Exact Quantification of Variance Error

Brett Ninness* Håkan Hjalmarsson†

Abstract

This paper establishes a method for quantifying variance error in cases where the input spectral density has a rational factorisation. Compared to previous work which has involved asymptotic-in-model-order techniques and yielded expressions which are only approximate for finite orders, the quantifications provided here are exact for finite model order, although they still only apply asymptotically in the observed data length. The key principle employed here is the use of a reproducing kernel in order to characterise the model class, and this implies a certain geometric-type approach to the problem.

Draft Version of Technical Report EE2001050, Department of Electrical and Computer Engineering, University of Newcastle, AUSTRALIA

1 Introduction

There has been much work over the last several years directed at quantifying the noise induced error (variance error) in frequency response estimates [7, 6, 11, 14, 15, 10]. In these contributions, in the interests of tractable analysis, a common theme has been to consider the limiting value of the variance error as the model order grows, and then assume that this same value can be approximately used for finite model order.

This approach leaves open the question of the accuracy of the ensuing approximate quantification, and indeed recent works [11, 14, 15, 10] have specifically addressed the issue by using certain orthonormal basis techniques designed to maximise the finite model order accuracy.

The paper here takes a different approach, and shows how in some circumstances it is possible to quantify the variance error in a manner that is *exact* even for finite model order, although it is still asymptotic in the data length. The key technique used to achieve this is a geometric one, in which recognition of certain subspace invariants (reproducing kernels) provides a means for the exact quantification.

The importance of this issue of non-asymptotic in model order variance quantification has been recognised by other authors [17], who have employed an analysis technique that is quite different to that of this work. The penultimate section of this paper profiles the results of this approach versus the one taken here.

---

*This work was supported by the Australian Research Council. This author is with the Department of Electrical & Computer Engineering, University of Newcastle, Australia and can be contacted at email:`brett@ee.newcastle.edu.au` or FAX: +61 49 21 69 93. Part of this work was completed while visiting S3-Automatic Control, The Royal Institute of Technology, Stockholm, Sweden.

†This author is with the Department of Sensors, Signals and Systems (Automatic Control), The Royal Institute of Technology, S-100 44 Stockholm, Sweden and may be contacted at email:`hakan.hjalmarsson@s3.e.kth.se` or FAX: +46 8 790 7329

2 Problem Setting

The first estimation problem to be considered here is one in which observed input-output data $\{y_t\}$, $\{u_t\}$ is generated according to

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

Here $\{e_t\}$ is a zero-mean white noise sequence that satisfies $\mathbf{E}\{e_t^2\} = \sigma^2$. As well, $G(q)$ is assumed to be an n'th order rational transfer function with poles at $\{\xi_0, \cdots, \xi_{n-1}\}$, and $\{u_t\}$ is taken to be a second order stationary process with associated spectral density $0 < \Phi_u(\omega) < \infty$.

For the purposes of estimating the dynamics $G(q)$, it is supposed that the following parameterised model structure is used

$$y_t = G(q,\theta)u_t + e_t, \qquad G(q,\theta) = A^{-1}(q)\sum_{k=0}^{n-1}\theta_k q^k, \qquad A(q) = \prod_{k=0}^{n-1}(q-\xi_k). \tag{1}$$

Here, the fixed poles $\{\xi_0, \cdots, \xi_{m-1}\}$ are meant to be chosen via a-priori knowledge [14, 15, 4, 11], and the special case of $\xi_k = 0$ renders (1) as the common FIR model structure.

With this in mind, an estimate $\widehat{\theta}_N$ of $\theta \triangleq [\theta_0, \cdots, \theta_{n-1}]^T$ based on N observations of $\{y_t\}$ and $\{u_t\}$ may be found via a "least squares" approach:

$$\widehat{\theta}_N = \arg\min_\theta V_N(\theta), \qquad V_N(\theta) = \frac{1}{N}\sum_{t=1}^N \left(y_t - G(q,\theta)u_t\right)^2. \tag{2}$$

Furthermore, since $G(q)$ is assumed to be n'th order, then there exists a true parameter vector $\theta_\circ$ such that $G(q,\theta_\circ) = G(q)$ and hence with

$$\psi_t = \Lambda(q)u_t, \qquad \Lambda(q) \triangleq A^{-1}(q)[q^{n-1}, \cdots, q, 1]^T \tag{3}$$

it is clear that the expression

$$\widehat{\theta}_N - \theta_\circ = \left(\frac{1}{N}\sum_{t=1}^N \phi_t\phi_t^T\right)^{-1}\frac{1}{N}\sum_{t=1}^N \phi_t e_t \tag{4}$$

is a quantification of the parameter space estimation error. Furthermore, since the estimated frequency response $G(e^{j\omega}, \widehat{\theta}_N)$ depends linearly on the parameter estimate $\widehat{\theta}_N$ according to

$$G(e^{j\omega}, \widehat{\theta}_N) = A^{-1}(e^{j\omega})\Lambda^T(e^{j\omega})\widehat{\theta}_N$$

then the frequency response estimation error may be quantified from (4) as

$$G(e^{j\omega}, \widehat{\theta}_N) - G(e^{j\omega}) = A^{-1}(e^{j\omega})\Lambda^T(e^{j\omega})\left(\frac{1}{N}\sum_{t=1}^N \phi_t\phi_t^T\right)^{-1}\frac{1}{N}\sum_{t=1}^N \phi_t e_t. \tag{5}$$

