15th Triennial World Congress, Barcelona, Spain

BIAS AND VARIANCE OF THE PARAMETER ESTIMATES FOR A ONE DIMENSIONAL HEAT DIFFUSION SYSTEM

Bharath Bhikkaji and **Torsten Söderström**

Department of Systems and Control,
Information Technology, Uppsala University,
P O Box 27, SE 75121, Uppsala, Sweden.
Email: ts@SysCon.uu.se, bb@SysCon.uu.se

Abstract: In the recent past, we considered some parameter estimation problems, arising from a one dimensional heat diffusion system. In such problems one normally uses finite order models to approximate the infinite order heat diffusion system, and estimate the parameters involved. Therefore, all the parameter estimates are bound to be biased, even if one considers the availability of infinite data sequences. In this paper we perform an analysis of the bias as well as the variance of the parameter estimates, obtained by using certain approximations. We derive some expressions, which we support with simulations.

Keywords: Diffusion, System Identification, Bias, Variance, Model approximation.

1. INTRODUCTION

In the recent past, we did some work in estimating the parameters of a one dimensional heat diffusion system, see (Bhikkaji, 2000), (Bhikkaji and Söderström, 2001) and (Remle, 2000). In all the above mentioned references, we considered a one dimensional heat diffusion system, which was modeled by a linear PDE involving some unknown parameters. To simulate this one dimensional heat diffusion system, one has to estimate these unknown parameters. Since the model is of infinite order, one has to first approximate it by a finite order model and then to estimate the parameters by using the finite order approximation. In this paper we do not go into the details of constructing finite order approximations, as they had already been done in the above mentioned references, but perform an analysis of the bias and the variance of parameter estimates obtained by using the approximate models suggested in the above mentioned references. It must be mentioned that similar analysis has been done for many other problems in the past, see (D. Kundu and A. Mitra, 1996), (R. Pintelon and J. Schoukens, 2001) and (Söderström and Stoica, 1989), to name a few.

In section 2 of this paper we define the problem and our goal, along with a general framework. In section 3 we derive some bounds for the bias and the variance in the parameter estimates.

2. A GENERAL FRAMEWORK

In (Bhikkaji, 2000), (Bhikkaji and Söderström, 2001) and (Remle, 2000), we have considered a simple case of a one dimensional heat diffusion across a homogeneous wall. The dynamics is described by the PDE,

$$\frac{\partial T(x,t)}{\partial t} = \alpha_0 \frac{\partial^2 T(x,t)}{\partial x^2} \tag{1}$$

with boundary constraints

$$q_i(t) = -\kappa_0 \frac{\partial T(x,t)}{\partial x} \mid_{x=0}, \tag{2}$$
$$T_e(t) = T(d,t), \tag{3}$$

where $T(x,t)$ is the temperature profile across the wall, which is of length d units, and α_0 and κ_0 are the parameters which characterise the wall. The problem considered in the above mentioned references is as follows. Given the boundary data $\{q_i(kh)\ ,\ T_e(kh)\}_{k=0}^{N}$ and the corresponding internal temperature, $\{T(0,kh)\}_{k=0}^{N}$, for a known sampling interval of h units, estimate the parameters α_0 and κ_0.

To estimate the parameter vector

$$\theta_0 = (\alpha_0\ ,\ \kappa_0)^T, \tag{4}$$

we converted the above problem into a standard system identification problem, by first transforming the PDE model, (1), (2) and (3), into a standard LTI system of the form,

$$y(t) = G^{\infty}(s,\theta_0)u(t) + v(t), \tag{5}$$

where $G^{\infty}(s,\theta_0)u(t)$ is a notation used, for convenience, to denote the time domain output of the LTI system with a transfer function $G^{\infty}(s,\theta_0)$ and input $u(t)$, $y(t)$ is the net system output, i.e, the output of the LTI system, $T(0,t)$, plus the noise $v(t)$ and

