

# Asymptotic Variance Expressions for Estimated Frequency Functions

Liang-Liang Xie and Lennart Ljung, *Fellow, IEEE*

**Abstract**—Expressions for the variance of an estimated frequency function are necessary for many issues in model validation and experiment design. A general result is that a simple expression for this variance can be obtained asymptotically as the model order tends to infinity. This expression shows that the variance is inversely proportional to the signal-to-noise ratio frequency by frequency. Still, for low order models the actual variance may be quite different. This has also been pointed out in several recent publications. In this contribution we derive an exact expression for the variance, which is not asymptotic in the model order. This expression applies to a restricted class of models: AR-models, as well as fixed pole models with a polynomial noise model. It brings out the character of the simple approximation and the convergence rate to the limit as the model order increases. It also provides nonasymptotic lower bounds for the general case. The calculations are illustrated by numerical examples.

**Index Terms**—Accuracy, asymptotic variance, FIR models, system identification.

## I. INTRODUCTION

EXPRESSIONS for the accuracy of an estimated frequency function are important for several purposes. They are needed to understand the reliability of the model, e.g., for control design. They are also important for various experimental design issues. See, e.g., [3]–[5]. For parametric estimation methods, expressions that are asymptotic in the number of observed data, can always be calculated from the asymptotic parameter covariance matrix for any given estimate. These expressions can easily be plotted or tabulated, but are normally not suitable for analytic calculations. It has therefore been of interest to find analytic expressions for the frequency function accuracy, if necessary at the price of approximations.

The basic such expression is as follows [6]–[9]: Consider a parameterized set of transfer function models:

$$y(t) = G(q, \theta)u(t) + H(q, \theta)e(t). \quad (1)$$

Let the order of the model be  $n$ , and suppose that the input–output data can be described by

$$y(t) = G_0(q)u(t) + v(t) \quad (2)$$

Manuscript received August 8, 2000; revised March 5, 2001. Recommended by Associate Editor Q. Zhang. The work of L.-L. Xie was completed while visiting Linköping University as a Guest Researcher.

L.-L. Xie is with the Institute of Systems Science, Chinese Academy of Sciences, 100080, Beijing, China.

L. Ljung is with the Department of Electrical Engineering, Linköping University, 581 83, Linköping, Sweden (e-mail: ljung@isy.liu.se).

Publisher Item Identifier S 0018-9286(01)11112-8.

where the spectra of  $u$  and  $v$  are  $\Phi_u(\omega)$  and  $\Phi_v(\omega)$ , respectively. Then the basic result is that the variance of the estimated frequency function  $\hat{G}_N(e^{j\omega})$  is given by

$$\text{Var} \left\{ \hat{G}_N(e^{j\omega}) \right\} \approx \frac{n}{N} \frac{\Phi_v(\omega)}{\Phi_u(\omega)} \quad (3)$$

where the expression is asymptotic in both the model order  $n$ , and the number of data  $N$ . For this result to hold, it is essentially only required that the model parameterization in (1) has a block shift structure, so that the gradient w.r.t. the parameter block  $k$ ,  $\theta_{\{k\}}$ , is a shifted version of the gradient w.r.t.  $\theta_{\{1\}}$  (see [7, p. 293]). This structure is at hand for most of the typical black box model parameterizations, like ARMAX, Box–Jenkins, etc.

A similar result holds if the frequency function is estimated by spectral analysis, [2], [7, Ch. 6]:

$$\text{Var} \left\{ \hat{G}_N(e^{j\omega}) \right\} \approx \frac{W(\gamma)}{N} \frac{\Phi_v(\omega)}{\Phi_u(\omega)} \quad (4)$$

where  $W(\gamma)$  is a number that depends on the type of window used, as well as on its size  $\gamma$ .

These results are intuitively appealing, and mean that the variance of the estimate is inversely proportional to the signal-to-noise ratio at the frequency in question.

However, it is important to realize that they are asymptotic in the model order  $n$ . Simulations in [7, Fig. 9.1], indicate that they may give a reasonable picture of the true variance, also for moderately large  $n$ , but that the goodness of the approximation becomes worse for colored inputs.

In the recent paper [12] model parameterizations like (1) with fixed denominators of  $G$  are studied. The asymptotic result (3) applies also to such structures. However, in [12] it is shown that this simple expression may be quite misleading for low order models, i.e., that the convergence in  $n$  to the right-hand side of (3) could be quite slow. The authors also derived an alternative asymptotic expression that shows better accuracy in these cases. See also [11].

In this paper, we will derive exact, nonasymptotic (in  $n$ ) expressions for the variance for the following special case of (1):

$$y(t) = \frac{B(q, \theta)}{F^\dagger(q)} u(t) + C^\dagger(q)e(t) \quad (5)$$

with  $F^\dagger$  and  $C^\dagger$  fixed polynomials,  $B$  the polynomial with the uncertain parameters and  $u$  being an autoregressive process. The arguments also apply to the AR-modeling

$$A(q, \theta)y(t) = e(t). \quad (6)$$

This exact expression will bring out the character of the approximation in (3), as well the general convergence rate. It will also be useful for more exact analytical calculations, involving the model accuracy. We will also comment on how the results relate to more general parameterizations.

The remainder of the paper is structured as follows. In Section II, we give some preliminary observations on the plausibility and limitation of the asymptotic (in  $n$  and  $N$ ) variance expression (3). In Section III, a nonasymptotic (in model order  $n$ ) expression is presented for the model structure (5), which improves over (3) especially in the case of low order models. The detailed proof is presented in Section IV. In Section V, other model structures are considered, where normally only lower bounds are obtained, except for the AR-model, for which an exact expression can be derived in the same spirit. In Section VI, some numerical illustrations are given. Finally, we conclude this paper in Section VII.

## II. SOME PRELIMINARY OBSERVATIONS

Consider the model (5) and assume that  $u$  is an autoregression

$$u(t) = \frac{1}{L^\dagger(q)} v(t) \quad (7)$$

where  $v$  is white noise with variance  $\sigma_v^2$ . Suppose the true system can be described by

$$y(t) = \frac{B_0(q)}{F^\dagger(q)} u(t) + C_0(q) e_0(t) \quad (8)$$

where  $B_0$ ,  $C_0$  are polynomials and  $e_0$  is white noise with variance  $\sigma_e^2$ . In other words, comparing (5) and (8), our assumption is that the *a priori* guess on the poles of the input–output dynamics  $F^\dagger$  is correct, while the guess of the noise model could be wrong, i.e.,  $C^\dagger(q) \neq C_0(q)$ .

The model (5) is a linear regression for estimating  $B$

$$\tilde{y}(t) = B(q, \theta) \tilde{u}(t) + e(t) \quad (9)$$

with

$$\begin{aligned} \tilde{y}(t) &= \frac{1}{C^\dagger(q)} y(t) \\ \tilde{u}(t) &= \frac{1}{C^\dagger(q)F^\dagger(q)} u(t) = \frac{1}{C^\dagger(q)F^\dagger(q)L^\dagger(q)} v(t). \end{aligned}$$

The true system corresponds to

$$\tilde{y}(t) = B_0(q) \tilde{u}(t) + \frac{C_0(q)}{C^\dagger(q)} e_0(t). \quad (10)$$

Apply the LS method to estimate

$$\theta = [b_1, b_2, \dots, b_n]^T.$$

It is well known that the covariance matrix of the estimate  $\hat{\theta}_N$  is given by (see, e.g., [7, (9.42)])

$$\lim_{N \rightarrow \infty} N \cdot \text{Cov} \hat{\theta}_N = F_1^{-1} F_2 F_1^{-1} \quad (11)$$

where the elements of the matrix  $F_1$  is given by the correlation function of  $\tilde{u}$

$$(F_1)_{k,j} = R_{\tilde{u}}(k-j) \triangleq E \tilde{u}(t) \tilde{u}(t + (k-j))$$

and the elements of the matrix  $F_2$  is given by the correlation function of  $u_F$

$$(F_2)_{k,j} = R_{u_F}(k-j) \triangleq E u_F(t) u_F(t + (k-j))$$

where

$$u_F(t) \triangleq \sum_{k=0}^{\infty} h(k) \tilde{u}(t+k)$$

with

$$\sigma_e \frac{C_0(q)}{C^\dagger(q)} = \sum_{k=0}^{\infty} h(k) q^{-k}.$$

From this follows immediately one important observation: *The covariance matrix of the estimate  $\hat{\theta}_N$ , depends on the polynomials,  $F^\dagger$ ,  $L^\dagger$ ,  $C^\dagger$ ,  $B_0$ , and  $C_0$  only in the combinations*

$$C^\dagger \cdot F^\dagger \cdot L^\dagger \quad (12)$$

and

$$\frac{C_0}{C^\dagger}. \quad (13)$$

The latter expression describes the discrepancy between the true noise description in (8) and the noise model in (5). We shall refer to this ratio as the *noise model error*.

From the general asymptotic result (3), it follows that the variance of:

$$\hat{B}_N(e^{j\omega}) = \sum_{k=1}^n \hat{b}_k(N) e^{-jk\omega} = W(e^{j\omega}) \hat{\theta}_N \quad (14)$$

is asymptotically in  $N$  and  $n$  given by

$$\text{Var} \left\{ \hat{B}_N(e^{j\omega}) \right\} \approx \frac{n}{N} \frac{\sigma_e^2}{\sigma_v^2} \frac{\frac{|C_0(e^{j\omega})|^2}{|C^\dagger(e^{j\omega})|^2}}{\frac{1}{|C^\dagger(e^{j\omega})|^2 |F^\dagger(e^{j\omega})|^2 |L^\dagger(e^{j\omega})|^2}}. \quad (15)$$

Since the transfer function estimate is  $\hat{G} = \hat{B}/F^\dagger$ , this gives the asymptotic variance of the frequency function

$$\begin{aligned} \text{Var} \left\{ \hat{G}_N(e^{j\omega}) \right\} &\approx \frac{n}{N} \frac{\sigma_e^2 \frac{|C_0(e^{j\omega})|^2}{|C^\dagger(e^{j\omega})|^2}}{\sigma_v^2 \frac{1}{|C^\dagger(e^{j\omega})|^2 |F^\dagger(e^{j\omega})|^2 |L^\dagger(e^{j\omega})|^2}} \\ &= \frac{n}{N} \frac{\sigma_e^2}{\sigma_v^2} |C_0(e^{j\omega})|^2 |L^\dagger(e^{j\omega})|^2. \end{aligned} \quad (16)$$

The asymptotic variance thus *does not depend on either*  $C^\dagger$  *or*  $F^\dagger$ . This may be counter-intuitive for two reasons.