Of course, this expression (5) depends on the exact noise realisation $\{e_t\}$ which cannot be known, so in order to quantify the estimation error $G(e^{j\omega}, \widehat{\theta}_N) - G(e^{j\omega})$ it is usual to consider its average over the ensemble of possible noise realisations. That is, consideration is given to $\mathrm{Var}\left\{G(e^{j\omega}, \widehat{\theta}_N)\right\}$, which by (5) and assuming open-loop data collection, may be computed as

$$\mathrm{Var}\left\{G(e^{j\omega}, \widehat{\theta}_N)\right\} = \frac{\sigma^2}{N}\cdot\frac{1}{|A(e^{j\omega})|^2}\cdot\Lambda^\star(e^{j\omega})R_N^{-1}\Lambda(e^{j\omega}), \qquad R_N \triangleq \frac{1}{N}\sum_{t=1}^N \phi_t\phi_t^T \tag{6}$$

where $\cdot^\star$ represents the operation of 'conjugate transpose'.

Now, while (6) gives an exact quantification of the noise induced estimation error, it does not expose very much about how experiment design choices (or other factors) might affect it. In consideration of this, the approach pioneered in works such as [16, 3, 7] is to notice that according to the assumptions made on the input

$$R \triangleq \lim_{N\to\infty} R_N = T_n\left(\frac{\Phi_u}{|A|^2}\right) \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} \Lambda(e^{j\lambda})\Lambda^\star(e^{j\lambda})\frac{\Phi_u(\lambda)}{|A(e^{j\lambda})|^2}\,\mathrm{d}\lambda. \tag{7}$$

Here, the notation $T_n(\Phi_u)$ is chosen to denote a Toeplitz matrix, of size $n \times n$, that is completely characterised by the spectral density function Φ_u. In fact, a well known feature of such matrices is that for large n, then

$$T_n^{-1}\left(\frac{\Phi_u}{|A|^2}\right) \approx T_n\left(\frac{|A|^2}{\Phi_u}\right)$$

where the approximation sign means that the Hilbert–Schmidt matrix norm [1] between the two quantities can be made arbitrarily small for arbitrarily large model roder n [2]. In this case, for the same large n it is inviting to approximate $\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\}$ from (6) as

$$\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\} \approx \frac{\sigma^2}{N}\cdot\frac{1}{|A(e^{j\omega})|^2}\cdot\Lambda^\star(e^{j\omega})T_n\left(\frac{|A|^2}{\Phi_u}\right)\Lambda(e^{j\omega}). \tag{8}$$

Finally, note that

$$\frac{1}{n}\Lambda^\star(e^{j\omega})T_n\left(\frac{|A|^2}{\Phi_u}\right)\Lambda(e^{j\omega}) = \frac{1}{2n\pi}\int_{-\pi}^{\pi}|\Lambda^\star(e^{j\omega})\Lambda(e^{j\lambda})|^2\frac{|A(e^{j\lambda})|^2}{\Phi_u(\lambda)}\,\mathrm{d}\lambda = \sum_{\tau=1-N}^{N-1}\left(1-\frac{|\tau|}{N}\right)c_\tau e^{j\omega\tau} \tag{9}$$

where the $\{c_\tau\}$ are the Fourier co-efficients of $|A|^2/\Phi_u$:

$$c_\tau \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{|A(e^{j\lambda})|^2}{\Phi_u(\lambda)}\,e^{-j\lambda\tau}\,\mathrm{d}\lambda. \tag{10}$$

Since the Fourier series in (9) can be expected to converge to the function $|A|^2/\Phi_u$ determining its co-efficients via (10), it is arguable via (8), (9) that for large model order n

$$\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\} \approx \frac{\sigma^2 n}{N}\cdot\frac{1}{|A(e^{j\omega})|^2}\cdot\frac{|A(e^{j\omega})|^2}{\Phi_u(\omega)} = \frac{n}{N}\frac{\sigma^2}{\Phi_u(\omega)}. \tag{11}$$

In contrast to (6), this approximate expression (11) very clearly exposes the key factors that contribute to the noise induced estimation error. As such, its use has become a fundamental part of the science of System Identification [5] since it was first presented for the FIR case ($\{\xi_k\} = 0$) in [7] and then extended to much more general model structures in [6, 11, 14, 15].

However, as argued in [10, 8, 9], this reliance can be misplaced. For example, the 'Achillee's Heel' of (11) is that it clearly depends on the model order n being large in order for the approximating steps to be accurate. In recognition of this, the work [10] has suggested that a quantification that can be more reliable for low model order n is

$$\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\} \approx \frac{1}{N}\cdot\frac{\sigma^2}{\Phi_u(\omega)}\cdot\sum_{k=0}^{n-1}\frac{1-|\xi_k|^2}{|e^{j\omega}-\xi_k|}. \tag{12}$$

Note that as pointed out in [10, 17], the greater the proportion of the poles $\{\xi_k\}$ located at the origin, the closer the approximation (12) is to the seminal one (11), and when all the poles are at the origin, then (11) becomes a special case of (12). Nevertheless, the approximation (12) still depends on high model order n in order to obtain high accuracy.

