International Journal of Control
Vol. 80, No. 2, February 2007, 256–267

Taylor & Francis
Taylor & Francis Group

Extended least-correlation estimates for errors-in-variables non-linear models

B.-E. JUN*† and D. S. BERNSTEIN‡

†Guidance & Control Department, Agency for Defense Development, Youseong, Daejeon 305-600, Korea
‡Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109-2140, USA

(*Received 12 April 2005; in final form 8 September 2006*)

This paper introduces a method of parameter estimation working on errors-in-variables polynomial non-linear models in which all measurements are corrupted by noise. The first step is to develop the linear regression models which are equivalent to polynomial non-linear systems. A main idea is to extend the parameter vector by even-order components of noise and to augment the regression vector by appropriate constants or measurements. Applying the method of least correlation, which has a capability to cope with errors-in-variables linear models, to the equivalent model with extended parameters and augmented regressors yields an extended least-correlation estimator. Analysis shows that, for non-linear systems with third or lower order polynomials, the parameters estimated by the proposed method asymptotically converge to the true values. Numerical examples also support analytical results. Applications of the approach to Volterra models, Hammerstein models and Weiner non-linear systems are included.

1. Introduction

The method of least squares works on linear systems provided that they are described by the linear regression model and measurements included in the regressors are free from noise (Rugh 1981, Johansson 1993, Doyle *et al*. 2002). A generalized approach to modelling noise is to view all variables as contaminated by noise, called errors-in-variables (EIV) models (Scherrer and Deistler 1998, Ljung 1999, Soderstrom *et al*. 2002, van Huffel and Lemmerling 2002). Noises included in regressors make identification problems challenging. Nonlinearities in system models make the problems more challenging.

There are previous contributions for EIV non-linear models which are described by polynomials (Vajk and Hetthessy 2003), Volterra models (Mattera and Paura 1998, Mattera 1999), Wiener–Hammerstein models (Tan and Godfrey 2002), and general non-linear functions (Vandersteen *et al*. 1996, Fazekas and Kukush 1997, Hermey and Watson 1999, Baran 2000, Li 2002). Vajk and Hetthessy (2003) generalize the classical eigenvalue-decomposition method to polynomial non-linear systems with *a priori* knowledge about the structure of the noise covariance matrix. Works for Volterra models assume that the input signal is amplitude-modulated cyclostationary (Mattera and Paura 1998) or consider the polyperiodic non-linear models (Mattera 1999). Tan and Godfrey (2002) identify the linear subsystems of a Wiener–Hammerstein model through the measurements of second-order Volterra kernels in frequency domain. Vandersteen *et al*. (1996), Fazekas and Kukush (1997), and Baran (2000) introduce estimators for systems which are described by general non-linear functions including polynomials, but they need *a priori* knowledge of every moment of noise. Li (2002) identifies parameters based on the estimated statistics of input variable. Hermey and Watson (1999) treat a kind of fitting problem to Huber function.

The method of least correlation (Jun and Bernstein 2006) has a capability to find out the best fit to a given structure of EIV linear models without a priori knowledge about noise covariance. Direct application of the

*Corresponding author. Email: bejun@add.re.kr

International Journal of Control
ISSN 0020–7179 print/ISSN 1366–5820 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00207170600999884

method to EIV polynomial non-linear systems, however, yields error-prone estimates affected by noises included in regressors. In the present paper, we introduce a procedure to extend the parameter vector by even-order components of noise and to augment the regressor vector by appropriate constants or measurements. Applying the method of least correlation to the new models with the augmented regressors and the extended parameters gives extended least-correlation (ELC) estimates.

The ELC estimator does not need *a priori* information about the noise covariance, but previous estimators (Vandersteen *et al*. 1996, Fazekas and Kukush 1997, Hermey and Watson 1999, Baran 2000, Li 2002, Vajk and Hetthessy 2003) available for polynomial non-linear systems need the knowledge. Analysis shows that the ELC estimates used for systems with third or lower non-linear degree asymptotically converge to the true values. Monte Carlo simulations for simple examples support the analytical results. For fourth or higher order polynomials, unfortunately, the ELC estimates tend to include a bit of bias unless the augmented regressors are completely decoupled from the extended parameters. We applied the method of ELC to Volterra models, Hammerstein non-linear systems, and Wiener models.