- We know that the best choice of noise model  $C^\dagger$  is to let it be equal to the true one  $C_0$  (see, e.g., [7, p. 478]). This minimizes the variance, and makes it equal to the Cramer–Rao bound. Yet, the asymptotic (in  $n$ ) variance is independent of  $C^\dagger$ ! To understand this paradox, it is necessary to realize the role of the noise model filter: It weighs together information about the system from different frequency points to merge it into one model. The noise filter will make sure that information where there supposedly is much noise [i.e., frequencies  $\omega$  where  $C^\dagger(e^{j\omega})$  is large] carries less weight in the final estimate. As the model order increases, the estimates of  $G$  become more and more decoupled at different frequencies, so this global frequency weighting becomes less pronounced. As a consequence, the influence of the noise model filter on the estimate becomes less important.
- The prior knowledge about system poles, represented by  $F^\dagger$  should be beneficial for the variance of the estimate (provided it is correct). However, again as the model order increases, the value of this knowledge decreases. In other words, for large  $n$ , the set of models described by  $B/F^\dagger$  as  $B$  ranges over all  $n$ -order polynomials comes close to describe all “nice” linear systems, regardless of  $F^\dagger$ .

It is also clear that for small values of  $n$ , the influence of  $C^\dagger$  and  $F^\dagger$  should be more pronounced. The paper [12] contains several striking examples of how much the variance for small  $n$  may differ from the asymptotic expression.

In the next section, we shall demonstrate this by deriving a nonasymptotic expression for the variance.

### III. AN EXACT EXPRESSION FOR THE VARIANCE

First, let us explicitly write out the polynomials  $C_0$ ,  $C^\dagger$ ,  $F^\dagger$ ,  $L^\dagger$  in the following form:

$$\begin{aligned} C_0(q) &= \prod_{k=1}^{r_1} (1 - c_k q^{-1}) \\ C^\dagger(q) &= \prod_{k=1}^{r_1} (1 - \ell_k q^{-1}) \\ F^\dagger(q)L^\dagger(q) &= \prod_{k=r_1+1}^r (1 - \ell_k q^{-1}) \end{aligned} \quad (17)$$

with  $0 \leq r_1 \leq r \leq n$ ,  $\max_{1 \leq k \leq r} |\ell_k| < 1$ ,  $\max_{1 \leq k \leq r_1} |c_k| < 1$ . Note that since  $c_k$  or  $\ell_k$  can be zero,  $C_0(q)$  and  $C^\dagger(q)$  are not necessarily of the same order.

The main technical result of the paper is the following one.

*Theorem 3.1:* Consider the problem of estimating the frequency function in the model structure (5), where the input is given by (7). Assume that the true system is given by (8). Suppose that the model order  $n$  is no less than the order of  $B_0$  and also the order of the polynomial  $L^\dagger \cdot C^\dagger \cdot F^\dagger$ . Then the vari-

ance of the frequency function estimate, asymptotically in  $N$ , but exact in  $n$ , is given by

$$\begin{aligned} \lim_{N \rightarrow \infty} N \text{Var} \left\{ \hat{G}_N(e^{j\omega}) \right\} &= P(\omega) \\ &= \frac{\sigma_e^2}{\sigma_v^2} |C_0(e^{j\omega})|^2 |L^\dagger(e^{j\omega})|^2 \\ &\quad \times \left[ (n-r) + R_1(C^\dagger \cdot F^\dagger \cdot L^\dagger) + \frac{|C^\dagger(e^{j\omega})|^2}{|C_0(e^{j\omega})|^2} R_2(C_0/C^\dagger) \right]. \end{aligned} \quad (18)$$

Here

- $n$  is the order of  $B$ , and  $r$  is the order of  $L^\dagger \cdot C^\dagger \cdot F^\dagger$ ; (Note that  $n \geq r$ );
- $R_1(K)$  for a polynomial  $K(q) = \prod_{m=1}^r (1 - k_m q^{-1})$  is given by

$$R_1(K) = \sum_{m=1}^r \frac{1 - |k_m|^2}{|e^{j\omega} - k_m|^2}; \quad (19)$$

- $R_2(\beta)$  for a rational function  $\beta(q)$  is given by

$$R_2(\beta) = -2 \text{Re} \left\{ \frac{1}{2\pi j} \oint_{\mathbf{T}} \frac{z e^{j\omega}}{(e^{j\omega} - z)^2} \beta(z) \beta^*(z) \frac{dz}{z} \right\} \quad (20)$$

where,  $\beta^*(z) \triangleq \overline{\beta(\bar{z}^{-1})}$  is the mirror of  $\beta(z)$  and  $\mathbf{T}$  is any simple closed positively oriented curve on the complex plane with all the poles of  $\beta(z)$  inside and the point  $e^{j\omega}$  outside.

*Some Observations About the Expression for  $P(\omega)$ :*

- The factor

$$\frac{\sigma_e^2}{\sigma_v^2} |C_0(e^{j\omega})|^2 |L^\dagger(e^{j\omega})|^2 = \sigma_e^2 |C_0(e^{j\omega})|^2 \left/ \frac{\sigma_v^2}{|L^\dagger(e^{j\omega})|^2} \right.$$

actually is the noise-to-signal ratio (SNR)  $\Phi_v(\omega)/\Phi_u(\omega)$  as in (3).

- Note that (18) can be rearranged as

$$\begin{aligned} P(\omega) &= n \frac{\Phi_v(\omega)}{\Phi_u(\omega)} \left[ 1 + \frac{1}{n} \left( -r + R_1(C^\dagger \cdot F^\dagger \cdot L^\dagger) \right. \right. \\ &\quad \left. \left. + \frac{|C^\dagger(e^{j\omega})|^2}{|C_0(e^{j\omega})|^2} R_2(C_0/C^\dagger) \right) \right]. \end{aligned} \quad (21)$$

As the model order  $n$  increases (and the polynomials  $F^\dagger$  and  $C^\dagger$  remain fixed) the expression for the covariance,  $P(\omega)/n$  thus converges to the basic expression (3) at the rate  $1/n$ .

- When  $C^\dagger \cdot F^\dagger \cdot L^\dagger = 1$ , and the noise model is correct, i.e.,  $C_0/C^\dagger = 1$ , that is, we have a FIR model with white input and noise, it follows from (19) and (20) that  $R_1(1) = r$ ,  $R_2(1) = 0$  and (18) reduces to

$$\lim_{N \rightarrow \infty} N \text{Var} \left\{ \hat{G}_N(e^{j\omega}) \right\} = n \frac{\sigma_e^2 |C_0(e^{j\omega})|^2}{\sigma_v^2 |L^\dagger(e^{j\omega})|^2}$$

which agrees with (3).

- When the noise model is correct, i.e.,  $C_0/C^\dagger = 1$ , the expression (18) reduces to

$$\lim_{N \rightarrow \infty} N \text{Var} \left\{ \hat{G}_N(e^{j\omega}) \right\} = \frac{\sigma_e^2 |C_0(e^{j\omega})|^2}{\sigma_v^2} \left[ (n-r) + \sum_{m=1}^r \frac{1 - |\ell_m|^2}{|e^{j\omega} - \ell_m|^2} \right]$$

which basically is the main result in [12]. However, apparently it is not realized in that paper that this is the exact expression when the noise model is correct and the input is an autoregression.

- The main new contribution of (18) is the third term  $(|C^\dagger(e^{j\omega})|^2/|C_0(e^{j\omega})|^2)R_2(C_0/C^\dagger)$ , which obviously is accounting for noise model error [it is easy to see that  $R_2(1) = 0$ ]. In Section VI, some interpretations are given on how it depends on the filtered noise spectrum  $|C_0(e^{j\omega})|^2/|C^\dagger(e^{j\omega})|^2$ . It should be noted that this term is bounded since  $\max_{1 \leq k \leq r_1} |\ell_k| < 1$  and  $\max_{1 \leq k \leq r_1} |c_k| < 1$  by (17).

#### IV. PROOF OF MAIN RESULTS

##### A. An Extension of Laurent Expansion

First, on a more general orthonormal basis than the trigonometric one  $\{z^n, n \in \text{all integers}\}$ , we establish an extension of the Laurent Expansion, which plays a key role in the proof of Theorem 3.1.

For complex numbers  $\{\xi_n, n \geq 0\}$  with  $\sup_{n \geq 0} |\xi_n| < 1$ , it is easy to see that the functions  $B_n(z)$  given by

$$B_n(z) \triangleq \frac{\sqrt{1 - |\xi_n|^2}}{1 - \xi_n z} \prod_{k=0}^{n-1} \left( \frac{z - \bar{\xi}_k}{1 - \xi_k z} \right), \quad n \geq 0 \quad (22)$$

are orthonormal on the unit circle  $\{z \in C: |z| = 1\}$ , i.e.,  $\langle B_m(z), B_n(z) \rangle = 0$  for  $m \neq n$ , with

$$\begin{aligned} \langle f(z), g(z) \rangle &\triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{j\omega}) \overline{g(e^{j\omega})} d\omega \\ &= \frac{1}{2\pi j} \oint_{|z|=1} f(z) \overline{g(z)} \frac{dz}{z} \end{aligned} \quad (23)$$

where “ $\bar{\cdot}$ ” denotes “complex conjugate.” By [10, Th. 1],  $B_n(z)$ ,  $n \geq 0$  form a complete orthonormal basis for the analytic functions. Specially, if  $\xi_n \equiv 0, \forall n \geq 0$ , then  $B_n(z) = z^n$ .

Similarly, it is easy to see that  $z^{-1}B_n(z^{-1}), n \geq 0$  are also orthonormal and  $z^{-1}B_n(z^{-1}) = z^{-(n+1)}$ , if  $\xi_n \equiv 0, n \geq 0$ .

