Fundamental Limitations on the Variance of Estimated Parametric Models

Cristian R. Rojas, James S. Welsh, *Member, IEEE* and Juan C. Agüero, *Member, IEEE*

***Abstract*—In this paper fundamental integral limitations are derived on the variance of estimated parametric models, for both open and closed loop identification. As an application of these results we show that, for multisine inputs, a well known asymptotic (in model order) variance expression provides upper bounds on the actual variance of the estimated models for finite model orders. The fundamental limitations established here give rise to a 'water-bed' effect, which is illustrated in an example.**

I. INTRODUCTION

Fundamental Limitations are of importance since they quantify the possible and the impossible. In feedback control, the development of fundamental limitations has given insight and understanding of the achievable performance of a feedback control system [1]–[3]. Knowledge of these limitations also allows informed decisions to be made regarding the tradeoffs between conflicting performance criteria, e.g. the Bode integral shows that increasing performance in a particular frequency region will reduce performance in another. This is known as the *water-bed effect* [3].

The original motivation for the study of fundamental limitations was in feedback design for control systems. However there have been a number of limitations developed in other areas. For example, the Cramér-Rao Bound is an important relationship in estimation theory [4], [5]. In information theory there is the Shannon Theorem [6], which is sometimes known as the fundamental theorem of information theory. Again the limitations described by these two results give inescapable performance bounds.

To date, there has been relatively few publications dealing with fundamental limitations in system identification. Previous work in this area has examined integral constraints on systematic errors (bias) for least-squares estimators [7]–[10]. In spectral estimation, a fundamental limitation has been developed in [11]–[13]. Specifically, an integral constraint on the relative variance was established for the parametric estimation of a signal spectrum. This result was used to demonstrate the 'water-bed' effect in spectral estimation [13].

The current paper differs from [11]–[13] in that we establish fundamental limitations on the variance of estimated parametric models possessing exogenous inputs. Specifically we show the relationship between the results presented in [11]–[13], which obtain a lower bound on the variance of parametric spectral estimators, with the new results obtained in this paper. Fundamental limitations are obtained for both open and closed loop identification. With respect to closed loop identification, both direct and indirect methods are considered. For the case of direct identification, bounds are established in lieu of an exact expression.

The paper is organised as follows. Section II describes the problem set-up. Section III establishes a fundamental limitation for open loop identification. In Section IV we show the relationship between existing results and the new results established in this paper. Fundamental limitations are established for closed loop identification in Section V. In Section VI, bounds on the variance of estimated models are obtained. Section VII shows, via an example, the tradeoffs dictated by the fundamental limitations. Finally, Section VIII presents conclusions.

II. PROBLEM SET-UP

Consider a single-input single-output (SISO) linear system given by

$$y(t) = G_0(q)u(t) + H_0(q)w(t),$$

where $\{u(t)\}$ is a quasi-stationary signal [14] and $\{w(t)\}$ is a zero mean Gaussian white noise sequence with variance σ^2. The operator q is the unit shift and H_0 is assumed to be a stable minimum phase transfer function with $H_0(\infty) = 1$. For simplicity we denote $H_0(q)w(t)$ by $v(t)$.

Given the input-output data pairs $\{u(t), y(t)\}_{t=1}^{N}$, a model,

$$y(t) = G(q,\theta)u(t) + H(q,\theta)\epsilon(t),$$

is inferred. We assume no undermodeling, i.e. there exists a $\theta = \theta_0$ such that $G_0(q) = G(q,\theta_0)$ and $H_0(q) = H(q,\theta_0)$. Furthermore, we assume that the estimators of G_0 and H_0 are asymptotically efficient (e.g. Maximum Likelihood (ML), or Prediction Error Methods (PEM) for Gaussian disturbances).

We define the spectrum of a quasi-stationary signal $\{x(t)\}$, according to [14], as

$$\Phi_x(e^{j\omega}) := \sum_{\tau=-\infty}^{\infty} R_x(\tau)e^{-j\omega\tau}, \quad \omega \in [-\pi, \pi],$$

where $R_x(\tau) := \bar{E}\{x(t)x(t-\tau)\}$ is the autocovariance of $\{x(t)\}$, and $\bar{E}\{f(t)\} := \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} E\{f(t)\}$.

To reduce notation, we omit the argument $e^{j\omega}$ in the functions, and assume that all integrals are taken with respect to $\omega \in [-\pi, \pi]$. Integrals in the sequel are also assumed to exist and to be finite. Estimators are denoted by a superscript '$\widehat{\ }$' and implicitly depend on the data length, N. Covariance

C. R. Rojas is with the ACCESS Linnaeus Center, Electrical Engineering, KTH-Royal Institute of Technology, S-100 44 Stockholm, Sweden. E-mail: cristian.rojas@ee.kth.se.

