Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Exact quantification of the variance of estimated zeros

Jonas Mårtensson and Håkan Hjalmarsson

Abstract— **This paper is concerned with quantification of noise induced errors in estimates of zeros of dynamic systems. Preceding work on this problem has provided variance expressions that are asymptotic in data length *and* model order for *non-minimum phase* zeros. This paper presents expressions that are asymptotic only in data length and they are therefore 'exact' for arbitrarily small true model orders. These expressions are also valid for both minimum phase and non-minimum phase zeros. A key insight is that the variance error quantification problem is equivalent to deriving a reproducing kernel for a space that depends on the employed model structure.**

I. Introduction

Quantification of the variance of estimated poles and zeros is closely related to variance quantification for estimated frequency functions as all these problems deal with quadratic forms based on the covariance matrix of the underlying parameter estimate. The latter problem has received significant interest. In the mid-eighties, an expression for the asymptotic (as the sample size grows) variance of estimated frequency functions was presented in [1], [2]. It showed that the variance increases proportionally to the model order regardless of model structure as the model order becomes large. An alternative asymptotic variance expression with, for many model structures, improved accuracy was proposed in [3]. In [4] and [5] expressions that are exact for finite model orders were derived for the variance of estimated frequency functions. Closed loop estimation is treated in the same framework in [6].

Parallel to this, there has been a series of results regarding the accuracy of estimated non-minimum phase zeros. As mentioned above, the variance of an estimate usually increases proportionally with the model order, but estimates of non-minimum phase zeros only suffer from a moderate increase in the variance. This was shown for FIR-models in [7] and ARX-models in [8]. More general models, such as output-error and Box-Jenkins, were treated in [9]. Closed loop identification of non-minimum phase zeros and unstable poles was treated in [10]. These contributions provide variance expressions that are asymptotic in model order and the key observation here is that the variance of estimated non-minimum phase zeros and unstable poles, *asymptotically*, does not depend on the model order.

This paper presents variance expressions that are exact for finite model orders and it uses the same methods as in [5]. The variance quantification problem is shown to be equal to that of deriving a reproducing kernel for a space that depends on the model structure. Estimation of both minimum phase and non-minimum phase zeros in general model structures as output-error and Box-Jenkins can be treated in this framework and the variance expressions presented here are more accurate than the asymptotic expression derived earlier. Exact variance expressions can only be found for certain model structures and input excitations. As a complement to this we show that upper bounds can be found for more general model structures and input signals.

The outline of this paper is as follows. Parametric system identification and its statistical properties is presented in Section II. In Section III this is related to the properties of the corresponding zero estimates. Reproducing kernels are introduced in Section IV and it is shown how a reproducing kernel can be used to quantify the variance of an estimated zero. For some specific model structures these methods can be used to derive a variance expression that is exact for finite model orders. This is presented in Section V. More general model structures, for which an upper bound of the variance can be derived (also for finite model orders), are treated in Section VI. An asymptotic (in model order) expression is also presented here. Some simulations are used in Section VII to illustrate the improved accuracy obtained with the 'exact' expression compared to the asymptotic. Finally, some concluding remarks are given in Section VIII.

II. System identification

In this section the settings of the system identification is briefly outlined and some well-known statistical properties of the estimated model is presented. The method used is the standard prediction error method, see e.g [11]. In this contribution we focus on systems where the noise model is independently parameterized so that distinction is made from start.

The model structure is parameterized by the two vectors θ and η and it can be described with the rational transfer functions $G(q,\theta)$ and $H(q,\eta)$ in the input-output relation

$$y_t = G(q,\theta)u_t + H(q,\eta)e_t \tag{1}$$

where $H(q,\eta)$ is monic and $\{e_t\}$ is a zero-mean white noise sequence. Here q is the forward shift operator $qu_t = u_{t+1}$, ($q^{-1}u_t = u_{t-1}$). The parameter vectors are estimated by minimizing the sum of squared prediction errors,

$$[\hat{\theta}_N, \hat{\eta}_N] = \arg\min_{\theta,\eta} \frac{1}{N}\sum_{t=1}^{N} \varepsilon_t^2(\theta,\eta), \tag{2}$$

where the prediction error is given by

$$\varepsilon_t(\theta,\eta) = \frac{1}{H(q,\eta)}\left(y_t - G(q,\theta)u_t\right). \tag{3}$$

Assume that the true system can be described with vectors θ^o and η^o and a white noise sequence $\{e_t^o\}$ with variance λ_0, such that

$$y_t = G(q,\theta^o)u_t + H(q,\eta^o)e_t^o. \tag{4}$$

This work was supported by the Swedish Research Council

J. Mårtensson and H. Hjalmarsson are affiliated with the Department of Signals, Sensors and System, KTH, SE-100 44 Stockholm, Sweden
jonas.martensson@s3.kth.se

0-7803-9568-9/05/$20.00 ©2005 IEEE

The true system will sometimes be denoted without the second argument as $G_o(q) \triangleq G(q,\theta^o)$. An analogue notation will be used for other functions that regard the true system.

