Systems & Control Letters 22 (1994) 111–121
North-Holland

A note on continuous-time ELS*

L. Guo
Institute of Systems Science, Chinese Academy of Sciences, Beijing 100080, People's Republic of China

Received 13 February 1993
Revised 13 May 1993

Abstract: The parameter estimation of continuous-time finite-dimensional linear stochastic systems is a problem of long-standing interest. The method usually used is the extended least-squares (ELS) algorithm described by a nonlinear stochastic differential equation (SDE), with the existence of the global strong solution assumed. This paper shows that the ELS estimate does exist in $[0, \infty)$, and at the same time presents a number of convergence results paralleling those for the discrete-time case.

Keywords: Stochastic systems; extended least squares; stochastic differential equation.

1. Introduction

Consider the following standard linear state space model:

$$\mathrm{d}x_t = Ax_t\,\mathrm{d}t + Bu_t^0\,\mathrm{d}t + D\,\mathrm{d}w_t, \qquad \mathrm{d}y_t^0 = Cx_t\,\mathrm{d}t + \mathrm{d}w_t,$$

where y_t^0 and u_t^0 are the scalar output and input, respectively, x_t is the r-dimensional state vector, $\{w_t, \mathscr{F}_t\}$ is a Wiener process on the basic probability space $(\Omega, \mathscr{F}, P)$, and A, B, C and D are unknown matrices of compatible dimensions. Without loss of generality, assume that $y_t^0 = u_t^0 = 0$, $x_t = 0$, $\forall t \leq 0$. Then the input–output relationship can be written as (cf. [3, pp. 403–404])

$$A(S)y_t^0 = SB(S)u_t^0 + C(S)w_t,$$

where $A(S), B(S)$ and $C(S)$ are polynomials in integral operator S (i.e. $Sw_t \triangleq \int_0^t w_z\,\mathrm{d}z$):

$$A(S) = 1 + a_1 S + \cdots + a_p S^p, \quad p \geq 1,$$

$$B(S) = b_1 + b_2 S + \cdots + b_q S^{q-1}, \quad q \geq 1,$$

$$C(S) = 1 + c_1 S + \cdots + c_r S^r, \quad r \geq 0,$$

with unknown real coefficients a_i, b_j and c_k and with known upper bounds p, q and r for orders.

As noted by Moore [12, p. 197], it is usual to introduce a prefilter $D(S) = 1 + d_1 S + \cdots + d_r S^r$, which is exponentially stable, giving rise to prefiltered variables y_t and u_t defined from $y_t \triangleq D^{-1}(S)y_t^0$, and $u_t \triangleq D^{-1}(S)u_t^0$. Thus we obtain the following relationship between y_t and u_t:

$$A(S)y_t = SB(S)u_t + C(S)v_t, \quad t \geq 0, \tag{1.1}$$

$$D(S)v_t = w_t. \tag{1.2}$$

Let us denote the unknown parameter by

$$\theta = [-a_1 \ \ldots \ -a_p \ b_1 \ \ldots \ b_q \ c_1 \ \ldots \ c_r]^{\mathrm{T}} \tag{1.3}$$

Correspondence to: L. Guo, Academia Sinica, Institute of Systems Science, Chinese Academy of Sciences, Beijing 100080, People's Republic of China.
* Work supported by the National Natural Science Foundation of China.

SSDI 0167-6911(93)E0067-Q

and set

$$\varphi_t^0 = [y_t, Sy_t, \ldots S^{p-1}y_t \; u_t, \ldots S^{q-1}u_t \; v_t, \ldots S^{r-1}v_t]^{\mathrm{T}} \tag{1.4}$$

$$\bar{D}(S) = [D(S) - 1]/S, \qquad \bar{C}(S) = [C(S) - 1]/S. \tag{1.5}$$

Then (1.1) and (1.2) can be rewritten as

$$\mathrm{d}y_t = \theta^{\mathrm{T}}\varphi_t^0 \,\mathrm{d}t + \mathrm{d}v_t, \tag{1.6}$$

$$\mathrm{d}v_t = \mathrm{d}w_t - [\bar{D}(S)v_t]\,\mathrm{d}t. \tag{1.7}$$

The commonly used extended least-squares (ELS) estimate $\hat{\theta}_t$ for θ is defined by the following stochastic differential equation (SDE) (cf. [1, 5]):

$$\mathrm{d}\hat{\theta}_t = P_t\varphi_t[\mathrm{d}y_t - \hat{\theta}_t^{\mathrm{T}}\varphi_t\mathrm{d}t + (\bar{D}(S)\hat{v}_t)\mathrm{d}t], \tag{1.8}$$

$$\mathrm{d}\hat{v}_t = \mathrm{d}y_t - \hat{\theta}_t^{\mathrm{T}}\varphi_t\,\mathrm{d}t, \tag{1.9}$$

$$\varphi_t = [y_t, Sy_t, \ldots S^{p-1}y_t \; u_t, \ldots S^{q-1}u_t \; \hat{v}_t, \ldots S^{r-1}\hat{v}_t]^{\mathrm{T}} \tag{1.10}$$

$$P_t = \left(\int_0^t \varphi_s\varphi_s^{\mathrm{T}}\,\mathrm{d}s + aI\right)^{-1}, \quad a > 0, \tag{1.11}$$

where $\hat{\theta}_0$ is deterministic and arbitrarily chosen, and $\hat{v}_0 = 0$.

Clearly, (1.8)–(1.11) is a nonlinear SDE for $(\hat{\theta}_t, \hat{v}_t)$. This SDE may be regarded as the continuous-time analogue of the discrete-time extended least-squares (ELS) algorithm (e.g. [1–5, 12]).

The above ELS algorithm has attracted much research interest over the past decade. However, to the best of the author's knowledge, the basic existence and uniqueness problem for ELS still remains open in the literature. Indeed, almost all of the existing results build on the assumption that the ELS estimate $\hat{\theta}_t$ exists in $[0, \infty)$ (see, e.g. [1, p. 131; 2, p. 515; 5, p. 267; 12, p. 199]). In this paper, we will first study this long-standing problem, proving that under very mild conditions on the input process and the noise model, the ELS estimate $\hat{\theta}_t$ does exist in $[0, \infty)$ and is the unique strong solution of the SDE (1.8)–(1.11), and then we will discuss the convergence rate of ELS together with the related excitation problem.