J. S. Welsh and J. C. Agüero are with the School of Electrical Engineering and Computer Science, The University of Newcastle, NSW 2308, Australia. E-mails: james.welsh@newcastle.edu.au and juan.aguero@newcastle.edu.au.

expressions are valid as $N \to \infty$ [14] (i.e. they are correct up to order[1] $1/N$). Note that all quantities involved in the fundamental limitations must be evaluated at their true values.

III. A Fundamental Limitation in Open Loop

We first consider the open loop case, i.e. when $\{u(t)\}$ and $\{w(t)\}$ are independent, and develop an integral constraint on the variance of an estimated parametric model.

Theorem 1 (Limitations in Open Loop Identification): In open loop identification, where G and H are independently parameterised with n_G and n_H parameters respectively, and $(G(q,\theta_G), H(q,\theta_H))$ is *parameter identifiable* under Φ_u for the ML method [15], then

$$\frac{1}{2\pi}\int \frac{\Phi_u}{\Phi_v}\mathrm{Var}[\hat{G}] = \frac{n_G}{N}, \tag{1}$$

$$\frac{1}{2\pi}\int \frac{\sigma^2}{\Phi_v}\mathrm{Var}[\hat{H}] = \frac{n_H}{N}. \tag{2}$$

Proof: We start with the covariance expressions [14, Section 9.4],

$$\mathrm{Var}[\hat{\theta}_G] = \left[\frac{N}{2\pi}\int \Gamma_G \Gamma_G^H \frac{\Phi_u}{\Phi_v}\right]^{-1},$$
$$\mathrm{Var}[\hat{\theta}_H] = \left[\frac{N}{2\pi}\int \Gamma_H \Gamma_H^H \frac{\sigma^2}{\Phi_v}\right]^{-1},$$

where θ_G and θ_H are the parameter vectors of G and H, respectively, and

$$\Gamma_G := \frac{\partial G}{\partial \theta_G}, \quad \Gamma_H := \frac{\partial H}{\partial \theta_H},$$

By the *Gauss' approximation formula* [14],

$$\mathrm{Var}[\hat{G}] = \Gamma_G^H \mathrm{Var}[\hat{\theta}_G]\Gamma_G, \quad \mathrm{Var}[\hat{H}] = \Gamma_H^H \mathrm{Var}[\hat{\theta}_H]\Gamma_H.$$

Therefore,

$$\begin{aligned}
&\frac{1}{2\pi}\int \frac{\Phi_u}{\Phi_v}\mathrm{Var}[\hat{G}] \\
&= \frac{1}{N}\mathrm{Tr}\left\{\left[\frac{1}{2\pi}\int \Gamma_G\Gamma_G^H\frac{\Phi_u}{\Phi_v}\right]\left[\frac{1}{2\pi}\int \Gamma_G\Gamma_G^H\frac{\Phi_u}{\Phi_v}\right]^{-1}\right\} \\
&= \frac{n_G}{N}.
\end{aligned}$$

A similar argument applies to the integral of $\mathrm{Var}[\hat{H}]$. ■

Remark 1: In general, if Φ_u is not persistently exciting of order n_G, the integral of $\mathrm{Var}[\hat{G}]$ will not be proportional to n_G. However, the integral will be proportional to the rank of the information matrix of θ_G, that is, to the number of spectral lines of Φ_u.

It can be seen from (1) that, under the assumption of no undermodelling, a 'water-bed' effect exists on the variance of $\hat{G}$, since if $\hat{G}$ is small for some frequencies, it must necessarily be large for others, in order to satisfy (1).

[1]Loosely speaking, this means that all expressions in the sequel which involve variances, of the form $A = (\geq) B$, should be interpreted as $\lim_{N\to\infty} NA = \lim_{N\to\infty} NB$.

IV. Relationship to Previous Results

Theorem 1 establishes a fundamental limitation on the variance of estimators of the transfer functions G_0 and H_0. A result has been derived for the variance of spectral estimators in [11]–[13], which essentially establishes (in the notation of Theorem 1) that

$$\frac{1}{2\pi}\int \frac{1}{|H_0|^4}\mathrm{Var}[|\hat{H}|^2] = \frac{2n_H}{N}. \tag{3}$$

In [11], [12] the term on the left side of (3) is considered as a measure of the accuracy of a spectral estimator, hence (3) provides a lower bound on the spectral accuracy. Note that, according to [13], (3) imposes a water-bed effect on the variance of an asymptotically efficient spectral estimator.