Under some mild conditions the parameter estimates have the following statistical properties. The estimates $\hat{\theta}_N$ and $\hat{\eta}_N$ are asymptotically uncorrelated and $\hat{\theta}_N$ has an asymptotic distribution

$$\sqrt{N}(\hat{\theta}_N - \theta^o) \in \mathrm{AsN}(0, \lambda_0 P) \tag{5}$$

where

$$P^{-1} = \mathbf{E}\psi_t(\theta^o,\eta^o)\psi_t^T(\theta^o,\eta^o) \tag{6}$$

and

$$\psi_t(\theta,\eta) = \frac{1}{H(q,\eta)} \frac{\partial G(q,\theta)}{\partial \theta} u_t. \tag{7}$$

Let the input u_t have a spectrum $\phi_u(e^{i\omega}) = Q(e^{i\omega})Q^*(e^{i\omega})$ for some minimum phase filter $Q(q)$ and denote

$$\Psi(q) \triangleq \frac{Q(q)}{H(q,\eta^o)} \left.\frac{\partial G(q,\theta)}{\partial \theta}\right|_{\theta=\theta^o}. \tag{8}$$

Now the parameter covariance can be expressed as

$$P^{-1} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Psi(e^{i\omega})\Psi^*(e^{i\omega})d\omega. \tag{9}$$

The inverse exists provided that no pole-zero cancellations occur in $G(q,\theta^o)$ or $H(q,\eta^o)$.

III. Identification of zeros

The main interest in this contribution is the variance of the zeros of the estimated system $G(q,\hat{\theta}_N)$. The system is assumed to be a linear time-invariant (LTI) rational transfer function defined by the two polynomials $A(q,\theta)$ and $B(q,\theta)$ as

$$G(q,\theta) = \frac{B(q,\theta)}{A(q,\theta)}. \tag{10}$$

The polynomials are given by

$$\begin{aligned} A(q,\theta) &= 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}, \\ B(q,\theta) &= b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b} \end{aligned} \tag{11}$$

where the parameter vector is

$$\theta = [a_1, \cdots, a_{n_a}, b_1, \cdots, b_{n_b}]^T. \tag{12}$$

The noise model $H(q,\eta)$ and the vector η will be defined in a similar way as

$$\begin{aligned} H(q,\eta) &= \frac{C(q,\eta)}{D(q,\eta)} = \frac{1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c}}{1 + d_1 q^{-1} + \cdots + d_{n_d} q^{-n_d}}, \\ \eta &= [c_1, \cdots, c_{n_c}, d_1, \cdots, d_{n_d}]^T. \end{aligned} \tag{13}$$

The $n_b - 1$ zeros of the system, denoted $\{\mathrm{z}_i\}_{i=1}^{n_b-1}$, are defined as the roots of the polynomial

$$p(\mathrm{z},\theta) \triangleq b_1 \mathrm{z}^{n_b-1} + \cdots + b_{n_b} \tag{14}$$

where $b_{n_b} \neq 0$ which implies that $\mathrm{z}_i \neq 0$. A zero at $z = 0$ is a pure time delay and here, time delays are not included in the zeros. The system is modelled with one time delay, cf. (11), and any additional time delays can be modelled by shifting the input signal. The zeros will be functions of the parameter vector θ. The estimated zeros will be denoted $\hat{\mathrm{z}}_i \triangleq \mathrm{z}_i(\hat{\theta}_N)$ and the true zeros will be denoted $\mathrm{z}_i^o \triangleq \mathrm{z}_i(\theta^o)$. The asymptotic variance of the estimated zero will in this presentation be denoted as $\overline{var}\ \hat{\mathrm{z}}_i \triangleq \lim_{N\to\infty} N\mathbf{E}(\hat{\mathrm{z}}_i - \mathrm{z}_i^o)^2$. Assume that all zeros are distinct and that the system has no zeros on the unit circle. Consider one particular zero z_k. In [9] it is shown that the asymptotic variance of the estimated zero $\hat{\mathrm{z}}_k$ can be expressed as

$$\overline{var}\ \hat{\mathrm{z}}_k = \frac{\lambda_0 |\mathrm{z}_k^o|^2}{|\widetilde{B}_o(\mathrm{z}_k^o)|^2} \Gamma^*(\mathrm{z}_k^o) P \Gamma(\mathrm{z}_k^o) \tag{15}$$

where

$$\Gamma(q) = [0, \cdots, 0, q^{-1}, \cdots, q^{-n_b}]^T \tag{16}$$

and

$$\widetilde{B}_o(q) = \frac{B_o(q)}{1 - \mathrm{z}_k^o q^{-1}}. \tag{17}$$

In the following sections it will be shown how the quantity $\Psi^* P \Psi$ can be evaluated. That result will be used to evaluate the variance of estimated zeros, but in order to do so, (15) must be reformulated since it involves $\Gamma(z)$ and not $\Psi(z)$. First we look at the gradient of $G(q,\theta)$ with respect to θ

$$\begin{aligned} \left.\frac{\partial G(q,\theta)}{\partial a_i}\right|_{\theta=\theta^o} &= -\frac{B_o(q)}{A_o^2(q)} q^{-i}, \\ \left.\frac{\partial G(q,\theta)}{\partial b_i}\right|_{\theta=\theta^o} &= \frac{1}{A_o(q)} q^{-i}. \end{aligned} \tag{18}$$