$$u(t) = [q_i(t)\ ,T_e(t)]^T, \tag{6}$$

is the input. Further, $v(t)$ denotes a Gaussian white noise sequence with mean zero and variance σ^2. Note that, the symbol s in the transfer function $G^{\infty}(s,\theta_0)$ denotes the Laplace transform of the operator $\frac{d}{dt}$. The parameter vector θ_0 is then estimated by using the standard least squares technique,

$$\hat{\theta} = \min_{\theta} \frac{1}{N} \sum_{k=1}^{N} (y(kh) - y^m(kh))^2, \tag{7}$$

where the model output $y^m(kh)$ is

$$y^m(t) = G^{\infty}(s,\theta)u(t) \tag{8}$$

sampled at the time instants $t = h, 2h, \ldots, Nh$, and $\hat{\theta}$ denotes the estimate of the parameter vector θ_0.

In the above mentioned references, we have shown that the dynamics (1) - (3) can be converted to the transfer function

$$G^{\infty}(s,\theta_0) = \left[\frac{\tanh(d\sqrt{\frac{s}{\alpha_0}})}{\kappa_0\sqrt{\frac{s}{\alpha_0}}}\ ,\ \frac{1}{\cosh(d\sqrt{\frac{s}{\alpha_0}})} \right] \tag{9}$$

which is a transcendental function (and an infinite order transfer function). To numerically compute the parameter estimate, $\hat{\theta}$ in (7), one has to approximate the infinite order transfer function $G^{\infty}(s,\theta)$ by a finite order transfer function, when computing the model output $y^m(t)$, (8).

Therefore, in the earlier references we constructed some finite order approximations of (5), by using certain standard numerical PDE solvers like Finite-Difference scheme in (Remle, 2000), Chebyshev-Galerkein, Chebyshev-Tau, Chebyshev-Collocation and Chebyshev Interpolation, in (Bhikkaji, 2000). These approximate models were typically of the form

$$y^m(t) = \hat{G^n}(s,\theta)u(t), \tag{10}$$

where

$$\hat{G^n}(s,\theta) = [G_1^n(s,\theta)\ ,\ G_2^n(s,\theta)], \tag{11}$$

with

$$G_1^n(s,\theta) = \frac{1}{\kappa}D_1 + C(sI - \alpha A)^{-1}\frac{1}{\kappa}\alpha B_1 \tag{12}$$

and

$$G_2^n(s,\theta) = D_2 + C(sI - \alpha A)^{-1}\alpha B_2, \tag{13}$$

$\theta = (\alpha\ ,\ \kappa)^T$ and $u(t)$ as defined in (6). In (12) and (13) A is an $n \times n$ matrix, B_1 and B_2 are $n \times 1$ column vectors, C is a $1 \times n$ row vector and D_1 and D_2 are scalars. For the sake of notational convenience we rewrite (12) and (13) in a more compact form,

$$\hat{G^n}(s,\theta) = D + C(sI - \alpha A)^{-1}\alpha B, \tag{14}$$

where

$$B = [\frac{1}{\kappa}B_1\ ,B_2], \tag{15}$$
$$D = [\frac{1}{\kappa}D_1\ ,D_2]. \tag{16}$$

Note that, the scaling done in the first column of B and D, by a factor $\frac{1}{\kappa}$, is consistent with the scaling done in the first column of $G^{\infty}(s,\theta_0)$ by κ_0.

From the results we obtained in (Bhikkaji, 2000), (Bhikkaji and Söderström, 2001) and (Remle, 2000), we assume the following

(1) The convergence

$$\hat{G^n}(s,\theta) \to G^{\infty}(s,\theta), \tag{17}$$

is uniform in both the variables s and θ, as $n \to \infty$.

(2) If

$$\hat{G^n}(s,\theta_1) = \hat{G^n}(s,\theta_2) \quad \forall s \tag{18}$$

then

$$\theta_1 = \theta_2, \tag{19}$$

and the same holds for $G^{\infty}(s,\theta)$ also.