The results presented in Theorem 1 differ from those established in [11]–[13]. This difference is highlighted by the fact that Theorem 1 is based on Ljung's covariance expression [14]:

$$\mathrm{Cov}[\hat{\theta}_H] = \left[\frac{N}{2\pi}\int\left(\frac{\partial H}{\partial\theta_H}\right)\left(\frac{\partial H}{\partial\theta_H}\right)^H\frac{\sigma^2}{\Phi_v}\right]^{-1},$$

which is the base to develop both exact and asymptotic (in model order) variance expressions [16], [17], whilst [11]–[13] rely on '*Whittle's formula*' for the asymptotic covariance of an asymptotically efficient estimator $\hat{\theta}_H$ of θ_H [18], [19]:

$$\mathrm{Cov}[\hat{\theta}_H] = \left[\frac{N}{4\pi}\int\left(\frac{\partial \Phi_v}{\partial\theta_H}\right)\left(\frac{\partial \Phi_v}{\partial\theta_H}\right)^T\frac{1}{\Phi_v^2}\right]^{-1}.$$

Now by the Gauss' Approximation Formula,

$$\begin{aligned}
&\mathrm{Var}[|\hat{H}|^2] \\
&= \frac{1}{\sigma^4}\left(\frac{\partial \Phi_v}{\partial\theta_H}\right)^T\left[\frac{N}{4\pi}\int\left(\frac{\partial \Phi_v}{\partial\theta_H}\right)\left(\frac{\partial \Phi_v}{\partial\theta_H}\right)^T\frac{1}{\Phi_v^2}\right]^{-1}\frac{\partial \Phi_v}{\partial\theta_H}
\end{aligned} \tag{4}$$

and

$$\mathrm{Var}[\hat{H}] = \left(\frac{\partial H}{\partial\theta_H}\right)^H\left[\frac{N}{2\pi}\int\left(\frac{\partial H}{\partial\theta_H}\right)\left(\frac{\partial H}{\partial\theta_H}\right)^H\frac{\sigma^2}{\Phi_v}\right]^{-1}\frac{\partial H}{\partial\theta_H}. \tag{5}$$

Since $\Phi_v = \sigma^2|H_0|^2$, it might seem straightforward to relate (4) and (5) by establishing a connection between $\partial\Phi_v/\partial\theta_H$ and $\partial H/\partial\theta_H$. However, due to the complex-valued nature of H_0, this is not possible in general (even though the respective expressions for $\mathrm{Cov}[\hat{\theta}_H]$, Ljung's expression and Whittle's formula, are in fact equivalent [15, Problem 7.19]). Nonetheless, we can relate their integrals, as shown in the following theorem, which establishes the relationship between our result (Theorem 1) and the results in [11]–[13].

Theorem 2 (Relationship to Previous Results): Let H_0 be a stable minimum phase transfer function such that $H_0(\infty) = 1$ and $H_0(\overline{z}) = \overline{H_0(z)}$ for all $z \in \mathbb{C}$. Also, let $\hat{H}_N$ be an asymptotically efficient estimator of H_0 (subject to the same constraints imposed on H_0), where N is the number of samples. Then,

$$\frac{1}{2\pi}\int\frac{1}{|H_0|^2}\mathrm{Var}[\hat{H}_N] = \frac{1}{4\pi}\int\frac{1}{|H_0|^4}\mathrm{Var}[|\hat{H}_N|^2].$$

Proof: See Appendix A. ■

Theorem 2 links the results of Theorem 1 for the noise transfer function with the results of [11]–[13], thus showing that they can be considered, in some sense, equivalent. However, Theorem 1 also establishes a similar result for G, which has no resemblance with previous results in the literature, since [11]–[13] do not consider exogenous signals.

V. Fundamental limitations in Closed Loop Identification

In closed loop identification, we consider the input $\{u(t)\}$ to be generated as

$$u(t) = r(t) - C(q)y(t),$$

where $\{r(t)\}$ is a quasi-stationary reference signal, independent of $\{w(t)\}$.

