

**A STRONG APPROXIMATION THEOREM FOR  
ESTIMATOR PROCESSES IN CONTINUOUS TIME**

By

**László Gerencsér**

and

**Zsuzsanna Vágó**

**IMA Preprint Series # 698**

September 1990

**A STRONG APPROXIMATION THEOREM  
FOR  
ESTIMATOR PROCESSES IN CONTINUOUS TIME**

László Gerencsér\*, Zsuzsanna Vágó\*\*

Department of Electrical Engineering,  
McGill University,  
Montréal, Québec, Canada.

**Abstract**

We prove that the parameter estimation error of continuous time linear stochastic system can be written as the mean of a stochastic integral plus a residual term, the moments of which decay as  $T^{-1}$  where  $[0, T]$  is the observation period.

---

\* On leave from the Computer and Automation Institute of the Hungarian Academy of Sciences, Budapest.  
\*\* Technical University of Budapest, Hungary

## 1. Introductions and the main theorem

Let us consider a linear time-invariant stochastic system given by the state space equations

$$dx_t(\theta^*) = A(\theta^*)x_t(\theta^*)dt + K(\theta^*)dw_t \quad (1.1)$$

$$dy_t(\theta^*) = C(\theta^*)x_t(\theta^*)dt + dw_t \quad (1.2)$$

where  $dw_t$  is a standard Gaussian white noise process in  $\mathbb{R}^m$ . The process  $(y_t(\theta^*))$  is called the output process, while  $(x_t(\theta^*))$  is the state-vector. We need the following conditions:

**Condition 1.1** The matrices  $A(\theta^*), K(\theta^*), C(\theta^*)$  are defined for  $\theta^* \in D \subset \mathbb{R}^p$  where  $D$  is an open domain and they are  $C^\infty$  function of  $\theta^*$ . Moreover  $A(\theta^*)$  and  $A(\theta^*) - K(\theta^*)C(\theta^*)$  are stable for  $\theta^* \in D$ .

The last part of the condition means that  $A(\theta^*)$  and  $A(\theta^*) - K(\theta^*)C(\theta^*)$  have all their eigenvalues in the left half-plane. It follows that  $dy_t$  is stationary process the innovation process of which  $dw_t$ .

To estimate  $\theta^*$  on the basis of the observation process  $y_t$  we proceed as follows: fix a  $\theta \in D$  and invert the system (1.1), (1.2) assuming that  $\theta = \theta^*$  to get  $dw_t$ . The inverse system is given by the equations:

$$dx_t(\theta^*) = (A(\theta^*) - K(\theta^*)C(\theta^*))x_t(\theta^*)dt + K(\theta^*)dy_t(\theta^*) \quad (1.5)$$

$$dw_t = dy_t(\theta^*) - C(\theta^*)x_t(\theta^*)dt \quad (1.6)$$

Now if  $\theta$  is chosen arbitrarily we still can use the equation above to generate a process  $d\epsilon_t$  which is an estimation of  $dw_t$ . The governing equations are

$$dx_t(\theta, \theta^*) = (A(\theta) - K(\theta)C(\theta))x_t(\theta, \theta^*)dt + K(\theta)dy_t(\theta^*). \quad (1.7)$$

$$d\epsilon_t(\theta, \theta^*) = dy_t(\theta^*) - C(\theta)x_t(\theta, \theta^*)dt. \quad (1.8)$$

In practice we set zero initial state, i.e.  $x_0(\theta, \theta^*) = 0$ . However for analysis purposes we assume that  $(x_t(\theta, \theta^*))$  is the stationary solution of (1.7) (1.8). Since the difference between

the zero-initial state and the stationary solution is known to be  $O_M(e^{-\lambda t})$  with some  $\lambda > 0$  it is easy to check that all the results we derive using the latter process will hold if we use the former one.

Here we used the following convention: if  $\xi_n$  is a sequence of random variables and  $c_n$  is a sequence of positive numbers then we write  $\xi_n = O_M(c_n)$  if  $\xi_n/c_n = O_M(1)$ , which in turn means that the sequence  $\xi_n/c_n$  is  $M$ -bounded, as defined in the Appendix.

It is well known that the negative log-likelihood function can be written as

$$V_T(\theta, \theta^*) = \frac{1}{2} \int_0^T |C(\theta)x_t(\theta, \theta^*)|^2 dt - \int_0^T x_t^T(\theta, \theta^*)C^T(\theta)dw_t \quad (1.10)$$

(c.f e.g. Arató (1984)), and from this we get

$$\frac{\partial}{\partial \theta} V_T(\theta, \theta^*) = \int_0^T \frac{d}{dt} \frac{\partial}{\partial \theta} \epsilon_t(\theta, \theta^*) \cdot d\epsilon_t(\theta, \theta^*) \quad (1.11)$$

(c.f. e.g. Gerencsér, Gyöngy, Michaletzky (1984)). Here  $\frac{\partial}{\partial \theta}$  means differentiation both in the  $M$ -sense and almost surely (c.f. the Appendix).