(3) For any given s and θ

$$\| G^{\infty}(s,\theta) - \hat{G^n}(s,\theta) \| \leq \frac{M_1}{n^k}, \quad (20)$$

where M_1 is a positive constant and k is an integer ≥ 1. The value of k depends on the choice of approximation.

(4) For any given s and θ

$$\frac{\partial \hat{G^n}(s,\theta)}{\partial \theta} \rightarrow \frac{\partial G^{\infty}(s,\theta)}{\partial \theta}. \quad (21)$$

It must be mentioned here that, the (17) and (20) have not been proved for the all approximate models constructed in the above mentioned references. In most cases they are justified by using simulations. Nevertheless we accept them here. Using (17) and (20) one can prove (21).

Since we use the finite order approximation $\hat{G^n}(s,\theta)$, (14), to estimate the parameters, the parameter estimate $\hat{\theta}$, (7), is replaced by

$$\hat{\theta}_N^n \triangleq \arg\min_{\theta} V_N^n(\theta), \quad (22)$$

where

$$V_N^n(\theta) = \frac{1}{N}\sum_{k=1}^{N}(y(kh) - y^m(kh))^2, \quad (23)$$

and $\{y^m(kh)\}_{k=1}^N$ is $y^m(t)$, (10), sampled at $t = h, 2h, \ldots, Nh$. Note that, from here on for technical ease, we refer to the time samples $y(kh)$ and $y^m(kh)$ by $G^{\infty}(s,\theta)u(k)$ and $\hat{G^n}(s,\theta)u(k)$, respectively.

Here, our goal is to derive bounds on the bias (for a large n) in the parameter estimates and also obtain an expression for the variance (for a large N) of the parameter estimates. In the ideal case where both $n \rightarrow \infty$ and $N \rightarrow \infty$, it can be easily shown the parameter estimate that

$$\hat{\theta}_{\infty}^{\infty} = \theta_0, \quad (24)$$

provided the input $u(t)$ is second order persistently exciting, see (Söderström and Stoica, 1989) for details.

3. BIAS ANALYSIS

In this section, we do an analysis of the bias in the parameter estimates, (22), as $N \rightarrow \infty$. It must be mentioned here that, a similar analysis of the bias in the parameter estimates has been done in (Remle, 2000).

Let

$$V_{\infty}^n(\theta) \triangleq \lim_{N\rightarrow\infty} V_N^n(\theta), \quad (25)$$

where $V_N^n(\theta)$ is as defined in (23), and define the asymptotic parameter estimate as

$$\hat{\theta}_{\infty}^n \triangleq \arg\min_{\theta} V_{\infty}^n(\theta). \quad (26)$$

A standard argument, see (Söderström and Stoica, 1989), shows that

$$V_{\infty}^n(\theta) = E\{(y(k) - \hat{G^n}(s,\theta)u(k))^2\}. \quad (27)$$

Note that

$$\begin{aligned} V_{\infty}^n(\theta) &= E\{([G^{\infty}(s,\theta_0) - \hat{G^n}(s,\theta)]u(k))^2\} + \sigma^2, \\ &\triangleq W(\theta) + \sigma^2. \end{aligned} \quad (28)$$

Since

$$\lim_{n\rightarrow\infty} \hat{G^n}(s,\theta) \rightarrow G^{\infty}(s,\theta), \quad (29)$$

uniformly in θ, one can see that

$$\lim_{n\rightarrow\infty} V_{\infty}^n(\theta) \rightarrow V_{\infty}^{\infty}(\theta) \quad (30)$$

uniformly in θ. This implies that

$$\lim_{n\rightarrow\infty} \min_{\theta} V_{\infty}^n(\theta) \rightarrow \min_{\theta} V_{\infty}^{\infty}(\theta). \quad (31)$$

Since θ_0 is a unique minimum of $V_{\infty}^{\infty}(\theta)$,

$$\lim_{n\rightarrow\infty} \hat{\theta}_{\infty}^n \rightarrow \theta_0. \quad (32)$$

Thus the convergence of the parameter estimates $\hat{\theta}_{\infty}^n$ to θ_0 has been rather easily established. To obtain a bound on the bias present in the parameter estimates, we do the following.