In order to derive fundamental limitations, analogous to those in the open loop case, for closed-loop identification, we let:

$$\begin{aligned} S &= \frac{1}{1+GC}, \qquad G_{cl} = \frac{G}{1+GC}, \qquad H_{cl} = \frac{H}{1+GC}, \\ \Phi_u^r &= \Phi_r |S|^2, \qquad \Phi_v = \sigma^2 |H|^2. \end{aligned} \tag{6}$$

Then

$$\frac{\partial G_{cl}}{\partial \theta} = S^2 \frac{\partial G}{\partial \theta}, \qquad \frac{\partial H_{cl}}{\partial \theta} = -HCS^2 \frac{\partial G}{\partial \theta} + S \frac{\partial H}{\partial \theta}.$$

Thus,

$$\Gamma_{cl} = \Gamma_{ol} \begin{bmatrix} S^2 & -HCS^2 \\ 0 & S \end{bmatrix}, \tag{7}$$

where

$$\Gamma_{cl} := \begin{bmatrix} \frac{\partial G_{cl}}{\partial \theta} & \frac{\partial H_{cl}}{\partial \theta} \end{bmatrix}, \quad \Gamma_{ol} := \begin{bmatrix} \frac{\partial G}{\partial \theta} & \frac{\partial H}{\partial \theta} \end{bmatrix}.$$

Remark 2: In the case where a reference prefilter, say $F(q)$, is present, the expressions of this section can be easily adapted accordingly, by replacing Φ_r with $|F|^2\Phi_r$.

In the sequel we assume that $(G(q,\theta_G), H(q,\theta_H))$ is *parameter identifiable* under Φ_r for the ML method [15].

A. General Case

Theorem 3 (Limitations in Closed Loop Identification): In the closed loop case, i.e. where $\{u(t)\}$ and $\{w(t)\}$ are not necessarily independent, and G with H are not necessarily independently parameterised, with a common parameter vector $\theta \in \mathbb{R}^{n_\theta}$, then

$$\frac{1}{2\pi}\int \mathrm{Tr}\left\{ \begin{bmatrix} \frac{\Phi_u^r}{\Phi_v} + |CS|^2 & -\frac{CS}{\overline{H}} \\ -\frac{\overline{CS}}{H} & \frac{1}{|H|^2} \end{bmatrix} \mathrm{Cov}\begin{bmatrix} \hat{G} \\ \hat{H} \end{bmatrix} \right\} = \frac{n_\theta}{N},$$

Proof: Note that

$$\begin{aligned} \mathrm{Cov}[\hat{\theta}] &= \frac{2\pi}{N}\left\{ \int \frac{1}{|S|^2\Phi_v} \Gamma_{cl} \begin{bmatrix} \Phi_r & 0 \\ 0 & \sigma^2 \end{bmatrix} \Gamma_{cl}^H \right\}^{-1} \\ &= \frac{2\pi}{N}\left\{ \int \frac{1}{|S|^2\Phi_v} \Gamma_{ol} \begin{bmatrix} S^2 & -HCS^2 \\ 0 & S \end{bmatrix} \begin{bmatrix} \Phi_r & 0 \\ 0 & \sigma^2 \end{bmatrix} \right. \\ &\qquad \left. \cdot \begin{bmatrix} S^2 & -HCS^2 \\ 0 & S \end{bmatrix}^H \Gamma_{ol}^H \right\}^{-1}. \end{aligned}$$

Now,

$$\begin{aligned} &\frac{1}{|S|^2\Phi_v} \begin{bmatrix} S^2 & -HCS^2 \\ 0 & S \end{bmatrix} \begin{bmatrix} \Phi_r & 0 \\ 0 & \sigma^2 \end{bmatrix} \begin{bmatrix} S^2 & -HCS^2 \\ 0 & S \end{bmatrix}^H \\ &\qquad = \begin{bmatrix} \frac{\Phi_u^r}{\Phi_v} + |CS|^2 & -\frac{CS}{\overline{H}} \\ -\frac{\overline{CS}}{H} & \frac{1}{|H|^2} \end{bmatrix}. \end{aligned}$$

By utilising the *Gauss' approximation formula* [14] we have that

$$\mathrm{Cov}\begin{bmatrix} \hat{G} \\ \hat{H} \end{bmatrix} = \Gamma_{ol}^H \mathrm{Cov}[\hat{\theta}] \Gamma_{ol}.$$

The rest of the proof follows similar lines to Theorem 1. ■

B. Indirect Identification

Theorem 4 (Limitations in Indirect Identification): In the indirect closed loop identification case [20], i.e. where $\{u(t)\}$ and $\{w(t)\}$ are not necessarily independent, however G_{cl} and H_{cl} are independently parameterised with $n_{G_{cl}}$ and $n_{H_{cl}}$ parameters, respectively, so that the closed loop is described by $y(t) = G_{cl}(q)r(t) + H_{cl}(q)w(t)$, we have that

$$\frac{1}{2\pi}\int \frac{\Phi_u^r}{\Phi_v} \mathrm{Var}[\hat{G}] = \frac{n_{G_{cl}}}{N}.$$

