Strong Consistency of Recursive Identification for Hammerstein Systems With Discontinuous Piecewise-Linear Memoryless Block

Han-Fu Chen

***Abstract*—This note deals with identification of Hammerstein systems with discontinuous piecewise-linear memoryless block followed by a linear subsystem. Recursive algorithms are proposed for estimating coefficients of the linear subsystem and six unknown parameters contained in the nonlinear static block. By taking a sequence of iid random variables with uniform distribution to serve as the system input, strong consistency is proved for all estimates given in the note. The theoretical results are verified by computer simulation.**

***Index Terms*—Hammerstein system, least squares, parametric approach, recursive estimation, strong consistency.**

I. INTRODUCTION

The system consisting of a static nonlinear block followed by a linear dynamic system is a useful model for many areas and is called the Hammerstein system. Because of the importance of Hammerstein systems in many control applications, its identification issue has been attracting many scientists. The block diagram of the Hammerstein system is presented in Fig. 1.

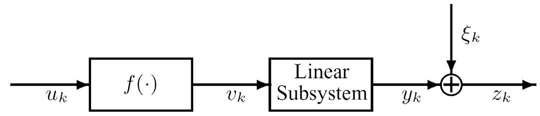

Fig. 1. Hammerstein system.

This is a single-input–single-output (SISO) system with input u_k, output y_k, and the memoryless nonlinearity $f(\cdot)$. The output y_k is observed with additive noise ξ_k, and the observation is z_k. For identifying the nonlinearity, there are parametric [1], [3], [4], [13], [16], [17] and nonparametric [2], [6], [9]–[12], [15] approaches. In the nonparametric approach, the unknown function $f(u)$ at an arbitrarily fixed u, where $f(\cdot)$ is continuous, may be directly identified [6], [10], [11], but $f(\cdot)$ may also be identified with the help of its approximation by smooth functions, e.g., by an approximating polynomial in [12] and by series expansion in [15].

In the parametric approach the estimates for unknown parameters are usually obtained by minimizing some loss function formed from data of fixed size. In this case the estimates are nonrecursive.

As pointed in [1], [17], the discontinuous nonsmooth nonlinearities, for example, the two-segment piecewise-linear with preloads and dead zones, are common in engineering practice. The nonlinearity $f(\cdot)$ considered in this note is similar to but more general than those discussed in [1], [17]. To be precise, it is a discontinuous piecewise-linear function containing six unknown parameters $c^+, c^-, b^+, b^-, d^+, d^-$ expressed as

$$f(u) = \begin{cases} c^+(u-d^+)+b^+, & u > d^+ \\ 0, & -d^- \le u \le d^+ \\ c^-(u+d^-)-b^-, & u < -d^-. \end{cases} \tag{1}$$

Since a stable autoregression and moving average (ARMA) model can be approximated by a moving average (MA) model with any accuracy if the order of the MA model is sufficiently high, let us assume the linear subsystem to be given by the MA system as follows:

$$y_{k+1} = D(z)v_k \tag{2}$$

where

$$D(z) = 1 + d_1 z + \cdots + d_q z^q \quad zy_k = y_{k-1}. \tag{3}$$

The purpose of the note is to recursively estimate the coefficients $[d_1, \ldots, d_q]$ of the linear subsystem and six parameters c^+, c^-, b^+, b^-, d^+, and d^- contained in the nonlinear block on the basis of the observation $\{u_k, z_k\}$ and to prove the strong consistence of the estimates. The system input $\{u_k\}$ is at our disposal.

It is worth noting that because of discontinuity of $f(\cdot)$ for its identification many methods consisting in representing $f(\cdot)$ by a finite combination of continuous functions cannot be used, and the methods proposed in [6] and [10] can neither be applied because in [6] and [10] $f(\cdot)$ is estimated only at u where $f(\cdot)$ is continuous.

The rest of the note is arranged as follows. The estimation algorithms are given in Section II and their convergence to true parameters is proved in Section III. A numerical example is given in Section IV and some concluding remarks are included in Section V.

II. RECURSIVE ESTIMATION ALGORITHMS

Before defining estimation algorithms for $[d_1, \ldots, d_q]$, c^+, c^-, b^+, b^-, d^+, and d^-, we first list conditions to be used later on.

A1) $D(z)$ is stable, i.e., all roots of $D(z)$ are outside the closed unit disk.

A2) The upper bound U for d^+ and d^- is available

$$0 \le d^+ < U \quad \text{and} \quad 0 \le d^- < U.$$

A3) The observation noise $\{\xi_k\}$ is a sequence of mutually independent random variables with $E\xi_k = 0$ and $\sup_k E\xi_k^2 < \infty$.

As the system input, let us take $\{u_k\}$ to be a sequence of mutually independent and identically distributed (iid) random variables with uniform distribution over $[-2U, 2U]$ and independent of $\{\xi_k\}$.

A. Estimation Algorithms for the Linear Subsystem

Since without the strictly positive realness condition on $D(z)$ the extended least squares (ELS) estimate for $[d_1, \ldots, d_q]$ may be inconsistent [7], in lieu of ELS we apply the stochastic approximation algorithm with expanding truncations [5] to estimate the coefficients of the linear subsystem.