*Lemma 4.1 (An Extension of the Laurent Expansion):* Suppose  $\{\xi_n, n \geq 0\}, \{\eta_n, n \geq 0\}$  are complex numbers with  $\sup_{n \geq 0} |\xi_n| < 1/\rho_1 < 1$ , and  $\sup_{n \geq 0} |\eta_n| < \rho_2 < 1$ . Consider the region  $A \triangleq \{z \in C | \rho_2 \leq |z| \leq \rho_1\}$ . Let  $f(z)$  be analytic on the region  $A$ . Then, we have the following expansion for  $z \in A^0$  (the interior of  $A$ )

$$\begin{aligned} f(z) &= \sum_{n=0}^{\infty} \langle f(\zeta), B_n(\zeta) \rangle B_n(z) \\ &\quad + \sum_{n=0}^{\infty} \langle f(\zeta), \zeta^{-1}B'_n(\zeta^{-1}) \rangle z^{-1}B'_n(z^{-1}) \end{aligned} \quad (24)$$

where,  $B_n(z)$  is defined as in (22),  $B'_n(z)$  is similarly defined with  $\xi_n$  replaced by  $\eta_n$ .

*Proof:* By Cauchy's Integral Formula, we have for  $z \in A^0$

$$f(z) = \frac{1}{2\pi j} \oint_{|\zeta|=\rho_1} \frac{f(\zeta)}{\zeta - z} d\zeta - \frac{1}{2\pi j} \oint_{|\zeta|=\rho_2} \frac{f(\zeta)}{\zeta - z} d\zeta. \quad (25)$$

For  $|\zeta| = \rho_1$ ,  $1/(\zeta - z)$  is analytic on  $|z| < \rho_1$ . Hence, by [10, Th. 1], we have for  $|z| < \rho_1$

$$\frac{1}{\zeta - z} = \sum_{n=0}^{\infty} \left\langle \frac{1}{\zeta - z}, B_n(z) \right\rangle B_n(z)$$

with

$$\begin{aligned} \overline{\left\langle \frac{1}{\zeta - z}, B_n(z) \right\rangle} &= \left\langle B_n(z), \frac{1}{\zeta - z} \right\rangle \\ &= \frac{1}{2\pi j} \oint_{|z|=1} B_n(z) \overline{\frac{1}{\zeta - z^{-1}}} \frac{dz}{z} \\ &= \frac{1}{\zeta} \frac{1}{2\pi j} \oint_{|z|=1} B_n(z) \frac{1}{z - 1/\bar{\zeta}} dz \\ &= \frac{1}{\bar{\zeta}} B_n(1/\bar{\zeta}). \end{aligned}$$

Hence, we have

$$\frac{1}{\zeta - z} = \sum_{n=0}^{\infty} \frac{1}{\zeta} \overline{B_n(1/\bar{\zeta})} B_n(z). \quad (26)$$

Hence, the first term in RHS of (25) can be written into

$$\begin{aligned} &\frac{1}{2\pi j} \oint_{|\zeta|=\rho_1} \frac{f(\zeta)}{\zeta - z} d\zeta \\ &= \sum_{n=0}^{\infty} \frac{1}{2\pi j} \oint_{|\zeta|=\rho_1} f(\zeta) \overline{B_n(1/\bar{\zeta})} \frac{d\zeta}{\zeta} B_n(z) \\ &= \sum_{n=0}^{\infty} \frac{1}{2\pi j} \oint_{|\zeta|=1} f(\zeta) \overline{B_n(\zeta)} \frac{d\zeta}{\zeta} B_n(z) \\ &= \sum_{n=0}^{\infty} \langle f(\zeta), B_n(\zeta) \rangle B_n(z) \end{aligned} \quad (27)$$

where the second equation follows from Cauchy's Integral Formula and  $\zeta = 1/\bar{\zeta}$  for  $|\zeta| = 1$ .

For the second term in the right-hand side of (25), let

$$g(z) \triangleq -\frac{1}{2\pi j} \oint_{|\zeta|=\rho_2} \frac{f(\zeta)}{\zeta - z} d\zeta. \quad (28)$$

Then

$$g(z^{-1}) = \frac{1}{2\pi j} \oint_{|\zeta|=\rho_2} \frac{zf(\zeta)}{1 - \zeta z} d\zeta. \quad (29)$$

For  $|\zeta| = \rho_2$ ,  $1/(1 - \zeta z)$  is analytic on  $|z| < 1/\rho_2$ . Hence by [10, Th. 1], we have for  $|z| < 1/\rho_2$

$$\frac{1}{1 - \zeta z} = \sum_{n=0}^{\infty} \left\langle \frac{1}{1 - \zeta z}, B'_n(z) \right\rangle B'_n(z)$$

with

$$\begin{aligned} \left\langle \frac{1}{1-\zeta z}, B'_n(z) \right\rangle &= \left\langle B'_n(z), \frac{1}{1-\zeta z} \right\rangle \\ &= \frac{1}{2\pi j} \oint_{|z|=1} B'_n(z) \frac{1}{1-\bar{\zeta}z^{-1}} \frac{dz}{z} \\ &= B'_n(\bar{\zeta}). \end{aligned}$$

Hence, we have

$$\frac{1}{1-\zeta z} = \sum_{n=0}^{\infty} \overline{B'_n(\bar{\zeta})} B'_n(z). \quad (30)$$

Then, by (29) and (30), we have

$$\begin{aligned} g(z^{-1}) &= \sum_{n=0}^{\infty} \frac{1}{2\pi j} \oint_{|\zeta|=\rho_2} f(\zeta) \overline{B'_n(\bar{\zeta})} d\zeta z B'_n(z) \\ &= \sum_{n=0}^{\infty} \frac{1}{2\pi j} \oint_{|\zeta|=\rho_2} f(\zeta) \overline{B'_n(\bar{\zeta})} \frac{d\zeta}{\zeta} z B'_n(z) \\ &= \sum_{n=0}^{\infty} \frac{1}{2\pi j} \oint_{|\zeta|=1} f(\zeta) \overline{B'_n(\zeta^{-1})} \frac{d\zeta}{\zeta} z B'_n(z) \\ &= \sum_{n=0}^{\infty} \langle f(\zeta), \zeta^{-1} B'_n(\zeta^{-1}) \rangle z B'_n(z) \end{aligned}$$

where the third equation follows from Cauchy's Integral Formula and  $\zeta = 1/\bar{\zeta}$  for  $|\zeta| = 1$ . Hence

$$g(z) = \sum_{n=0}^{\infty} \langle f(\zeta), \zeta^{-1} B'_n(\zeta^{-1}) \rangle z^{-1} B'_n(z^{-1})$$

which together with (27) lead to (24).  $\diamond$

### B. Proof of Theorem 3.1

It is obvious from (5), (7), (8), and (17) that the true model for  $y'(t) \triangleq (1/C^\dagger)y(t)$  and  $v(t)$  is

$$y'(t) = \frac{B(q, \theta)}{L(q)} v(t) + \frac{C(q)}{L(q)} e(t) \quad (31)$$

where

$$\begin{aligned} L(q) &\triangleq C^\dagger(q) F^\dagger(q) L^\dagger(q) \\ &= \prod_{k=1}^r (1 - \ell_k q^{-1}) \\ C(q) &\triangleq C_0(q) F^\dagger(q) L^\dagger(q) \\ &= \prod_{k=1}^{r_1} (1 - c_k q^{-1}) \prod_{k=r_1+1}^r (1 - \ell_k q^{-1}) \end{aligned}$$

and

$$B(q, \theta) = \theta_1 q^{-1} + \theta_2 q^{-2} + \dots + \theta_n q^{-n}.$$

By reparameterization using the orthonormal basis (see [10])

$$\Gamma_{n,r}(q) = [\gamma_1(q), \gamma_2(q), \dots, \gamma_n(q)]^T \quad (32)$$

where

$$\gamma_i(q) \triangleq \begin{cases} q^{-i}, & 1 \leq i \leq n-r; \\ b_{i-(n-r)}(q) q^{-(n-r)}, & n-r < i \leq n \end{cases} \quad (33)$$

with

$$b_k(q) \triangleq \frac{\sqrt{1 - |\ell_k|^2}}{q - \ell_k} \prod_{m=1}^{k-1} \frac{1 - \bar{\ell}_m q}{q - \ell_m}, \quad 1 \leq k \leq r \quad (34)$$

we have the new equivalent (in the sense of frequency function) model

$$y'(t) = G(q, \theta') v(t) + \frac{C(q)}{L(q)} e(t) \quad (35)$$

with

$$G(q, \theta') \triangleq \theta'_1 q^{-1} + \dots + \theta'_{n-r} q^{-(n-r)} + \theta'_{n-r+1} b_1(q) q^{-(n-r)} + \dots + \theta'_n b_r(q) q^{-(n-r)} \quad (36)$$

and there exists some linear relationship between  $\theta$  and  $\theta'$ . Note that  $\theta'$  could be complex numbers.

By the same reason as deriving (11), applying [7, (9.42)] to the model (35) and (36), we have

$$\lim_{N \rightarrow \infty} N \cdot \text{Cov} \hat{\theta}'_N = F_1'^{-1} F_2' F_1'^{-1} \quad (37)$$

with

$$\begin{aligned} F_1' &\triangleq E [\Gamma_{n,r}(q) v(t) \cdot \Gamma_{n,r}^*(q) v(t)] \\ F_2' &\triangleq E \left[ \sigma_e \frac{C(q^{-1})}{L(q^{-1})} \Gamma_{n,r}(q) v(t) \cdot \sigma_e \frac{C(q^{-1})}{L(q^{-1})} \Gamma_{n,r}^*(q) v(t) \right] \end{aligned}$$

where  $\Gamma_{n,r}(q)$  is defined in (32)–(34).