2. Existence of ELS

For convenience of analysis, we first derive the error equation associated with (1.8) and (1.9). Set

$$\tilde{\theta}_t = \theta - \hat{\theta}_t, \qquad \tilde{v}_t = v_t - \hat{v}_t. \tag{2.1}$$

Then substituting (1.6) and (1.7) into (1.8) we have

$$\begin{aligned} \mathrm{d}\tilde{\theta}_t &= -\,\mathrm{d}\hat{\theta}_t = -P_t\varphi_t[\theta^{\mathrm{T}}(\varphi_t^0 - \varphi_t)\mathrm{d}t - \bar{D}(S)\tilde{v}_t\mathrm{d}t + \tilde{\theta}_t^{\mathrm{T}}\varphi_t\mathrm{d}t + \mathrm{d}w_t] \\ &= -P_t\varphi_t[(\bar{C}(S) - \bar{D}(S))\tilde{v}_t\mathrm{d}t + \tilde{\theta}_t^{\mathrm{T}}\varphi_t\mathrm{d}t + \mathrm{d}w_t] \end{aligned} \tag{2.2}$$

Next, by (1.6) and (1.9),

$$\begin{aligned} \mathrm{d}\tilde{v}_t &= \mathrm{d}v_t - \mathrm{d}\hat{v}_t = \hat{\theta}_t^{\mathrm{T}}\varphi_t\,\mathrm{d}t - \theta^{\mathrm{T}}\varphi_t^0\,\mathrm{d}t \\ &= -[\tilde{\theta}_t^{\mathrm{T}}\varphi_t + \theta^{\mathrm{T}}(\varphi_t^0 - \varphi_t)]\,\mathrm{d}t = -\tilde{\theta}_t^{\mathrm{T}}\varphi_t\,\mathrm{d}t - [\bar{C}(S)\tilde{v}_t]\,\mathrm{d}t. \end{aligned} \tag{2.3}$$

Consequently, $C(S)(\mathrm{d}\tilde{v}_t/\mathrm{d}t) = -\tilde{\theta}_t^{\mathrm{T}}\varphi_t$. By this we may rewrite (2.2) as

$$\begin{aligned} \mathrm{d}\tilde{\theta}_t &= -P_t\varphi_t[(\bar{C}(S) - \bar{D}(S))(-SC^{-1}(S)\tilde{\theta}_t^{\mathrm{T}}\varphi_t)\mathrm{d}t + \tilde{\theta}_t^{\mathrm{T}}\varphi_t\mathrm{d}t + \mathrm{d}w_t] \\ &= -P_t\varphi_t[(D(S)C^{-1}(S)\tilde{\theta}_t^{\mathrm{T}}\varphi_t)\mathrm{d}t + \mathrm{d}w_t] \end{aligned} \tag{2.4}$$

We now write (2.3) and (2.4) in the following compact form:

$$\mathrm{d}x_t = a(t,x)\,\mathrm{d}t + b(t,x)\,\mathrm{d}w_t, \qquad x_0 = [\tilde{\theta}_0, 0, \ldots, 0]^{\mathrm{T}}, \tag{2.5}$$

where

$$x = (x_t)_{t \geq 0}, \quad x_t = [\tilde{\theta}_t^{\mathrm{T}}, \tilde{v}_t, S\tilde{v}_t, \ldots S^{r-1}\tilde{v}_t]^{\mathrm{T}} \tag{2.6}$$

$$a(t,x) = \begin{pmatrix} -P_t\varphi_t f_t \\ -g_t - \bar{C}(S)\tilde{v}_t \\ \tilde{v}_t \\ \vdots \\ S^{r-2}\tilde{v}_t \end{pmatrix}, \qquad b(t,x) = \begin{pmatrix} -P_t\varphi_t \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{2.7}$$

$$g_t = \varphi_t^{\mathrm{T}}\tilde{\theta}_t, \qquad f_t = [D(S)C^{-1}(S)]g_t \tag{2.8}$$

and

$$\varphi_t = [y_t, \ldots S^{p-1}y_t,\ u_t, \ldots S^{q-1}u_t,\ (v_t - \tilde{v}_t), \ldots S^{r-1}(v_t - \tilde{v}_t)]^{\mathrm{T}} \tag{2.9}$$

and where P_t is defined by (1.11), and the initial value $\tilde{\theta}_0$ is deterministic and arbitrarily chosen.

Clearly, the existence of a global solution of the SDE (1.8)–(1.11) implies that of the SDE (2.5)–(2.9) and vice versa. So we need only study the SDE (2.5)–(2.9).

The main result of this section is as follows:

Theorem 2.1. *For System* (1.1) *and* (1.2), *assume that the input process* $\{u_t\}$ *is continuous and adapted to* $\{\mathscr{F}_t\}$, *and that the transfer function* $D(S)C^{-1}(S) - 1/2$ *is strictly positive real. Then the SDE* (2.5)–(2.9) *has a unique strong solution* $\{x_t, \mathscr{F}_t\}$ *on* $[0, \infty)$.

An immediate consequence of Theorem 2.1 is that the ELS estimate $\hat{\theta}_t$ exists on $[0, \infty)$ and is the unique strong solution of the SDE (1.8)–(1.11). It should be noted that the continuity of the input process is assumed only for simplicity of discussions; it can be further weakened, as can easily be seen from the proof. Also, the positive real condition on the noise model is a standard one in the literature of recursive system identification.