In order to define the increasing truncations, let $\{M_k\}$ be a sequence of positive real numbers with $M_{k+1} > M_k, \forall k$ and $M_k \xrightarrow[k\to\infty]{} \infty$. With arbitrary initial values $\theta_0(i),\ i = 0, 1, \ldots, q$, define

$$\theta_{k+1}(i) = (\theta_k(i) - a_k(\theta_k(i) - u_k z_{k+1+i})) \cdot I_{[|\theta_k(i)-a_k(\theta_k(i)-u_k z_{k+1+i})| \le M_{\sigma_k(i)}]} \tag{4}$$

$$\sigma_k(i) = \sum_{j=1}^{k-1} I_{[|\theta_j(i)-a_j(\theta_j(i)-u_j z_{j+1+i})| > M_{\sigma_j(i)}]}$$

$$\sigma_0(i) = 0 \quad a_k = \frac{1}{k} \tag{5}$$

Manuscript received May 13, 2004; revised February 25, 2005 and May 17, 2005. Recommended by Associate Editor L. Ljung. This work was supported by the National Science Foundation of China under Grants G0221301, 60334040, and 60474004.

The author is with the Key Laboratory of Systems and Control, Institute of Systems Science, AMSS, Chinese Academy of Sciences, Beijing 100080, China (e-mail: hfchen@iss.ac.cn).

Digital Object Identifier 10.1109/TAC.2005.856658

where I_A denotes the indicator of a random event A with $I_A = 1$ if $\omega \in A$, and $I_A = 0$, if $\omega \notin A$. Note that $\theta_k(i)$ is used to estimate $\rho \triangleq Eu_1v_1$, and $\theta_k(0)$ for $\rho d_i, i = 1, \ldots, q$, respectively.

B. Estimation Algorithms for c^+, h^+, c^-, and h^-

We now estimate c^+, c^-, h^+, and h^-, where $h^+ \triangleq c^+d^+ - b^+$, and $h^- \triangleq c^-d^- - b^-$.

In order to avoid the possible division by zero we modify $\theta_k(0)$ as follows:

$$\rho_k \triangleq \begin{cases} \theta_k(0), & \text{if } |\theta_k(0)| \geq \frac{1}{k} \\ (\text{sign}\theta_k(0))\frac{1}{k}, & \text{if } |\theta_k(0)| < \frac{1}{k}. \end{cases} \tag{6}$$

Define the estimate d_{ik} for d_i

$$d_{ik} \triangleq \frac{\theta_k(i)}{\rho_k}. \tag{7}$$

Further define $q \times q$-matrices D and D_k and q-dimensional vector H as follows:

$$D \triangleq \begin{pmatrix} -d_1 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ \vdots & & & \\ \vdots & & & 1 \\ -d_q & 0 & & 0 \end{pmatrix}$$

$$D_k \triangleq \begin{pmatrix} -d_{1k} & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ \vdots & & & \\ \vdots & & & 1 \\ -d_{qk} & 0 & & 0 \end{pmatrix}$$

$$H^T \triangleq (1 \quad 0 \quad \cdots \quad 0). \tag{8}$$

Recursively define $\hat{x}_k$ with an arbitrary initial $\hat{x}_0$

$$\hat{x}_k = D_k\hat{x}_{k-1} + Hz_{k+1}. \tag{9}$$

and define the estimate $\hat{v}_k$ for v_k, the output of the nonlinear block

$$\hat{v}_k \triangleq H^T\hat{x}_k. \tag{10}$$

Let

$$w_k^+ \triangleq \hat{v}_k I_{[u_k \geq U]}. \tag{11}$$

and

$$\mu^+ \triangleq [c^+, h^+]^T \quad \phi_k^+ \triangleq [u_k, -1]^T I_{[u_k \geq U]}. \tag{12}$$

Noticing that $f(u) = c^+u - h^+$ for $u \geq U$ or $v_k = \mu^{+T}\phi_k^+$ for $u_k \geq U$, it is natural to estimate μ^+ by the least squares (LS) algorithm [7], [14]

$$\mu_k^+ = \mu_{k-1}^+ + a_k^+P_k^+\phi_k^+\left(w_k^+ - \phi_k^{+T}\mu_{k-1}^+\right) \tag{13}$$

$$P_{k+1}^+ = P_k^+ - a_k^+P_k^+\phi_k^+\phi_k^{+T}P_k^+$$

$$a_k^+ = \left(1 + \phi_k^{+T}P_k^+\phi_k^+\right)^{-1}. \tag{14}$$

The estimation for $\mu^- \triangleq [c^-, h^-]^T$ is carried out in a similar way. Defining

$$w_k^- \triangleq \hat{v}_k I_{[u_k \leq -U]} \quad \phi_k^- \triangleq [u_k \quad 1]^T I_{[u_k \leq -U]} \tag{15}$$

we estimate μ^- by the recursive LS algorithm

$$\mu_k^- = \mu_{k-1}^- + a_k^-P_k^-\phi_k^-\left(w_k^- - \phi_k^{-T}\mu_{k-1}^-\right) \tag{16}$$

$$P_{k+1}^- = P_k^- - a_k^-P_k^-\phi_k^-\phi_k^{-T}P_k^-$$

$$a_k^- = \left(1 + \phi_k^{-T}P_k^-\phi_k^-\right)^{-1}. \tag{17}$$

C. Estimates for d^+, b^+, d^-, and b^-

Set $\bar{w}_0^+ = 0$, and recursively define the time average of $\{\hat{v}_kI_{[u_k \geq 0]}\}$

$$\bar{w}_k^+ = \frac{k-1}{k}\bar{w}_{k-1}^+ + \frac{\hat{v}_kI_{[u_k \geq 0]}}{k}. \tag{18}$$

