \dfrac{\Phi_\varepsilon(\omega,\theta_\circ^n)}{\Phi_u(\omega)} & 0 \\ 0 & \dfrac{\Phi_\varepsilon(\omega,\theta_\circ^n)}{\Phi_u(\lambda)} \end{bmatrix} = 0$$

Here $\Phi_\varepsilon(\omega, \theta_\circ^n)$ is to be interpreted as the spectral density of the prediction error residual sequence $\{\varepsilon_t(\theta^n)\}$ defined in (4) evaluated at $\theta^n = \theta_\circ^n$, the latter being defined in (6).

Corollary 5.1 *Under the same conditions as the previous theorem, but with a strengthened requirement that $\mathsf{E}\{e_t^8\} < \infty$ then*

$$\lim_{n\to\infty} \lim_{N\to\infty} \frac{N}{K_n(\omega,\omega)} \mathsf{E}\left\{ |G(e^{j\omega}, \widehat{\theta}_N^n) - G(e^{j\omega}, \theta_\circ^n)|^2 \right\} - \frac{\Phi_\varepsilon(\omega,\theta_\circ^n)}{\Phi_u(\omega)} = 0 \tag{30}$$

Proof: Follows using the methods in Appendix 9B of [7]. ∎

By drawing on the precedent of [2, 1] the new asymptotic result (30) suggests the approximation

$$\mathrm{Var}\left\{ G(e^{j\omega}, \widehat{\theta}_N^n) \right\} \approx \frac{1}{N} \frac{\Phi_\varepsilon(\omega,\theta_\circ^n)}{\Phi_u(\omega)} \sum_{k=0}^{n-1} |\mathcal{B}_k(e^{j\omega})|^2. \tag{31}$$

Typically the limit

$$\Phi_{\varepsilon^\star}(\omega) \triangleq \lim_{n\to\infty} \Phi_\varepsilon(\omega, \theta_\circ^n)$$

exists. The spectral density $\Phi_{\varepsilon^\star}(\omega)$ is the spectrum of the prediction error which is obtained when $\lim_{N\to\infty} \mathsf{E}\left\{ \frac{1}{N} \sum_{t=1}^N (y_t - \widehat{G}(q)u_t)^2 \right\}$ is minimized over all causal stable transfer functions $\widehat{G}(q)$. The model $\widehat{G}(q)$ that gives the smallest criterion is called the limit model. In the open loop case, the limit model equals the true system, i.e. $\widehat{G}(q) = G(q)$. This means that $\Phi_{\varepsilon^\star}(\omega) = \Phi_v(\omega)$ and the classical result [1] (1) is recovered for FIR models since then $|\mathcal{B}_k(e^{i\omega})| = 1$. The new result thus encompasses the existing theory.

6 Numerical Illustration

To illustrate the result (31), we will study the system

$$y_t = G(q)u_t + H(q)e_t \tag{32}$$

$$G(q) = \frac{q^{-1}}{(1-0.9q^{-1})(1-0.8q^{-1})(1-0.7q^{-1})} \tag{33}$$

$$H(q) = \frac{1}{(1-0.9q^{-1})(1+0.9q^{-1})(1+0.9q^{-1})} \tag{34}$$

The Bode diagrams of G and H are given in Figure 1. The noise e is Gaussian with zero mean and variance $\sigma^2 = 0.01$.

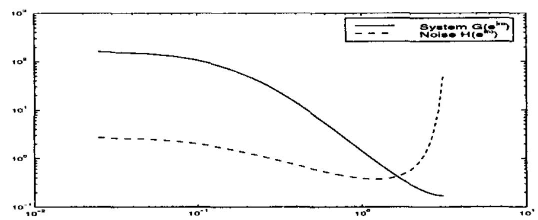

Figure 1: Bode plot of true system frequency function $G(e^{i\omega})$ and the noise model frequency function $H(e^{i\omega})$.

The system (32)–(34) is controlled by the following proportional controller

$$u_t = K\,(r_t - y_t) \tag{35}$$

where the gain $K = 0.03$ and where the reference signal r_t is zero mean Gaussian white noise with unit variance. This gives a closed loop system with two poles at 0.94 and one pole at 0.57. The response to a square wave reference r is given in Figure 2.

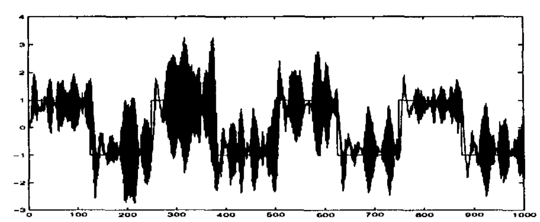

Figure 2: Output response to a square wave reference.