By Parseval's formula, we have

$$F_1' = M_{n,r}(\sigma_v^2), \quad F_2' = M_{n,r} \left( \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \sigma_e^2 \sigma_v^2 \right) \quad (38)$$

where  $M_{n,r}(f(e^{j\omega}))$  is an  $n \times n$  symmetric Toeplitz matrix for any positive function  $f(\cdot)$ :

$$M_{n,r}(f(e^{j\omega})) \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} \Gamma_{n,r}(e^{j\omega}) \Gamma_{n,r}^*(e^{j\omega}) f(e^{j\omega}) d\omega. \quad (39)$$

It follows immediately that  $M_{n,r}(\lambda) = \lambda \cdot I_n$  for any constant  $\lambda$  since  $\Gamma_{n,r}$  form an orthonormal basis. Hence, by (35)–(38), we have

$$\begin{aligned} &\lim_{N \rightarrow \infty} N \text{Var} \left\{ \hat{G}_{v \rightarrow y'} \right\} \\ &= \lim_{N \rightarrow \infty} N \Gamma_{n,r}^*(e^{j\omega}) \text{Cov} \hat{\theta}'_N \Gamma_{n,r}(e^{j\omega}) \\ &= \Gamma_{n,r}^*(e^{j\omega}) M_{n,r}^{-1}(\sigma_v^2) M_{n,r} \left( \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \sigma_e^2 \sigma_v^2 \right) \\ &\quad \times M_{n,r}^{-1}(\sigma_v^2) \Gamma_{n,r}(e^{j\omega}) \\ &= \frac{\sigma_e^2}{\sigma_v^2} \Gamma_{n,r}^*(e^{j\omega}) M_{n,r} \left( \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \right) \Gamma_{n,r}(e^{j\omega}). \quad (40) \end{aligned}$$

Equation (40) can also be derived by the arguments in [12]. The only difference is that undermodeling is not our concern in this paper since  $B_0(q) \in \{B(q, \theta): \theta \in \mathbb{R}^n\}$  by our assumption. Thus, the term containing  $\Delta_n$  in [12, (38)] will drop out and leave the same equation as (40).

So now the left task is to make (40) into a simpler and more explicit form. The method used in [12] is to apply an asymptotic result which holds as the model order  $n \rightarrow \infty$ . But in this paper, to get a nonasymptotic expression, we will exactly calculate (40), which is feasible due to the specially chosen orthonormal basis (32).

An application of Lemma 4.1 in our problem is as follows. For any  $1 \leq v \leq n - r$ , let

$$\xi_i \triangleq \begin{cases} 0, & 0 \leq i \leq n - r - v; \\ \bar{\ell}_{i-(n-r-v)}, & n - r - v < i \leq n - v; \\ 0, & i > n - v \end{cases}$$

and

$$\eta_i \triangleq \begin{cases} 0, & 0 \leq i \leq n - r - v - 1; \\ \ell_{i-(n-r-v-1)}, & n - r - v - 1 < i \leq n - v - 1; \\ 0, & i > n - v - 1 \end{cases}$$

in Lemma 4.1. Then it is easy to check that

$$B_i(z) = \overline{\gamma_{i+v}(z)} \gamma_v(z), \quad 0 \leq i \leq n - v; \quad (41)$$

$$z^{-1} B'_i(z^{-1}) = \gamma_{i+v+1}(z) \overline{\gamma_v(z)}, \quad 0 \leq i \leq n - v - 1 \quad (42)$$

with  $\gamma_i(\cdot)$  defined in (33) and that

$$B_i(z) = z^{i-r} \prod_{k=1}^r \frac{z - \ell_k}{1 - \bar{\ell}_k z}, \quad i > n - v; \quad (43)$$

$$z^{-1} B'_i(z^{-1}) = z^{-(i-r)} \prod_{k=1}^r \frac{1 - \bar{\ell}_k z}{z - \ell_k}, \quad i > n - v - 1. \quad (44)$$

By (43), (44) and using Cauchy's Integral Formula, it is obvious that for any  $1 \leq v \leq n - r$

$$\left\langle \frac{|C(z)|^2}{|L(z)|^2}, B_i(z) \right\rangle = 0, \quad i > n - v$$

$$\left\langle \frac{|C(z)|^2}{|L(z)|^2}, z^{-1} B'_i(z^{-1}) \right\rangle = 0, \quad i > n - v - 1.$$

Hence by (41), (42), and Lemma 4.1 for  $f(z) = |C(z)|^2 / |L(z)|^2$  with  $z = e^{j\omega}$ , we have for any  $1 \leq v \leq n - r$

$$\frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} = \sum_{u=v}^n \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \overline{\gamma_u(z)} \gamma_v(z) \right\rangle \overline{\gamma_u(e^{j\omega})} \gamma_v(e^{j\omega})$$

$$+ \sum_{u=v+1}^n \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \gamma_u(z) \overline{\gamma_v(z)} \right\rangle \gamma_u(e^{j\omega}) \overline{\gamma_v(e^{j\omega})}. \quad (45)$$

Before going further, let us introduce some notations. Corresponding to  $b_k(z)$  defined in (34), let

$$\bar{b}_k(z) \triangleq \frac{z \sqrt{1 - |\ell_k|^2}}{1 - \bar{\ell}_k z} \prod_{m=1}^{k-1} \frac{z - \ell_m}{1 - \bar{\ell}_m z}, \quad 1 \leq k \leq r \quad (46)$$

and

$$R(z) \triangleq \prod_{m=1}^r \frac{1 - \bar{\ell}_m z}{z - \ell_m}, \quad \bar{R}(z) \triangleq \prod_{m=1}^r \frac{z - \ell_m}{1 - \bar{\ell}_m z}. \quad (47)$$

It is not difficult to prove the following two equations for  $1 \leq q \leq r$ :

$$\sum_{k=1}^r \bar{b}_k(\ell_q) b_k(e^{j\omega}) = \frac{\ell_q}{e^{j\omega} - \ell_q}; \quad (48)$$

and

$$\sum_{k=1}^r \lim_{z \rightarrow \ell_q} b_k(z) \bar{R}(z) \bar{b}_k(e^{j\omega}) R(e^{j\omega}) = \frac{e^{j\omega}}{e^{j\omega} - \ell_q}. \quad (49)$$

In fact, for any  $|x| < 1, |y| < 1$ , let

$$g(x) \triangleq \frac{x}{e^{j\omega} - x}$$

$$f(x, y) \triangleq \frac{x \sqrt{1 - |y|^2}}{1 - x \bar{y}} \frac{\sqrt{1 - |y|^2}}{e^{j\omega} - y} + g(x) \frac{x - y}{1 - x \bar{y}} \frac{1 - \bar{y} e^{j\omega}}{e^{j\omega} - y}.$$

It is not difficult to establish the fact that  $f(x, y) = g(x)$ . Hence by the definition of  $\bar{b}_k(z)$ , we have

$$\sum_{k=1}^r \bar{b}_k(\ell_q) b_k(e^{j\omega})$$

$$= \sum_{k=1}^q \bar{b}_k(\ell_q) b_k(e^{j\omega})$$

$$= \sum_{k=1}^{q-1} \bar{b}_k(\ell_q) b_k(e^{j\omega})$$

$$+ g(\ell_q) \prod_{m=1}^{q-1} \left( \frac{\ell_q - \ell_m}{1 - \bar{\ell}_m \ell_q} \right) \prod_{m=1}^{q-1} \left( \frac{1 - \bar{\ell}_m e^{j\omega}}{e^{j\omega} - \ell_m} \right)$$

$$= \sum_{k=1}^{q-2} \bar{b}_k(\ell_q) b_k(e^{j\omega})$$

$$+ f(\ell_q, \ell_{q-1}) \prod_{m=1}^{q-2} \left( \frac{\ell_q - \ell_m}{1 - \bar{\ell}_m \ell_q} \right) \prod_{m=1}^{q-2} \left( \frac{1 - \bar{\ell}_m e^{j\omega}}{e^{j\omega} - \ell_m} \right)$$

$$= \sum_{k=1}^{q-2} \bar{b}_k(\ell_q) b_k(e^{j\omega})$$

$$+ g(\ell_q) \prod_{m=1}^{q-2} \left( \frac{\ell_q - \ell_m}{1 - \bar{\ell}_m \ell_q} \right) \prod_{m=1}^{q-2} \left( \frac{1 - \bar{\ell}_m e^{j\omega}}{e^{j\omega} - \ell_m} \right)$$

$$= \dots$$

$$= g(\ell_q)$$

which is (48). By similar arguments, we can prove (49).

Now another application of Lemma 4.1 with

$$\begin{aligned} \xi_0 &= 0; & \xi_i &= \bar{\ell}_i, & 1 \leq i \leq r; & & \xi_{r+i} &= \bar{\ell}_i, & 1 \leq i \leq r \\ \xi_i &= 0, & i &> 2r; & \eta_{i-1} &= \ell_i, & 1 \leq i \leq r \\ \eta_{r+i-1} &= \ell_i, & 1 \leq i \leq r; & & \eta_i &= 0, & i > 2r - 1 \end{aligned}$$

gives

$$\begin{aligned} B_0(z) &= 1; & B_i(z) &= \bar{b}_i(z), & 1 \leq i \leq r \\ B_{r+i}(z) &= \bar{b}_i(z) \cdot \bar{R}(z), & 1 \leq i \leq r \\ B_i(z) &= \bar{R}(z) \cdot \bar{R}(z) \cdot z^{i-2r}, & i > 2r \\ z^{-1} B'_{i-1}(z^{-1}) &= b_i(z), & 1 \leq i \leq r \\ z^{-1} B'_{r+i-1}(z^{-1}) &= b_i(z) \cdot R(z), & 1 \leq i \leq r \\ z^{-1} B'_i(z^{-1}) &= R(z) \cdot R(z) \cdot z^{-(i-2r+1)}, & i > 2r - 1 \end{aligned}$$

with  $b_i(z)$ ,  $\bar{b}_i(z)$ ,  $R(z)$ ,  $\bar{R}(z)$  defined in (34), (46) and (47). Then, for

$$f(z) = \frac{|C(z)|^2}{|L(z)|^2} b_k(z), \quad 1 \leq k \leq r$$

by Cauchy's Integral Formula, we have

$$\begin{aligned} \langle f(z), B_{r+i}(z) \rangle &= 0, & 1 \leq i \leq r \\ \langle f(z), B_i(z) \rangle &= 0, & i > 2r \\ \langle f(z), z^{-1} B'_i(z^{-1}) \rangle &= 0, & i > 2r - 1. \end{aligned}$$