To prove Theorem 2.1, it is necessary to introduce some notations. Denote by $C^d[0,T]$, the space of $\mathscr{R}^d$-valued continuous functions on the interval $[0,T]$, $T > 0$, and by $(C^d[0,T], \mathscr{B}_T)$ the measurable space of continuous functions $x = (x_t, 0 \leq t \leq T)$ with the σ-algebra $\mathscr{B}_T = \sigma\{x: x_t, 0 \leq t \leq T\}$. Also, set $\mathscr{B}_t = \sigma\{x: x_s, s \leq t\}$, $\forall t \leq T$. As usual, $\lambda_{\min}(X)$ and $\lambda_{\max}(X)$ denote the minimum and maximum eigenvalues of a real matrix X, respectively, and the norm of X is defined as $\|X\| = \{\lambda_{\max}(XX^{\mathrm{T}})\}^{1/2}$. When $x = (x_t)_{t \leq T}$ is a $C^d[0,T]$ process, we set $\|x\|_{[0,T]} = \max_{t \leq T}\|x_t\|$.

We first present some preliminary results on the following general vector SDE:

$$\mathrm{d}x_t = a(t,x)\,\mathrm{d}t + b(t,x)\,\mathrm{d}w_t, \quad x_0 = \eta \tag{2.10}$$

where η is an $\mathscr{F}_0$-measurable random vector, $\{w_t, \mathscr{F}_t\}$ is a standard Wiener process, $a(t,x)$ and $b(t,x)$, $t \leq T, x \in C^d[0,T]$, are $\mathscr{F}_t \times \mathscr{B}_t$-measurable vector functionals of dimension $d (\geq 1)$.

Lemma 2.1. *Assume that for each* $n \geq 1$ *there exists a continuous process* $\{L_t^{(n)}, \mathscr{F}_t\}_{t \leq T}$, *such that for* $t \leq T$, $x, y \in C^d[0,T]$ *and* $n \geq 1$,

$$\begin{aligned} &[\|a(t,x) - a(t,y)\|^2 + \|b(t,x) - b(t,y)\|^2]\, I(\|x\|_{[0,T]} \leq n, \|y\|_{[0,T]} \leq n) \\ &\quad \leq L_t^{(n)}\left\{\|x_t - y_t\|^2 + \int_0^t \|x_s - y_s\|^2\,\mathrm{d}s\right\}, \quad \text{a.s. } \forall n \geq 1, \end{aligned} \tag{2.11}$$

and

$$[\|a(t,x)\|^2 + \|b(t,x)\|^2] I(\|x\|_{[0,T]} \le n) \le L_t^{(n)} \left\{ 1 + \|x_t\|^2 + \int_0^t \|x_s\|^2 \, ds \right\} \quad a.s., \tag{2.12}$$

where $I(A)$ is the indicator function of a set A. Then there exists an $\mathscr{F}_t$-Markov time $\sigma_T > 0$ such that the SDE (2.10) *has a unique strong solution $x = (x_t)$ on $[t < \sigma_T]$, and*

$$\sup_{t<\sigma_T} \|x_t\| = \infty \quad a.s., \quad on\ [\sigma_T < T].$$

This lemma is a minor extension of the existing results on local solutions (see, e.g. [8, Theorem 3.1]), for a proof, see Appendix A. The Markov time σ_T is usually called the explosion time of the SDE (2.10). If in (2.12) the process $\{L_t^{(n)}\}$ does not depend on n, i.e. $L_t^{(n)} \equiv L_t$, then it can be shown that $P(\sigma_T = T) = 1$, which means that the solution of the SDE (2.10) is a global one on $[0, T]$. Related results under certain nonlinear growth conditions may also be found in the literature (e.g. [7, 9]). However, direct applications of these results to the SDE (1.8)–(1.11) are found to be difficult. The main reason is that we do not know how to verify the growth conditions in, e.g. [7, Theorem 1]. Therefore, we present another lemma on the existence of the global solution by using a different growth condition, which can be directly applied to the SDE (1.8)–(1.11), giving a simple and straightforward existence proof of ELS.

Lemma 2.2. *Let $a(t,x)$ and $b(t,x)$, $t \ge 0$, and $x \in C^d[0,\infty)$ (the space of d-dimensional continuous functions on $[0,\infty)$), satisfy* (2.11) *and* (2.12) *for any $T > 0$, where $(L_t^{(n)}, \mathscr{F}_t)_{t\ge 0}$ is a continuous nonnegative process for each $n \ge 1$. Suppose there exists a symmetric matrix functional $Q_t(x): [0,\infty) \times C^d[0,\infty) \times \Omega \to \mathscr{R}^{d\times d}$ with $Q_t(x)$ and $dQ_t(x)/dt$ measurable $\mathscr{F}_t \times \mathscr{B}_t$, and with $\inf_{t,x} \lambda_{\min}[Q_t(x)] > 0$ a.s., such that for any $t > 0, x \in C^d[0,\infty)$,*

$$\int_0^t \mathscr{L}_s(x)\, ds \le F_t + G_t \int_0^t \|x_s\|^2 \, ds - \varepsilon \int_0^t \|b^T(s,x) Q_s(x) x_s\|^2 \, ds \tag{2.13}$$

where $\varepsilon > 0$ is a constant, $\{\mathscr{F}_t, \mathscr{F}_t\}_{t\ge 0}$ and $\{G_t, \mathscr{F}_t\}_{t\ge 0}$ are continuous nonnegative adapted processes, and $\mathscr{L}_t(x)$ is defined by

$$\mathscr{L}_t(x) = x_t^T \left[\frac{dQ_t(x)}{dt} \right] x_t + 2x_t^T Q_t(x) a(t,x) + b^T(t,x) Q_t(x) b(t,x). \tag{2.14}$$

Then the SDE (2.10) *has a unique strong solution on $[0,\infty)$.*

The proof is given in Appendix B.

We are now in a position to prove Theorem 2.1.