Proof: With G_{cl} and H_{cl} independently parameterised, we can essentially consider the open loop case (with Φ_r, $|S|^2\Phi_v$, G_{cl} and H_{cl} instead of Φ_u, Φ_v, G and H, respectively). Thus, from Theorem 1 we have that

$$\frac{1}{2\pi}\int \frac{\Phi_r}{|S|^2\Phi_v} \mathrm{Var}[\hat{G_{cl}}] = \frac{n_{G_{cl}}}{N}. \tag{8}$$

From (7) and the *Gauss' approximation formula* we relate the variance of G_{cl} and G as

$$\mathrm{Var}[\hat{G}_{cl}] = |S|^4 \mathrm{Var}[\hat{G}]. \tag{9}$$

Substituting (9) into (8) gives

$$\frac{1}{2\pi}\int \frac{|S|^2\Phi_r}{\Phi_v} \mathrm{Var}[\hat{G}] = \frac{n_{G_{cl}}}{N},$$

which completes the proof. ■

Corollary 1 (Tailor-made parametrisation): For indirect closed loop identification with a tailor-made parametrisation [21] (i.e. when $\{u(t)\}$ and $\{w(t)\}$ are not necessarily independent), but when G and H_{cl} are independently parameterised with n_G and $n_{H_{cl}}$ parameters respectively (as in (6)), we have that

$$\frac{1}{2\pi}\int \frac{\Phi_u^r}{\Phi_v} \mathrm{Var}[\hat{G}] = \frac{n_G}{N}.$$

Proof: This follows from Theorem 4 and the fact that G_{cl} and H_{cl} are independently parameterised, with $n_{G_{cl}} = n_G$ and $n_{H_{cl}}$ parameters, respectively. ■

C. Direct Identification

In direct closed-loop identification it is difficult to establish an exact integral constraint for the fundamental limitation. However, based on results from [22], [23], the following bounds are established:

Theorem 5 (Limitations in Direct Identification): In direct closed loop identification [20], i.e. when $\{u(t)\}$ and $\{w(t)\}$ are not necessarily independent, however G and H are independently parameterised with n_G and n_H parameters respectively, then

$$\frac{1}{2\pi}\int \frac{\Phi_u}{\Phi_v}\mathrm{Var}[\hat{G}] \geq \frac{n_G}{N},$$
$$\frac{1}{2\pi}\int \frac{1}{\Phi_v}\left(\Phi_u - \frac{\Phi_{uw}}{\sigma^2}\right)\mathrm{Var}[\hat{G}] \leq \frac{n_G}{N}.$$

Proof: By applying the Cauchy-Schwarz inequality, the following inequalities can be obtained [23]:

$$\mathrm{Var}[\hat{\theta}_G] \geq \left[\frac{N}{2\pi}\int \Gamma_G\Gamma_G^H \frac{\Phi_u}{\Phi_v}\right]^{-1},$$
$$\mathrm{Var}[\hat{\theta}_G] \leq \left[\frac{N}{2\pi}\int \Gamma_G\Gamma_G^H \frac{1}{\Phi_v}\left(\Phi_u - \frac{\Phi_{uw}}{\sigma^2}\right)\right]^{-1}.$$

By the *Gauss' approximation formula*,

$$\mathrm{Var}[\hat{G}] = \Gamma_G^H \mathrm{Var}[\hat{\theta}_G]\Gamma_G.$$

The rest of the proof follows similar lines to Theorem 1. ■

Remark 3: Notice that the inequalities of Theorem 5 are valid even if the controller C is nonlinear and/or time varying, provided $\{u(t)\}$ is quasi-stationary. If the controller is linear and time invariant, the expression $\Phi_u - \Phi_{uw}/\sigma^2$ in the second inequality of Theorem 5 corresponds to Φ_u^r, as defined in (6).

Remark 4: In the open-loop case, i.e. when $\Phi_{uw} = 0$, the combination of both inequalities of Theorem 5 gives the result of Theorem 1.

VI. Bounds on the Variance

As an application of the above results, we show that for an input comprising multisines, the asymptotic covariance expression [14]

$$\mathrm{Var}[\hat{G}(e^{j\omega})] = \frac{n}{N}\frac{\Phi_v(e^{j\omega})}{\Phi_u(e^{j\omega})},$$

provides an upper bound on the variance of G, irrespective of the model structure.