Similar to (6), we modify c_k^+ as follows:

$$\bar{c}_k^+ \triangleq \begin{cases} c_k^+, & \text{if } |c_k^+| \geq \frac{1}{k} \\ (\text{sign}c_k^+)\frac{1}{k}, & \text{if } |c_k^+| < \frac{1}{k}. \end{cases} \tag{19}$$

Then d^+ and b^+ are estimated by d_k^+ and b_k^+, respectively, where

$$d_k^+ = \frac{h_k^+ - \text{sign}(h_k^+)\left(h_k^{+2} + 4\bar{c}_k^+U(\bar{c}_k^+U - h_k^+ - 2\bar{w}_k^+)\right)^{\frac{1}{2}}}{\bar{c}_k^+} \tag{20}$$

and

$$b_k^+ \triangleq c_k^+d_k^+ - h_k^+. \tag{21}$$

Similarly, set

$$w_0^- = 0 \tag{22}$$

and define

$$\bar{w}_k^- = \frac{k-1}{k}\bar{w}_{k-1}^- + \frac{\hat{v}_kI_{[u_k \leq 0]}}{k}. \tag{23}$$

After modifying c_k^- to $\bar{c}_k^-$

$$\bar{c}_k^- \triangleq \begin{cases} c_k^-, & \text{if } |c_k^-| \geq \frac{1}{k} \\ (\text{sign}c_k^-)\frac{1}{k}, & \text{if } |c_k^-| < \frac{1}{k} \end{cases} \tag{24}$$

d^- and b^- are, respectively, estimated by

$$d_k^- = \frac{h_k^- - \text{sign}(h_k^-)((h_k^-)^2 + 4\bar{c}_k^-U(\bar{c}_k^-U - h_k^- + 2\bar{w}_k^-))^{\frac{1}{2}}}{\bar{c}_k^-} \tag{25}$$

and

$$b_k^- \triangleq c_k^-d_k^- - h_k^-. \tag{26}$$

III. Strong Consistency

We now show that all estimates defined in Section II converge to the corresponding true values with probability one.

Theorem 1: Assume conditions A1)–A3) hold and $\rho = Eu_1v_1 \neq 0$. Then the coefficients of the linear subsystem are strongly consistently estimated by (4)–(7)

$$\theta_k(0) \underset{k\to\infty}{\longrightarrow} \rho \quad \text{and} \quad d_{ik} \underset{k\to\infty}{\longrightarrow} d_i \quad \text{a.s.,} \qquad i = 1, \ldots, q. \tag{27}$$

Proof: The conclusions of the theorem coincide with those given in [6, Th. 1]. Notice only that $f(\cdot)$ in [6] is assumed to be continuous at u where $f(u)$ is estimated, but this continuity is not used there in the proof of Theorem 1. ∎

Theorem 2: Under the assumptions of Theorem 1, μ_k^+ and μ_k^- given by (13), (14) and (16), (17), respectively, are strongly consistent

$$\mu_k^+ \underset{k\to\infty}{\longrightarrow} [c^+, h^+]^T \quad \text{a.s.} \quad \mu_k^- \underset{k\to\infty}{\longrightarrow} [c^-, h^-]^T \quad \text{a.s.}$$

Proof: Define

$$D_{ni} \triangleq D_n D_{n-1} \dots D_i, \quad \text{for} \quad n \geq i \quad D_{ji} \triangleq I, \quad \text{for} \quad j < i \tag{28}$$

$$x_k = Dx_{k-1} + Hy_{k+1} \quad \text{and} \quad \hat{\hat{x}}_k = D_k \hat{\hat{x}}_{k-1} + Hy_{k+1}. \tag{29}$$

It is clear that

$$v_k = H^T x_k \tag{30}$$

and both $\hat{x}_k$ given by (9) and $\hat{\hat{x}}_k$ given by (29) serve as estimates for x_k.

Denote the estimation error for the output of the nonlinear block by

$$\nu_k = \hat{v}_k - v_k. \tag{31}$$

The proof is essentially based on the following fact, which will also be used in the proof of Theorem 3

$$\frac{1}{n}\sum_{k=1}^{n} \nu_k I_{[u_k \in B]} \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s.} \quad \text{for any Borel set } B. \tag{32}$$

We now prove (32). Noticing

$$\hat{x}_k - \hat{\hat{x}}_k = D_k(\hat{x}_{k-1} - \hat{\hat{x}}_{k-1}) + H\xi_{k+1} \tag{33}$$

and

$$\hat{\hat{x}}_k - x_k = D_k(\hat{\hat{x}}_{k-1} - x_{k-1}) + (D_k - D)x_{k-1} \tag{34}$$

by (10), (28), (31), (33), and (34), we have

$$\begin{aligned} \nu_k I_{[u_k\in B]} &= H^T \hat{x}_k I_{[u_k\in B]} - H^T x_k I_{[u_k\in B]} \\ &= H^T(\hat{x}_k - \hat{\hat{x}}_k) I_{[u_k\in B]} + H^T(\hat{\hat{x}}_k - x_k) I_{[u_k\in B]} \\ &= H^T D_{k1}(\hat{x}_0 - \hat{\hat{x}}_0) I_{[u_k\in B]} \\ &\quad + H^T \sum_{i=1}^{k} D_{k,i+1} H\xi_{i+1} I_{[u_k\in B]} \\ &\quad + H^T D_{k1}(\hat{\hat{x}}_0 - x_0) I_{[u_k\in B]} \\ &\quad + H^T \sum_{i=1}^{k} D_{k,i+1}(D_i - D)x_{i-1} I_{[u_k\in B]}. \end{aligned} \tag{35}$$