Proof of Theorem 2.1. In order to apply Lemma 2.2, we only need to verify (2.11) and (2.13), since $a(t,0)$ and $b(t,0)$ are continuous random processes. To emphasize the dependence of $\varphi_t, P_t, f_t, \ldots$ on x defined by (2.6), we will write them as $\varphi_t(x), P_t(x), f_t(x), \ldots$. For any $T > 0$, $x, z \in C^d[0,\infty)$, $(d = p + q + 2r)$, $\|x\|_{[0,T]} \le n$, $\|z\|_{[0,T]} \le n$, $n \ge 1$, by (2.6)–(2.9) we can derive that

$$\|b(t,x) - b(t,z)\| = \|P_t(x)\varphi_t(x) - P_t(z)\varphi_t(z)\| \le L_t^{(n)} \left\{ \|x_t - z_t\| + \int_0^t \|x_s - z_s\| \, ds \right\}, \quad \forall t \le T, \tag{2.15}$$

where $L_t^{(n)}$ is $\mathscr{F}_t$-measurable and continuous because $\{y_t, u_t\}$ is continuous. We now proceed to consider $a(t,x)$. By (2.9), $\forall t \le T$,

$$\|g_t(x) - g_t(z)\| = \|\varphi_t(x)^T \tilde{\theta}_t(x) - \varphi_t(z)^T \tilde{\theta}_t(z)\| \le N_t^{(n)} \|x_t - z_t\| \tag{2.16}$$

holds for some continuous $N_t^{(n)}$ depending on t, n and $\{y_s, u_s, s \le t\}$.

Set $h_t = C^{-1}(S)x_t, x \in C^d[0, \infty)$, $H_t = [h_t, \ldots, S^{r-1}h_t]^{\mathrm{T}}$ and

$$F_c = \begin{Bmatrix} -c_1 & \cdots & -c_{r-1} & -c_r \\ 1 & & & \\ & \ddots & & \\ & & 1 & 0 \end{Bmatrix}.$$

Then we have $H_t = F_c S H_t + [x_t, 0 \ldots 0]^{\mathrm{T}}$, and so

$$H_t = F_c \int_0^t \exp\{F_c(t-s)\}[x_s, 0 \ldots 0]^{\mathrm{T}} \mathrm{d}s + [x_t, 0 \ldots 0]^{\mathrm{T}}.$$

Hence by the fact that all eigenvalues of F_c are negative, we get

$$\begin{aligned} \|D(S)C^{-1}(S)x_t\| &= \|h_t + D_1 S h_t + \cdots + D_r S^r h_t\| \\ &\le \mathrm{O}(\|H_t\| + \|SH_t\|) = \mathrm{O}(\|H_t\| + \|F_c^{-1}(H_t - [x_t, 0 \ldots 0]^{\mathrm{T}})\|) \\ &\le \mathrm{O}(\|H_t\| + \|x_t\|) \le \mathrm{O}\left(\int_0^t \|x_s\| \mathrm{d}s + \|x_t\|\right). \end{aligned} \tag{2.17}$$

Replacing x_t by $g_t(x) - g_t(z)$ and noting (2.8) and (2.16) we have

$$\begin{aligned} \|f_t(x) - f_t(z)\| &= \|D(S)C^{-1}(S)[g_t(x) - g_t(z)]\| \\ &= K_t^{(n)}\left(\int_0^t \|x_s - z_s\| \mathrm{d}s + \|x_t - z_t\|\right) \end{aligned} \tag{2.18}$$

where $K_t^{(n)}$ is a continuous process adapted to $\{\mathscr{F}_t\}$.

Combining (2.15), (2.16) and (2.18) it is easy to see from definition (2.7) that $a(t,x)$ and $b(t,x)$ satisfy the local Lipschitz condition (2.11).

We now proceed to verify (2.13).

Since $D(S)C^{-1}(S) - \frac{1}{2}$ is strictly positive real, by (2.8) there are constants $\varepsilon > 0$ and $K_0 \ge 0$ such that

$$\int_0^t \left[f_s - \left(\frac{1}{2} + \varepsilon\right) g_s\right] g_s \mathrm{d}s + K_0 \ge 0, \quad \forall t \ge 0, \ x \in C^d[0, \infty). \tag{2.19}$$

Define

$$Q_t(x) = \begin{pmatrix} P_t^{-1} & 0 \\ 0 & \varepsilon I_r \end{pmatrix}. \tag{2.20}$$

It is easy to see that $Q_t(x)$ is symmetric, uniformly positive definite, $\mathscr{F}_t \times \mathscr{B}_t$-measurable, and such that

$$\|b^{\mathrm{T}}(t,x)Q_t(x)x_t\|^2 = \|\varphi_t^{\mathrm{T}}\tilde{\theta}_t\|^2 = g_t^2. \tag{2.21}$$

By (2.14), (2.21) and the definitions for $a(t,x)$, $b(t,x)$ and $Q_t(x)$ we have

$$\begin{aligned} \mathscr{L}_t(x) &= g_t^2 - 2\tilde{\theta}_t^{\mathrm{T}}\varphi_t f_t - 2\varepsilon\tilde{v}_t g_t + 2\varepsilon(\tilde{v}_t, \ldots, S^{r-1}\tilde{v}_t)\begin{pmatrix} -\bar{C}(S)\tilde{v}_t \\ \tilde{v}_t \\ \vdots \\ S^{r-2}\tilde{v}_t \end{pmatrix} + \varphi_t^{\mathrm{T}} P_t \varphi_t \\ &\le g_t^2 - 2g_t f_t + \varepsilon g_t^2 + \varepsilon\|\tilde{v}_t\|^2 + \varepsilon K_1\|x_t\|^2 + \varphi_t^{\mathrm{T}} P_t \varphi_t \end{aligned}$$

$$\leq -2\left[f_t - \left(\frac{1}{2} + \varepsilon\right) g_t\right] g_t - \varepsilon g_t^2 + \varepsilon(1 + K_1)\|x_t\|^2 + \|\varphi_t\|^2 a^{-1}$$

$$\leq -2\left[f_t - \left(\frac{1}{2} + \varepsilon\right) g_t\right] g_t - \varepsilon \|b^{\mathrm{T}}(t,x) Q_t(x) x_t\|^2$$

$$+ 2a^{-1}\|\varphi_t^0\|^2 + [\varepsilon(1 + K_1) + 2a^{-1}]\|x_t\|^2$$

where K_1 is some constant. Hence by (2.19) we have

$$\int_0^t \mathscr{L}_s(x)\,\mathrm{d}s \leq 2K_0 + 2a^{-1}\int_0^t \|\varphi_t^0\|^2\,\mathrm{d}s - \varepsilon\int_0^t \|b^{\mathrm{T}}(s,x)Q_s(x)x_s\|^2\,\mathrm{d}s + K_2\int_0^t \|x_s\|^2\,\mathrm{d}s,$$

where K_2 is a constant. Threfore, (2.13) is verified and the proof of Theorem 2.1 is complete. □

3. Convergence and Excitation

Having established the existence and uniqueness of the ELS estimate, we discuss some of its asymptotic properties in this section. The following results are a continuous-time analogue of those for discrete-time systems (see [4, 10]).