Let  $D^o \subset D$  be a compact domain such that  $\theta^* \in \text{int}D^o$ . Then the maximum-likelihood estimator  $\hat{\theta}_T$  of  $\theta^*$  is defined as the solution of the equation

$$\frac{\partial}{\partial \theta} V_T(\theta, \theta^*) = 0 \quad (1.12)$$

if the solution is unique in  $D^o$ . Otherwise we define  $\hat{\theta}_T$  arbitrarily subject to the condition that  $\hat{\theta}_T \in D^o$  a.s. and  $\hat{\theta}_T$  must be a random variable.

Let us define the asymptotic cost function  $W(\theta, \theta^*)$

$$W(\theta, \theta^*) = \frac{1}{T} E V_T(\theta, \theta^*). \quad (1.13)$$

It is known and easily proved that

$$\frac{\partial}{\partial \theta} W(\theta, \theta^*)|_{\theta=\theta^*} = 0. \quad (1.14)$$

**Condition 1.2** We assume that the asymptotic log-likelihood equation (1.14) has a unique solution  $\theta = \theta^*$  in  $D$ . Moreover we assume that  $R^* = (\partial^2/\partial\theta^2)W(\theta, \theta^*) \Big|_{\theta=\theta^*}$  is positive definite.

This may seem a slightly restrictive condition, however without this we can not hope to get anything close to a consistent estimator. For discrete time ARMA-systems this condition was verified by Åström and Söderström (1974). For multivariate MA-processes Condition 1.2 was verified in Söderström and Stoica (1982). For both discrete and continuous-time systems a partial result was given in G. Vágó and Gerencsér (1985). To make the notations simpler we shall use  $\theta$  subscript to denote derivative w.r.t.  $\theta$ .

The main result of the paper is that the analysis of  $\widehat{\theta}_T - \theta^*$  can be reduced to that of a stochastic integral as described by the following theorem:

**Theorem 1.1.** Under Conditions 1.1.-1.2. we have

$$\widehat{\theta}_T - \theta^* = -(R^*)^{-1} \frac{1}{T} \int_0^T \dot{\epsilon}_{\theta t}(\theta^*, \theta^*) dw_t + O_M(T^{-1/2}).$$

This theorem is an extension of an earlier result of Gerencsér(1989b), and has many interesting corollaries. E.g. using this theorem and another result on the strong approximation of multidimensional integrals we derived the following result in Gerencsér(1989c):

**Theorem A..** Assume that the underlying probability space is sufficiently rich. Then under Conditions 1.1.-1.2. we have for every  $\epsilon > 0$

$$\widehat{\theta}_T - \theta^* = \frac{1}{T}(R^*)^{\frac{1}{2}} \overline{w}_T + O_M(T^{-3/5+\epsilon})$$

where  $(\overline{w}_T)$  is a standard Wiener-process in  $\mathbb{R}^p$ .

## 2. The proof of the theorem

A basic tool in the analysis of the maximum-likelihood estimator is the following theorem.

**Theorem 2.1.** We have

$$\sup_{\theta \in D^o} \left| \frac{1}{T} V_T(\theta, \theta^*) - W(\theta, \theta^*) \right| = O_M(T^{-\frac{1}{2}}) \quad (2.1)$$

and similar estimates hold for all derivatives of  $V_T(\theta, \theta^*)$ .

**Proof:** In the proofs we shall use the properties of  $L$ -mixing processes, derived in Gerencsér(1989a).

We shall summarize some of the more important properties in the Appendix.

Since  $y_t(\theta^*)$  is the output of a stable linear state-space system, it is  $L$ -mixing w.r.t.  $(\mathcal{F}_t, \mathcal{F}_t^+)$  by Theorem 3.2 where

$$\mathcal{F}_t = \sigma\{w_s : s \leq t\}, \quad \mathcal{F}_t^+ = \sigma\{w_s - w_{s'} : s, s' > t\}.$$