Now, since z_k^o is a zero of $B_o(q)$, $\Psi(\mathrm{z}_k^o)$ can be expressed by using (8) as

$$\Psi(\mathrm{z}_k^o) = \frac{Q(\mathrm{z}_k^o)}{H_o(\mathrm{z}_k^o) A_o(\mathrm{z}_k^o)} \Gamma(\mathrm{z}_k^o) \tag{19}$$

and together with (15) we get

$$\overline{var}\ \hat{\mathrm{z}}_k = \frac{\lambda_0 |\mathrm{z}_k^o|^2 |H_o(\mathrm{z}_k^o)|^2}{|\widetilde{G}_o(\mathrm{z}_k^o)|^2 |Q(\mathrm{z}_k^o)|^2} \Psi^*(\mathrm{z}_k^o) P \Psi(\mathrm{z}_k^o) \tag{20}$$

where

$$\widetilde{G}_o(z) = \frac{\widetilde{B}_o(z)}{A_o(z)}. \tag{21}$$

IV. Reproducing kernels

In this section the concept of reproducing kernels is introduced and some properties of the reproducing kernel is presented. The reason for involving reproducing kernels is that they give a means to evaluate the quantity $\Psi^* P \Psi$ which is part of the expression for the variance of an estimated zero, see (20).

Suppose that $\mathcal{X}_n$ is a complex vector space with elements being complex valued functions

$$\mathcal{X}_n = \mathrm{span}\{g_1(z), \cdots, g_n(z)\} \tag{22}$$

with an inner product defined as

$$\langle f, g \rangle \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\omega}) g^*(e^{i\omega}) d\omega. \tag{23}$$

The reproducing kernel for the space $\mathcal{X}_n$, denoted $\varphi_n(z,\mu)$, is a family of functions with the properties

$$\varphi_n(\cdot,\mu) \in \mathcal{X}_n, \quad \mu \in \mathbf{C} \tag{24}$$
$$\langle f(\cdot), \varphi_n(\cdot,\mu)\rangle = f(\mu), \quad \forall f \in \mathcal{X}_n. \tag{25}$$

The reproducing kernel for a space is unique and hence independent of which basis functions are used to describe the space. The following lemma shows how the reproducing kernel can be expressed in terms of the basis functions.

Lemma 4.1: The reproducing kernel of the space spanned by the functions $\{g_1(z),\cdots,g_n(z)\}$ is given by

$$\varphi_n(z,\mu) = \Psi_n^*(\mu)P_n\Psi_n(z) \tag{26}$$

where

$$\Psi_n(z) = [g_1(z),\cdots,g_n(z)]^T, \tag{27}$$
$$P_n = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\Psi_n(e^{i\omega})\Psi_n^*(e^{i\omega})d\omega\right)^{-1}. \tag{28}$$

Proof: See [5]. ■

Now the variance of an estimated zero can be expressed in terms of a reproducing kernel as

$$\overline{var}\ \hat{z}_k = \frac{\lambda_0|z_k^o|^2|H_o(z_k^o)|^2}{|\widetilde{G}_o(z_k^o)|^2|Q(z_k^o)|^2}\varphi(z_k^o,z_k^o) \tag{29}$$

where $\varphi(\cdot,\cdot)$ is the reproducing kernel for the space spanned by the elements of the prediction error gradient (8).

The following lemma shows how the reproducing kernel can be expressed if an *orthonormal* basis for the space is found.

Lemma 4.2: Suppose that $\{\mathcal{B}_1,\cdots,\mathcal{B}_n\}$ is an orthonormal basis for the space $\mathcal{X}_n$. Then the reproducing kernel for $\mathcal{X}_n$ can be expressed as

$$\varphi_n(z,\mu) = \sum_{j=1}^{n}\mathcal{B}_j^*(\mu)\mathcal{B}_j(z). \tag{30}$$

Proof: See [5]. ■

Now consider the space spanned by the n_a+n_b elements of the vector

$$\frac{Q(z)}{H(z,\eta^o)}\left.\frac{\partial G(z,\theta)}{\partial\theta}\right|_{\theta=\theta^o} \tag{31}$$

and suppose that $\{\mathcal{B}_k\}_{k=1}^{n_a+n_b}$ is an orthonormal basis for that space. Then Lemma 4.1 and 4.2 can be used to express the variance of an estimated zero as

$$\lim_{N\to\infty} N\mathbf{E}(\hat{z}_k - z_k^o)^2 = \frac{\lambda_0|z_k^o|^2|H_o(z_k^o)|^2}{|\widetilde{G}_o(z_k^o)|^2|Q(z_k^o)|^2}\sum_{j=1}^{n_a+n_b}|\mathcal{B}_j(z_k^o)|^2. \tag{32}$$

A. Spaces induced by model structure

The model structure (1) in consideration has an independently parameterized noise model and includes commonly used models as FIR, output-error and Box-Jenkins. We will show that all these model structures can be treated in the same framework and to do so we start with the most general one, Box-Jenkins. The space that is spanned by the elements of the prediction error gradient (8) is given by

$$\text{span}\left\{\frac{Q(z)G_o(z)z^{-1}}{H_o(z)A_o(z)},\cdots,\frac{Q(z)G_o(z)z^{-n_a}}{H_o(z)A_o(z)},\frac{Q(z)z^{-1}}{H_o(z)A_o(z)},\cdots,\frac{Q(z)z^{-n_b}}{H_o(z)A_o(z)}\right\} \tag{33}$$