Theorem 3.1. *Let the conditions of Theorem 2.1 be satisfied. Then as $t \to \infty$ the estimation error $\hat{\theta}_t - \theta$ produced by the ELS estimate* (1.8)–(1.11) *has the following convergence rate*:

$$\|\hat{\theta}_t - \theta\|^2 = 0\left(\frac{\log r_t^0}{\lambda_{\min}\left(\int_0^t \varphi_s^0 \varphi_s^{0\mathrm{T}}\,\mathrm{d}s + aI\right)}\right) \text{ a.s.,} \tag{3.1}$$

provided that

$$\frac{\log r_t^0}{\lambda_{\min}\left(\int_0^t \varphi_s^0 \varphi_s^{0\mathrm{T}}\,\mathrm{d}s + aI\right)} \to 0 \text{ a.s.}$$

where φ_t^0 is defined by (1.4) *and $r_t^0 = e + \int_0^t \|\varphi_s^0\|^2\,\mathrm{d}s$.*

Proof. First, note that (cf. [2, Lemma 2])

$$\int_0^t \|\tilde{\theta}_s^{\mathrm{T}}\varphi_s\|^2\,\mathrm{d}s = 0(\log r_t) \text{ a.s.} \tag{3.2}$$

Set $\tilde{V}_t = [\mathrm{d}\tilde{v}_t/\mathrm{d}t, \tilde{v}_t, \ldots, S^{r-2}\tilde{v}_t]^{\mathrm{T}}$. Then by (2.3) we have $\tilde{V}_t = F_c S \tilde{V}_t + [-\tilde{\theta}_t^{\mathrm{T}}\varphi_t, 0 \ldots 0]^{\mathrm{T}}$, and so

$$\tilde{V}_t = F_c \int_0^t \exp\{F_c(t-s)\}[-\tilde{\theta}_s^{\mathrm{T}}\varphi_s, 0 \ldots 0]^{\mathrm{T}}\,\mathrm{d}s + [-\tilde{\theta}_t^{\mathrm{T}}\varphi_t, 0 \ldots 0]^{\mathrm{T}},$$

where F_c is defined in the proof of Theorem 2.1. By the stability of F_c and (3.2) it is obvious that $\int_0^t \|\tilde{V}_s\|^2\,\mathrm{d}s = 0(\log r_t)$, and so

$$\int_0^t \|S\tilde{V}_s\|^2\,\mathrm{d}s = \int_0^t \|F_c^{-1}[\tilde{V}_s - (\tilde{\theta}_s^{\mathrm{T}}\varphi_s 0 \ldots 0)^{\mathrm{T}}]\|^2\,\mathrm{d}s = 0(\log r_t). \tag{3.3}$$

Hence

$$\int_0^t [\|\tilde{v}_s\|^2 + \cdots + \|S^{r-1}\tilde{v}_s\|^2]\,\mathrm{d}s = 0(\log r_t).$$

By this and the definitions for φ_t and φ_t^0 we have

$$\int_0^t \|\varphi_s - \varphi_s^0\|^2 \, ds = 0(\log r_t) \text{ a.s.},$$

and consequently it follows that

$$r_t = 0(r_t^0) \text{ a.s.} \tag{3.4}$$

On the other hand, for any $x \in \mathscr{R}^{p+q+r}$, $\|x\| = 1$, we have

$$\begin{aligned}\lambda_{\min}\left(\int_0^t \varphi_s^0 \varphi_s^{0\mathrm{T}} \, ds + aI\right) &\le 2\int_0^t [\|\varphi_s^0 - \varphi_s\|^2 + \|x^{\mathrm{T}}\varphi_s\|^2] \, ds + a \\ &\le 2x^{\mathrm{T}}\left(\int_0^t \varphi_s\varphi_s^{\mathrm{T}} \, ds + aI\right)x + 0(\log r_t).\end{aligned}$$

From this, (3.4) and the assumption it is easy to get

$$\lambda_{\min}\left(\int_0^t \varphi_s^0 \varphi_s^{0\mathrm{T}} \, ds + aI\right) = 0\left[\lambda_{\min}\left(\int_0^t \varphi_s\varphi_s^{\mathrm{T}} \, ds + aI\right)\right].$$

Finally, the desired conclusion follows from this, (3.4) and [2, Remark 1]. $\square$

Theorem 3.1 shows that the strong consistency of the ELS is closely related to the growth rate of $\lambda_{\min}(\int_0^t \varphi_s^0 \varphi_s^{0\mathrm{T}} \, ds)$. We now give an explicit connection between $\lambda_{\min}(\int_0^t \varphi_s^0 \varphi_s^{0\mathrm{T}} \, ds)$ and the 'input' process $\{u_t, v_t\}$.

Lemma 3.1. *For the system* (1.1), *assume that the polynomials $A(S)$, $B(S)$ and $C(S)$ have no common factor, and $|a_p| + |b_q| + |c_r| > 0$. Then there exists a constant $c > 0$ such that*

$$\lambda_{\min}\left(\int_0^t Z_t Z_t^{\mathrm{T}} \, dt\right) \le c\lambda_{\min}\left(\int_0^t \varphi_t^0 \varphi_t^{0\mathrm{T}} \, dt\right) \quad \text{a.s.}, \ \forall T > 0, \tag{3.5}$$

where Z_t is defined by ($k = p + \max\{q, r\}$), and

$$Z_t = [z_t^{\mathrm{T}}, \ldots, S^{k-1} z_t^{\mathrm{T}}]^{\mathrm{T}}, \ z_t = \frac{1}{E(S)}[u_t, v_t]^{\mathrm{T}} \tag{3.6}$$

with $E(S)$ being any monic stable polynomial of S and $\deg\{E(S)\} \ge p$.