I.e. the  $\sigma$ -algebras  $(\mathcal{F}_t, \mathcal{F}_t^+)$  represent the past and the future of the Gaussian white noise process  $dw_t$ , respectively. Also since the system-matrices are smooth in  $\theta^*$  it follows that all derivatives of  $y_t(\theta^*)$  w.r.t.  $\theta^*$  are  $L$ -mixing. Furthermore it is easy to see that  $y_t(\theta^*)$  and its derivatives are uniformly  $L$ -mixing w.r.t.  $\theta^*$  when  $\theta^* \in D$ . Similarly  $x_t(\theta, \theta^*)$  and its derivatives w.r.t.  $\theta$  and  $\theta^*$  are uniformly  $L$ -mixing. Let us now introduce the notation

$$\delta V_T(\theta, \theta^*) = \frac{1}{T} V_T(\theta, \theta^*) - W(\theta, \theta^*)$$

which can also be written as

$$\begin{aligned} \delta V_T(\theta, \theta^*) &= \frac{1}{2T} \int_0^T (|C(\theta)x_t(\theta, \theta^*)|^2 - E|C(\theta)x_t(\theta, \theta^*)|^2) dt - \frac{1}{T} \int_0^T \dot{x}_t^T(\theta, \theta^*) C^T(\theta) dw_t \\ &\triangleq \frac{1}{2T} \int_0^T u_t(\theta, \theta^*) dt - \frac{1}{T} \int_0^T v_t(\theta, \theta^*) dw_t. \end{aligned}$$

Here  $u_t(\theta, \theta^*), v_t(\theta, \theta^*)$  are  $L$ -mixing processes uniformly in  $\theta, \theta^*$ , and the same holds true for all their derivatives w.r.t.  $\theta$  and  $\theta^*$ . Therefore we can apply Theorems 3.1.-3.3. to get that

$$\sup_{\theta \in D^o} \frac{1}{\sqrt{T}} \int_0^T u_t(\theta, \theta^*) dt = O_M(1), \quad \text{and} \quad \sup_{\theta \in D^o} \frac{1}{\sqrt{T}} \int_0^T v_t(\theta, \theta^*) dw_t = O_M(1)$$

which implies that

$$\sup_{\theta \in D^o} T^{1/2} \delta V_T(\theta, \theta^*) = O_M(1).$$

For the derivatives of  $\delta V_T(\theta, \theta^*)$  we can use similar arguments, thus the proof of Theorem 2.1. is complete.

**Lemma 2.2.** For any  $d > 0$  the equation  $V_{\theta T}(\theta, \theta^*) = 0$  has a unique solution in  $\{\theta : |\theta - \theta^*| \leq d\}$  with probability at least  $1 - O(T^{-s})$  for any  $s > 0$ .

**Proof:** We shall apply a simple analytic lemma which we present in the Appendix (Lemma 3.4). Choose

$$G(\theta) = W_{\theta}(\theta, \theta^*),$$

$$\delta G(\theta) = \delta V_{\theta}(\theta, \theta^*) \triangleq \frac{1}{T} V_{\theta}(\theta, \theta^*) - W_{\theta}(\theta, \theta^*),$$

and let us consider the set  $A_T$  defined by

$$A_T = \left\{ \omega : \sup_{\theta \in D^o} |\delta V_{\theta}(\theta, \theta^*)| < d' \quad \text{and} \quad \sup_{\theta \in D^o} |\delta V_{\theta\theta}(\theta, \theta^*)| < d'' \right\}$$

where  $d', d''$  are sufficiently small positive numbers. Then by Lemma 3.4. equation (1.12) has a unique solution  $\widehat{\theta}_T$  in  $D^o$  and  $|\widehat{\theta}_T - \theta^*| \leq d$ , where  $d$  is any fixed positive number. To estimate the probability of the event  $A_T$  we can apply Theorem 2.1. , from which we get using Chebishev's inequality that  $P(A_T^c) = O(T^{-s})$  for any  $s > 0$ .

**Proof of Theorem 1.1.:** Let us now consider equation (1.12) and write it as

$$0 = V_{\theta T}(\widehat{\theta}_T, \theta^*) = V_{\theta T}(\theta^*, \theta^*) + \overline{V}_{\theta\theta T}(\widehat{\theta}_T - \theta^*) \quad (2.2)$$

where

$$\overline{V}_{\theta\theta T} = \int_0^1 V_{\theta\theta T}((1 - \lambda)\theta^* + \lambda\widehat{\theta}_T, \theta^*) d\lambda$$

After a simple rearrangement we get from (2.2):

$$R^*(\widehat{\theta}_T - \theta^*) = -\frac{1}{T} V_{\theta T}(\theta^*, \theta^*) + (R^* - \frac{1}{T} \overline{V}_{\theta\theta T})(\widehat{\theta}_T - \theta^*). \quad (2.3)$$

Thus it is enough to prove that the second term on the right hand side is  $O_M(T^{-1})$ .