Consider the open-loop case, with a multisine input of the form

$$\Phi_u(e^{j\omega}) = \sum_{i=1}^{m} 2\pi U_i \delta(\omega - \omega_i),$$

where $\omega_i \in [-\pi, \pi]$, $U_i > 0$ for every $i = 1, \ldots, m$, and where Φ_u is even. For identifiability reasons, we assume that $m \geq n_G$, the number of parameters in G.

By Theorem 1, we have that

$$\sum_{i=1}^{m} \frac{U_i}{\Phi_v(e^{j\omega_i})}\mathrm{Var}[\hat{G}(e^{j\omega_i})] = \frac{n_G}{N}.$$

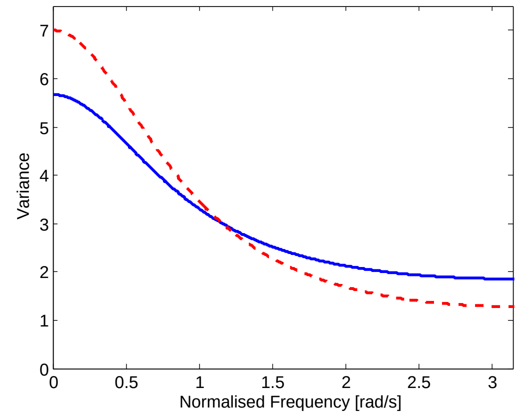

Fig. 1. (Normalised) variance of the transfer function estimators of G, based on the model structures $(G_1(q,\theta_1), H_1(q,\theta_1))$ (solid) and $(G_2(q,\theta_2), H_2(q,\theta_2))$ (dashed), as functions of ω.

Since all terms in the sum are nonnegative, we obtain

$$\mathrm{Var}[\hat{G}(e^{j\omega_i})] \leq \frac{n_G}{N}\frac{\Phi_v(e^{j\omega_i})}{U_i}, \quad i = 1, \ldots, m.$$

Similarly, in the closed-loop case (either direct or indirect, assuming the controller is linear and time invariant), we have that

$$\mathrm{Var}[\hat{G}(e^{j\omega_i})] \leq \frac{n_{G_{cl}}}{N}\frac{\Phi_v(e^{j\omega_i})}{U_i^r}, \quad i = 1, \ldots, m.$$

where $n_{G_{cl}}$ is the number of parameters in G or G_{cl}, (dependent on whether direct or indirect identification is used),

$$\Phi_r(e^{j\omega}) = \sum_{i=1}^{m} 2\pi R_i \delta(\omega - \omega_i),$$

where $\omega_i \in [-\pi, \pi]$, $R_i > 0$ for every $i = 1, \ldots, m$, Φ_r is even, and $U_i^r := |S(e^{j\omega_i})|^2 R_i$, for $i = 1, \ldots, m$. Again, for identifiability reasons we also assume that $m \geq n_G$, the number of parameters in G (or G_{cl}).

Hence, for multisine inputs (or reference signals), Ljung's asymptotic (in model order) variance expressions provide an upper bound on the true variance of the parametric models.

VII. Example

Consider a system described by

$$G_0(q) = \frac{q^{-1}}{1 - a^0 q^{-1}}, \quad H_0(q) = 1,$$

where $a^0 = 0.4$, and the model structures

$$G_1(q,\theta_1) = \frac{b_1 q^{-1} + b_2 q^{-2}}{1 + a_1 q^{-1}}, \qquad H_1(q,\theta_1) = 1,$$
$$G_2(q,\theta_2) = \frac{b_1 q^{-1} + b_2 q^{-2} + b_3 q^{-3}}{(1 - 0.4q^{-1})^3}, \qquad H_2(q,\theta_2) = 1,$$

where $\theta_1 := [b_1 \ \ b_2 \ \ a_1]^T$ and $\theta_2 := [b_1 \ \ b_2 \ \ b_3]^T$. Notice that both model structures, $(G_1(q,\theta_1), H_1(q,\theta_1))$ and $(G_2(q,\theta_2), H_2(q,\theta_2))$, have 3 parameters and include the true plant.