In recognition of this last point, the purpose of this paper is to derive an expression for $\lim_{N\to\infty} \mathrm{Var}\left\{G(e^{j\omega}, \widehat{\theta}_N)\right\}$ that is *exact* for finite model order n.

3 Mathematical Background

The key tool used in this paper is what might be called a geometric one, and it depends on the idea of a 'Reproducing Kernel'. To explain this, note that any frequency response of interest can be assumed to lie in a certain Hilbert Space H_2, where any geometric ideas such as orthogonality are obtained from the inner product between two functions $f, g \in H_2$ which is defined as

$$\langle f, g\rangle_\mu \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\lambda)\overline{g(\lambda)}\,\mu(\lambda)\,\mathrm{d}\lambda \tag{13}$$

with $\mu(\lambda)$ being some positive definite 'weighting' function.

Now, suppose that f is an element of an n-dimensional subspace X_n defined by certain *orthonormal* basis elements $\varphi_0, \cdots, \varphi_{n-1}$ as

$$f \in X_n \triangleq \mathrm{Span}\{\varphi_0, \cdots, \varphi_{n-1}\}.$$

Then it is a consequence of the Riesz Representation Theorem[12, 13] that for any fixed ω, a further element $K_n(\lambda, \omega) \in X_n$ exists such that

$$f(\omega) = \langle f(\lambda), K_n(\lambda, \omega)\rangle_\mu \qquad \forall f \in X_n. \tag{14}$$

This reproducing kernel will be vital to the arguments of this paper, so it is important to present its fundamental features. Firstly it is 'Hermitian Symmetric' in that, since for any fixed λ the kernel $K_n(\omega, \lambda) \in X_n$, then (integration in the inner product is with respect to the common variable)

$$K_n(\omega,\lambda) = \langle K_n(\sigma,\lambda), K_m(\sigma,\omega)\rangle_\mu = \overline{\langle K_n(\sigma,\omega), K_n(\sigma,\lambda)\rangle_\mu} = \overline{K_n(\lambda,\omega)}.$$

This then implies that $K_n(\omega, \lambda)$ is the *unique* element in X_n that has the property (14), since if another function $H_n(\omega, \sigma)$ also satisfied (14), then it would hold that

$$H_n(\omega,\lambda) = \overline{\overline{H_n(\omega,\lambda)}} = \overline{\langle H_n(\sigma,\lambda), K_n(\sigma,\omega)\rangle_\mu} = \langle K_n(\sigma,\omega), H_n(\sigma,\lambda)\rangle_\mu = K_n(\lambda,\omega).$$

Finally, although the reproducing kernel is unique, there may (of course) be many different ways of expressing it. One obvious one uses the orthonormal basis $\{\varphi_k\}$ for X_n to express the quantity as

$$K_n(\lambda,\omega) = \sum_{k=0}^{n-1} \varphi_k(\lambda)\overline{\varphi_k(\omega)}. \tag{15}$$

This formulation can be verified by noting that if $X_n \ni f = \sum_\tau c_\tau\varphi_\tau$ for some constants c_τ, then

$$\left\langle \sum_{\tau=0}^{n-1} c_\tau\varphi_\tau(\lambda), \sum_{k=0}^{n-1}\varphi_k(\lambda)\overline{\varphi_k(\omega)}\right\rangle_\mu = \sum_{\tau=0}^{n-1} c_\tau \sum_{k=0}^{n-1}\varphi_k(\omega)\langle\varphi_\tau(\lambda),\varphi_k(\lambda)\rangle_\mu = \sum_{\tau=0}^{n-1} c_\tau\varphi_\tau(\omega) = f(\omega). \tag{16}$$

4 Main Result

With these preliminary ideas in mind, then as just argued via (6) and (8)

$$\lim_{N\to\infty} N\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\} = \frac{\sigma^2}{|A(e^{j\omega})|^2}\cdot\Lambda^\star(e^{j\omega})T_n^{-1}\left(\frac{\Phi_u}{|A|^2}\right)\Lambda(e^{j\omega}). \tag{17}$$

The challenge then is to quantify this quadratic form in an exact manner, as opposed to the previously profiled approximation schemes based on asymptotic analysis. This ambition may be achieved by the following Theorem, which is the main result of the paper.