**Lemma 2.3.:** We have

$$\widehat{\theta}_T - \theta^* = O_M(T^{-1/2}). \quad (2.4)$$

**Proof:** Let us investigate  $\bar{V}_{\theta\theta T}$ . Define

$$\bar{W}_{\theta\theta T} = \int_0^1 W_{\theta\theta T}((1-\lambda)\theta^* + \lambda\hat{\theta}_T, \theta^*)d\lambda. \quad (2.5)$$

We show that  $\bar{W}_{\theta\theta T} > c'$  with some positive  $c'$  on  $A_T$  if  $d$  is sufficiently small. Indeed since  $W$  is smooth we have for  $0 \leq \lambda \leq 1$

$$\|W_{\theta\theta}((1-\lambda)\theta^* + \lambda\hat{\theta}_T, \theta^*) - W_{\theta\theta}(\theta^*, \theta^*)\| \leq |\hat{\theta}_T - \theta^*| < Cd. \quad (2.6)$$

where  $C$  is a system constant. Hence if  $d$  is sufficiently small then the positive definiteness of  $W_{\theta\theta}(\theta^*, \theta^*)$  and (2.4) imply that  $\bar{W}_{\theta\theta T} > c'I$  with some positive  $c'$ . Since we have on  $A_T$

$$\|\frac{1}{T}\bar{V}_{\theta\theta T} - \bar{W}_{\theta\theta T}\| < d''$$

it follows that if  $d''$  is sufficiently small then

$$\lambda_{\min}(\frac{1}{T}\bar{V}_{\theta\theta T}) > c > 0 \quad (2.7)$$

on  $A_T$  where in general  $\lambda_{\min}(B)$  denotes the minimal eigenvalue of the matrix  $B$ . Then by (2.2) we get

$$\chi_{A_T}|\hat{\theta}_T - \theta^*| \leq c^{-1}T^{-1}\|V_{\theta T}(\theta^*, \theta^*)\|. \quad (2.8)$$

It is easy to see that  $V_{\theta T}(\theta^*, \theta^*) = O_M(T^{1/2})$ . Indeed we have  $d \epsilon_t(\theta^*, \theta^*) = dw_t + O_M(e^{-\lambda t})dt$  where the second term is due to nonstationary initial conditions, the contribution of which to  $V_{\theta T}(\theta^*, \theta^*)$  is obviously  $O_M(1)$ . On the other hand we have for any  $q \geq 1$  that

$$E^{1/2q}|\int_0^T \dot{\epsilon}_{\theta t}(\theta^*, \theta^*)dw_t|^{2q} \leq C_q E^{1/2q}(\int_0^T \dot{\epsilon}_{\theta t}(\theta^*, \theta^*)dt)^q. \quad (2.9)$$

(c.f. Krylov (1977)). Using Hölder's inequality we can continue (2.9) as

$$\leq C_q(\int_0^T E^{1/q}|\dot{\epsilon}_{\theta t}(\theta^*, \theta^*)|^{2q}dt)^{1/2} = O(T^{1/2}).$$



Substituting into (2.8) we get

$$\chi_{A_T}(\widehat{\theta}_T - \theta^*) = O_M(T^{-1/2}). \quad (2.10)$$

Combining this inequality with the previous inequality  $P(A_T^c) = O(T^{-s})$  where  $A_T^c$  denotes the complement of  $A_T$  and using the fact that  $|\widehat{\theta}_T - \theta^*|$  is bounded we get for any  $s > 0$  that

$$\chi_{A_T^c}(\widehat{\theta}_T - \theta^*) = O_M(T^{-s}). \quad (2.11)$$

Adding this equality to (2.10) we get the lemma.

Now we can complete the analysis of (2.5) as follows: First we can write

$$\|W_{\theta\theta}(\theta^*, \theta^*) - \frac{1}{T}\overline{V}_{\theta\theta T}\| \leq \|W_{\theta\theta}(\theta^*, \theta^*) - \overline{W}_{\theta\theta T}\| + \|\overline{W}_{\theta\theta T} - \frac{1}{T}\overline{V}_{\theta\theta T}\|. \quad (2.12)$$

Using the previous lemma we get

$$\begin{aligned} \|W_{\theta\theta}(\theta^*, \theta^*) - \overline{W}_{\theta\theta T}\| &\leq \int_0^1 \|W_{\theta\theta}(\theta^*, \theta^*) - W_{\theta\theta}((1-\lambda)\theta^* + \lambda\widehat{\theta}_T, \theta^*)\| d\lambda \leq \\ &\leq C \int_0^1 |(1-\lambda)\theta^* + \lambda\widehat{\theta}_T - \theta^*| d\lambda \leq C|\widehat{\theta}_T - \theta^*| = O_M(T^{-1/2}). \end{aligned} \quad (2.13)$$

On the other hand we have by Theorem 2.1.