Since $\hat{G^n}(s,\theta)$ is of the form (14), it is infinitely differentiable with respect to the parameter θ and hence, $V_{\infty}^n(\theta)$ is also infinitely differentiable. Therefore by applying the mean value theorem on $V_{\infty}^{n\,\prime}(\theta)$ in a closed and bounded domain $\mathcal{D}$ containing both $\hat{\theta}_{\infty}^n$ and θ_0, we have

$$V_{\infty}^{n\,\prime}(\hat{\theta}_{\infty}^n) = V_{\infty}^{n\,\prime}(\theta_0) + (\hat{\theta}_{\infty}^n - \theta_0)^T V_{\infty}^{n\,\prime\prime}(\theta^*) (33)$$

where θ^* is a unique element contained in $\mathcal{D}$, and is a convex combination of θ_0 and $\hat{\theta}_{\infty}^n$.

Note that, from (32), for an n "large enough" $\hat{\theta}_{\infty}^n \approx \theta_0$, and so $\theta^* \approx \theta_0$. Therefore, since $V_{\infty}^{n\,\prime\prime}(\theta)$ varies continuously with respect to θ, we have

$$V_{\infty}^{n\,\prime\prime}(\theta^*) \approx V_{\infty}^{n\,\prime\prime}(\theta_0). \quad (34)$$

Note that by definition, (26),

$$V_{\infty}^{n\,\prime}(\hat{\theta}_{\infty}^n) = 0. \quad (35)$$

Hence using (34) and (35) in (33), we have

$$(\hat{\theta}_{\infty}^n - \theta_0) = -[V_{\infty}^{n\,\prime\prime}(\theta_0)]^{-1}\left(V_{\infty}^{n\,\prime}(\theta_0)\right)^T. \quad (36)$$

Therefore to obtain an estimate for the bias (36), one has to compute the first and the second derivatives of $V_{\infty}^n(\theta_0)$.

Note that from (28),

$$V_\infty^{n\,\prime}(\theta_0) = W'(\theta_0) \quad (37)$$

$$V_\infty^{n\,\prime\prime}(\theta_0) = W''(\theta_0). \quad (38)$$

Further, from (28) we have

$$W'(\theta_0) = E\{2(-\frac{\partial \hat{G^n}(s,\theta_0)}{\partial \theta}u(k)) \times ((G^\infty(s,\theta_0) - \hat{G^n}(s,\theta_0))u(k))\} \quad (39)$$

The Hessian (38) can be evaluated as follows:

$$W''(\theta_0) = 2E\{(-\frac{\partial \hat{G^n}(s,\theta_0)}{\partial \theta}u(k)) \times (-\frac{\partial \hat{G^n}(s,\theta_0)}{\partial \theta}u(k))^T\} + 2E\{\Big((G^\infty(s,\theta_0) - \hat{G^n}(s,\theta_0))u(k)\Big) \times (-\frac{\partial^2 \hat{G^n}}{\partial \theta^2}u(k))\}. \quad (40)$$

If one chooses a large n, the second term of (40) is much smaller than the first term. Hence, for a large n, (40) can be approximated by

$$W''(\theta_0) \approx 2E\{(-\frac{\partial \hat{G^n}(s,\theta_0)}{\partial \theta}u(k)) \times (-\frac{\partial \hat{G^n}(s,\theta_0)}{\partial \theta}u(k))^T\}, \triangleq 2\Sigma_n \quad (41)$$

which is a 2×2 positive definite matrix.