TABLE I

$A_\dagger$ AND n FOR DIFFERENT MODEL STRUCTURES

	FIR	OE	BJ
$A_\dagger$	$1/Q$	A_o^2/Q	$H_oA_o^2/Q$
n	n_b	n_a+n_b	n_a+n_b

and this space can be shown to be equal to

$$\text{span}\left\{\frac{Q(z)z^{-1}}{H_o(z)A_o^2(z)},\cdots,\frac{Q(z)z^{-(n_a+n_b)}}{H_o(z)A_o^2(z)}\right\} \tag{34}$$

provided that $A_o(z)$ and $B_o(z)$ are co-prime [5]. The FIR and output-error structures are special cases of (34) and all cases can be expressed as

$$\text{span}\left\{\frac{z^{-1}}{A_\dagger(z)},\cdots,\frac{z^{-n}}{A_\dagger(z)}\right\} \tag{35}$$

where $A_\dagger$ and n take on different roles for each model structure, see Table I. The reproducing kernel for the space (35) will be further studied in the following sections.

V. Spaces with fixed poles

In this section we treat the space (35) where the elements have a fixed set of poles, and zeros only at the origin. This imposes some restrictions on the system $\{G_o(q),H_o(q)\}$ and the input filter $Q(q)$ but we will come back to that later.

Consider the space

$$\mathcal{X}_n \triangleq \text{span}\left\{\frac{z^{-1}}{M(z)},\cdots,\frac{z^{-n}}{M(z)}\right\} \tag{36}$$

where

$$M(z) \triangleq \prod_{i=1}^{m}(1-\xi_i z^{-1}), \quad |\xi_i| < 1 \tag{37}$$

and $n \geq m$. In [12] it is shown that an orthonormal basis $\{\mathcal{B}_i(z)\}_{i=1}^n$ for the space $\mathcal{X}_n$ can be formed by the functions

$$\mathcal{B}_i(z) \triangleq \frac{\sqrt{1-|\xi_i|^2}}{z-\xi_i}\phi_{i-1}(z), \quad \phi_i(z) \triangleq \prod_{j=1}^{i}\frac{1-\overline{\xi}_j z}{z-\xi_j} \tag{38}$$

where $\xi_i = 0$ for $i > m$ and $\phi_0(z) = 1$. Now it is also possible to express the reproducing kernel for the space $\mathcal{X}_n$ in terms of the function $\phi_n(z)$ in what is called a Christoffel-Darboux formula [13]. The reproducing kernel for $\mathcal{X}_n$ is given by

$$\varphi_n(z,\mu) = \sum_{i=1}^{n}\mathcal{B}_i^*(\mu)\mathcal{B}_i(z) = \frac{1-\overline{\phi_n(\mu)}\phi_n(z)}{z\overline{\mu}-1} \tag{39}$$

and it can also be expressed in terms of the function $M(z)$ by noting that

$$\phi_n(z) = \frac{M(z^{-1})}{M(z)}z^{-n}. \tag{40}$$

Especially, this means that the reproducing kernel can be written as

$$\varphi_n(z,z) = \frac{|z|^{-2}}{1-|z|^{-2}}\left(1-\frac{|M(1/z)|^2}{|M(z)|^2}|z|^{-2n}\right). \tag{41}$$

The results above can now be used to form an expression for the variance of an estimated zero. Unlike the expressions presented in [9] this is valid for finite model orders and for both minimum and non-minimum phase zeros.

Theorem 5.1: Assume that the input filter $Q(q)$ is an AR-filter given by

$$Q(q) = \frac{\sigma}{F(q)}, \quad F(q) \triangleq 1 + f_1 q^{-1} \cdots + f_{n_f} q^{-n_f} \tag{42}$$

and that the model order constraints $n_b \geq n_a + n_c + n_f$ and $n_d = 0$ hold. Then the variance of an estimated zero is given by

$$\boxed{\overline{var}\ \hat{z}_k = \frac{\lambda_0 |H_o(z_k^o)|^2}{(1-|z_k^o|^{-2})|\widetilde{G}_o(z_k^o)|^2|Q(z_k^o)|^2}\left(1 - \frac{|A_\ddagger(1/z_k^o)|^2|z_k^o|^{-2(n_a+n_b)}}{|A_\ddagger(z_k^o)|^2}\right)} \tag{43}$$

where $A_\ddagger(z) = A_o^2(z)C_o(z)F(z)$.

Proof: With this choice of system we get the function space

$$\mathcal{X}_n = \text{span}\left\{\frac{z^{-1}}{A_\ddagger(z)}, \cdots, \frac{z^{-(n_a+n_b)}}{A_\ddagger(z)}\right\}, \tag{44}$$

c.f (34), where $A_\ddagger(z)$ is a polynomial of degree $2n_a + n_c + n_f$. The expression (43) is formed by combining (29) and (41). *Remark:* For FIR models it holds that $n_a = n_c = n_d = 0$ and for output-error models it holds that $n_c = n_d = 0$. ■

Theorem 5.1 gives an expression for the variance of an estimated zero and the expression is asymptotic only in the number of data and not in model order. However, it does suffer from the restriction that the noise and input signals must be described as AR-filtered white noise.