Hence, by Lemma 4.1, we have

$$\begin{aligned} & \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} b_k(e^{j\omega}) \\ &= \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), b_m(z) \right\rangle b_m(e^{j\omega}) \\ &+ \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), b_m(z) R(z) \right\rangle b_m(e^{j\omega}) R(e^{j\omega}) \\ &+ \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), 1 \right\rangle \\ &+ \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), \bar{b}_m(z) \right\rangle \bar{b}_m(e^{j\omega}). \end{aligned}$$

Then it follows from the definition (23) and the summation of the above equation over  $1 \leq k \leq r$  that:

$$\begin{aligned} & \sum_{k=1}^r \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \bar{b}_k(z) b_m(z) \right\rangle \bar{b}_k(e^{j\omega}) b_m(e^{j\omega}) \\ &= \sum_{k=1}^r \bar{b}_k(e^{j\omega}) \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), b_m(z) \right\rangle b_m(e^{j\omega}) \\ &= \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \sum_{k=1}^r |b_k(e^{j\omega})|^2 \\ &- \sum_{k=1}^r \bar{b}_k(e^{j\omega}) \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), \bar{b}_m(z) \right\rangle \bar{b}_m(e^{j\omega}) \end{aligned}$$

$$\begin{aligned} & - \sum_{k=1}^r \bar{b}_k(e^{j\omega}) \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), 1 \right\rangle \\ & - \sum_{k=1}^r \bar{b}_k(e^{j\omega}) \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), b_m(z) R(z) \right\rangle \\ & \times b_m(e^{j\omega}) R(e^{j\omega}). \end{aligned} \quad (50)$$

Now, to calculate the inner products in the above equation, first let us suppose  $\ell_q$ ,  $1 \leq q \leq r_1$  are distinct in order to use the Residue Theorem for single poles. Let

$$\alpha(z) \triangleq \prod_{k=1}^{r_1} \frac{(z - c_k)(1 - \bar{c}_k z)}{(z - \ell_k)(1 - \bar{\ell}_k z)}$$

in the following calculation. By the definition of inner product, we have

$$\begin{aligned} & \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), \bar{b}_m(z) \right\rangle \\ &= \frac{1}{2\pi j} \oint_{|z|=1} \frac{|C(z)|^2}{|L(z)|^2} \bar{b}_k(z) \bar{b}_m(z) \frac{dz}{z} \\ &= \frac{1}{2\pi j} \oint_{|z|=1} \alpha(z) \bar{b}_k(z) \bar{b}_m(z) \frac{dz}{z} \\ &= \sum_{q=1}^{r_1} \lim_{z \rightarrow \ell_q} \alpha(z) \bar{b}_k(z) \bar{b}_m(z) \frac{z - \ell_q}{z} \\ &= \sum_{q=1}^{r_1} \overline{\bar{b}_k(\ell_q) \bar{b}_m(\ell_q)} \lim_{z \rightarrow \ell_q} \alpha(z) \frac{z - \ell_q}{z}. \end{aligned}$$

Similarly, we can calculate that

$$\begin{aligned} & \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), 1 \right\rangle \\ &= \sum_{q=1}^{r_1} \overline{\bar{b}_k(\ell_q)} \lim_{z \rightarrow \ell_q} \alpha(z) \frac{z - \ell_q}{z} \\ & \left\langle \frac{|C(z)|^2}{|L(z)|^2} b_k(z), b_m(z) R(z) \right\rangle \\ &= \sum_{q=1}^{r_1} \lim_{z \rightarrow \ell_q} b_k(z) \bar{R}(z) \bar{b}_m(\ell_q) \lim_{z \rightarrow \ell_q} \alpha(z) \frac{z - \ell_q}{z}. \end{aligned}$$

Then, it follows from (50), (48), and (49) that

$$\begin{aligned} & \sum_{k=1}^r \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \bar{b}_k(z) b_m(z) \right\rangle \bar{b}_k(e^{j\omega}) b_m(e^{j\omega}) \\ &= \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \sum_{k=1}^r |b_k(e^{j\omega})|^2 \\ &- \sum_{q=1}^{r_1} \left[ \sum_{k=1}^r \sum_{m=1}^r \overline{\bar{b}_k(\ell_q) \bar{b}_m(\ell_q)} \bar{b}_k(e^{j\omega}) \bar{b}_m(e^{j\omega}) \right. \\ & \quad \left. + \sum_{k=1}^r \overline{\bar{b}_k(\ell_q) \bar{b}_k(e^{j\omega})} \right] \times \lim_{z \rightarrow \ell_q} \alpha(z) \frac{z - \ell_q}{z} \\ &- \sum_{q=1}^{r_1} \sum_{k=1}^r \sum_{m=1}^r \lim_{z \rightarrow \ell_q} b_k(z) \bar{R}(z) \bar{b}_k(e^{j\omega}) \\ & \times R(e^{j\omega}) \bar{b}_m(\ell_q) b_m(e^{j\omega}) \times \lim_{z \rightarrow \ell_q} \alpha(z) \frac{z - \ell_q}{z} \end{aligned}$$

$$\begin{aligned}
&= \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \sum_{k=1}^r |b_k(e^{j\omega})|^2 \\
&\quad - \sum_{q=1}^{r_1} \left[ \left( \frac{\bar{\ell}_q e^{j\omega}}{1 - \bar{\ell}_q e^{j\omega}} \right)^2 + \frac{\bar{\ell}_q e^{j\omega}}{1 - \bar{\ell}_q e^{j\omega}} \right] \times \overline{\lim_{z \rightarrow \ell_q} \alpha(z) \frac{z - \ell_q}{z}} \\
&\quad - \sum_{q=1}^{r_1} \left( \frac{e^{j\omega}}{e^{j\omega} - \ell_q} \right) \left( \frac{\ell_q}{e^{j\omega} - \ell_q} \right) \times \lim_{z \rightarrow \ell_q} \alpha(z) \frac{z - \ell_q}{z} \\
&= \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \sum_{k=1}^r |b_k(e^{j\omega})|^2 \\
&\quad - \sum_{q=1}^{r_1} \overline{\left( \frac{\ell_q e^{j\omega}}{(e^{j\omega} - \ell_q)^2} \times \lim_{z \rightarrow \ell_q} \alpha(z) \frac{z - \ell_q}{z} \right)} \\
&\quad - \sum_{q=1}^{r_1} \left( \frac{\ell_q e^{j\omega}}{(e^{j\omega} - \ell_q)^2} \times \lim_{z \rightarrow \ell_q} \alpha(z) \frac{z - \ell_q}{z} \right).
\end{aligned}$$

Then, using the Residue Theorem, we have for any simple closed positively oriented curve  $\mathbf{T}$  with all the poles of  $\alpha(z)$  inside and the point  $e^{j\omega}$  outside, such that

$$\begin{aligned}
&\sum_{k=1}^r \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \bar{b}_k(z) b_m(z) \right\rangle \bar{b}_k(e^{j\omega}) b_m(e^{j\omega}) \\
&= \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \sum_{k=1}^r |b_k(e^{j\omega})|^2 \\
&\quad - 2 \operatorname{Re} \left\{ \frac{1}{2\pi j} \oint_{\mathbf{T}} \frac{z e^{j\omega}}{(e^{j\omega} - z)^2} \alpha(z) \frac{dz}{z} \right\}. \quad (51)
\end{aligned}$$

For the case where there are multiple poles in  $\ell_q$ ,  $1 \leq q \leq r_1$ , we can perturb  $\ell_q$  a little in order to apply the above analysis. Then because the complex integrations on both sides of (51) are continuous on  $\ell_q$ , we have (51) hold for the general case.

Now we are well prepared to calculate (40). By the definitions (32) and (39), we have

$$\begin{aligned}
&\Gamma_{n,r}^*(e^{j\omega}) M_{n,r} \left( \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \right) \Gamma_{n,r}(e^{j\omega}) \\
&= \sum_{u=1}^n \sum_{v=1}^n \gamma_u(e^{j\omega}) \frac{1}{2\pi} \int_{-\pi}^{\pi} \gamma_u(e^{j\omega}) \overline{\gamma_v(e^{j\omega})} \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \\
&\quad \times d\omega \gamma_v(e^{j\omega}) \\
&= \sum_{u=1}^n \sum_{v=1}^n \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \overline{\gamma_u(z)} \gamma_v(z) \right\rangle \overline{\gamma_u(e^{j\omega})} \gamma_v(e^{j\omega}) \\
&= \left( \sum_{v=1}^{n-r} \sum_{u=v}^n + \sum_{u=1}^{n-r} \sum_{v=u+1}^n + \sum_{u=n-r+1}^n \sum_{v=n-r+1}^n \right) \\
&\quad \times \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \overline{\gamma_u(z)} \gamma_v(z) \right\rangle \overline{\gamma_u(e^{j\omega})} \gamma_v(e^{j\omega}) \\
&= \sum_{v=1}^{n-r} \sum_{u=v}^n \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \overline{\gamma_u(z)} \gamma_v(z) \right\rangle \overline{\gamma_u(e^{j\omega})} \gamma_v(e^{j\omega}) \\
&\quad + \sum_{v=1}^{n-r} \sum_{u=v+1}^n \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \overline{\gamma_u(z)} \gamma_v(z) \right\rangle \gamma_u(e^{j\omega}) \overline{\gamma_v(e^{j\omega})} \\
&\quad + \sum_{k=1}^r \sum_{m=1}^r \left\langle \frac{|C(z)|^2}{|L(z)|^2}, \bar{b}_k(z) b_m(z) \right\rangle \bar{b}_k(e^{j\omega}) b_m(e^{j\omega})
\end{aligned}$$

$$\begin{aligned}
&= (n-r) \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} + \frac{|C(e^{j\omega})|^2}{|L(e^{j\omega})|^2} \sum_{k=1}^r |b_k(e^{j\omega})|^2 \\
&\quad - 2 \operatorname{Re} \left\{ \frac{1}{2\pi j} \oint_{\mathbf{T}} \frac{z e^{j\omega}}{(e^{j\omega} - z)^2} \alpha(z) \frac{dz}{z} \right\} \quad (52)
\end{aligned}$$

where, the last equation follows from (45) and (51).