$$\begin{aligned} \|\overline{W}_{\theta\theta T} - \frac{1}{T}\overline{V}_{\theta\theta T}\| &\leq \int_0^1 \|W_{\theta\theta}((1-\lambda)\theta^* + \lambda\widehat{\theta}_T, \theta^*) - \frac{1}{T}V_{\theta\theta}((1-\lambda)\theta^* + \lambda\widehat{\theta}_T, \theta^*)\| d\lambda \\ &\leq \sup_{\theta \in D^o} \|W_{\theta\theta}(\theta, \theta^*) - \frac{1}{T}V_{\theta\theta}(\theta, \theta^*)\| = O_M(T^{-1/2}). \end{aligned} \quad (2.14)$$

Summarizing (2.13) and (2.14) we get that

$$(W_{\theta\theta}(\theta^*, \theta^*) - \frac{1}{T}\overline{V}_{\theta\theta T}) = O_M(T^{-1}).$$

Together with (2.4) it yields that

$$R^*(\widehat{\theta}_T - \theta^*) = -\frac{1}{T}V_{\theta\theta T}(\theta^*, \theta^*) + O_M(T^{-1}).$$

### 3. Appendix

In this section we summarize a few definitions and theorems we need for this paper. A detailed exposition is given in Gerencsér (1989a).

Let  $D \subset \mathbb{R}^p$  be a compact domain and let the stochastic process  $(x_t(\theta))$  be defined on the parameter set  $\mathbb{R}^+ \times D$ .

Let  $(\mathcal{F}_s), s \geq 0$  be a family of monotone increasing  $\sigma$ -algebras, and  $(\mathcal{F}_s^+)s \geq 0$  be a monotone decreasing family of  $\sigma$ -algebras. We assume that  $(\mathcal{F}_s^+)$  is continuous from the right, i.e.  $\mathcal{F}_s^+ = \sigma\{\cup_{0 < \epsilon} \mathcal{F}_{s+\epsilon}^+\}$ . Furthermore assume that for all  $s \geq 0, \mathcal{F}_s$  and  $\mathcal{F}_s^+$  are independent. For  $s < 0 \mathcal{F}_s^+ = \mathcal{F}_0^+$ . A typical example is provided by the  $\sigma$ -algebras

$$\mathcal{F}_s = \sigma\{w_t : t \leq s\} \quad \mathcal{F}_s^+ = \sigma\{w_t - w_{t'} : t, t' > s\}$$

where  $(w_t, t \geq 0)$  is a standard Wiener-process.

**Definition 3.1.** A stochastic process  $(x_t), t \geq 0$  is  $L$ -mixing with respect to  $(\mathcal{F}_t, \mathcal{F}_t^+)$  if it is  $\mathcal{F}_t$ -progressively measurable,  $M$ -bounded and with any  $q$  such that  $1 \leq q < \infty$  and with

$$\gamma_q(\tau, x) = \gamma_q(\tau) = \sup_{t \geq \tau} \mathbb{E}^{1/q} |x_t - \mathbb{E}(x_t | \mathcal{F}_{t-\tau}^+)|^q \quad \tau \geq 0$$

we have

$$\Gamma_q = \Gamma_q(x) = \int_0^\infty \gamma_q(\tau) d\tau < \infty. \quad (3.1)$$

It can be shown that  $\gamma_q(\tau)$  is measurable and thus the integral (3.1) makes sense.

**Example 1.** The process  $(x_t)$  given by

$$dx_t = Ax_t dt + Bdw_t \quad x_0 = 0$$

where  $w_t$  is an  $m$ -dimensional Wiener-process,  $A$  is  $n \times n, B$  is  $n \times m$  matrix and  $A$  is stable is  $L$ -mixing with respect to  $(\mathcal{F}_t, \mathcal{F}_t^+)$  given in (3.1).

The above definition extends to discrete-time processes in a natural way.

An important property of  $L$ -mixing processes is that if  $(x_t), (y_t)$  are  $L$ -mixing then  $(z_t)$  with  $z_t = x_t y_t$  is also  $L$ -mixing. This is seen by direct calculations.

**Theorem 3.1** Let  $u_t(\theta)$  be a parameter dependent stochastic process such that itself and the process  $\Delta u / \Delta^\alpha \theta$  are both  $L$ -mixing and  $\mathbb{E}u_t(\theta) = 0$  for all  $\theta \in D$ . Let

$$x_t^* = \sup_{\theta \in D^o} \left| \frac{1}{\sqrt{t}} \int_0^t u_s(\theta) ds \right|.$$

Then the process  $x_t^*$  is  $M$ -bounded, and we have for all  $1 \leq q < \infty$  and  $s > p$

$$M_{2q}(x^*) \leq C(M_{2qs}(u) + M_{2qs}(\Delta u / \Delta^\alpha \theta))(\Gamma_{2qs}(u) + \Gamma_{2qs}(\Delta u / \Delta^\alpha \theta))$$

where  $C$  depends only on  $p, q, s, D$  and  $D^o$ .