Since, by Theorem 1, D_k converges to the stable matrix D as k tends to infinity, there exist constants $c > 0$ and $\lambda \in (0,1)$ such that

$$\|D_{kj}\| \leq c\lambda^{k-j} \quad \text{and} \quad \|D^{k-j}\| \leq c\lambda^{k-j} \qquad \forall k \geq j. \tag{36}$$

Noticing that $\{u_k\}$ is bounded, we find that both $\{v_k\}$ and $\{y_k\}$ are bounded. Therefore, by (35), (36) for (32) it suffices to show

$$\frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{k} D_{k,i+1} H\xi_{i+1} I_{[u_k\in B]} \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s.} \tag{37}$$

and

$$\frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{k} \|D_{k,i+1}(D_i - D)\| \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s.} \tag{38}$$

The left-hand side of (37) can be written as

$$\begin{aligned} &\frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{k} D^{k-i} H\xi_{i+1} I_{[u_k\in B]} \\ &\quad + \frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{k} (D_{k,i+1} - D^{k-i}) H\xi_{i+1} I_{[u_k\in B]} \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{k=i}^{n} D^{k-i} I_{[u_k\in B]}\right) H\xi_{i+1} \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{k=i}^{n} (D_{k,i+1} - D^{k-i}) I_{[u_k\in B]}\right) H\xi_{i+1}. \end{aligned} \tag{39}$$

Since $\{u_k\}$ is independent of $\{\xi_k\}$ and uniformly bounded, by stability of D the sum $\sum_{k=i}^{n} D^{k-i} I_{[u_k\in B]}$ as n tends to infinity a.s. converges to a finite random matrix G_i, which is independent of $\{\xi_k\}$ and $\|G_i\| < \|D^{-1}\| c/(1-\lambda)$. Then, the first term on the right-hand side of (39) is estimated as follows:

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n}\left(\sum_{k=i}^{n} D^{k-i} I_{[u_k\in B]}\right) H\xi_{i+1} \\ &\quad = \frac{1}{n}\sum_{i=1}^{n} G_i H\xi_{i+1} \\ &\qquad + \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{k=i}^{n} D^{k-i} I_{[u_k\in B]} - G_i\right) H\xi_{i+1} \\ &\quad \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s.} \end{aligned} \tag{40}$$

where on the right-hand side the first term tends to zero as $n \to \infty$ by [7, Th. 2.8], while the last term tends to zero because

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} |\xi_{i+1}| < \infty \quad \text{and} \quad \left|\sum_{k=i}^{n} D^{k-i} I_{[u_k\in B]} - G_i\right| \underset{n\to\infty}{\longrightarrow} 0.$$

We now estimate the last term in (39) as follows:

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n}\left(\sum_{k=i}^{n} (D_{k,i+1} - D^{k-i}) I_{[u_k\in B]}\right) H\xi_{i+1} \\ &\quad = \frac{1}{n}\sum_{i=1}^{n_1}\left(\sum_{k=i}^{n} (D_{k,i+1} - D^{k-i}) I_{[u_k\in B]}\right) H\xi_{i+1} \\ &\qquad + \frac{1}{n}\sum_{i=n_1+1}^{n}\left(\sum_{k=i}^{i+n_2} (D_{k,i+1} - D^{k-i}) I_{[u_k\in B]}\right) H\xi_{i+1} \\ &\qquad + \frac{1}{n}\sum_{i=n_1+1}^{n}\left(\sum_{k=i+n_2+1}^{n} (D_{k,i+1} - D^{k-i}) I_{[u_k\in B]}\right) H\xi_{i+1}. \end{aligned} \tag{41}$$

For a given $\epsilon > 0$, we first take a sufficiently large n_2 so that the norm of the last term in (41) is less than $\epsilon/3$. This is possible because

$$\left\|\sum_{k+i+n_2+1}^{n} (D_{k,i+1} - D^{k-i})\right\| \leq \frac{2c\lambda^{n_2}}{1-\lambda}.$$

For this n_2 by $D_k \underset{k\to\infty}{\longrightarrow} D$ we can take a large n_1 such that the norm of the last but one term in (41) is less than $\epsilon/3$. Finally, for large enough n the norm of the first term on the right-hand side of (41) is also less than $\epsilon/3$ by (36) and the boundedness of $\{u_k\}$. This verifies (37).

For (38), it suffices to note that by (36) the left-hand side of (38) is bounded by

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n}\left(1 + \sum_{k=i}^{n} c\lambda^{k-i}\right)\|D_i - D\| \\ &\qquad \leq \left(1 + \frac{c}{1-\lambda}\right)\frac{1}{n}\sum_{i=1}^{n}\|D_i - D\| \underset{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

since $D_i \underset{i\to\infty}{\longrightarrow} D$.

Thus, we have proved (37) and (38) and, hence, (32).