The variance expression (43) shows how the variance depends on the model order. For a *minimum phase* zero, ($|z| < 1$), the variance grows as $|z|^{-2n_b}$. For a *non-minimum phase* zero, ($|z| > 1$), the variance approaches a constant value with the rate $|z|^{-2n_b}$ when the model order increases. Also note that the variance increases as $\frac{1}{1-|z|^{-2}}$ when the zero approaches the unit circle

VI. Spaces with fixed poles and zeros

In this section we will derive an upper bound on the variance of an estimated zero in the case when the model structure $\{G_o(q), H_o(q)\}$ and input filter $Q(q)$ induce a space of functions that have both specified poles and zeros. There will still be a model order requirement for this result to hold. We will also present results that are asymptotic in model order which will confirm the work in [7] and [9].

But first we present two important lemmas that relate reproducing kernels of different spaces to each other.

Lemma 6.1: Let the two spaces $\mathcal{X}$ and $\mathcal{X}^+$ have reproducing kernels $\varphi(z,\mu)$ and $\varphi^+(z,\mu)$ and let $\mathcal{X} \subseteq \mathcal{X}^+$. Then it holds that

$$\varphi(z,z) \leq \varphi^+(z,z). \tag{45}$$

Proof: Let $\{\mathcal{B}_k\}_{k=1}^n$ be an orthonormal basis for the space $\mathcal{X}$. By Gram-Schmidt orthonormalization these functions can form the n first basis functions in an orthonormal basis for $\mathcal{X}^+$ so that $\{\mathcal{B}_k\}_{k=1}^{n+m}$ is an orthonormal basis for $\mathcal{X}^+$. Now it is clear, by using Lemma 4.2, that

$$\varphi(z,z) = \sum_{k=1}^{n} |\mathcal{B}_k(z)|^2 \leq \sum_{k=1}^{n+m} |\mathcal{B}_k(z)|^2 = \varphi^+(z,z). \quad ■ \tag{46}$$

Lemma 6.2: Let the spaces $\mathcal{X}$ and $\widetilde{\mathcal{X}}$ have reproducing kernels $\varphi(z,\mu)$ and $\widetilde{\varphi}(z,\mu)$. Suppose that there exists a constant δ such that for every function $f \in \mathcal{X}$ there exists a function $g \in \widetilde{\mathcal{X}}$ that fulfills

$$|f(z) - g(z)| < \delta, \quad \forall z. \tag{47}$$

Then it holds that

$$\sqrt{\varphi(z,z)} \leq \sqrt{\widetilde{\varphi}(z,z)} + \delta + \sqrt{\delta}. \tag{48}$$

Proof: See the appendix. ■

A. Upper bound on the zero estimate variance

Here we derive an upper bound on the variance of estimated zeros. The only restriction is a model order requirement on $B(q,\theta)$. This result is presented in the following theorem.

Theorem 6.1: Suppose that the input filter is given by

$$Q(q) = \frac{E(q)}{F(q)} = \frac{e_0 + e_1 q^{-1} + \cdots e_{n_e} q^{-n_e}}{1 + f_1 q^{-1} + \cdots f_{n_f} q^{-n_f}} \tag{49}$$

and that the model order constraint $n_b + n_d + n_e \geq n_a + n_c + n_f$ holds. Then the variance of an estimated zero is bounded above by

$$\boxed{\overline{var}\ \hat{z}_k \leq \frac{\lambda_0 |H_o(z_k^o)|^2}{(1-|z_k^o|^{-2})|\widetilde{G}_o(z_k^o)|^2|Q(z_k^o)|^2}\left(1 - \frac{|A_\dagger(1/z_k^o)|^2|z_k^o|^{-2n_\dagger}}{|A_\dagger(z_k^o)|^2}\right)} \tag{50}$$

where $A_\dagger(z) = A_o^2(z)C_o(z)F(z)$ and $n_\dagger = n_a + n_b + n_d + n_e$.

Proof: First consider the space

$$\mathcal{X}_n \triangleq \text{span}\left\{\frac{L(z)}{M(z)}z^{-1}, \cdots, \frac{L(z)}{M(z)}z^{-n}\right\} \tag{51}$$

where $M(z)$ is given by (37) and

$$L(z) \triangleq l_0 + l_1 z^{-1} + \cdots + l_{n_l} z^{-n_l}. \tag{52}$$

A larger space that contains $\mathcal{X}_n$ can be defined as

$$\mathcal{X}_n^+ \triangleq \text{span}\left\{\frac{z^{-1}}{M(z)}, \cdots, \frac{z^{-(n+n_l)}}{M(z)}\right\} \tag{53}$$

where now $\mathcal{X}_n \subseteq \mathcal{X}_n^+$. With $\varphi_n(z,\mu)$ and $\varphi_n^+(z,\mu)$ denoting the reproducing kernels for $\mathcal{X}_n$ and $\mathcal{X}_n^+$, Lemma 6.1 gives that $\varphi_n(z,z) \leq \varphi_n^+(z,z)$ where $\varphi_n^+(z,z)$ can be expressed by using the results in Section V.