Theorem 4.1. *Suppose that $\Phi_u(\omega)$ has a finite dimensional rational spectral factorisation*

$$\Phi_u(\omega) = \kappa^2\Psi(e^{j\omega})\overline{\Psi(e^{j\omega})}, \qquad \Psi(z) \triangleq \prod_{t=0}^{\nu-1}\left(\frac{z-\beta_t}{z-\alpha_t}\right).$$

Then

$$\frac{1}{|A(e^{j\omega})|^2}\Lambda^\star(e^{j\omega})T_n^{-1}\left(\frac{\Phi_u}{|A|^2}\right)\Lambda(e^{j\omega}) = \frac{1}{\kappa^2}\sum_{k=0}^{n-1}|\varphi_k(e^{j\omega})|^2 \tag{18}$$

where $\{\varphi_k(z)\}$ is such that $\{\Psi(z)\varphi_0(z),\cdots,\Psi(z)\varphi_{n-1}(z)\}$ is an orthonormal set of functions satisfying

$$X_n \triangleq \mathrm{Span}\left\{\Psi(z)\varphi_0(z),\cdots,\Psi(z)\varphi_{n-1}(z)\right\} = \mathrm{Span}\left\{\Psi(z)\frac{z^{n-1}}{A(z)},\cdots,\Psi(z)\frac{z}{A(z)},\Psi(z)\frac{1}{A(z)}\right\}. \tag{19}$$

Proof. Note that as a simple consequence of the definition (7)

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\overline{\Psi(e^{j\omega})}}{A(e^{j\omega})}\Lambda_n^\star(e^{j\omega})T_n^{-1}\left(\frac{\Phi_u}{|A|^2}\right)\Lambda_n(e^{j\lambda})\kappa^2\frac{\Psi(e^{j\lambda})\overline{\Psi(e^{j\lambda})}}{A(e^{j\lambda})\overline{A(e^{j\lambda})}}\Lambda_n^\star(e^{j\lambda})\,\mathrm{d}\lambda = \frac{\overline{\Psi(e^{j\omega})}}{A(e^{j\omega})}\Lambda_n^\star(e^{j\omega}).$$

Therefore, the reproducing kernel $K_n(\lambda,\omega)$ for the space X_n defined in (19) is given by the following quadratic form:

$$K_n(\lambda,\omega) = \kappa^2\frac{\Psi(e^{j\lambda})\overline{\Psi(e^{j\omega})}}{A(e^{j\lambda})\overline{A(e^{j\omega})}}\Lambda^\star(e^{j\omega})T_n^{-1}\left(\frac{\Phi_u}{|A|^2}\right)\Lambda(e^{j\lambda}). \tag{20}$$

However, as established in (16), if $\{\Psi(z)\varphi_0(z),\cdots,\Psi(z)\varphi_{n-1}(z)\}$ is an orthonormal basis for X_n then

$$K_n(\lambda,\omega) = \sum_{k=0}^{n-1}\Psi(e^{j\lambda})\varphi_k(e^{j\lambda})\overline{\Psi(e^{j\omega})\varphi_k(e^{j\omega})} = \Psi(e^{j\lambda})\overline{\Psi(e^{j\omega})}\sum_{k=0}^{n-1}\varphi_k(e^{j\lambda})\overline{\varphi_k(e^{j\omega})}. \tag{21}$$

Equating (21) with (20) when $\lambda=\omega$ then gives (18). □

On account of (17), this result immediately gives the following exact (in n) formulation for variance error

$$\lim_{N\to\infty} N\cdot\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\} = \frac{\sigma^2}{\kappa^2}\sum_{k=0}^{n-1}|\varphi_k(e^{j\omega})|^2 \tag{22}$$

which implies the following error quantification, which is not asymptotic in model order n:

$$\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\} \approx \frac{1}{N}\cdot\frac{\sigma^2}{\kappa^2}\sum_{k=0}^{n-1}|\varphi_k(e^{j\omega})|^2. \tag{23}$$

Of course, for this to produce an explicit quantification it is necessary to compute the basis $\{\Psi(z)\varphi_k(z)\}$ for the space X_n defined in (19). The following examples will demonstrate how this may be achieved analytically.

5 First Order Example

In order to illustrate the preceding ideas, consider the simplest and lowest order example possible in which the observed data is obtained from the first order system

$$y_t = \left(\frac{1-\xi_0}{q-\xi_0}\right)u_t + e_t$$

where the spectral factor $\Psi(z)$ of Φ_u is of the form

$$\Psi(z) = \frac{z-\beta}{z-\alpha}.$$

Then an elementary calculation shows that

$$\left\|\Psi(z)\frac{1}{(z-\xi_0)}\right\|^2 = \frac{(1+\alpha\xi_0)(1+\beta^2)-2\beta(\alpha+\xi_0)}{(1-\alpha^2)(1-\alpha\xi_0)(1-|\xi_0|^2)}$$

and hence the unit norm version $\Psi(z)\varphi_0(z)$ of $\Psi(z)(z-\xi_0)^{-1}$ is

$$\Psi(z)\varphi_0(z) = \left[\frac{(1-\alpha^2)(1-\alpha\xi_0)}{(1+\alpha\xi_0)(1+\beta^2)-2\beta(\alpha+\xi_0)}\right]\Psi(z)\frac{1}{(z-\xi_0)}.$$