Note that

$$\frac{\partial \hat{G^n}(s,\theta_0)}{\partial \theta}u(k) = [\frac{\partial \hat{G^n}(s,\theta_0)}{\partial \alpha}u(k)\ , \frac{\partial \hat{G^n}(s,\theta_0)}{\partial \kappa}u(k)]^T. \quad (42)$$

Let

$$X(k) = \frac{\partial \hat{G^n}(s,\theta_0)}{\partial \alpha}u(k), \quad (43)$$

$$Y(k) = \frac{\partial \hat{G^n}(s,\theta_0)}{\partial \kappa}u(k) \quad (44)$$

and

$$Z(k) = (G^\infty(s,\theta_0) - \hat{G^n}(s,\theta_0))u(k). \quad (45)$$

Hence, from (39), we have

$$W'(\theta_0) = E\{[X(k)\ , Y(k)]^T Z(k)\}. \quad (46)$$

Since $Z(k)$, (45), is a scalar, by using the triangle inequality on (46) we have

$$\| W'(\theta_0) \| \leq | E\{X(k)Z(k)\} | + | E\{Y(k)Z(k)\} | . \quad (47)$$

By Cauchy-Schwartz inequality,

$$| E\{X(k)Z(k)\} | \leq \sqrt{E\{| X(k) |^2\}} \times \sqrt{E\{| Z(k) |^2\}} \quad (48)$$

and

$$| E\{Y(k)Z(k)\} | \leq \sqrt{E\{| Y(k) |^2\}} \times \sqrt{E\{| Z(k) |^2\}}. \quad (49)$$

Hence, form (47), (48) and (49), we have

$$\| W'(\theta_0) \| \leq 2(\sqrt{E\{| X(k) |^2\}} + \sqrt{E\{| Y(k) |^2\}}) \times \sqrt{E\{| Z(k) |^2\}}. \quad (50)$$

Note that the expected values $E\{X(k)Z(k)\}$, $E\{Y(k)Z(k)\}$, $E\{| X(k) |^2\}$ and $E\{| Y(k) |^2\}$ exist, since $\frac{\partial \hat{G^n}(s,\theta_0)}{\partial \theta}$ is a stable transfer function and $u(t)$ is a bounded input.

Note that, from (41),

$$\| [V_\infty^{n\,\prime\prime}(\theta_0)]^{-1} \| \leq \frac{1}{2} \| \Sigma_n^{-1} \| . \quad (51)$$

Since, from (36),

$$\| \theta_0 - \hat{\theta}_\infty^n \| \leq \| [V_\infty^{n\,\prime\prime}(\theta_0)]^{-1} \| \| V_\infty^{n\,\prime}(\theta_0) \| \quad (52)$$

by using (50) and (51) in (52) we have

$$\| \theta_0 - \hat{\theta}_\infty^n \| \leq \| \Sigma_n^{-1} \| (\sqrt{E\{| X |^2\}} + \sqrt{E\{| Y |^2\}}) \times \sqrt{E\{| Z |^2\}}. \quad (53)$$

To obtain an expression for the rate of convergence, we estimate $E\{| Z(k) |^2\}$. Note that $E\{| Z(k) |^2\}$ is a sampled version of

$$E\{\Big((G^\infty(s,\theta_0) - \hat{G^n}(s,\theta_0))u(t)\Big)^2\}. \quad (54)$$