Now with $M = A_\dagger \triangleq A_o^2 C_o F$ and $L = D_o E$ we get that $\mathcal{X}_n$ is the space induced by the Box-Jenkins structure, c.f (34), and the upper bound on the variance of an estimated zero is formulated by combining Theorem 5.1 and Lemma 6.1. ■

B. Asymptotic results for non-minimum phase zeros

We have previously presented a number of results ([7], [8] and [9]) where the variance of estimated *non-minimum phase* zeros has been investigated. Those results are asymptotic in model order and many model structures have been treated. For output-error and Box-Jenkins models there has been a flaw in the proof, but here, by using reproducing kernels, this will be rectified.

The result is stated in the following theorem.

Theorem 6.2: Let z_k be a *non-minimum phase* zero of the system G and let the model order n_b go to infinity. Then the variance of the estimated zero is given by

$$\boxed{\lim_{n_b\to\infty} \overline{var}\ \hat{z}_k = \frac{\lambda_0|H_o(z_k^o)|^2}{(1-|z_k^o|^{-2})|\widetilde{G}_o(z_k^o)|^2|Q(z_k^o)|^2}.} \tag{54}$$

Proof: Consider the two spaces $\mathcal{X}_n$ and $\mathcal{X}_n^+$ given by (51)-(53) and let $\varphi_n(z,\mu)$ and $\varphi_n^+(z,\mu)$ be the associated reproducing kernels. Take an arbitrary function $f \in \mathcal{X}_n^+$ which can be written as $f(z) = \frac{1}{M(z)}\sum_{k=1}^{n+n_l}\alpha_k z^{-k}$ for some parameters α_k. A function $g \in \mathcal{X}_n$ can be written as $g(z) = \frac{1}{M(z)}\sum_{k=1}^{n+n_l}\beta_k z^{-k}$ where only n of the $n+n_l$ parameters β_k can be assigned arbitrarily. To approximate the function $f \in \mathcal{X}_n^+$ with a function $g \in \mathcal{X}_n$ the first n parameters can be chosen as $\beta_k = \alpha_k$ and the difference between the functions is bounded by

$$\begin{aligned} |f(z)-g(z)| &\le \frac{1}{|M(z)|}\left|\sum_{k=n+1}^{n+n_l}(\alpha_k-\beta_k)z^{-k}\right| \\ &\le \frac{|z|^{-n}}{|M(z)|}\left|\sum_{k=1}^{n_l}(\alpha_{n+k}-\beta_{n+k})z^{-k}\right| \le c|z|^{-n} \end{aligned} \tag{55}$$

for some constant c. Lemma 6.2 can now be applied and we get (for any n)

$$\sqrt{\varphi_n^+(z,z)} \le \sqrt{\varphi_n(z,z)} + c|z|^{-n} + \sqrt{c|z|^{-n}}. \tag{56}$$

Let z be a non-minimum phase zero, i.e $|z|>1$, and let the model order n go to infinity. This means that $\lim_{n\to\infty} c|z|^{-n} = 0$ and that gives

$$\lim_{n\to\infty}\varphi_n^+(z,z) \le \lim_{n\to\infty}\varphi_n(z,z). \tag{57}$$

Together with Lemma 6.1 which states that

$$\lim_{n\to\infty}\varphi_n(z,z) \le \lim_{n\to\infty}\varphi_n^+(z,z) \tag{58}$$

we finally get that

$$\lim_{n\to\infty}\varphi_n^+(z,z) = \lim_{n\to\infty}\varphi_n(z,z). \tag{59}$$

Letting the model order n go to infinity in (41) gives that

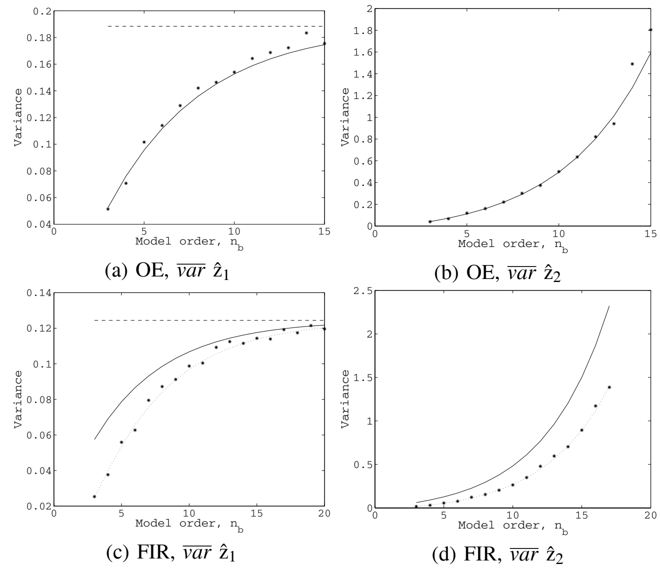
$$\lim_{n\to\infty}\varphi_n(z,z) = \frac{|z|^{-2}}{(1-|z|^{-2})} \tag{60}$$

The asymptotic expression for the variance of an estimated non-minimum phase zero (54) is achieved by letting $\mathcal{X}_n$ be the space induced by the Box-Jenkins structure, i.e by letting $M = A_\dagger \triangleq A_o^2 C_o F$ and $L = D_o E$ and then combining (29) and (60). ∎

VII. Simulations

In this section, the relevance of the variance expressions (43), (50) and (54) will be evaluated with Monte Carlo simulations. Two simulation experiments will be presented here: one example for which Theorem 5.1 is applicable to give an 'exact' expression of the variance and one where the upper bound in Theorem 6.1 must be used instead.