At last, Theorem 3.1 follows from (40), (52) and the fact that

$$\operatorname{Var} \left\{ \hat{G}_{u \rightarrow y} \right\} = |C^\dagger(e^{j\omega}) L^\dagger(e^{j\omega})|^2 \operatorname{Var} \left\{ \hat{G}_{v \rightarrow y'} \right\}.$$

## V. APPLICATIONS TO OTHER MODEL STRUCTURES

### A. Estimated Noise Model

Consider again the model structure (5), but assume that the noise model polynomial  $C$  is estimated, rather than being fixed

$$y(t) = \frac{B(q, \theta)}{F^\dagger(q)} u(t) + C(q, \eta) e(t). \quad (53)$$

If  $B$  and  $C$  are estimated by a standard prediction error method, and the system operates in open loop, the estimates of  $C$  and  $B$  are asymptotically (in the number of data  $N$ ) uncorrelated (see, e.g., [7, (9.42)]). That means that the variance of  $B$  will be the same as if  $C$  was fixed at the resulting estimate. If the order of the model polynomial  $C$  is at least as large as that of the true noise description  $C_0$ , the estimate  $\hat{C}$  will converge to  $C_0$ . All this means that the variance of the estimated frequency function  $\hat{B}/F^\dagger$  will be given by (18) with  $R_2 = 0$  in case the noise model order is sufficiently large.

### B. AR-Models

Consider now the case of estimating a time-series AR-model

$$A(q, \theta) y(t) = e(t). \quad (54)$$

Assume that the true time series is described by

$$A_0(q) y(t) = e(t). \quad (55)$$

Define

$$\hat{A}_N(e^{j\omega}) = 1 + \sum_{k=1}^n \hat{a}_k(N) e^{-j\omega k} = 1 + W(e^{j\omega}) \hat{\theta}_N \quad (56)$$

where  $\hat{\theta}_N$  contains the LS-estimated AR-parameters. It is well known that the covariance matrix of these estimated AR-parameters is given by

$$\lim_{N \rightarrow \infty} N \cdot \operatorname{Cov} \hat{\theta}_N = \sigma_e^2 R^{-1} \quad (57)$$

where  $R$  is the correlation matrix of  $y$ :

$$(R)_{k,j} \triangleq E y(t-k) y(t-j).$$

Thus, note that  $\sigma_e^2 R^{-1}$  is the same matrix as the right-hand side of (11) for  $F^\dagger = C^\dagger = C_0 = 1$ ,  $L^\dagger = A_0$ , and  $\sigma_v = \sigma_e$ . This means that the variance of the  $A$ -function estimate is the same as a special case of Theorem 3.1.

We formulate the result as a separate theorem.

**Theorem 5.1:** Consider the AR-process (55) of order  $r$ . Let the  $A$ -polynomial be estimated as an  $n$ th-order AR model (54)



and assume that  $n \geq r$ . Then the variance of the  $\hat{A}_N(e^{j\omega})$ -function is given by

$$\lim_{N \rightarrow \infty} N \text{Var} \left\{ \hat{A}_N(e^{j\omega}) \right\} = |A_0(e^{j\omega})|^2 \left[ (n-r) + \sum_{m=1}^r \frac{1 - |\alpha_m|^2}{|e^{j\omega} - \alpha_m|^2} \right] \quad (58)$$

where  $\alpha_m$ ,  $1 \leq m \leq r$  are the roots of the  $A_0(z)$ -polynomial.

### C. Lower Bounds for Output Error (OE)-Models

It seems quite difficult to derive an analytic expression of the variance like (18) for the general model (1), where also the poles are estimated. However, (18) actually provides a lower bound for more general models.

Consider first the OE model

$$y(t) = \frac{B(q)}{F(q)} u(t) + e(t). \quad (59)$$

Suppose the true system can be described by

$$y(t) = \frac{B_0(q)}{F_0(q)} u(t) + e(t). \quad (60)$$

It is easy to see that if we partition a parameter vector into one part that is estimated and one part that is replaced by the true values, then the covariance matrix of the estimated part cannot be larger than if the whole vector is estimated. This follows from [7, (9.59)] and the fact that for any  $\alpha \in \mathbb{C}^n$ ,  $\beta \in \mathbb{C}^n$ ,  $\Pi \in \mathbb{C}^{n \times n}$ , and symmetric and positive matrices  $A, B \in \mathbb{R}^{n \times n}$ ,

$$[\alpha^*, \beta^*] \begin{bmatrix} A & \Pi \\ \Pi^* & B \end{bmatrix}^{-1} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \geq \beta^* B^{-1} \beta$$

if  $\begin{bmatrix} A & \Pi \\ \Pi^* & B \end{bmatrix} > 0$ . (Where “ $\cdot$ ” denotes “conjugate transpose”).

In the present case, this means that the variance of the estimate  $\hat{B}/\hat{F}$  must be at least as large as if the denominator polynomial is fixed to its true value:

$$y(t) = \frac{B(q)}{F_0(q)} u(t) + e(t). \quad (61)$$

The variance of the frequency function estimate in (61) is given by Theorem 3.1, which leads to the following result.

*Theorem 5.2:* Consider the output error model (59), where the input is given by (7). Assume that the true system is given by (60). Suppose that the order of  $B$  is no less than the order of  $B_0$  and also the order of  $F_0 \cdot L^\dagger$ . Then the variance of the frequency function estimate is bounded from below by

$$\lim_{N \rightarrow \infty} N \text{Var} \left\{ \hat{G}_N(e^{j\omega}) \right\} \geq \frac{\sigma_e^2}{\sigma_v^2} |L^\dagger(e^{j\omega})|^2 \left[ (n-r) + R_1(F_0 \cdot L^\dagger) \right] \quad (62)$$

where  $R_1$  is defined in Theorem 3.1.

### D. Lower Bounds for ARMAX-Models

Now, consider the ARMAX model

$$A(q, \theta)y(t) = B(q, \theta)u(t) + C(q, \theta)e(t). \quad (63)$$

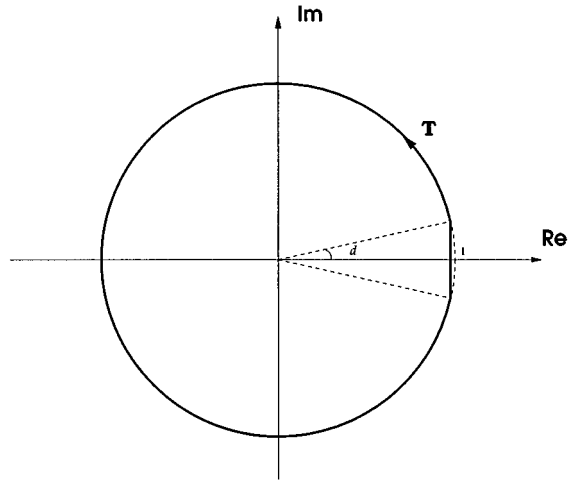


Fig. 1.  $\mathbf{T}$  is chosen close to the circle  $|z| = 1$  and avoiding the singular point  $e^{j\omega}$  with  $\omega = 0$ .

Suppose the true system can be described by

$$A_0(q)y(t) = B_0(q)u(t) + C_0(q)e(t). \quad (64)$$

As above, it follows that the variance of the frequency function estimate using the model (63) is no less than that using the following model:

$$A_0(q)y(t) = B(q, \theta)u(t) + C_0(q)e(t). \quad (65)$$

Define  $\tilde{y}(t) = A_0(q)y(t)$  and apply Theorem 3.1 to the estimation of  $B$  in  $\tilde{y}(t) = B(q, \theta)u(t) + C_0(q)e(t)$ . The variance of  $\hat{B}$  according to (18) can then be translated to the variance of the frequency function estimate  $\hat{B}/A_0$  to yield the following result.

*Theorem 5.3:* Consider the problem of estimating the frequency function in the model structure (63), where the input is given by (7). Assume that the true system is given by (64). Suppose that the order of  $B$  is no less than the order of  $B_0$  and also the order of  $C_0 \cdot L^\dagger$ . Then the variance of the frequency function estimate is bounded from below by

$$\lim_{N \rightarrow \infty} N \text{Var} \left\{ \hat{G}_N(e^{j\omega}) \right\} \geq \frac{\sigma_e^2}{\sigma_v^2} \frac{|C_0(e^{j\omega})|^2 |L^\dagger(e^{j\omega})|^2}{|A_0(e^{j\omega})|^2} \times \left[ (n-r) + R_1(C_0 \cdot L^\dagger) \right]. \quad (66)$$

## VI. NUMERICAL ILLUSTRATION

### A. What Does $R_2(C_0/C^\dagger)$ Look Like?

It would be instructive to see how  $R_2(C_0/C^\dagger)$  depends on

$$\phi(\eta) \triangleq \left| \frac{C_0(e^{j\eta})}{C^\dagger(e^{j\eta})} \right|^2$$

which is the filtered noise spectrum. To this end, we'd like to choose  $\mathbf{T}$  as the circle  $|z| = 1$  in the complex integration (20) since then  $z = e^{j\eta}$ ,  $\beta(z)\beta^*(z) = |\beta(e^{j\eta})|^2 = \phi(\eta)$  and (20) becomes a weighted integration of the filtered noise spectrum  $\phi(\eta)$ . But  $z = e^{j\omega}$  is a singular point in the complex integration

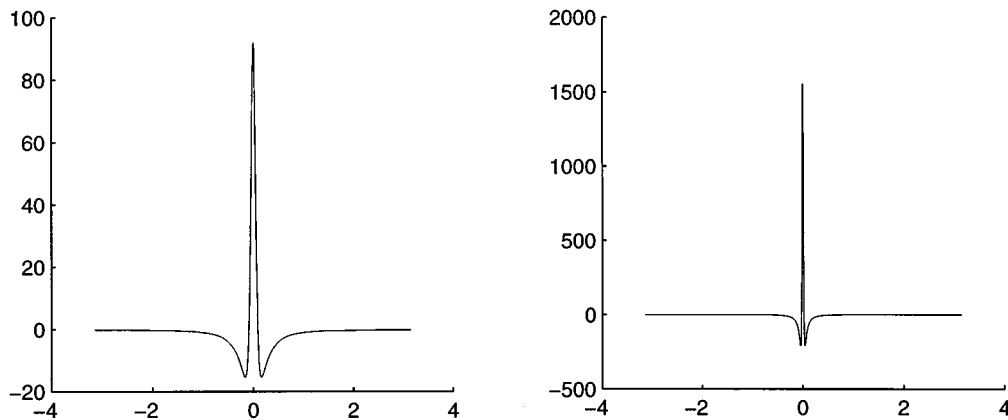


Fig. 2. The real part of the weighting function  $W(\omega, \eta)$  as a function of  $\eta$  for  $\omega = 0$ . Left:  $d = \pi/7$ . Right:  $d = \pi/14$ .