Proof. *Set* $\bar{\varphi}_t^0 = [A(S)/E(S)]\,\varphi_t^0$. Then by (1.1) and (1.4) we have

$$\bar{\varphi}_t^0 = \frac{1}{E(S)}\left\{\begin{array}{l} SB(S)u_t + C(S)v_t \\ \vdots \\ S^pB(S)u_t + S^{p-1}C(S)v_t \\ A(S)u_t \\ \vdots \\ S^{q-1}A(S)u_t \\ A(S)v_t \\ \vdots \\ S^{r-1}A(S)v_t \end{array}\right\} = \left\{\begin{array}{ll} SB(S) & C(S) \\ \vdots & \vdots \\ S^pB(S) & S^{p-1}C(S) \\ A(S) & 0 \\ \vdots & \vdots \\ S^{q-1}A(S) & 0 \\ 0 & A(S) \\ \vdots & \vdots \\ 0 & S^{r-1}A(S) \end{array}\right\} z_t$$

$$\triangleq H(S)z_t,$$

where $H(S)$ is a $(p+q+r)\times 2$ polynomial matrix defined in an obvious way.

For any $x\in\mathscr{R}^{p+q+r}$, $\|x\|=1$, let us define h_i, g_i, $0\le i\le k-1$, via $x^{\mathrm{T}}H(S)=\sum_{i=0}^{k-1}[h_i,g_i]S^i$. Since $A(S)$, $B(S)$ and $C(S)$ have no common factor, by a similar argument to that used in [3, pp. 209–211] it is easy to show that $\inf_{\|x\|=1}\sum_{i=0}^{k-1}(h_i^2+g_i^2)>0$. Hence we have for some $c_0>0$,

$$\lambda_{\min}\left(\int_0^{\mathrm{T}}\bar{\varphi}_t^0\bar{\varphi}_t^{0\mathrm{T}}\,\mathrm{d}t\right)=\inf_{\|x\|=1}\int_0^{\mathrm{T}}\left[\sum_{i=0}^{k-1}(h_i,g_i)S^iz_t\right]^2\mathrm{d}t$$

$$\ge\inf_{\|x\|=1}\sum_{i=0}^{k-1}(f_i^2+g_i^2)\,\lambda_{\min}\left(\int_0^{\mathrm{T}}Z_tZ_t^{\mathrm{T}}\,\mathrm{d}t\right)\ge c_0\lambda_{\min}\left(\int_0^{\mathrm{T}}Z_tZ_t^{\mathrm{T}}\,\mathrm{d}t\right). \tag{3.7}$$

Next, applying Lemma 1 in [2], we have for some $c_1>0$,

$$\lambda_{\min}\left(\int_0^{\mathrm{T}}\bar{\varphi}_t^0\bar{\varphi}_t^{0\mathrm{T}}\,\mathrm{d}t\right)=\inf_{\|x\|=1}\int_0^{\mathrm{T}}\left(\frac{A(S)}{E(S)}x^{\mathrm{T}}\varphi_t^0\right)^2\mathrm{d}t$$

$$\le c_1\inf_{\|x\|=1}\int_0^{\mathrm{T}}(x^{\mathrm{T}}\varphi_t^0)^2\,\mathrm{d}t=c_1\lambda_{\min}\left(\int_0^{\mathrm{T}}\varphi_t^0\varphi_t^{0\mathrm{T}}\,\mathrm{d}t\right). \tag{3.8}$$

Finally, the desired result follows from (3.7) and (3.8). □

Theorem 3.2. *Assume that the input process of system* (1.1) *is a Gaussian* (*asymptotically*) *stationary process with rational spectrum density, which has the form* $u_t=[P(S)/Q(S)]B_t$, $u_0=0$, *where* $\{B_t\}$ *is a Brownian motion independent of* $\{w_t\}$, *and* $P(S)$ *and* $Q(S)$ *are monic stable polynomials with* $\deg P(S)+1\le\deg Q(S)$. *If* $D(S)C^{-1}(S)-\frac{1}{2}$ *is strictly positive real,* $A(S)\ne 0$, $\forall\,\mathrm{Re}\{S\}>0$, $A(S)$, $B(S)$ *and* $C(S)$ *have no common factor and* $|a_p|+|b_q|+|c_r|>0$, *then the ELS estimate has the following convergence rate*:

$$\|\hat{\theta}_t-\theta\|^2=\mathrm{O}\left(\frac{\log t}{t}\right)\quad a.s.\quad as\ t\to\infty.$$

Proof. Since $\{u_t\}$ and $\{v_t\}$ are asymptotically stationary processes and $A(S)$ has no zeros in the open right-half plane, it is easy to convince oneself that

$$\log r_t^0=\mathrm{O}(\log t)\ \text{a.s.}, \tag{3.9}$$

where r_t^0 is defined in Theorem 3.1.

Take $E(S)=P(S)E^0(S)$, with $E^0(S)$ monic, stable and satisfying

$$\deg E^0(S)\ge p+\max(q,r)-\deg P(S).$$

The last requirement guarantees that both the degrees of $E^0(S)Q(S)$ and $E(S)D(S)$ are not less than $p+\max(q,r)$.