A system $G(q)$ with a *non-minimum phase* zero at $z_1 = -1.1$ and a *minimum phase* zero at $z_2 = -0.9$ is used to generate an input/output data record with 10,000 samples. The output $\{y_t\}$ is corrupted by white Gaussian noise with variance $\lambda_0 = 10^{-4}$ and the input $\{u_t\}$ is generated as filtered unit-variance Gaussian white noise. The system $G(q)$ and the input filter $Q(q)$ for the two experiments and the simulation results are described further in Sections VII-A and VII-B.



(a) OE, $\overline{var}\ \hat{z}_1$ (b) OE, $\overline{var}\ \hat{z}_2$ (c) FIR, $\overline{var}\ \hat{z}_1$ (d) FIR, $\overline{var}\ \hat{z}_2$

Fig. 1. Simulation results that show the variability of zero estimates. Fig.1a-b shows the simulation results for the OE system and Fig.1c-d for the FIR system. The solid lines show the 'true' variance from (43) in Theorem 5.1 for the OE system and the upper bound from (50) in Theorem 6.1 for the FIR system. The dotted lines in Fig.1c-d show the 'true' variance calculated from (15). For the non-minimum phase zero, z_1, the asymptotic variance (54) in Theorem 6.2 is indicated with a dashed line. The dots show the sample variances from the simulated experiments.

The observed data is used to estimate a model with the method described in Section II and the corresponding zero estimates, $\hat{z}_1$ and $\hat{z}_2$, are calculated. The sample variance of 1,000 identification experiments, with different noise and input realizations, is used as an estimate of $\overline{var}\ \hat{z}_1$ and $\overline{var}\ \hat{z}_2$ and they are plotted as dots in Fig.1. The experiment is repeated for different model orders n_b.

A. Simulation of an output error system

The simulated output error (OE) system is given by $G(q) = \frac{q^{-1}+2q^{-2}+0.99q^{-3}}{1-0.8q^{-1}+0.25q^{-2}}$ and the input filter is $Q(q) = \frac{1}{1-0.6q^{-1}}$. The sample variance of the estimated zeros are plotted as dots in Fig.1a-b and the 'true' variance (43) from Theorem 5.1 is shown as a solid line. The agreement between the simulated variance and the theoretic variance is very good in these simulations, but if more noise is added or the model order is increased further, the approximation (15) is less accurate, which results in an underestimation of the variance. See also [14] for more comments on the accuracy of (15).

For the non-minimum phase zero, z_1, the asymptotic variance expression (54) is plotted as a dashed line. The simulations show that, for moderate model orders, the 'exact' expression (43) is significantly more accurate than the asymptotic expression.

Remark: To ensure that the model is globally identifiable, the model order n_a is always equal to the order of true system and only the model order n_b is allowed to vary.

B. Simulation of a FIR system

The simulated FIR system is given by $G(q) = q^{-1} + 2q^{-2} + 0.99q^{-3}$ and the input filter is $Q(q) = \frac{1+0.4q^{-1}}{1-0.6q^{-1}}$. In this example, Theorem 5.1 can not be applied to get an expression for the variance, but the variance expression (15) is calculated and plotted as dotted lines in Fig.1c-d. The upper bounds from Theorem 6.1 are plotted as solid lines and the asymptotic variance for the non-minimum phase zero, from Theorem 6.2, is indicated with a dashed line. The sample variance from the simulations are plotted as dots.

The simulations show that the variance of an estimated non-minimum phase zero may be approximated with the upper bound given in Theorem 6.1. The approximation gets better for higher model orders, but even for low model orders it is significantly more accurate than the asymptotic variance from Theorem 6.2. For minimum phase zeros, the upper bound is perhaps not a very good approximation of the true variance, but at least it gives an indication of the *magnitude* of the variability.

VIII. Conclusions

The main theme of this paper was to show how variance expressions for estimates of zeros of dynamic systems can be derived by finding a reproducing kernel of a specific space that depends on the model structure. The method is applicable to model structures that have an independently parameterized noise model, e.g the output-error and Box-Jenkins model structures.

Variance expressions that are exact for arbitrarily small model orders have been derived for certain model structures and input excitations. In addition to that, we have also derived an expression for an upper bound on the variance. This expression is valid for general models and inputs.

The asymptotic variance expression for non-minimum phase zeros, first presented in [9], has also been proved by using reproducing kernels.

References

[1] L. Ljung and Z. Yuan, "Asymptotic properties of black-box identification of transfer functions," *IEEE Trans. Automatic Control*, vol. 30, no. 6, pp. 514–530, 1985.