Let us consider a linear filter described by

$$x_t = \int_0^t \phi(t-s) u_s ds$$

where  $\phi(\tau)$  is locally integrable.

We say that the filter is stable if

$$\Phi^* = \int_0^\infty \phi(\tau) d\tau < \infty.$$

The definition extends to vector valued processes in an obvious way.

**Theorem 3.2** Let us consider a stable linear filter described above. Assume that

$$\Phi^{**} = \int_0^\infty \tau \phi(\tau) d\tau < \infty.$$

Then if the input process  $u_t$  is  $L$ -mixing then the output process  $x_t$  is also  $L$ -mixing.

**Theorem 3.3.** Let  $u_t(\theta)$  be a progressively measurable  $\mathbb{R}^p$ -valued stochastic process which satisfies the conditions of the conditions of Theorem 3.1. and let

$$x_t(\theta) = \frac{1}{\sqrt{t}} \int_0^t u_s(\theta) dw_s$$

and take a separable version of  $x_t(\theta)$ . Then

$$x_t^* = \sup_{\theta \in D^o} |x_t(\theta)|$$

is  $M$ -bounded, and we have for  $q \geq 1, s > p/\alpha$

$$M_{2q}(x^*) \leq C(M_{2qs}(u) + M_{2qs}(\Delta u/\Delta^\alpha \theta))$$

where  $C$  depends only on  $p, q, s, D$  and  $D^o$ .

**Lemma 3.4.** Let  $G(\theta)$  and  $\delta G(\theta)$  be  $\mathbb{R}^p$ -valued continuously differentiable functions,  $\theta \in D \subset \mathbb{R}^p$ . Let for some  $\theta^* \in D^o$   $G(\theta^*) = 0$ , where  $D^o$  is a compact subset of  $D$ . Assume that  $G_\theta(\theta^*)$  is nonsingular. Then for any  $d > 0$  there exist  $d', d''$  positive numbers such that if

$$|\delta G(\theta)| < d' \quad \text{and} \quad |G_\theta(\theta)| < d''$$

for all  $\theta \in D^o$  then

$$G(\theta) + \delta G(\theta) = 0$$

has a unique solution  $\hat{\theta} \in D^o$ , moreover  $|\theta^* - \hat{\theta}| < d$ .

## REFERENCES

- Arató, M. (1984). Linear Stochastic Systems with Constant Coefficients. Lecture Notes in Control and Information Science, Springer.
- Åström, K.J. and Söderström, T. (1974). Uniqueness of the Maximum-Likelihood Estimates of the Parameters of an ARMA 1974 Model. IEEE Transactions on Automatic Control **AC-19** 769-773.
- Caines, P.E. (1988). Linear Stochastic Systems. Wiley.

- Gerencsér, L. (1989a). On a class of mixing processes. *Stochastics* **26** 165-191.
- Gerencsér, L. (1989b). On the martingale approximation of the estimation error of ARMA parameters. McGill Research Centre for Intelligent Machines, TR-CIM-89-6
- Gerencsér, L. (1989c). Strong approximation theorems for estimator processes in continuous time. *Proc. of Limit Theorems in Probability and Statistics*. Pécs, Hungary.
- Gerencsér, L., Gyöngy, Gy. and Michaletzky, Gy. (1984). Continuous-Time Recursive Maximum Likelihood Method. A New Approach to Ljung's Scheme. In: *A Bridge Between Control Science and Technology*. Proc. of the Ninth Triennial World Congress of IFAC., **2**, Identification, Adaptive and Stochastic Control (eds. Ljung, L. and Aström, K.J.). Pergamon Press, Oxford, 683-683.
- Hannan, E.J. and Deistler, M. (1988). *The Stastical Theory of Linear Systems*, Wiley.
- Krylov, N.V. (1977). *Controlled diffusion processes*. Nauka, Moscow. English translation, Springer-Verlag, Berlin(1980).
- Kutoyants, Yu. A. (1980, 1984). *Estimation of Stochastic Processes*. Publishing House of the Armenian Academy of Sciences, Yerevan (in Russian) and Heldermann, Berlin (in English).
- Söderström, T. and Stoica, P. (1982). Uniqueness of Prediction Error Estimates of Multi-variable Moving Average Models, *Automatica* **18** 617-620.

G. Vágó, Zs. and Gerencsér, L. (1985). Uniqueness of the Maximum-Likelihood Estimates of the Kalman-gain Matrix of a State Space Model, Proc. of the IFAC/IFORS Conference on Dynamic Modelling of National Economics (ed. M. Ziermann), Budapest.