Hence any bound that holds for (54) holds for $E\{| Z(k) |^2\}$ also. Note that

$$E\{\Big((G^\infty(s,\theta_0) - \hat{G^n}(s,\theta_0))u(t)\Big)^2\} = \int_{-\infty}^{\infty} (G^\infty(i\omega,\theta_0) - \hat{G^n}(i\omega,\theta_0))\Phi_u(\omega) \times (G^\infty(-i\omega,\theta_0) - \hat{G^n}(-i\omega,\theta_0))^T d\omega, \quad (55)$$

where $\Phi_u(\omega)$ is the power spectrum of the input $u(t)$, see (Söderström and Stoica, 1989) for a definition of power spectrum. Therefore

$$E\{| Z |^2\} \leq \int_{-\infty}^{\infty} \| G^\infty(i\omega,\theta_0) - \hat{G^n}(i\omega,\theta_0) \|^2 \times \| \Phi_u(\omega) \| d\omega, \leq \left(\frac{M_1}{n^k}\right)^2 \int_{-\infty}^{\infty} \| \Phi_u(\omega) \| d\omega. \quad (56)$$

Note that, the above inequality is obtained by using the assumption (20). Let

$$m_1 = \int_{-\infty}^{\infty} \| \Phi_u(\omega) \| d\omega < \infty. \quad (57)$$

Note that, if the input $u(t)$ is a zero mean stationary random process, then its variance is given by

$$E\{u^2(t)\} = \int_{-\infty}^{\infty} \Phi_u(\omega)d\omega$$
$$\leq m_1. \quad (58)$$

Hence m_1, in some sense, can be interpreted as a measure of the variance of the input $u(t)$. Note that, from (53),

$$\| \theta_0 - \hat{\theta}_\infty^n \| \leq \frac{K}{n^k}, \quad (59)$$

where K is a constant,

$$K = \| \Sigma_n^{-1} \| (\sqrt{E\{| X |^2\}} + \sqrt{E\{| Y |^2\}}) \times M_1\sqrt{m_1}. \quad (60)$$

In summary, the parameter estimates $\hat{\theta}_\infty^n$ converge to θ_0 at the same rate $O(\frac{1}{n^k})$ as the approximate transfer function $\hat{G}^n(s,\theta)$ approaches $\hat{G}^\infty(s,\theta)$, see (59) and (20).

4. VARIANCE ANALYSIS

In the last section, we have done an analysis on the asymptotic bias present in the parameters due to the use of a finite order model. Here, we look into the asymptotic variance of the parameters from the true value θ_0.

Note that the estimation error $(\hat{\theta}_N^n - \theta_0)$ can be written as

$$\hat{\theta}_N^n - \theta_0 = (\hat{\theta}_N^n - \hat{\theta}_\infty^n) + (\hat{\theta}_\infty^n - \theta_0). \quad (61)$$

The first term in the right hand side of (61) is referred to as the variance contribution, and the second term as the bias contribution. We have dealt with the bias contribution in the previous section. Here we consider the variance contribution $(\hat{\theta}_N^n - \hat{\theta}_\infty^n)$.

As before, see (33), by using the mean value theorem and the fact that $V_N^{n\,\prime}(\hat{\theta}_N^n) = 0$, we have

$$(\hat{\theta}_N^n - \hat{\theta}_\infty^n) = -[V_N^{n\,\prime\prime}(\theta^*)]^{-1} \left(V_N^{n\,\prime}(\hat{\theta}_\infty^n)\right)^T. (62)$$

where θ^* lies on the straight line between $\hat{\theta}_N^n$ and $\hat{\theta}_\infty^n$. If the data length N is large enough, then $\hat{\theta}_N^n \approx \hat{\theta}_\infty^n$, which in turn implies that $\theta^* \approx \hat{\theta}_\infty^n$. And since $\hat{\theta}_\infty^n \approx \theta_0$, we have

$$(\hat{\theta}_N^n - \hat{\theta}_\infty^n) \approx -[V_N^{n\,\prime\prime}(\theta_0)]^{-1} \left(V_N^{n\,\prime}(\theta_0)\right)^T. (63)$$

Therefore, to quantify $(\hat{\theta}_N^n - \hat{\theta}_\infty^n)$, one has the evaluate the first and the second derivatives of $V_N^n(\theta)$ at θ_0.