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

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

[4] L.-L. Xie and L. Ljung, "Asymptotic variance expressions for estimated frequency functions," *IEEE Trans. Automatic Control*, vol. 46, no. 12, pp. 1887–1899, 2001.

[5] 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, 2004.

[6] ——, "Accuracy of closed loop estimates revisited," in *IEEE Conference on Decision and Control*. Maui, Hawaii, USA: IEEE, 2003.

[7] K. Lindqvist, "On experiment design in identification of smooth linear systems," Licentiate thesis, Royal Institute of Technology, Stockholm, Sweden, 2001.

[8] H. Hjalmarsson and K. Lindqvist, "Identification of performance limitations in control using arx-models," in *Proceedings of The 15th IFAC World Congress*, 2002.

[9] J. Mårtensson and H. Hjalmarsson, "Identification of performance limitations in control using general siso-models," in *13th IFAC Symposium on System Identification*, 2003.

[10] ——, "Closed loop identification of unstable poles and non-minimum phase zeros," in *Proceedings of the 16th IFAC World Congress on Automatic Control*, July 2005.

[11] L. Ljung, *System Identification: Theory for the user*, 2nd ed. Prentice Hall, 1999.

[12] B. Ninness and F. Gustafsson, "A unifying construction of orthonormal bases for system identification," *IEEE Transactions on Automatic Control*, vol. 42, pp. 515–521, April 1997.

[13] B. Ninness, H. Hjalmarsson, and F. Gustafsson, "Generalized fourier and toeplitz results for rational orthonormal bases," *SIAM Journal on Control and Optimization*, vol. 37, no. 2, pp. 429–460, 1998.

[14] R. Vuerinckx, R. Pintelon, J. Schoukens, and Y. Rolain, "Obtaining accurate confidence regions for the estimated zeros and poles in system identification problems," *IEEE Trans. on Automatic Control*, vol. 46, no. 4, pp. 656–659, April 2001.

Appendix
Proof of Lemma 6.2

Proof: Let $\|\cdot\|$ be the norm induced by the inner product $\langle\cdot,\cdot\rangle$, i.e $\|\cdot\| \triangleq \sqrt{\langle\cdot,\cdot\rangle}$. Since $\varphi(z,\mu) \in \mathcal{X}$ there is a family of functions $g_\mu \in \widetilde{\mathcal{X}}$ such that

$$|\varphi(z,\mu) - g_\mu(z)| < \delta, \quad \forall z,\mu \tag{61}$$

and we can write

$$\varphi(z,\mu) = g_\mu(z) + \Delta_\mu(z), \quad |\Delta_\mu(z)| < \delta, \quad \forall z,\mu. \tag{62}$$

Now by using the Cauchy-Schwarz inequality and the fact that $\langle\varphi(\cdot,\mu),\varphi(\cdot,z)\rangle = \varphi(z,\mu)$ we get

$$\begin{aligned}
|\varphi(z,\mu)| &\le |g_\mu(z)| + \delta && (63)\\
&= |\langle g_\mu(\cdot), \widetilde{\varphi}(\cdot,z)\rangle| + \delta && (64)\\
&\le \|g_\mu(\cdot)\|\|\widetilde{\varphi}(\cdot,z)\| + \delta && (65)\\
&= \|g_\mu(\cdot)\|\sqrt{\widetilde{\varphi}(z,z)} + \delta && (66)\\
&= \|\varphi(\cdot,\mu) - \Delta_\mu(\cdot)\|\sqrt{\widetilde{\varphi}(z,z)} + \delta && (67)\\
&\le \left(\|\varphi(\cdot,\mu)\| + \|\Delta_\mu(\cdot)\|\right)\sqrt{\widetilde{\varphi}(z,z)} + \delta && (68)\\
&\le \left(\|\varphi(\cdot,\mu)\| + \delta\right)\sqrt{\widetilde{\varphi}(z,z)} + \delta && (69)\\
&= \left(\sqrt{\varphi(\mu,\mu)} + \delta\right)\sqrt{\widetilde{\varphi}(z,z)} + \delta. && (70)
\end{aligned}$$

For $\mu = z$ it holds that $\varphi(z,z) = |\varphi(z,z)|$ and the inequality above can be written as

$$\varphi(z,z) - \sqrt{\varphi(z,z)\widetilde{\varphi}(z,z)} + \frac{\widetilde{\varphi}(z,z)}{4} \le \delta\sqrt{\widetilde{\varphi}(z,z)} + \frac{\widetilde{\varphi}(z,z)}{4} + \delta \tag{71}$$

which is the same as

$$\begin{aligned}
\left(\sqrt{\varphi(z,z)} - \frac{1}{2}\sqrt{\widetilde{\varphi}(z,z)}\right)^2 &\le \left(\frac{1}{2}\sqrt{\widetilde{\varphi}(z,z)} + \delta\right)^2 + \delta - \delta^2\\
&\le \left(\frac{1}{2}\sqrt{\widetilde{\varphi}(z,z)} + \delta + \sqrt{\delta}\right)^2.
\end{aligned} \tag{72}$$

Now this implies that

$$\sqrt{\varphi(z,z)} \le \sqrt{\widetilde{\varphi}(z,z)} + \delta + \sqrt{\delta} \tag{73}$$

which concludes the proof. ■