Therefore, according to (22)

$$\lim_{N\to\infty}\mathrm{Var}\{G(e^{j\omega},\widehat{\theta}_N^n)\} = \frac{\sigma^2}{\kappa^2}\cdot\frac{(1-\alpha^2)(1-|\xi_0|^2)(1-\alpha\xi_0)}{(1+\alpha\xi_0)(1+\beta^2)-2\beta(\alpha+\xi_0)}\cdot\frac{1}{|e^{j\omega}-\xi_0|^2}. \tag{24}$$

For the case of $\alpha = 0.5$, $\beta = 0$, $\xi_0 = 0.9$, $\kappa = 1$, $\sigma^2 = 1e-4$ the results of using the ensuing quantification

$$\mathrm{Var}\{G(e^{j\omega},\widehat{\theta}_N^n)\} \approx \frac{\sigma^2}{N\kappa^2}\cdot\frac{(1-\alpha^2)(1-|\xi_0|^2)(1-\alpha\xi_0)}{(1+\alpha\xi_0)(1+\beta^2)-2\beta(\alpha+\xi_0)}\cdot\frac{1}{|e^{j\omega}-\xi_0|^2}. \tag{25}$$

are shown in figure 1. Note that the expression (25) is (essentially) exact, which provides one example validating the claim of this paper that it is possible to derive accurate variance error quantifications that apply for arbitrary low model order.

Here, and in what follows, the true variability $\mathrm{Var}\{G(e^{j\omega},\widehat{\theta}_N^n)\}$ is estimated via the sample average over 10000 different realisations of input and measurement noise, and is illustrated as a solid line. The exact quantification, such as given by (25) is shown as a dashed line, and the well known approximation (11) is shown as a dash-dot line.

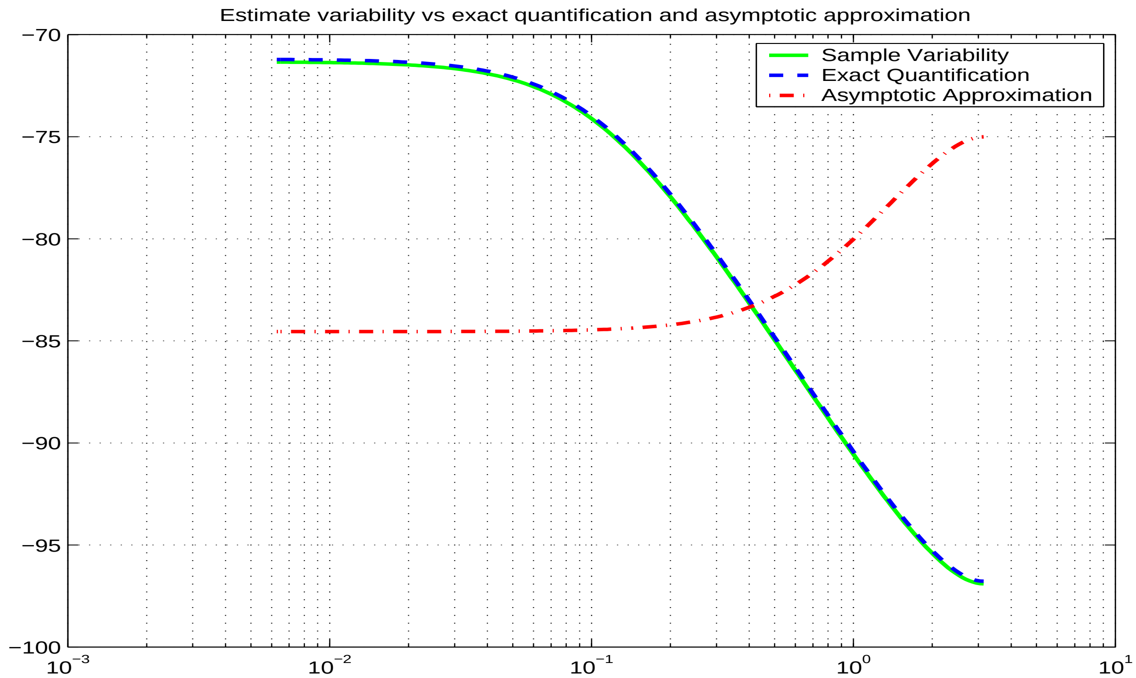

Figure 1: *Variability of Fixed Denominator Estimate - First Order Example. True variability vs. theoretically derived approximations. Solid line is Monte–Carlo estimate of true variability, dashed line is quantification (25), dash-dot line is asymptotic-based approximation (11).*

6 Second Order Example

Suppose we increase the model complexity in the previous example to that of

$$G(q) = \frac{(1-\xi_0)(1-\xi_1)}{(q-\xi_0)(q-\xi_1)}.$$

In order to compute the asymptotic variance associated with this model structure, it is necessary to compute the orthonormal basis $\{\Psi\varphi_0, \Psi\varphi_1\}$ that spans the same space as $\{\Psi(z)(z-\xi_0)^{-1}, \Psi(z)(z-\xi_1)^{-1}\}$.