For $\sigma = 1$ and $\Phi_u(e^{j\omega}) = 1$, the normalised (i.e. multiplied by N) variances of the transfer function estimators $\hat{G}_1(e^{j\omega})$ and $\hat{G}_2(e^{j\omega})$ are shown in Figure 1. From the figure, we see that the variances are different functions of frequency. In particular, $\text{Var}[\hat{G}_1]$ is smaller than $\text{Var}[\hat{G}_2]$ at low frequencies and larger at high frequencies. This is consistent with the fundamental limitation derived in Theorem 1, namely that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \text{Var}[\hat{G}_i(e^{j\omega})]d\omega = \frac{3}{N}, \quad i = 1, 2.$$

This means that it is not possible to reduce the variance of $\hat{G}$ at all frequencies by choosing a suitable model structure, since if we reduce the variance at some frequencies, it will necessarily increase at others, which is essentially the 'water-bed' effect.

VIII. Conclusions

In this paper we have established fundamental limitations on the variance of the frequency response of estimated parametric models, for both open and closed loop identification. Furthermore we have shown the relationship to previous results and established that the results presented in this paper hold for more general systems (with exogenous signals). For the closed loop case, we have obtained results for both the direct and indirect identification methods. Based on these results, we have shown that for multisine inputs, the well-known asymptotic (in model order) variance expressions provide upper bounds on the actual variance of estimated models for finite model orders. It can be clearly seen from the results that any over parameterisation results in an increase in the integrated variance of the transfer function estimators. Finally, we have presented an example which shows the tradeoffs imposed by the fundamental limitations derived in the paper, and also illustrates the 'water-bed' effect in system identification.

Appendix A
Proof of Theorem 2

We can assume without loss of generality that $\hat{H}_N = H(\hat{\theta}_N)$ is the ML estimator of H_0, where $\hat{\theta}_N$ is the ML estimator of $\theta_0 \in \mathbb{R}^n$. By the Gauss' Approximation Formula, we have that

$$\begin{aligned}
&\frac{1}{4\pi}\int \frac{1}{|H_0|^4}\text{Var}[|\hat{H}_N|^2] \\
&= \frac{1}{4\pi}\int \frac{1}{|H_0|^4}\left(\frac{\partial |H|^2}{\partial\theta}\right)^H \text{Cov}[\hat{\theta}_N]\frac{\partial |H|^2}{\partial\theta} \qquad (10)\\
&= \frac{1}{4\pi}\int \text{Tr}\left\{\text{Cov}[\hat{\theta}_N]\frac{1}{|H_0|^4}\frac{\partial |H|^2}{\partial\theta}\left(\frac{\partial |H|^2}{\partial\theta}\right)^H\right\} \\
&= \frac{1}{4\pi}\text{Tr}\left\{\text{Cov}[\hat{\theta}_N]\left[\int\left(\frac{1}{|H_0|^2}\frac{\partial |H|^2}{\partial\theta}\right)\left(\frac{1}{|H_0|^2}\frac{\partial |H|^2}{\partial\theta}\right)^H\right]\right\}.
\end{aligned}$$

and, similarly,

$$\begin{aligned}
&\frac{1}{2\pi}\int \frac{1}{|H_0|^2}\text{Var}[\hat{H}_N] \\
&\quad = \frac{1}{2\pi}\int \frac{1}{|H_0|^2}\left(\frac{\partial H}{\partial\theta}\right)^H \text{Cov}[\hat{\theta}_N]\frac{\partial H}{\partial\theta} \qquad (11)\\
&\quad = \frac{1}{2\pi}\text{Tr}\left\{\text{Cov}[\hat{\theta}_N]\left[\int\left(\frac{1}{H_0}\frac{\partial H}{\partial\theta}\right)\left(\frac{1}{H_0}\frac{\partial H}{\partial\theta}\right)^H\right]\right\}.
\end{aligned}$$

Now,

$$\frac{1}{|H_0|^2}\frac{\partial |H|^2}{\partial\theta} = \frac{1}{H_0}\frac{\partial H}{\partial\theta} + \frac{1}{\overline{H_0}}\frac{\partial \overline{H}}{\partial\theta}, \qquad (12)$$

where, due to the conditions of the Theorem, $H_0^{-1}(\partial H/\partial\theta)$ is stable, strictly proper, and has a Laurent series with real coefficients. Therefore,

$$\begin{aligned}
&\int\left(\frac{1}{H_0}\frac{\partial H}{\partial\theta}\right)\left(\frac{1}{\overline{H_0}}\frac{\partial \overline{H}}{\partial\theta}\right)^H = 0, \qquad (13)\\
&\int\left(\frac{1}{H_0}\frac{\partial H}{\partial\theta}\right)\left(\frac{1}{H_0}\frac{\partial H}{\partial\theta}\right)^H = \int\left(\frac{1}{\overline{H_0}}\frac{\partial \overline{H}}{\partial\theta}\right)\left(\frac{1}{\overline{H_0}}\frac{\partial \overline{H}}{\partial\theta}\right)^H.
\end{aligned}$$