The open loop system $G(q)$ is identified using a finite impulse response model

$$y_t = \sum_{k=1}^{n} \theta_k u_{t-k} + v_t = G(q,\theta)u_t + v_t \tag{36}$$

Since no noise model is included and the data are collected when the system is operating in closed loop, one should expect a biased estimate of $G(q)$ no matter how large n is used. An approximation of the limit model ($\lim_{n\to\infty} G(q,\theta_o^n)$) is given in Figure 3. Comparing with the true frequency response $G(e^{i\omega})$ we also see that there is a significant bias, both at low and at high frequencies. The approximate limit model has been obtained using identification with model order $n = 90$ and $N = 50000$.

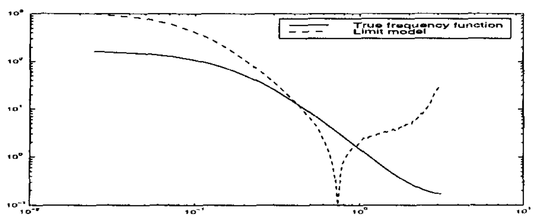

Figure 3: Gaussian experimental conditions: Bode plot of true system frequency function $G(e^{i\omega})$ and the limit model.

The variance

$$\mathrm{Var}\{G(e^{j\omega},\hat{\theta}_N^n)\} = \mathrm{E}\left[\left|G(e^{i\omega},\hat{\theta}_N^n) - G(e^{i\omega},\theta_0^n)\right|^2\right] \tag{37}$$

has been estimated using Monte Carlo simulations for $N = 1000$. The limit θ_o^n was taken as the mean of the $\hat{\theta}_N^n$ for the different simulations. In Figure 4, the Monte Carlo estimate (using 500 simulations) of the variance (37) when the model order $n = 80$ is plotted together with the standard variance expression (1) and the new expression

$$\mathrm{Var}\{G(e^{j\omega},\hat{\theta}_N^n)\} \approx \frac{n}{N}\,\frac{\Phi_\varepsilon(\omega,\theta_o^n)}{\Phi_u(\omega)} \tag{38}$$

Clearly, (38) reflects the true variance much better than (1).

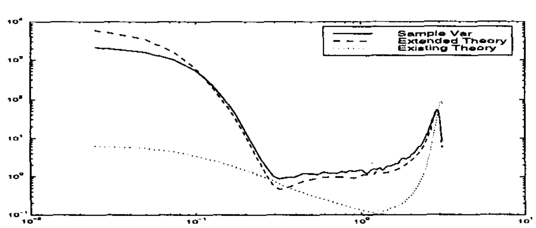

Figure 4: Gaussian experimental conditions: Variance of estimated frequency function for model order $n = 80$.

To test the sensitivity of (38) to the underlying assumptions, a set-up with non-Gaussian signals was tested. Instead of a white reference, a square wave with a period time of 250 and unit amplitude was used. Furthermore, the noise was taken to be binary with variance $\sigma^2 = 0.01$. For these experimental conditions, the limit model[1] is given in Figure 5.

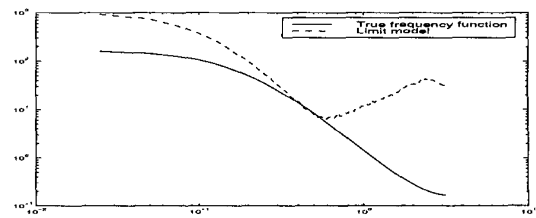

Figure 5: Non-Gaussian experimental conditions: Bode plot of true system frequency function $G(e^{i\omega})$ and the limit model.

Then, by Monte Carlo simulations, the estimated variance of the estimated frequency function is given in Figure 6 together with the theoretical high order variance expressions. Rather surprisingly, (38) is a good measure of the variance in this non-Gaussian set-up also.

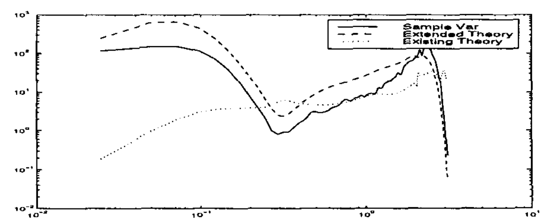

Figure 6: Non-Gaussian experimental conditions: Variance of estimated frequency function for model order $n = 80$.