In fact, this is quite difficult, primarily because of the restriction that all elements in this space have the same zeroes as $\Psi(z)$. Nevertheless it can be achieved via the well known Gramm–Schmidt procedure. Specifically, we have already established that

$$\Psi(z)\varphi_0(z) = \left[\frac{(1-\alpha^2)(1-\alpha\xi_0)}{(1+\alpha\xi_0)(1+\beta^2)-2\beta(\alpha+\xi_0)}\right]\Psi(z)\frac{1}{(z-\xi_0)}$$

is of unit norm, and hence a basis element $f(z)$ that is orthonormal to $\Psi(z)\varphi_0(z)$ and also such that $\text{Span}\{\Psi(z)\varphi_0, f\} = \text{Span}\{\Psi(z)(z-\xi_0)^{-1}, \Psi(z-\xi_1)^{-1}\}$ is given by

$$f(z) = \Psi(z)\frac{1}{(z-\xi_1)} - \left\langle \Psi(z)\frac{1}{(z-\xi_1)}, \Psi\varphi_0 \right\rangle \Psi(z)\varphi_0(z).$$

While this looks innocent enough, it evaluates to

$$f(z) = \frac{(z-\beta)(\gamma z+\delta)}{(z-\xi_0)(z-\xi_1)(z-\alpha)}$$

where

$$\gamma \triangleq -\frac{(|\xi_0|^2-1)(\beta-\alpha)(\alpha\beta-1)}{[(1+\alpha\xi_0)(1+\beta^2)-2\beta(\alpha+\xi_0)](\alpha\xi_1-1)} - \xi_0,$$

$$\delta \triangleq 1 + \frac{(|\xi_0|^2-1)(\beta-\alpha)(\alpha\beta-1)\xi_1}{[(1+\alpha\xi_0)(1+\beta^2)-2\beta(\alpha+\xi_0)](\alpha\xi_1-1)} \tag{26}$$

and hence,

$$\begin{aligned}\|f\|^2 \quad = \quad & \frac{(\xi_0-\beta)(\xi_0\gamma+\delta)(1-\beta\xi_0)(\gamma+\delta\xi_0)}{(\xi_0-\xi_1)(\xi_0-\alpha)(1-\xi_0^2)(1-\xi_0\xi_1)(1-\alpha\xi_0)} + \\ & \frac{(\xi_1-\beta)(\xi_1\gamma+\delta)(1-\beta\xi_1)(\gamma+\delta\xi_1)}{(\xi_1-\xi_0)(\xi_1-\alpha)(1-\xi_1^2)(1-\xi_0\xi_1)(1-\alpha\xi_1)} + \\ & \frac{(\alpha-\beta)(\alpha\gamma+\delta)(1-\beta\alpha)(\gamma+\delta\alpha)}{(\alpha-\xi_0)(\alpha-\xi_1)(1-\xi_0\alpha)(1-\xi_1\alpha)(1-\alpha^2)}.\end{aligned}$$

This is surprisingly complicated! Furthermore, it only applies for the case when all of ξ_0, ξ_1 and α are real valued and distinct. Nevertheless, with the definitions

$$K_0 \triangleq \frac{(1-\alpha^2)(1-\alpha\xi_0)(1-|\xi_0|^2)}{(1+\alpha\xi_0)(1+\beta^2)-2\beta(\alpha+\xi_0)}, \quad K_1 \triangleq \frac{1}{\|f\|^2} \tag{27}$$

it permits the computation in this situation for the second order case of

$$\mathrm{Var}\{G(e^{j\omega}, \widehat{\theta}_N^n)\} \approx \frac{\sigma^2}{N\kappa^2} \cdot \left[\frac{K_0}{|e^{j\omega}-\xi_0|^2} + \frac{K_1|\gamma e^{j\omega}+\delta|^2}{|e^{j\omega}-\xi_0|^2|e^{j\omega}-\xi_1|^2}\right] \tag{28}$$

The performance of this quantification is shown in figure 2, and again it is seen to be essentially exact.

7 Comparison with Other Approaches

The recent paper [17] has also tackled this problem of providing variance error quantification that is exact for finite model order. In that work, a completely different analysis approach is taken. Furthermore, some specific assumptions are made there that are different to those of this paper. Specifically, in [17] it is required that

1. The spectral factor $\Psi(z)$ is strictly all-pole (ie. $\{u_t\}$ is a strictly auto-regressive time series);
2. The model order chosen must be greater than or equal to the order n of the underlying dynamics *plus* the order ν of the spectral factor $\Psi(z)$;
3. In the model structure (1), all fixed poles ξ_k for $n < k \leq n+\nu$ are set to zero.