Note that, from (23), we have

$$V_N^{n\,\prime}(\hat{\theta}_N^n) = \frac{2}{N}\sum_{k=1}^{N}[(y(k) - \hat{G}^n(s,\theta_0))u(k)] \times[-\frac{\partial \hat{G}^n(s,\theta_0)}{\partial\theta}u(k)]$$
$$= \frac{2}{N}\sum_{k=1}^{N} v(k)[-\frac{\partial \hat{G}^n(s,\theta_0)}{\partial\theta}u(k)]^T$$
$$+\frac{2}{N}\sum_{k=1}^{N}((G^\infty(s,\theta_0) - \hat{G}^n(s,\theta_0))u(k)) \times[-\frac{\partial \hat{G}^n(s,\theta_0)}{\partial\theta}u(k)]. \quad (64)$$

If n is large enough, then the second term in the RHS of (64) can be neglected, i.e,

$$V_N^{n\,\prime}(\theta_0) \approx -\frac{2}{N}\sum_{k=1}^{N} v(k)(\frac{\partial \hat{G}^n(s,\theta_0)}{\partial\theta}u(k)). (65)$$

By using the central limit theorem, see (Söderström and Stoica, 1989), one can show that

$$\sqrt{N}V_N^{n\,\prime}(\theta_0) \to \mathcal{N}(0, 4\sigma^2\Sigma_n) \quad (66)$$

as $N \to \infty$, where Σ_n is as defined in (41).

From (23), one can see that

$$V_N^{n\,\prime\prime}(\theta_0) = 2\frac{1}{N}\sum_{k=1}^{N}\{(-\frac{\partial \hat{G}^n(s,\theta_0)}{\partial\theta}u(k))^T \times(-\frac{\partial \hat{G}^n(s,\theta_0)}{\partial\theta}u(k))\}$$
$$+2\frac{1}{N}\sum_{t=1}^{N}\{((G^\infty(s,\theta_0) - \hat{G}^n(s,\theta_0))u(t)) \times(-\frac{\partial^2 \hat{G}^n(s,\theta_0)}{\partial\theta^2}u(t))^T\}. \quad (67)$$

As before, by choosing a large n, we approximate (67) by

$$V_N^{n\,\prime\prime}(\theta_0) = 2\frac{1}{N}\sum_{t=1}^{N}\{(-\frac{\partial \hat{G}^n(s,\theta_0)}{\partial\theta}u(k)) \times(-\frac{\partial \hat{G}^n(s,\theta_0)}{\partial\theta}u(k))^T\}, \quad (68)$$

which implies that

$$\lim_{N\to\infty} V_N^{n\,\prime\prime}(\theta_0) = 2\Sigma_n. \quad (69)$$

Therefore

$$\sqrt{N}(\hat{\theta}_N^n - \hat{\theta}_\infty^n) \to \mathcal{N}(0, \sigma^2\Sigma_n^{-1}\Sigma_n\Sigma_n^{-1}) = \mathcal{N}(0, \sigma^2\Sigma_n^{-1}), \quad (70)$$

In summary, (70) implies that (for a large n) $\hat{\theta}_\infty^n$ is asymptotically (as $N \to \infty$) Gaussian distributed, with a covariance matrix that can be found from (70) and (41).

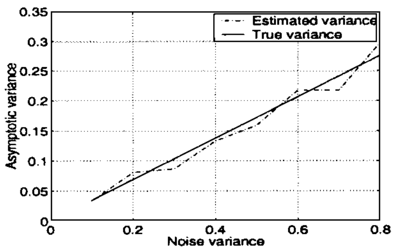

Fig. 1. Asymptotic variance of the random variable $\sqrt{N}(\hat{\alpha}_N^n - \alpha_0)$ along with its predicted variance.