The LS estimate μ_k^+ given by (13) and (14) equals

$$\mu_k^+ = \left(\sum_{i=1}^{k} \phi_i^+ \phi_i^{+T}\right)^{-1} \sum_{i=1}^{k} \phi_i^+ w_i^+ \tag{42}$$

whenever the matrix $\sum_{i=1}^{k} \phi_i^+ \phi_i^{+T}$ is nonsingular [7], [14].

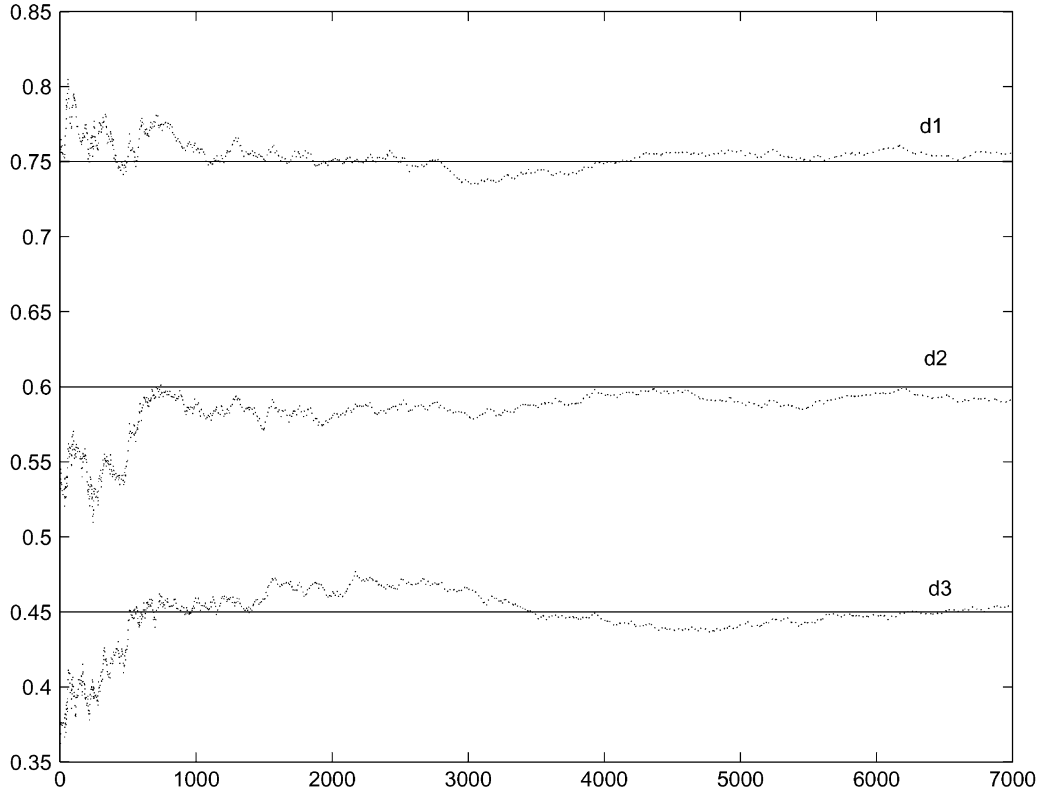

Fig. 2. Estimates for $\boldsymbol{d}_1$, $\boldsymbol{d}_2$, and $\boldsymbol{d}_3$.

Since $\{u_k\}$ is iid with uniform distribution over $[-2U, 2U]$, by the strong law of large numbers [8], we have

$$\frac{1}{k}\sum_{i=1}^{k}\phi_i^+\phi_i^{+T} = \frac{1}{k}\sum_{i=1}^{k}\begin{pmatrix} u_i^2 & -u_i \\ -u_i & 1\end{pmatrix} I_{[u_i \geq U]} \underset{k\to\infty}{\longrightarrow} \begin{pmatrix} \frac{7}{12}U^2 & -\frac{3}{8}U \\ -\frac{3}{8}U & \frac{1}{4}\end{pmatrix} \text{a.s.} \tag{43}$$

which is nondegenerate.

Noticing that $v_k I_{[u_k \geq U]} = \mu^{+T}\phi_k^+$, by (11), (32) we have

$$w_k^+ = \mu^{+T}\phi_k^+ + \nu_k I_{[u_k \geq U]}. \tag{44}$$

Consequently, from (42) and (43) by (32), we have

$$\mu_k^+ = \left(\sum_{i=1}^{k}\phi_i^+\phi_i^{+T}\right)^{-1}\sum_{i=1}^{k}\phi_i^+\left(\phi_i^{+T}\mu^+ + \nu_i I_{[u_i\geq U]}\right) \underset{k\to\infty}{\longrightarrow} \mu^+ \quad \text{a.s.}$$

The proof for the strong consistency of μ_k^- is completely the same as that for μ_k^+. ■

Theorem 3: Under the conditions of Theorem 1 d_k^+, b_k^+, d_k^-, and b_k^- given by (20), (21), (25), and (26) are strongly consistent

$$d_k^+ \underset{k\to\infty}{\longrightarrow} d^+ \text{a.s.} \quad b_k^+ \underset{k\to\infty}{\longrightarrow} b^+ \text{ a.s.}$$
$$d_k^- \underset{k\to\infty}{\longrightarrow} d^- \text{ a.s.} \quad \text{and} \quad b_k^- \underset{k\to\infty}{\longrightarrow} b^- \text{a.s.}$$