With these conditions in mind, the quantification proposed in [17] is

$$\lim_{N\to\infty} N\mathrm{Var}\left\{G(e^{j\omega}, \widehat{\theta}_N)\right\} = \frac{\sigma^2}{\Phi_u(\omega)}\left(\sum_{k=0}^{n-1}\frac{1-|\xi_k|^2}{|e^{j\omega}-\xi_k|^2} + \sum_{k=0}^{\nu-1}\frac{1-|\alpha_k|^2}{|e^{j\omega}-\alpha_k|^2}\right). \tag{29}$$

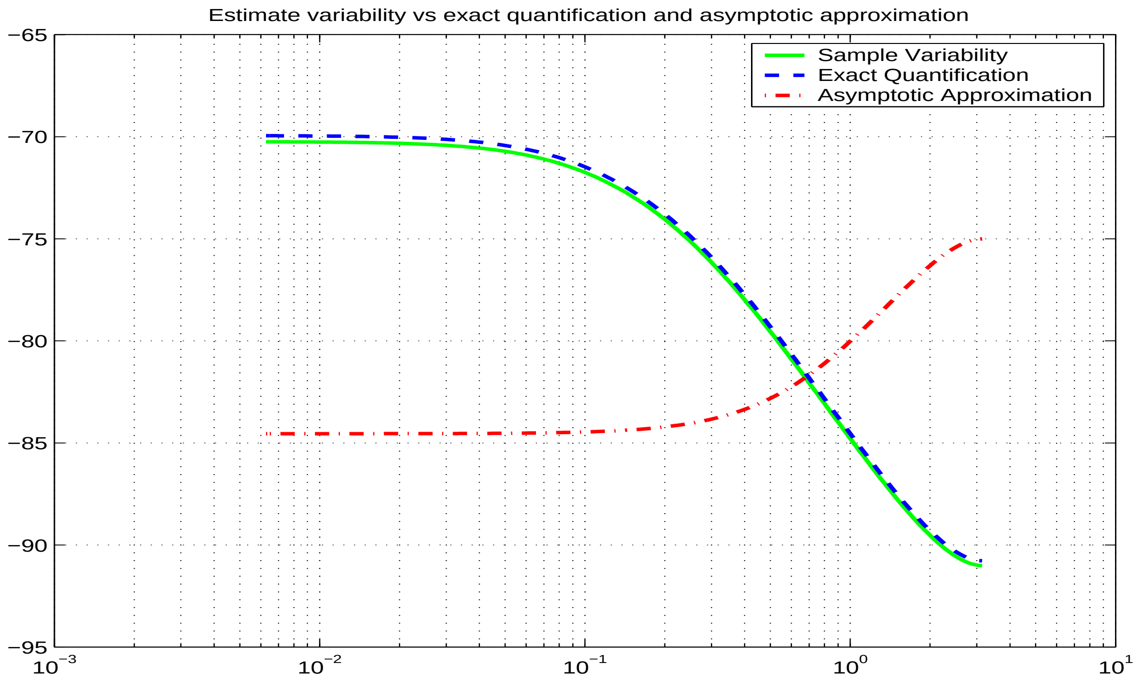

Figure 2: *Variability of Fixed Denominator Estimate: True variability vs. theoretically derived approximations. Solid line is Monte–Carlo estimate of true variability, dashed line is approximation (28), dash-dot line is asymptotic-based approximation (11)*

Clearly, it is important to reconcile this expression with that derived here. For this purpose, consider the first order example of §5 in which case, since $\Psi(z)$ is also first order in that example, to satisfy the above requirements of [17] a second order model of the form

$$G(q,\theta) = \frac{\theta_0 + \theta_1 q}{q(q-\xi_0)}$$

needs to be employed. In this case, substituting $\xi_1 = 0$, $\beta = 0$ into the analysis of §6 leads to a quantification of

$$\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\} \approx \frac{\sigma^2}{N\kappa^2}\cdot\frac{K_0 + K_1|\gamma e^{j\omega}-\delta|^2}{|e^{j\omega}-\xi_0|^2} \tag{30}$$

with γ, δ, K_0 and K_1 given by (6), (27). On the other hand, according to (29), the analysis of [17] gives a quantification of:

$$\mathrm{Var}\left\{G(e^{j\omega},\widehat{\theta}_N)\right\} \approx \frac{\sigma^2}{N\kappa^2}\cdot\frac{(1-|\xi_0|^2)|e^{j\omega}-\alpha|^2 + (1-|\alpha|^2)|e^{j\omega}-\xi_0|^2}{|e^{j\omega}-\xi_0|^2} \tag{31}$$

The performance of both these quantifications for the case of $\xi_0 = 0.9, \alpha = 0.5$ is shown in figure 3. Clearly, the quantification (30) seems to be exact, while there must be some sort of error in the quantification (31), since it is very different to the true variability.

8 Conclusions

The good news presented here is that in cases of rational input spectral density, it is possible to obtain variance quantifications that apply for arbitrarily low model order. The bad news is that the ana-