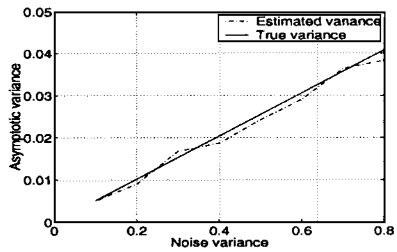

Fig. 2. Asymptotic variance of the random variable $\sqrt{N}(\hat{\kappa}_N^n - \kappa_0)$ along with its predicted variance.

5. NUMERICAL ILLUSTRATION

In this section, we present some numerical illustration of the variance result (70).

To make the simulations comprehensive, the following input sequence was generated for our identification experiment. First we generate two independent white noise sequences of variance one and pass them thorough the low pass filter

$$H(q^{-1}) = \frac{1}{1 - 1.88q^{-1} + 0.9732q^{-2}}, \qquad (71)$$

which is of finite bandwidth and has a resonant peak. The output of the filter is further added with white noise sequences of variance 0.2, so that the net signal is reasonably frequency rich within the filter band. The net signal is then chosen as the input $u(k)$.

In Figures 1 and 2 we have plotted the variance of the random variables

$$\sqrt{N}(\hat{\alpha}_N^n - \alpha_0) \qquad (72)$$

and

$$\sqrt{N}(\hat{\kappa}_N^n - \kappa_0), \qquad (73)$$

along with their respective predicted variances, for different values of the noise variance (or noise power, σ^2). Note that the predicted variances of (72) and (73) are the diagonal elements of the covariance matrix $\sigma^2\Sigma_n^{-1}$. The variance of the random variables (72) and (73) were estimated using a standard Monte Carlo technique with the number of noise realisations $N_r = 300$, the number of data points $N = 1000$ and the model order $n = 100$.

Note that the predicted variance is reasonably close to the estimated variance.

6. CONCLUDING DISCUSSION

We conclude this paper by briefly discussing some results presented in this paper. First, note that the bias, (59), converges at a rate, which is at least as fast as the convergence rate of the approximate transfer functions, (20).

In the case of the bias, (59), note that the constant K is finite iff either $\phi(\omega)$ has compact support or has an exponentially decaying tail. This is not surprising, since all the approximate models constructed in the above mentioned references converge to the D matrix, (16), where $| D |> 0$, as $\omega \rightarrow \infty$. Therefore the sum of the output-error is not finite, if the input has high frequency components. Hence, for this identification experiment it is always advisable to choose those inputs, for which the spectrum $\phi(\omega)$ diminishes beyond a particular band-width, where the approximation of the transfer-function deteriorates.

Note that the expression for the covariance $\sigma^2\Sigma_n^{-1}$, (70), is similar to the expression obtained in chapter 7 of (Söderström and Stoica, 1989). More importantly, note that the predicted covariance, $\sigma^2\Sigma_n^{-1}$, depends directly on the input. The larger the input spectrum, the larger the value of Σ_n, and hence the lower the predicted variance, $\sigma^2\Sigma_n^{-1}$.

7. REFERENCES

Bhikkaji, B. (2000). *Reduced Order Models for Diffusion Systems.* Licentiate thesis, Information Technology, Systems and Control. Uppsala University, Sweden.

Bhikkaji, B. and T. Söderström (2001). Reduced Order Models for Diffusion Systems. *International Journal of Control* **74**(15), 1543–1557.

D. Kundu and A. Mitra (1996). Asymptotic theory of Least Squares Estimates of a Non-Linear Time Series Regression Model. *Communications in Statistics: Theory and Methods* **25**(1), 133 - 141.

R. Pintelon and J. Schoukens (2001). *System Identification, A Frequency Domain Approach.* IEEE Press. New York.

Remle, S. (2000). *Modeling and Parameter Estimation of the Diffusion Equation.* Licentiate thesis, Information Technology, Systems and Control. Uppsala University, Sweden.

Söderström, T. and P. Stoica (1989). *System Identification.* Prentice Hall International. Hemel Hempstead, UK.