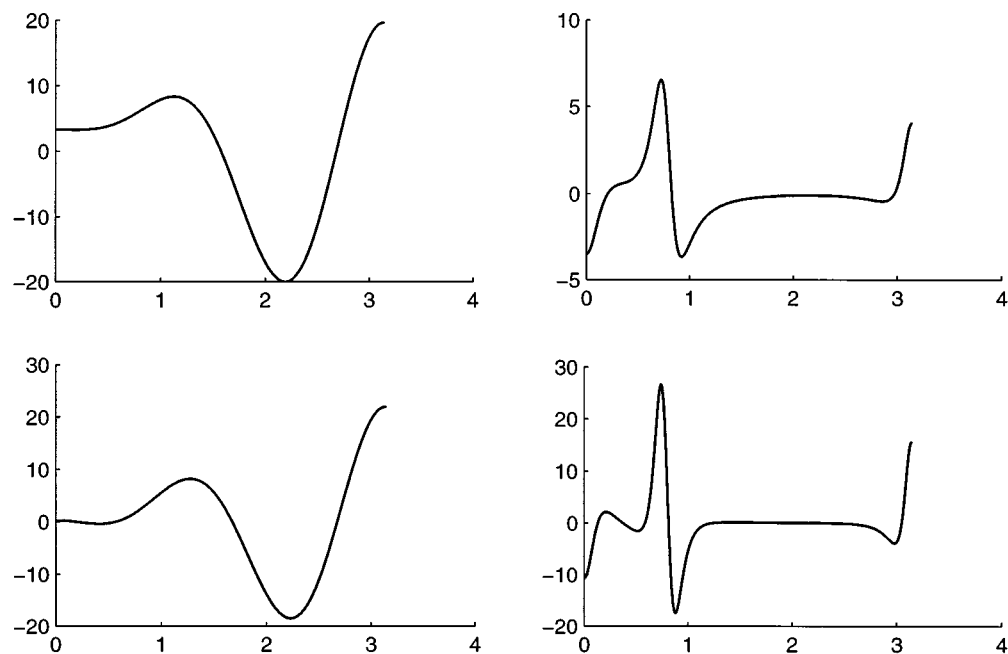


Fig. 3. Two comparisons between  $R_2(C_0/C^\dagger)$  (top) and  $(1/\pi)\phi''(\omega)$  (bottom) with  $\phi(\omega) = |C_0(e^{j\omega})/C^\dagger(e^{j\omega})|^2$ , as functions of  $\omega$ . Left:  $C_0$  zeros: 0.1, 0.6, 0.7, -0.9;  $C^\dagger$  zeros: 0.2, 0.5, 0.8, -0.85. Right: Two more zeros to  $C_0$ :  $0.7 - 0.6i$ ,  $0.7 + 0.6i$ ; Two more zeros to  $C^\dagger$ :  $0.6 - 0.6i$ ,  $0.6 + 0.6i$ .

(20). So  $\mathbf{T}$  must avoid it by indenting to the inside a little (e.g., in Fig. 1, where setting  $e^{j\omega} = 1$ ). To see how sensitively the value of the complex integration depends on the local behavior of  $\phi(\eta)$  around  $z = e^{j\omega}$ , it is interesting to note that the integration value will increase by  $\phi'(\omega)$  if  $\mathbf{T}$  instead indents to the outside a little. [This can easily be proved using the Residue Theorem for the multiple pole  $e^{j\omega}$  in (20).]

In the following, we specially choose  $\mathbf{T}$  as the curve shown in Fig. 1, which is actually made up of two parts: an arch and a piece of line. As the angle  $d$  tends to zero,  $\mathbf{T}$  tends to the circle  $|z| = 1$ . Hence we can nearly regard  $R_2(C_0/C^\dagger)$  as a weighted integration of the filtered noise spectrum  $|C_0(e^{j\eta})/C^\dagger(e^{j\eta})|^2$ , namely,

$$R_2\left(\frac{C_0}{C^\dagger}\right) \approx -2 \operatorname{Re} \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega, \eta) \left| \frac{C_0(e^{j\eta})}{C^\dagger(e^{j\eta})} \right|^2 d\eta \right\} \tag{67}$$

where, the weighting function is defined as

$$W(\omega, \eta) \triangleq \begin{cases} \frac{1}{2 \cos(\eta - \omega) - 2}, & \text{if } |\eta - \omega| \geq d \\ \frac{[\cos d \cdot e^{j\omega} + j \cos d \cdot \tan(\eta - \omega)] e^{j\omega}}{[e^{j\omega} - \cos d \cdot e^{j\omega} + j \cos d \cdot \tan(\eta - \omega)]^2} & \text{if } |\eta - \omega| < d. \end{cases} \tag{68}$$

Since the noise spectrum is real, we only need to consider the real part of the weighting function, which is plotted in Fig. 2 for  $d = \pi/7$  and  $d = \pi/14$  respectively (where setting  $\omega = 0$ ).

We can see that the graph of the weighting function is quite typical, made up of three jumps around  $\omega$ , the frequency in question. As  $d$  tends to zero, these jumps become more narrow and

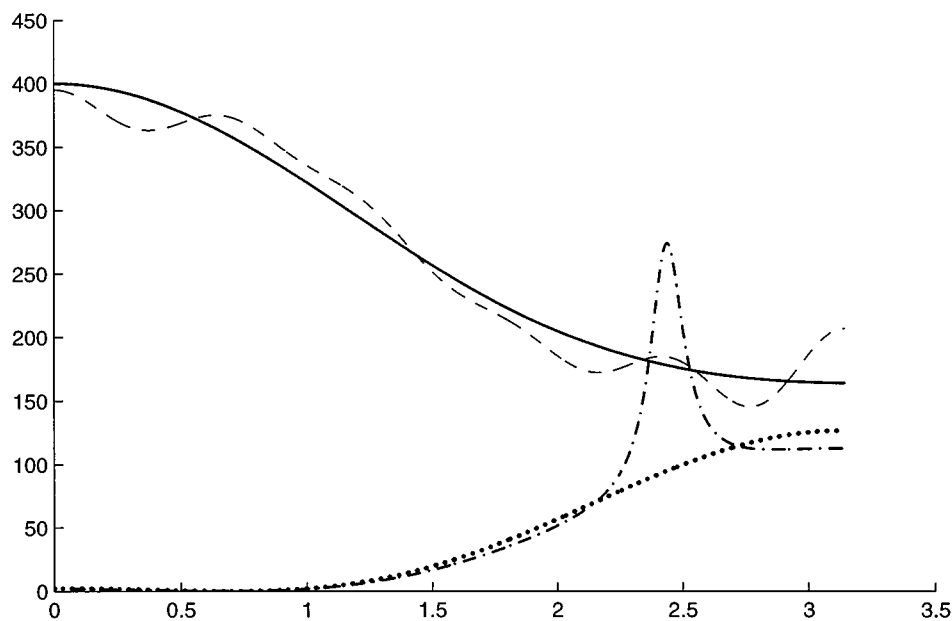


Fig. 4. Variance of the estimated frequency functions, as functions of frequency. The true system is given by (8) with  $B_0(q) = q^{-1}(1 + 0.5q^{-1})^{11}$ ,  $C_0(q) = 1 - 1.4q^{-1} + 0.85q^{-2}$  and  $F^\dagger = 1$ . The model structure is given by (5) with  $n = 12$  and  $C^\dagger(q) = 1 + 1.4q^{-1} + 0.85q^{-2}$ . The input is white noise. The dashed line shows the variance from a Monte Carlo experiment with 500 runs. The solid line is the expression (18). The dotted line is the “classical” expression (3), while the dash-dotted line is the expression from [12].

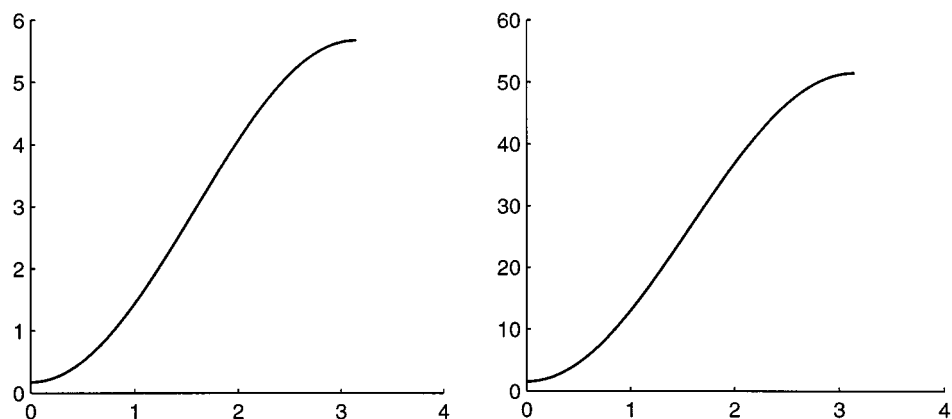


Fig. 5. Illustration of the effect of a bad choice of prefilter. The curves show the frequency function variance as a function of frequency. The expression is given by (18) with  $L^\dagger = F^\dagger = 1$ ,  $\sigma_s^2 = \sigma_v^2 = 1$  and  $n = r = 2$ . Left:  $C_0$  has zeros at 0.8, 0.9; while  $C^\dagger$  has zeros at 0.4, 0.5. Right: The converse situation:  $C_0$  has zeros at 0.4, 0.5; while  $C^\dagger$  has zeros at 0.8, 0.9. Note the different scales of the  $y$ -axis!

steeper and it can be proved that the area above zero is asymptotically proportional to the area below zero. By the formula for the second derivative:

$$f''(x) = \lim_{h \rightarrow 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}$$

one may guess that  $R_2(C_0/C^\dagger) \approx (1/\pi)\phi''(\omega)$ , with  $\phi(\omega) \triangleq |C_0(e^{j\omega})/C^\dagger(e^{j\omega})|^2$ . We present two examples in Fig. 3, which shows that the second derivative indeed can catch the spirit of the complex integration although the fitting is not exact and the scaling differs from case to case. Since from (67), it is quite clear that  $R_2(C_0/C^\dagger)$  depends on the global behavior of  $\phi(\eta)$ ,  $-\pi \leq \eta \leq \pi$ , we cannot expect the local behavior  $\phi''(\omega)$  provide the total information.