From (12) and (13), we obtain

$$\begin{aligned}
&\int\left(\frac{1}{|H_0|^2}\frac{\partial |H|^2}{\partial\theta}\right)\left(\frac{1}{|H_0|^2}\frac{\partial |H|^2}{\partial\theta}\right)^H \qquad (14)\\
&\qquad = 2\int\left(\frac{1}{H_0}\frac{\partial H}{\partial\theta}\right)\left(\frac{1}{H_0}\frac{\partial H}{\partial\theta}\right)^H.
\end{aligned}$$

Finally, the combination of (10), (11) and (14) yields the desired result. ■

Acknowledgment

The authors would like to thank the anonymous reviewers for their helpful comments.

References

[1] H. W. Bode, *Network Analysis and Feedback Amplifier Design*. D. Van Nostrand, 1945.
[2] G. C. Goodwin, S. F. Graebe, and M. E. Salgado, *Control System Design*. Upper Saddle River, New Jersey: Prentice Hall, 2001.
[3] M. M. Seron, J. H. Braslavsky, and G. C. Goodwin, *Fundamental Limitations in Filtering and Control*. Springer-Verlag, 1997.
[4] E. L. Lehmann and G. Casella, *Theory of Point Estimation, 2nd Edition*. New York: Springer, 1998.
[5] H. Cramér, *Mathematical Methods of Statistics*. Princeton, New Jersey: Princeton University Press, 1946.
[6] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 623–656, July, October 1948.
[7] M. E. Salgado, C. E. de Souza, and G. C. Goodwin, "Qualitative aspects of the distribution of errors in least squares estimation," *Automatica*, vol. 26, no. 1, pp. 97–101, January 1990.
[8] B. de Moor, M. Gevers, and G. C. Goodwin, "Overbiased, underbiased and unbiased estimation of transfer functions," in *Proceedings of the First European Control Conference*, vol. 2, 1991, pp. 1372–1377.
[9] ——, "L_2-overbiased, L_2-underbiased and L_2-unbiased estimation of transfer functions," *Automatica*, vol. 30, no. 5, pp. 893–898, 1994.
[10] B. Ninness, "Integral constraints on the accuracy of least squares estimation," *Automatica*, vol. 32, no. 3, pp. 391–397, 1996.
[11] W. E. Larimore, "A survey of some recent developments in system parameter estimation," in *Proceedings of the 6th IFAC Symposium on Identification and Parameter Estimation*, Arlington, VA, 1982, pp. 979–984.

[12] B. Friedlander and B. Porat, "A general lower bound for parametric spectrum estimation," *IEEE Transactions on Accoustics, Speech, and Signal Processing*, vol. 32(4), pp. 728–733, 1984.

[13] P. Stoica, J. Li, and B. Ninness, "The waterbed effect in spectral estimation," *IEEE Signal Processing Magazine*, vol. 21, pp. 88–89, May 2004.

[14] L. Ljung, *System Identification: Theory for the User, 2nd Edition*. Upper Saddle River, New Jersey: Prentice Hall, 1999.

[15] T. Söderström and P. Stoica, *System Identification*. Hertfordshire, United Kingdom: Prentice Hall, 1989.

[16] B. Ninness and H. Hjalmarsson, "Variance error quantifications that are exact for finite-model order," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1275–1291, August 2004.

[17] J. Mårtensson and H. Hjalmarsson, "Variance error quantification for identified poles and zeros," *Automatica (submitted)*, 2007.

[18] P. Whittle, "The analysis of multiple stationary time series," *Journal of the Royal Statistical Society*, vol. 15, pp. 125–139, 1953.

[19] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[20] U. Forssell and L. Ljung, "Closed-loop identification revisited," *Automatica*, vol. 35, no. 7, pp. 1215–1241, 1999.

[21] E. T. V. Donkelaar and P. M. J. Van den Hof, "Analysis of closed-loop identification with a tailor-made parameterization," *European Journal of Control*, vol. 6(1), pp. 54–62, 2000.

[22] J. C. Agüero and G. C. Goodwin, "On the optimality of open and closed loop experiments in system identification," in *45th IEEE Conference on Decision and Control (CDC)*, San Diego, CA, 2006.

[23] ——, "Choosing between open- and closed-loop experiments in linear system identification," *IEEE Transactions on Automatic Control*, vol. 52(8), pp. 1475–1480, August 2007.