7 Conclusions

In this contribution we have derived a novel high order variance expression for the estimated frequency function. The expression applies to models with fixed noise model under the assumption of Gaussian signals. Contrary to previously reported results, the expression holds regardless of whether the data is collected in closed loop or open loop. The expression is identical to the existing variance expression save that the noise spectrum is replaced by the *spectrum of the prediction errors corresponding to a model identified with an infinite amount of data.*

In simulations it has been shown that the novel variance expression provides useful information about the model variance also when the signals are non-Gaussian.

It is believed that this expression will be of importance when developing restricted complexity identifi-

[1]Notice that the limit model has changed compared to the previous experimental conditions, see Figure 3. The reason is that the limit model is the model which minimizes (5) over all causal transfer functions and the criterion (5) depends obviously on the experimental conditions such as the spectrum of the reference signal

cation methods for systems operating in closed loop.

A Auxiliary Results

Lemma A.1 *Suppose that $\{u_t\}$ and $\{e_t\}$ are both realisations of zero mean Gaussian distributed stationary stochastic processes and the orthonormal model structure (25) is employed. Then regardless of whether the data is collected in open or closed loop the matrix Q_n defined in (9) obeys*

$$Q_n \sim M_n(\Phi_u(\omega)\Phi_\varepsilon(\omega,\theta_\circ^n))$$

as $n \to \infty$ where the matrix formulation M_n is defined in (16) and $\Phi_\varepsilon(\omega,\theta_\circ^n)$ is the spectral density of the prediction residuals $\{\varepsilon_t(\theta^n)\}$ defined in (4) and evaluated at $\theta^n = \theta_\circ^n$ defined in (16).

Proof: From the definition (9) after using the stationarity assumption and the change of variable $\tau = k + m$ and regrouping terms

$$\begin{aligned} Q_n &= \lim_{N\to\infty} \sum_{\tau=-N}^{N} \left(1 - \frac{|\tau|}{N}\right) \mathbf{E}\left\{\psi_k\psi_{k-\tau}^T\varepsilon_k(\theta_\circ^n)\varepsilon_{k-\tau}(\theta_\circ^n)\right\}, \\ &= \sum_{\tau=-\infty}^{\infty} \mathbf{E}\left\{\psi_k\psi_{k-\tau}^T\varepsilon_k(\theta_\circ^n)\varepsilon_{k-\tau}(\theta_\circ^n)\right\}, \end{aligned}$$

where the properties of Cesàro means have been used in progressing to the last line. Using the Gaussianity assumption and the formula for fourth moments of jointly Gaussian random variables [12]

$$\begin{aligned} \mathbf{E}\left\{\psi_k\psi_{k-\tau}^T\varepsilon_k(\theta_\circ^n)\varepsilon_{k-\tau}(\theta_\circ^n)\right\} &= \mathbf{E}\left\{\psi_k\psi_{k-\tau}^T\right\}\mathbf{E}\left\{\varepsilon_k(\theta_\circ^n)\varepsilon_{k-\tau}(\theta_\circ^n)\right\} \\ &\quad + \mathbf{E}\left\{\psi_k\varepsilon_{k-\tau}(\theta_\circ^n)\right\}\mathbf{E}\left\{\psi_{k-\tau}^T\varepsilon_k(\theta_\circ^n)\right\} \end{aligned}$$

where use is made of the fact that by the definition of $\theta_\circ^n$, $\mathbf{E}\{\psi_k\varepsilon_k(\theta_\circ^n)\} = 0$. Furthermore, using Parseval's Theorem (suppressing the dependence on $\theta_\circ^n$)

$$\sum_{\tau=-\infty}^{\infty} \mathbf{E}\left\{\psi_k\psi_{k-\tau}^T\right\}\mathbf{E}\left\{\varepsilon_k\varepsilon_{k-\tau}\right\} = M_n(\Phi_u\Phi_\varepsilon).$$

Using an identical line of argument

$$\begin{aligned} &\sum_{\tau=-\infty}^{\infty} \mathbf{E}\left\{\psi_k\varepsilon_{k-\tau}\right\}\mathbf{E}\left\{\psi_{k-\tau}^T\varepsilon_k\right\} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \Gamma_n(\omega)\Gamma_n^T(\omega)\Phi_{u\varepsilon}^2(\omega)\,\mathrm{d}\omega \triangleq X_n. \end{aligned}$$