Proof: It is clear that

$$Ev_k I_{[u_k\geq 0]} = \frac{1}{4U}\int_{d^+}^{2U}(c^+u - h^+)\,du = -\frac{c^+}{8U}d^{+2} + \frac{h^+}{4U}d^+ + \frac{c^+U}{2} - \frac{h^+}{2}$$

and from here it follows that

$$d^+ = \begin{cases} \frac{1}{c^+}\left[h^+ - \operatorname{sign}(h^+)\left(h^{+2} + 4c^+U \cdot \left(c^+U - h^+ - 2Ev_kI_{[u_k\geq 0]}\right)\right)^{\frac{1}{2}}\right], & \text{if } |c^+| > 0 \\ \frac{4U}{h^+}\left(\frac{h^+}{2} + Ev_kI_{[u_k\geq 0]}\right), & \text{if } c^+ = 0\end{cases} \tag{45}$$

where "$-\operatorname{sign}(h^+)$" is taken to make d^+ to be continuous with respect to c^+ as $c^+ \to 0$ for a fixed $Ev_kI_{[u_k\geq 0]}$.

From (18) and (32), it follows that

$$\bar{w}_k^+ = \frac{1}{k}\sum_{i=1}^{k}\hat{v}_i I_{[u_i\geq 0]} = \frac{1}{k}\sum_{i=1}^{k}(v_i + \nu_i)I_{[u_i\geq 0]} \underset{k\to\infty}{\longrightarrow} Ev_1 I_{[u_1\geq 0]} \quad \text{a.s.} \tag{46}$$

By Theorem 2 and (19), we have $\bar{c}_k^+ \underset{k\to\infty}{\longrightarrow} c^+$ a.s. and $h_k^+ \underset{k\to\infty}{\longrightarrow} h^+$ a.s. This combining with (46) leads to that d_k^+ given by (20) tends to d^+ a.s. as $k \to \infty$ whenever c^+ is zero or not.

Since $b^+ = c^+d^+ - h^+, b_k^+$ given by (21) converges to b^+ a.s. as $k \to \infty$.

Finally, noticing that

$$Ev_kI_{[u_k\leq 0]} = \frac{1}{4U}\int_{-2U}^{-d^-}(c^-u + h^-)\,du = \frac{c^-(d^-)^2}{8U} - \frac{h^-d^-}{4U} - \frac{c^-U}{2} + \frac{h^-}{2}$$

and

$$d^- = \begin{cases} \frac{1}{c^-}\left[h^- - \operatorname{sign}(h^-)\left((h^-)^2 + 4c^-U \cdot \left(c^-U - h^- + 2Ev_kI_{[u_k\leq 0]}\right)\right)^{\frac{1}{2}}\right], & \text{if } |c^-| > 0 \\ \frac{4U}{h^-}\left(\frac{h^-}{2} - Ev_kI_{[u_k\leq 0]}\right), & \text{if } c^- = 0\end{cases}$$

we conclude that d_k^- given by (25) is strongly consistent, while strong consistency of b_k^- is obvious. ■

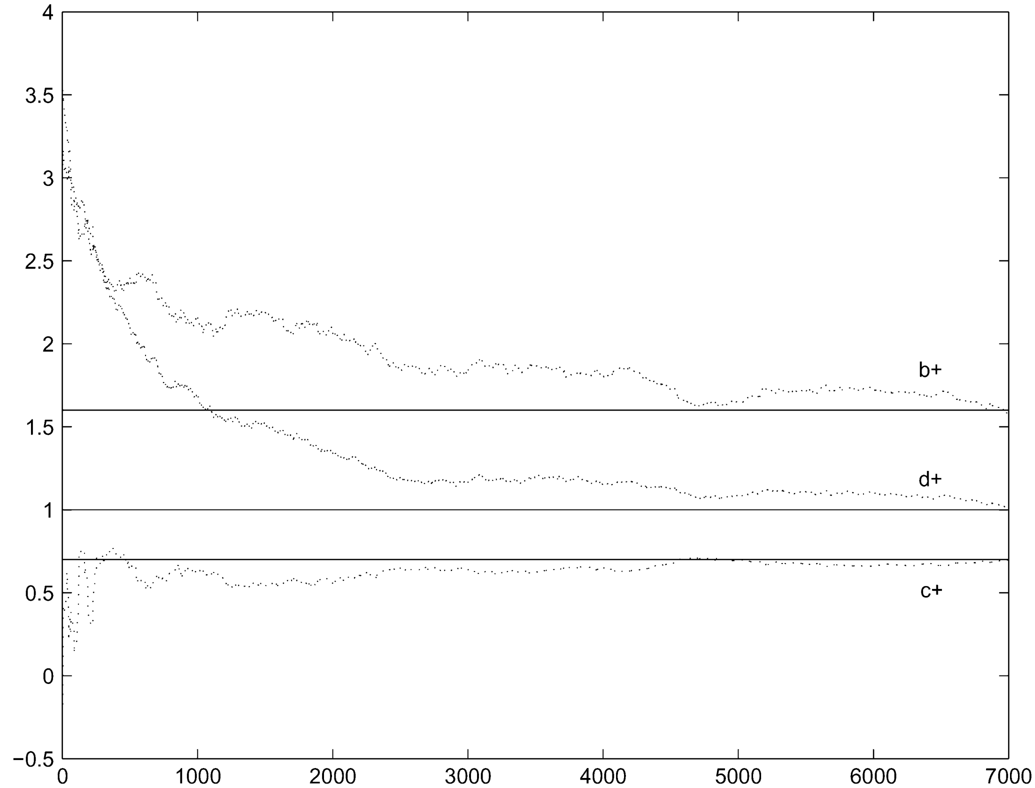

Fig. 3. Estimates for $\boldsymbol{b}^+, \boldsymbol{c}^+$, and $\boldsymbol{d}^+$.