### B. Some Comparisons

To see how much the new result may differ from the existing ones, we choose a  $n = 12$ th order model in Fig. 4. It is obvious that both the existing results [7] and [12] are quite misleading in this case.

From the model structure (10), it is clear that we can regard  $C^\dagger$  as a prefilter of the input and output data. It has been proved (see, e.g., [7]) that choosing  $C^\dagger = C_0$  will minimize the variance, i.e., achieve the Cramer–Rao bound. But the problem arises of course when  $C_0$  is not known exactly. In this case, one has to guess where the zeros of  $C_0$  are located and then choose the poles of the prefilter properly. But one must realize that bad poles of the prefilter will cause much larger variance than bad zeros of  $C_0$ . For example, suppose that it is believed that some zeros of  $C_0$  are close to the unit circle. Choosing a prefilter  $C^\dagger$

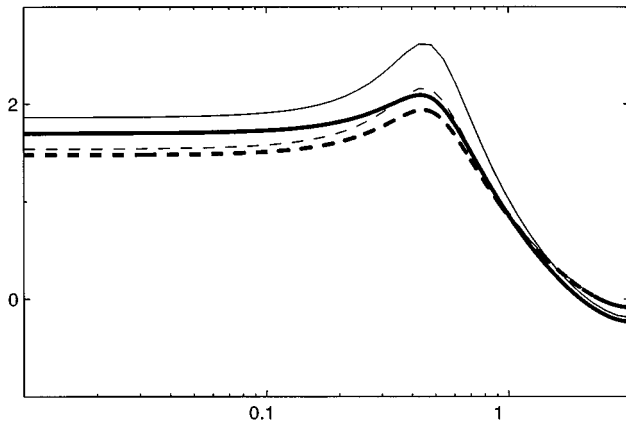


Fig. 6. Modification of (f) in [7, Fig. 9.1] with the new expression (66). Thin lines: Normalized true variances for  $n = 2$  (solid) and  $n = 10$  (dashed). Thick lines: Normalized lower bounds for  $n = 2$  (solid) and  $n = 10$  (dashed).

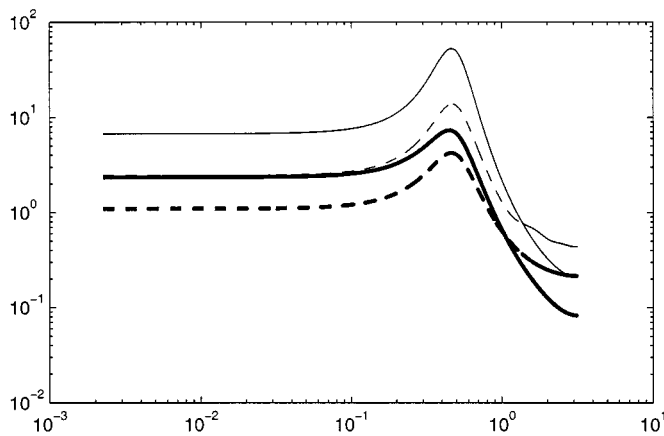


Fig. 7. Comparison between true variance and the lower bound (66) for the "Åström system" corrupted by a colored noise with two zeros at 0.5. Thick lines: normalized lower bounds; thin lines: normalized true variances by Monte Carlo estimate. Solid lines:  $n = 2$ ; dashed lines:  $n = 10$ .

of this kind could however lead to very high variance, if the true  $C_0$  does not have zeros close to the unit circle. The reverse situation that  $C_0$  indeed has zeros close to the unit circle leads to less variance. In other words, if you choose a prefilter with zeros close to the unit circle, you better be sure that this reflects a property of the true system. This is illustrated in Fig. 5.

This observation becomes more concrete when we consider the case when  $r_1 = 1$  in (17), since then by the Residue Theorem, (18) reduces to

$$\begin{aligned} \lim_{N \rightarrow \infty} N \text{Var } \hat{G}_N(e^{j\omega}) &= \frac{\sigma_e^2}{\sigma_v^2} |C_0(e^{j\omega})|^2 |L^\dagger(e^{j\omega})|^2 \times [(n-r) + R_1(F^\dagger \cdot L^\dagger)] \\ &+ \frac{\sigma_e^2}{\sigma_v^2} |L^\dagger(e^{j\omega})|^2 \times [(1+c_1^2)(1+\ell_1^2) - 4\ell_1 c_1] \frac{1}{1-\ell_1^2}. \end{aligned} \quad (69)$$

It is clear that it is very unfavorable for  $\ell_1$  to be close to 1 unless  $c_1$  is also close to 1.

### C. Lower Bounds for ARMAX Models

In the case where both the input and the noise are white, the expression (66) reduces to the expression (3), i.e., the result in [7]. In fact, [7, Fig. 9.1] has shown that it is really a lower bound for the ARMAX model and the true variance approaches the lower bound as the model order goes to infinity. The graph "f" is an exception because there the input data is not white. Hence, we need to use the formula (66) instead. The new result is shown in Fig. 6. (From now on, we use the log-log scale plot in accordance with that used in [7, Fig. 9.1].) It should be noted that in Fig. 6 the lower bounds for the second order and the tenth order are no longer the same due to the new expression (66).

Next, we consider the case where the noise is not white. We shall compare how much the true variance differs from the lower bound calculated from (66). Suppose the true system can be described by (64) with  $A_0(q) = 1 - 1.5q^{-1} + 0.7q^{-2}$ ,  $B_0(q) = q^{-1} + 0.5q^{-2}$ ,  $C_0(q) = 1 - q + 0.25q^{-2}$ , i.e., the "Åström system" corrupted by a colored noise with two zeros at 0.5. We use the ARMAX model (63) with order  $n = 2$  or  $n = 10$  respectively. The results are shown in Fig. 7.

## VII. CONCLUSION

It is a simple and natural asymptotic result that the covariance of an estimated frequency function is proportional to the noise-to-signal ratio at the frequency in question. However, due to erroneous noise models, colored inputs, fixed poles and other features, the actual variance of the frequency function estimate may differ substantially from the asymptotic value for low order models. This has been pointed out clearly in [12]. We have here developed an exact result for the covariance within a limited class of models and input spectra. This result points to the features that cause the deviations from the limit expression. It also establishes the convergence rate to the limit to be like 1 over the model order.

As a fringe benefit an exact result for the covariance of the AR-model frequency function is obtained, thus extending the classical asymptotic result by Berk, [1].

## REFERENCES

- [1] K. N. Berk, "Consistent autoregressive spectral estimates," *Ann. Stat.*, vol. 2, pp. 489–502, 1974.
- [2] D. R. Brillinger, *Time Series: Data Analysis and Theory*. San Francisco, CA: Holden-Day, 1981.
- [3] U. Forsell and L. Ljung, "Some results on optimal experiment design," *Automatica*, vol. 36, no. 5, pp. 749–756, May 2000.
- [4] M. Gevers and L. Ljung, "Optimal experiment designs with respect to the intended model application," *Automatica*, vol. 22, pp. 543–555, 1986.
- [5] G. C. Goodwin, M. Gevers, and B. Ninness, "Quantifying the error in estimated transfer functions with application to model order selection," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 913–929, July 1992.
- [6] L. Ljung, "Asymptotic variance expressions for identified black-box transfer function models," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 834–844, Sept. 1985.
- [7] —, *System Identification—Theory for the User*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.
- [8] L. Ljung and B. Wahlberg, "Asymptotic properties of the least-squares method for estimating transfer functions and disturbance spectra," *Adv. Appl. Prob.*, vol. 24, pp. 412–440, 1992.
- [9] L. Ljung and Z. D. Yuan, "Asymptotic properties of black-box identification of transfer functions," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 514–530, June 1985.

- [10] B. Ninness and F. Gustafsson, "A unifying construction of orthonormal bases for system identification," *IEEE Trans. Autom. Contr.*, vol. 42, pp. 515–521, Apr. 1999.
- [11] B. Ninness, H. Hjalmarsson, and F. Gustafsson, "Generalized Fourier and Toeplitz results for rational orthonormal bases," *SIAM J. Control Optim.*, vol. 37, no. 2, pp. 429–460, 1999.
- [12] —, "The fundamental role of general orthonormal bases in system identification," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 1384–1406, July 1999.



**Liang-Liang Xie** was born in Hunan, China, in 1974. He received the B.S. degree in mathematics from Shandong University and the Ph.D. degree in control theory from the Chinese Academy of Sciences in 1995 and 1999, respectively.

He joined the faculty of Institute of Systems Science, Chinese Academy of Sciences in 1999. He spent the academic year 1999–2000 as a Guest Researcher in the Automatic Control Group, Linköping University, Sweden. He is currently visiting Coordinated Science Laboratory, University

of Illinois at Urbana-Champaign as a Postdoctoral Research Associate. His research interests include stochastic control systems, adaptive control, system identification, network information theory, and wireless networks.

Dr. Xie received the Excellent Prize of President Scholarship from the Chinese Academy of Sciences in 1999.



**Lennart Ljung** (S'74–M'75–SM'83–F'85) received the Ph.D. degree in automatic control from Lund Institute of Technology, Sweden, in 1974.

Since 1976, he has been Professor of the Chair of Automatic Control in Linköping, Sweden, and is currently Director of the Competence Center "Information Systems for Industrial Control and Supervision" (ISIS). He has held visiting positions at Stanford University, Stanford, CA, and the Massachusetts Institute of Technology, Cambridge. He has written several books on system identification and estimation.

Dr. Ljung is an IFAC Advisor and a member of the Royal Swedish Academy of Sciences (KVA), a member of the Royal Swedish Academy of Engineering Sciences (IVA), and an Honorary Member of the Hungarian Academy of Engineering. He has received honorary doctorates from the Baltic State Technical University in St. Petersburg, and from Uppsala University, Sweden.