## Recent IMA Preprints

#	Author/s	Title
618	L.E. Fraenkel,	On a linear, partly hyperbolic model of viscoelastic flow past a plate
619	Stephen Schecter and Michael Shearer,	Undercompressive shocks for nonstrictly hyperbolic conservation laws
620	Xinfu Chen,	Axially symmetric jets of compressible fluid
621	J. David Logan,	Wave propagation in a qualitative model of combustion under equilibrium conditions
622	M.L. Zeeman,	Hopf bifurcations in competitive three-dimensional Lotka-Volterra Systems
623	Allan P. Fordy,	Isospectral flows: their Hamiltonian structures, Miura maps and master symmetries
624	Daniel D. Joseph, John Nelson, Michael Renardy, and Yuriko Renardy,	Two-Dimensional cusped interfaces
625	Avner Friedman and Bei Hu,	A free boundary problem arising in electrophotography
626	Hamid Bellout, Avner Friedman and Victor Isakov,	Stability for an inverse problem in potential theory
627	Barbara Lee Keyfitz,	Shocks near the sonic line: A comparison between steady and unsteady models for change of type
628	Barbara Lee Keyfitz and Gerald G. Warnecke,	The existence of viscous profiles and admissibility for transonic shocks
629	P. Szmolyan,	Transversal heteroclinic and homoclinic orbits in singular perturbation problems
630	Philip Boyland,	Rotation sets and monotone periodic orbits for annulus homeomorphisms
631	Kenneth R. Meyer,	Apollonius coordinates, the N-body problem and continuation of periodic solutions
632	Chjan C. Lim,	On the Poincaré-Whitney circuitspace and other properties of an incidence matrix for binary trees
633	K.L. Cooke and I. Györi,	Numerical approximation of the solutions of delay differential equations on an infinite interval using piecewise constant arguments
634	Stanley Minkowitz and Matthew Witten,	Periodicity in cell proliferation using an asynchronous cell population
635	M. Chipot and G. Dal Maso,	Relaxed shape optimization: The case of nonnegative data for the Dirichlet problem
636	Jeffery M. Franke and Harlan W. Stech,	Extensions of an algorithm for the analysis of nongeneric Hopf bifurcations, with applications to delay-difference equations
637	Xinfu Chen,	Generation and propagation of the interface for reaction-diffusion equations
638	Philip Korman,	Dynamics of the Lotka-Volterra systems with diffusion
639	Harlan W. Stech,	Generic Hopf bifurcation in a class of integro-differential equations
640	Stephane Laederich,	Periodic solutions of non linear differential difference equations
641	Peter J. Olver,	Canonical Forms and Integrability of BiHamiltonian Systems
642	S.A. van Gils, M.P. Krupa and W.F. Langford,	Hopf bifurcation with nonsemisimple 1:1 Resonance
643	R.D. James and D. Kinderlehrer,	Frustration in ferromagnetic materials
644	Carlos Rocha,	Properties of the attractor of a scalar parabolic P.D.E.
645	Debra Lewis,	Lagrangian block diagonalization
646	Richard C. Churchill and David L. Rod,	On the determination of Ziglin monodromy groups
647	Xinfu Chen and Avner Friedman,	A nonlocal diffusion equation arising in terminally attached polymer chains
648	Peter Gritzmann and Victor Klee,	Inner and outer $j$ - Radii of convex bodies in finite-dimensional normed spaces
649	P. Szmolyan,	Analysis of a singularly perturbed traveling wave problem
650	Stanley Reiter and Carl P. Simon,	Decentralized dynamic processes for finding equilibrium
651	Fernando Reitich,	Singular solutions of a transmission problem in plane linear elasticity for wedge-shaped regions
652	Russell A. Johnson,	Cantor spectrum for the quasi-periodic Schrödinger equation
653	Wenxiong Liu,	Singular solutions for a convection diffusion equation with absorption
654	Deborah Brandon and William J. Hrusa,	Global existence of smooth shearing motions of a nonlinear viscoelastic fluid
655	James F. Reineck,	The connection matrix in Morse-Smale flows II
656	Claude Baesens, John Guckenheimer, Seunghwan Kim and Robert Mackay,	Simple resonance regions of torus diffeomorphisms
657	Willard Miller, Jr.,	Lecture notes in radar/sonar: Topics in Harmonic analysis with applications to radar and sonar