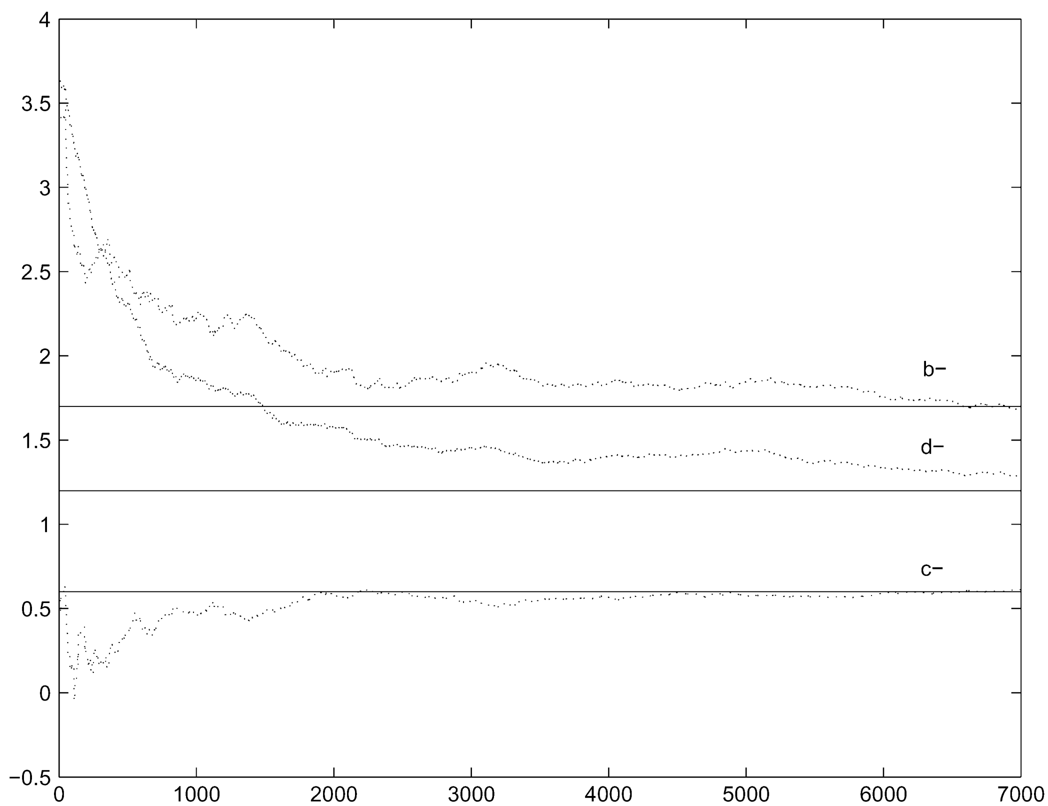

Fig. 4. Estimates for $\boldsymbol{b}^-, \boldsymbol{c}^-$, and $\boldsymbol{d}^-$.

IV. Numerical Example

To numerically demonstrate the strong consistency of the algorithms proposed in Section III, let the parameters appearing in (1), (3) of the system described by Fig. 1 take the following values:

$$d_1 = 0.75 \quad d_2 = 0.6 \quad d_3 = 0.45$$
$$d^+ = 1 \quad c^+ = 0.7 \quad b^+ = 1.6$$
$$d^- = 1.2 \quad c^- = 0.6 \quad b^- = 1.7 \quad \text{and} \quad U = 2.$$

It is direct to check that the polynomial

$$D(z) = 1 + 0.75z + 0.6z^2 + 0.45z^3$$

has roots equal to $4/3$ and $\pm i\sqrt{5/3}$. Hence, it is stable, but the SPR condition required for strong consistency of ELS [7] is not satisfied since $D^{-1}(e^{i\lambda}) + D^{-1}(e^{-i\lambda}) < 1$ at $\lambda = 0$.

Let u_k be uniformly distributed over $[-4, 4]$, and let $\xi_k \in \mathcal{N}(0, 1)$. Matlab is used to generate the iid sequences $\{u_k\}$ and $\{\xi_k\}$, and to carry out the recursive estimation according to (4), (5), (13), (14), (16), (17), (20), (21), and (25), (26).

Fig. 2 demonstrates the estimates for d_1, d_2, and d_3, while Figs. 3 and 4 give estimates for c^+, b^+, d^+, and c^-, b^-, d^-, respectively. In all figures, the solid lines represent the true values and the dotted lines are their estimates. It is seen that all estimates converge to their true values, i.e., the computer simulation justifies the theoretical conclusion concerning strong consistency.

V. Concluding Remarks

The recursive identification algorithms are proposed in the note for Hammerstein systems with memoryless nonlinearity being a discontinuous piecewise-linear function containing six unknown parameters. The strong consistency is proved for estimates for the unknown parameters contained in the static nonlinearity and for the coefficients of the linear subsystem as well.