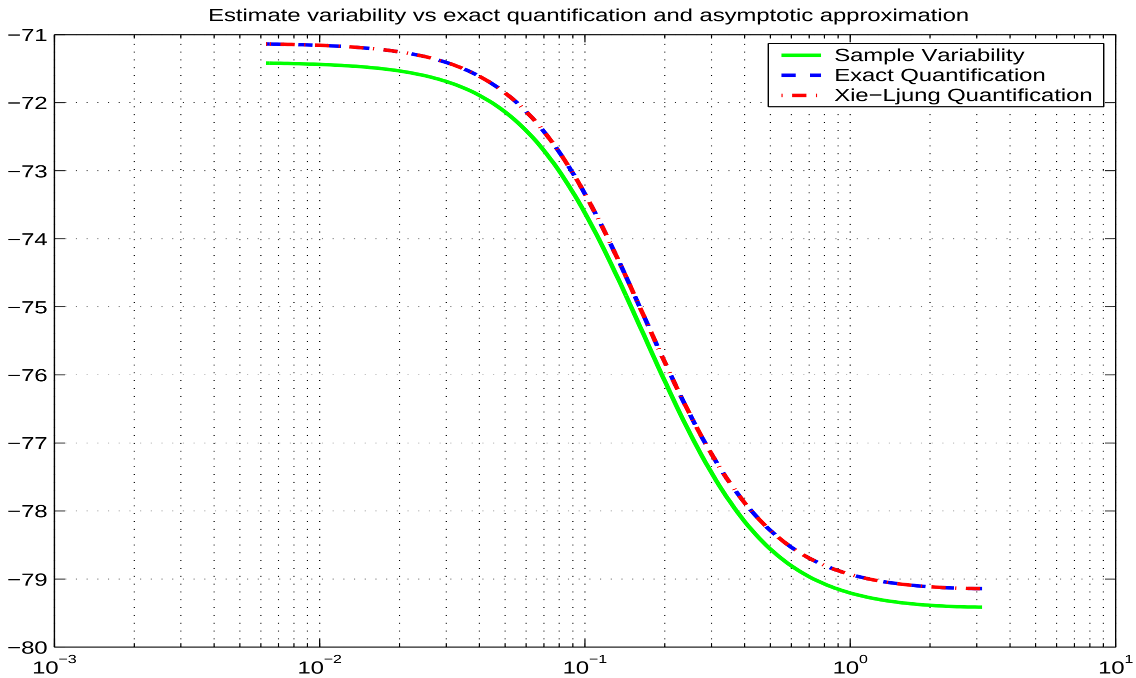

Figure 3: *Variability of Fixed Denominator Estimate - Second Order Case. True variability vs. theoretically derived approximations. Solid line is Monte–Carlo estimate of true variability, dashed line is approximation (30), dash-dot line is Xie–Ljung quantification (31)*

lytical manipulations required to compute the necessary orthonormal basis lead to very complicated expressions; this happens because some zeroes as well as all the poles are constrained in the required orthonormal spanning elements. This implies that some sort of numerical procedure may need to be developed to compute these quantities.

This work is related to work by other authors, but there are important differences. In particular, the results here are derived via new reproducing kernel techniques that avoid the need for the detailed residue-based computations employed in [17]. This approach allows for the inclusion of more general classes of input spectra and also avoids the requirement of model order being larger than the underlying true dynamics. However, because of this generality the quantifications here are more complicated than those of [17].

References

[1] G. Golub and C. V. Loan, *Matrix Computations*, Johns Hopkins University Press, 1989.

[2] U. Grenander and G. Szegö, *Toeplitz Forms and their Applications*, University of California Press, Berkeley, 1958.

[3] E. Hannan and D. Nicholls, *The estimation of the prediction error variance*, Journal of the American Statistical Association, 72 (1977), pp. 834–840.

[4] P. Heuberger, P.M.J. Van den Hof, and O. Bosgra, *A generalized orthonormal basis for linear dynamical systems*, IEEE Transactions on Automatic Control, AC-40 (1995), pp. 451–465.

[5] L. Ljung, *System Identification: Theory for the User, (2nd edition)*, Prentice-Hall, Inc., New Jersey, 1999.

[6] L.Ljung, *Asymptotic variance expressions for identified black-box transfer function models*, IEEE Transactions on Automatic Control, AC-30 (1985), pp. 834–844.

[7] L.Ljung and Z.D.Yuan, *Asymptotic properties of black-box identification of transfer functions*, IEEE Transactions on Automatic Control, 30 (1985), pp. 514–530.

[8] B. Ninness and H. Hjalmarsson, *The quantification of variance error part I: Fundamental principles*, Submitted to IEEE Transactions on Automatic Control, (2001).

[9] ——, *The quantification of variance error part II: General models and applications*, Submitted to IEEE Transactions on Automatic Control, (2001).

[10] B. Ninness, H. Hjalmarsson, and F. Gustafsson, *The fundamental role of general orthonormal bases in system identification*, IEEE Transactions on Automatic Control, 44 (1999), pp. 1384–1406.

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

[12] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Fredrick Ungar, New York, 1955.

[13] W. Rudin, *Real and Complex Analysis*, McGraw Hill, New York, 1966.

[14] B. Wahlberg, *System identification using Laguerre models*, IEEE Transactions on Automatic Control, AC-36 (1991), pp. 551–562.

[15] B. Wahlberg, *System identification using Kautz models*, IEEE Transactions on Automatic Control, AC-39 (1994), pp. 1276–1282.

[16] P. Whittle, *Estimation and information in stationary time series*, Arkiv För Matematik, 2 (1953), pp. 423–434.

[17] L.-L. Xie and L. Ljung, *Asymptotic variance expressions for estimated frequency functions*, Tech. Rep. LiTH-ISY-R-2277, Department of Electrical Engineering, Linkping University, SE-581 83 Linkping, Sweden, Aug 2000.