Let us set $z_t^{(1)}=E^{-1}(S)u_t$, $z_t^{(2)}=E^{-1}(S)v_t$. Then we have

$$z_t^{(1)}=\frac{1}{E^0(S)Q(S)}B_t,\qquad z_t^{(2)}=\frac{1}{E(S)D(S)}w_t.$$

Hence, by setting

$$Z_t^{(0)}=[Z_t^{(1)\mathrm{T}},Z_t^{(2)\mathrm{T}}]^{\mathrm{T}},\qquad W_t^0=[B_t,w_t]^{\mathrm{T}},$$

$$Z_t^{(1)}=[z_t^{(1)},\ldots,S^{k_1-1}z_t^{(1)}]^{\mathrm{T}},\quad k_1=\deg[E^0(S)Q(S)],$$

$$Z_t^{(2)}=[z_t^{(2)},\ldots,S^{k_2-1}z_t^{(2)}]^{\mathrm{T}},\quad k_2=\deg[E(S)D(S)],$$

we know that there is a controllable pair (F, G) with F stable such that $\mathrm{d}Z_t^0 = FZ_t^0\,\mathrm{d}t + G\,\mathrm{d}W_t^0$. Since W_t^0 is also a Wiener process, we have (cf. [2, Lemma 3])

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T Z_t^0 Z_t^{0\mathrm{T}}\,\mathrm{d}t = R > 0 \text{ a.s.,}$$

which, in conjunction with Lemma 3.1 and the fact that $\min\{k_1, k_2\} \geq p + \max\{q, r\}$, implies

$$\liminf_{T\to\infty} \frac{1}{T}\lambda_{\min}\left(\int_0^T \varphi_t^0\varphi_t^{0\mathrm{T}}\,\mathrm{d}t\right) > 0 \text{ a.s.}$$

Hence the desired result follows from this, (3.9) and Theorem 3.1. □

Appendix A

Proof of Lemma 2.1

First of all, we note that if in (2.11) and (2.12) the process $\{L_t^{(n)}\}$ does not depend on n, then by the standard truncation methods and the familiar results on the SDE (2.10) (e.g. [11, Theorem 4.6]), it can easily be shown that the SDE (2.10) has a unique strong solution x_t on $[0, T]$.

Next, let us define $n \geq 1$, $\xi \in \mathscr{R}^d$ and $x \in C^d[0, T]$,

$$g_n(\xi) = \xi \min\left(1, \frac{n}{\|\xi\|}\right), \qquad g_n(x) = [g_n(x_t)]_{t\leq T},$$

$$a_n(t, x) = a(t, g_n(x)), \qquad b_n(t, x) = b(t, g_n(x)).$$

Then it is obvious that

(i) $g_n(x) \in C^d[0, T]$ whenever $x \in C^d[0, T]$,

(ii) $\|g_n(x)\|_{[0,T]} \leq n$, $\forall x \in C^d[0, T]$,

(iii) $\|g_n(\xi) - g_n(z)\| \leq n\|\xi - z\|$, $\forall \xi, z \in \mathscr{R}^d$.

Hence we know that for each fixed $n \geq 1$, the SDE

$$\mathrm{d}x_t = a_n(t, x)\,\mathrm{d}t + b_n(t, x)\,\mathrm{d}w_t, \quad x_0 = \eta, \tag{A.1}$$

has a unique strong solution $x^{(n)} = (x_t^{(n)})_{t\leq T}$. Set

$$\sigma_n = \begin{cases} \inf\{t \leq T\colon \sup_{s\leq t}\|x_s^{(n)}\| \geq n\}, \\ T, \text{ if } \sup_{s\leq T}\|x_s^{(n)}\| < n. \end{cases}$$

Then by a standard treatment (cf. [11, p. 143]) it is not difficult to show that $\sigma_n \leq \sigma_{n+1}$ a.s., and $x_t^{(n+1)} = x_t^{(n)}$ *a.s., on* $[t \leq \sigma_n]$, $\forall n$. Let us define $\sigma_T \triangleq \lim_{n\to\infty}\sigma_n$, *and* $x_t \triangleq x_t^{(n)}$ *on* $[t \leq \sigma_n]$. Then it is easy to see that $x = (x_t)_{t<\sigma_T}$ is the unique local strong solution of the SDE (2.10). Moreover, since $x_t^{(n)}$ is continuous in t, we have $\sup_{s<\sigma_n}\|x_s^{(n)}\| = n$, *on* $[\sigma_n < T]$. Consequently, on the set $[\sigma_T < T]$, $\sup_{s<\sigma_T}\|x_s\| \geq \sup_{s<\sigma_n}\|x_s^{(n)}\| \to \infty$, as $n \to \infty$, which completes the proof.

Appendix B

Proof of Lemma 2.2

Let us keep the notations introduced in the proof of Lemma 2.1. For the desired result, it suffices to prove that $P(\sigma = \infty) = 1$, where $\sigma = \lim_{T\to\infty}\sigma_T$.

Since for any $T > 0$ and $n \geq 1$, $x^{(n)} = (x_t^{(n)})_{t \leq T}$ satisfies the SDE

$$\mathrm{d}x_t^{(n)} = a_n(t, x^{(n)})\,\mathrm{d}t + b_n(t, x^{(n)})\,\mathrm{d}w_t, \quad x_0^{(n)} = \eta,$$

by Ito's formula we have ($t \wedge \sigma_n \triangleq \min\{t, \sigma_n\}$),

$$x_{t\wedge\sigma_n}^{(n)\mathrm{T}} Q_{t\wedge\sigma_n}(x^{(n)}) x_{t\wedge\sigma_n}^{(n)} = \eta^{\mathrm{T}} Q_0(x^{(n)})\eta + \int_0^{t\wedge\sigma_n} \mathscr{L}_s^{(n)}(x^{(n)})\,\mathrm{d}s + 2\int_0^{t\wedge\sigma_n} b_n^{\mathrm{T}}(s, x^{(n)}) Q_s(x^{(n)}) x_s^{(n)}\,\mathrm{d}w_s$$

where

$$\mathscr{L}_t^{(n)}(x^{(n)}) = x_t^{(n)\mathrm{T}}\left[\frac{\mathrm{d}Q_t(x^{(n)})}{\mathrm{d}t}\right] x_t^{(n)} + 2x_t^{(n)\mathrm{T}} Q_t(x^{(n)}) a_n(t, x^{(n)})$$
$$+ b_n^{\mathrm{T}}(t, x^{(n)}) Q_t(x^{(n)}) b_n(t, x^{(n)}).$$