Now, in the case of all the poles $\{\xi_k\}$ being real, then since $\mathcal{B}_k(e^{j\omega}) = \overline{\mathcal{B}_k(e^{-j\omega})}$

$$\left|\frac{\Gamma_n^\star(\omega)X_n\Gamma_n(\omega)}{K_n(\omega,\omega)}\right| \le \frac{\|\Phi_{u\varepsilon}\|_\infty}{2\pi K_n(\omega,\omega)}\int_{-\pi}^{\pi} |K_n(\lambda,\omega)||K_n(-\omega,\lambda)|\,\mathrm{d}\lambda.$$

However, in [18] it is shown that expressions of this form tend to zero with increasing n provided all the poles $\{\xi_k\}$ are chosen within the open disc $\mathbf{D}$. This same result can also be shown to hold for the $\{\xi_k\}$ being complex, but at the expense of considerably more involved arithmetic which is not appropriate to document here. Therefore, since $Q_n = M_n(\Phi_u\Phi_\varepsilon) + X_n$ and it has just been established that $X_n \sim 0$ as $n \to \infty$, then $Q_n \sim M_n(\Phi_u\Phi_\varepsilon)$ as $n \to \infty$. ∎

References

[1] L. Ljung, "Asymptotic variance expressions for identified black-box transfer function models", *IEEE Trans. Automatic Control*, vol. AC-30, pp. 834–844, 1985.

[2] L. Ljung and Z.D. Yuan, "Asymptotic properties of black-box identification of transfer functions", *IEEE Trans. Automatic Control*, vol. AC-30, pp. 514–530, 1985.

[3] B. Wahlberg and L. Ljung, "Design variables for bias distribution in transfer function estimation", *IEEE Trans. Automatic Control*, vol. AC-31, pp. 134–144, 1986.

[4] G.C. Goodwin, M. Gevers, and D.Q. Mayne, "Bias and variance distribution in transfer function estimates", in *Proc. of IFAC/IFORS Symposium on Identification and Parameter Estimation*, 1991.

[5] B. Ninness and G. Goodwin, "Estimation of model quality", *Automatica*, vol. 31, pp. 32–74, 1995.

[6] M. Gevers and L. Ljung, "Optimal experiment designs with respect to the intended model application", *Automatica*, vol. 22, pp. 543–554, 1986.

[7] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, NJ, 1987.

[8] H. Hjalmarsson, M. Gevers, and F. De Bruyne, "For model based control design criteria, closed loop identification gives better performance", *Automatica*, vol. 32, pp. 1659–1673, December 1996.

[9] B. Wahlberg, "System identification using Laguerre models", *IEEE Trans. on Automatic Control*, vol. 36, pp. 551–562, 1991.

[10] P.M.J. Van den Hof, P.S.C. Heuberger, and J. Bokor, "System identification with generalized orthonormal basis functions", *Automatica*, vol. 31, pp. 1821–1834, 1995.

[11] B. Ninness, H. Hjalmarsson, and F. Gustafsson, "The fundamental role of general orthonormal bases in system identification", *IEEE Transactions on Automatic Control*, 1996, To appear.

[12] T. Söderström and P. Stoica, *System Identification*, Prentice-Hall International, Hemel Hempstead, Hertfordshire, 1989.

[13] M. Gevers, L. Ljung, and P.M.J. Van den Hof, "Asymptotic variance expressions for closed-loop identification and their relevance in identification for control", in *11th IFAC Symposium on System Identification*, Kitakyushu, Japan, 1997, vol. 1, pp. 1449–1454.

[14] L. Ljung, "Identification in closed loop: Some aspects on direct and indirect approaches", in *11th IFAC Symposium on System Identification*, Kitakyushu, Japan, 1997, vol. 1, pp. 141–146.

[15] R.R. Bitmead, M. Gevers, and V. Wertz, *Adaptive Optimal Control - The Thinking Man's GPC*, Prentice-Hall International, New York, 1990.

[16] P.S.C. Heuberger, P.M.J. Van den Hof, and O.H. Bosgra, "A generalized orthonormal basis for linear dynamical systems", *IEEE Transactions on Automatic Control*, vol. AC-40, no. 3, pp. 451–465, March 1995.

[17] B. Ninness and F. Gustafsson, "A unifying construction of orthonormal bases for system identification", *IEEE Transactions on Automatic Control*, vol. 42, no. 4, pp. 515–521, April 1997.

[18] B. Ninness, H. Hjalmarsson, and F. Gustafsson, "Generalised Fourier and Toeplitz results for rational orthonormal bases", *SIAM Journal on Control and Optimization*, 1996, To appear.