For further research, it is of interest to consider the multidimensional systems, to weaken conditions imposed on the linear subsystem, and to remove the availability assumption of an upper bound U for d^+ and d^-. It is important to connect the proposed identification method with control task, i.e., to solve adaptive control problems for systems described in the note.

Acknowledgment

The author would like to thank X. Hu and W. Zhao for their computer simulation.

References

[1] E. W. Bai, "Identification of linear systems with hard input nonlinearities of known structurexr," *Automatica*, vol. 38, no. 5, pp. 853–860, 2002.
[2] ——, "Frequency domain identification of Hammerstein models," *IEEE Trans. Autom. Control*, vol. 48, no. 4, pp. 530–542, Apr. 2003.
[3] H. Al-Duwaish and M. N. Karim, "A new method for the identification of Hammerstein model," *Automatica*, vol. 33, no. 10, pp. 1871–1875, 1997.
[4] F. H. I. Chang and R. Luus, "A noniterative method for identification using Hammerstein model," *IEEE Trans. Autom. Control*, vol. AC-16, no. 10, pp. 465–468, Oct. 1971.
[5] H. F. Chen, *Stochastic Approximation and Its Applications*. Dordrecht, The Netherlands: Kluwer, 2002.
[6] ——, "Path-wise convergence of recursive identification algorithms for Hammerstein systems," *IEEE Trans. Automat. Control*, vol. 49, no. 10, pp. 1641–1649, Oct. 2004.
[7] H. F. Chen and L. Guo, *Identification and Stochastic Adaptive Control*. Boston, MA: Birkhäuser, 1991.
[8] Y. S. Chow and H. Teicher, *Probability Theory*. New York: Springer-Verlag, 1978.
[9] P. Crama and J. Schoukens, "Initial estimates of Wiener and Hammerstein systems using multisine excitation," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 6, pp. 1791–1795, Dec. 2001.
[10] W. Greblicki, "Stochastic approximation in nonparametric identification of Hammerstein systems," *IEEE Trans. Autom. Control*, vol. 47, no. 11, pp. 1800–1810, Nov. 2002.
[11] W. Greblicki and M. Pawlak, "Nonparametric identification of Hammertein systems," *IEEE Trans. Inform. Theory*, vol. 35, no. 3, pp. 409–418, Mar. 1989.
[12] Z. Q. Lang, "A nonparametric polynomial identification algorithm for the Hammerstein system," *IEEE Trans. Autom. Control*, vol. 42, no. 10, pp. 1435–1441, Oct. 1997.
[13] H. E. Liao and W. A. Sethares, "Suboptimal identification of nonlinear ARMA models using an orthogonality approach," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 42, no. 1, pp. 14–22, Jan. 1995.
[14] L. Ljung, *System Identification*. Upper Saddle River, NJ: Prentice-Hall, 1987.
[15] M. Pawlak, "On the series expansion approach to the identification of Hammertein system," *IEEE Trans. Autom. Control*, vol. 36, no. 6, pp. 763–767, Jun. 1991.
[16] P. Stoica and T. Söderstrom, "Instrumental-variable methods for identification of Hammerstein systems," *Int. J. Control*, vol. 35, no. 3, pp. 459–476, 1982.
[17] J. Vörös, "Recursive identification of Hammerstein systems with discontinuous nonlinearities containing dead-zones," *IEEE Trans. Autom. Control*, vol. 48, no. 12, pp. 2203–2206, Dec. 2003.

Identification of IIR Wiener Systems With Spline Nonlinearities That Have Variable Knots

Matt C. Hughes and David T. Westwick

***Abstract*—An algorithm is developed for the identification of Wiener systems, linear dynamic elements followed by static nonlinearities. In this case, the linear element is modeled using a recursive digital filter, while the static nonlinearity is represented by a spline of arbitrary but fixed degree. The primary contribution in this note is the use of variable knot splines, which allow for the use of splines with relatively few knot points, in the context of Wiener system identification. The model output is shown to be nonlinear in the filter parameters and in the knot points, but linear in the remaining spline parameters. Thus, a separable least squares algorithm is used to estimate the model parameters. Monte-Carlo simulations are used to compare the performance of the algorithm identifying models with linear and cubic spline nonlinearities, with a similar technique using polynomial nonlinearities.**

***Index Terms*—Block structured models, cubic spline, Levenberg–Marquardt algorithm, nonlinear system identification, separable least squares optimization.**

I. Introduction

The Wiener system is a block oriented model consisting of a dynamic linear system followed by a memoryless nonlinearity, as shown in Fig. 1. Unlike more general models, such as a Volterra series or a multiple layer neural network, the Wiener model can represent systems with high-order nonlinearities using relatively few parameters, making it suitable for control applications.

The Wiener system is well suited to modeling linear processes measured by nonlinear sensors whose dynamics, if any, are orders of magnitude faster than the process being measured. As such, the Wiener model has found applications in the process control industry. Wiener

Manuscript received June 1, 2004; revised May 27, 2005. Recommended by Guest Editor A. Vicino. This work was supported by the Natural Sciences and Engineering Research Council of Canada.

M. Hughes is with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 3P6, Canada (e-mail: mhughe@uvic.ca).

D. Westwick is with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, AB T2N IN4, Canada (e-mail: dwestwic@ucalgary.ca).

Digital Object Identifier 10.1109/TAC.2005.856660