- 658 **Calvin H. Wilcox**, Lecture notes in radar/sonar: Sonar and Radar Echo Structure
- 659 **Richard E. Blahut**, Lecture notes in radar/sonar: Theory of remote surveillance algorithms
- 660 **D.V. Anosov**, Hilbert's 21st problem (according to Bolibruch)
- 661 **Stephane Laederich**, Ray–Singer torsion for complex manifolds and the adiabatic limit
- 662 **Geneviève Raugel and George R. Sell**, Navier-Stokes equations in thin 3d domains: Global regularity of solutions I
- 663 **Emanuel Parzen**, Time series, statistics, and information
- 664 **Andrew Majda and Kevin Lamb**, Simplified equations for low Mach number combustion with strong heat release
- 665 **Ju. S. Il'yashenko**, Global analysis of the phase portrait for the Kuramoto–Sivashinsky equation
- 666 **James F. Reineck**, Continuation to gradient flows
- 667 **Mohamed Sami Elbially**, Simultaneous binary collisions in the collinear N–body problem
- 668 **John A. Jacquez and Carl P. Simon**, Aids: The epidemiological significance of two different mean rates of partner-change
- 669 **Carl P. Simon and John A. Jacquez**, Reproduction numbers and the stability of equilibria of SI models for heterogeneous populations
- 670 **Matthew Stafford**, Markov partitions for expanding maps of the circle
- 671 **Ciprian Foias and Edriss S. Titi**, Determining nodes, finite difference schemes and inertial manifolds
- 672 **M.W. Smiley**, Global attractors and approximate inertial manifolds for abstract dissipative equations
- 673 **M.W. Smiley**, On the existence of smooth breathers for nonlinear wave equations
- 674 **Hitay Özbay and Janos Turi**, Robust stabilization of systems governed by singular integro-differential equations
- 675 **Mary Silber and Edgar Knobloch**, Hopf bifurcation on a square lattice
- 676 **Christophe Golé**, Ghost circles for twist maps
- 677 **Christophe Golé**, Ghost tori for monotone maps
- 678 **Christophe Golé**, Monotone maps of  $T^n \times R^n$  and their periodic orbits
- 679 **E.G. Kalnins and W. Miller, Jr.**, Hypergeometric expansions of Heun polynomials
- 680 **Victor A. Pliss and George R. Sell**, Perturbations of attractors of differential equations
- 681 **Avner Friedman and Peter Knabner**, A transport model with micro- and macro-structure
- 682 **E.G. Kalnins and W. Miller, Jr.**, A note on group contractions and radar ambiguity functions
- 683 **George R. Sell**, References on dynamical systems
- 684 **Shui-Nee Chow, Kening Lu and George R. Sell**, Smoothness of inertial manifolds
- 685 **Shui-Nee Chow, Xiao-Biao Lin and Kening Lu**, Smooth invariant foliations in infinite dimensional spaces
- 686 **Kening Lu**, A Hartman–Grobman theorem for scalar reaction-diffusion equations
- 687 **Christophe Golé and Glen R. Hall**, Poincaré's proof of Poincaré's last geometric theorem
- 688 **Mario Taboada**, Approximate inertial manifolds for parabolic evolutionary equations via Yosida approximations
- 689 **Peter Rejto and Mario Taboada**, Weighted resolvent estimates for Volterra operators on unbounded intervals
- 690 **Joel D. Avrin**, Some examples of temperature bounds and concentration decay for a model of solid fuel combustion
- 691 **Susan Friedlander and Misha M. Vishik**, Lax pair formulation for the Euler equation
- 692 **H. Scott Dumas**, Ergodization rates for linear flow on the torus
- 693 **A. Eden, A.J. Milani and B. Nicolaenko**, Finite dimensional exponential attractors for semilinear wave equations with damping
- 694 **A. Eden, C. Foias, B. Nicolaenko & R. Temam**, Inertial sets for dissipative evolution equations
- 695 **A. Eden, C. Foias, B. Nicolaenko & R. Temam**, Hölder continuity for the inverse of Mañé's projection
- 696 **Michel Chipot and Charles Collins**, Numerical approximations in variational problems with potential wells
- 697 **Huanan Yang**, Nonlinear wave analysis and convergence of MUSCL schemes
- 698 **László Gerencsér and Zsuzsanna Vágó**, A strong approximation theorem for estimator processes in continuous time
- 699 **László Gerencsér**, Multiple integrals with respect to  $L$ -mixing processes
- 700 **David Kinderlehrer and Pablo Pedregal**, Weak convergence of integrands and the Young measure representation
- 701 **Bo Deng**, Symbolic dynamics for chaotic systems
- 702 **P. Galdi, D.D. Joseph, L. Preziosi, S. Rionero**, Mathematical problems for miscible, incompressible fluids with Korteweg stresses
- 703 **Charles Collins and Mitchell Luskin**, Optimal order error estimates for the finite element approximation of the solution of a nonconvex variational problem
- 704 **Peter Gritzmann and Victor Klee**, Computational complexity of inner and outer  $j$ -radii of polytopes in finite-dimensional normed spaces