Note that on the set $[s \leq \sigma_n]$, we have $\| x_s^{(n)} \| \leq n$ and

$$a_n(s, x^{(n)}) = a(s, g_n(x^{(n)})) = a(s, x^{(n)}),$$
$$b_n(s, x^{(n)}) = b(s, g_n(x^{(n)})) = b(s, x^{(n)}).$$

Also note that $x_0^{(n)} = \eta$, $Q_0(x^{(n)}) \in \mathscr{F}_0$, which are actually independent on n. Hence, by assumption (2.13) we get

$$x_{t\wedge\sigma_n}^{(n)\mathrm{T}} Q_{t\wedge\sigma_n}(x^{(n)}) x_{t\wedge\sigma_n}^{(n)} \leq \eta^{\mathrm{T}} Q_0(x^{(n)})\eta + F_{t\wedge\sigma_n} + G_{t\wedge\sigma_n}\int_0^{t\wedge\sigma_n} \| x_s^{(n)} \|^2\,\mathrm{d}s$$
$$-\varepsilon\int_0^{t\wedge\sigma_n} \| b^{\mathrm{T}}(s, x^{(n)}) Q_s(x^{(n)}) x_s^{(n)} \|^2\,\mathrm{d}s + 2\int_0^{t\wedge\sigma_n} b^{\mathrm{T}}(s, x^{(n)}) Q_s(x^{(n)}) x_s^{(n)}\,\mathrm{d}w_s$$
$$\leq F_t^* + G_t^* \int_0^{t\wedge\sigma_n} \| x_s^{(n)} \|^2 I(s < \sigma_n)\,\mathrm{d}s + H_t(n), \tag{B.1}$$

where

$$F_t^* = \max_{0\leq s\leq t} F_t + \eta^{\mathrm{T}} Q_0(x^{(n)})\eta, \qquad G_t^* = \max_{0\leq s\leq t} G_s$$

$$H_t(n) = -\varepsilon\int_0^t \| b^{\mathrm{T}}(s, x^{(n)}) Q_s(x^{(n)}) x_s^{(n)} \|^2 I(s < \sigma_n)\,\mathrm{d}s$$
$$+ 2\int_0^t b^{\mathrm{T}}(s, x^{(n)}) Q_s(x^{(n)}) x_s^{(n)} I(s < \sigma_n)\,\mathrm{d}w_s.$$

By the assumption there is a random variable $\alpha > 0$ such that $\inf_{t,n} \lambda_{\min}(Q_t(x^{(n)}) \geq \alpha > 0$ a.s. Hence by (B.1) we know that

$$\| x_{t\wedge\sigma_n}^{(n)} \|^2 \leq \frac{1}{\alpha} x_{t\wedge\sigma_n}^{(n)\mathrm{T}} Q_{t\wedge\sigma_n}(x^{(n)}) x_{t\wedge\sigma_n}^{(n)}$$
$$\leq \frac{1}{\alpha}[F_t^* + H_t(n) + G_t^* \int_0^t \| x_{s\wedge\sigma_n}^{(n)} \|^2\,\mathrm{d}s]. \tag{B.2}$$

Note that by the well-known exponential inequality for stochastic integrals (see, e.g. [6, Chap. 4]) we know that

$$P\left(\sup_{0\leq s\leq t} H_s(n) \geq n\right) \leq \exp\left(-\frac{\varepsilon}{2} n\right).$$

Consequently, apply the Gronwall inequality to (B.2), we see that

$$P(\| x^{(n)}_{t\wedge\sigma_n} \|^2 \geq n^2) \to 0 \quad \text{as } n \to \infty \text{ for any fixed } t.$$

Therefore, for any fixed $t < T$,

$$\begin{aligned} P(\sigma_T \leq t) \leq P(\sigma_n \leq t) &= P(\sigma_n \leq t, \| x^{(n)}_{\sigma_n} \| \geq n) \\ &= P(\sigma_n \leq t, \| x^{(n)}_{t\wedge\sigma_n} \| \geq n) \leq P(\| x^{(n)}_{t\wedge\sigma_n} \| \geq n) \to 0, \quad \text{as } n \to \infty . \end{aligned}$$

Hence $P(\sigma_T < T) = 0$ *or* $P(\sigma_T = T) = 1$, and consequently $P(\sigma = \infty) = 1$. This completes the proof.

References

[1] H.F. Chen, *Recursive Estimation and Control for Stochastic Systems* (Wiley, New York, 1985).

[2] H.F. Chen and L. Guo, Continuous-time stochastic adaptive tracking: robustness and asymptotic properties, *SIAM J. Control Optim.* **28** (1990) 513–527.

[3] H.F. Chen and L. Guo, *Identification and Stochastic Adaptive Control* (Birkhäuser, Boston, 1991).

[4] H.F. Chen and L. Guo, Convergence rate of least squares identification and adaptive control for stochastic systems, *Int. J. Control* **44** (1986) 1459–1476.

[5] H.F. Chen and J.B. Moore, Convergence rate of continuous-time stochastic ELS parameter estimation, *IEEE Trans. Automat. Control* **32** (1987) 267–269.

[6] A. Friedman, *Stochastic Differential Equations and Applications*, Vol. 1 (Academic Press, New York, 1975).

[7] I. Gyongy and N.V. Krylov, On stochastic equations with respect to semimartingales I, *Stochastics* **4** (1980) 1–12.

[8] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* (North-Holland, Amsterdam, 1981).

[9] N.V. Krylov and B.L. Rozovskii, Ito equations in Banach spaces, *Soviet Math.* **16** (1991) 1233–1277.

[10] T.L. Lai and C.Z. Wei, Extended least squares and their application to adaptive control and prediction in linear systems, *IEEE Trans. Automat. Control* **AC-31** (1986) 898–906.

[11] R.S. Lipster and A.N. Shiryayev, *Statistics of Random Processes I, II* (Springer, New York, 1977).

[12] J.B. Moore, Convergence of continuous time stochastic ELS parameter estimation, *Stochastic Process. Appl.* **27** (1988) 195–215.