The next section states system models, definition of problem, and assumptions on systems and signals. Section 3 contains main idea and analysis for second, third and higher-order polynomial non-linear systems. We introduce applications of the ELC method to Volterra, Hammerstein and Wiener non-linear models in § 4. Section 5 shows numerical examples. In § 6 we discuss the capability and limitation of the ELC estimate. And a recursive version of the ELC estimate is also introduced in § 6. Some proofs of theorems are stated in Appendix.

2. System models and assumptions

Consider the discrete-time Volterra models (Rugh 1981, Doyle *et al*. 2002)

$$z(t) = z_0 + \sum_{\ell=1}^{L} \chi_M^\ell(t) + \eta_1(t), \tag{1}$$

$$\chi_M^\ell(t) = \sum_{i_1=0}^{M} \cdots \sum_{i_\ell=0}^{M} \sum_{j=1}^{n} \alpha_\ell^j(i_1, \ldots, i_\ell) \times u_j(t-i_1)\cdots u_j(t-i_\ell), \tag{2}$$

where $z(t) \in \mathbb{R}$ is the system response at tth sampling, z_0 is a constant, $\eta_1(t) \in \mathbb{R}$ denotes possible modelling errors, $u_j(t)$ is the jth element of input vector $u(t) \in \mathbb{R}^n$, L denotes the non-linear degree of model, and M is its dynamic order. The equations (1) and (2) state a general Volterra model, but we will impose some restrictions on it in later sections.

Suppose that both $z(t)$ and $u(t)$ are measured in noise as depicted in figure 1. Let $y(t) \in \mathbb{R}$ and $v(t) \in \mathbb{R}^n$ be the available measurements of $z(t)$ and $u(t)$ given by

$$y(t) = z(t) + \eta_2(t), \tag{3}$$

$$v(t) = u(t) + \nu(t), \tag{4}$$

respectively, where $\eta_2(t) \in \mathbb{R}$ and $\nu(t) \in \mathbb{R}^n$ denote measurement noises. Using (3) and (4) to (1) and (2) gives the EIV model

$$y(t) = y_0 + \sum_{\ell=1}^{L} \chi_M^\ell(t) + \eta(t), \tag{5}$$

$$\chi_M^\ell(t) = \sum_{i_1=0}^{M} \cdots \sum_{i_\ell=0}^{M} \sum_{j=1}^{n} \alpha_\ell^j(i_1, \ldots, i_\ell) \times \left\{ v_j(t-i_1)\cdots v_j(t-i_\ell) - w_j(t, i_1, \ldots, i_\ell) \right\}, \tag{6}$$

$$\begin{aligned} w_j(\cdot) = {} & v_j(t-i_1)\cdots v_j(t-i_{\ell-1})\nu_j(t-i_\ell) \\ & + v_j(t-i_1)\cdots v_j(t-i_{\ell-2})\nu_j(t-i_{\ell-1}) \\ & \times [v_j(t-i_\ell) - \nu_j(t-i_\ell)] \\ & \vdots \\ & + \nu_j(t-i_1)[v_j(t-i_2) - \nu_j(t-i_2)]\cdots \\ & \times [v_j(t-i_\ell) - \nu_j(t-i_\ell)], \end{aligned} \tag{7}$$

where $\eta(t) \triangleq \eta_1(t) + \eta_2(t)$. Now let us state the estimation problem.

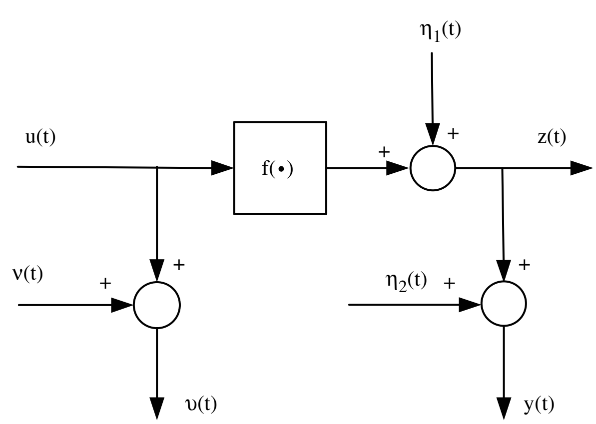

Figure 1. Description of EIV nonlinear dynamic models.

Problem 1: Given the system model (1) and (2) and the measurement model (3) and (4), determine an estimate of the system parameters $\alpha^j_\ell(i_1,\ldots,i_\ell)$ based on the available measurements $v(t)$ and $y(t)$.

Identification problems frequently work with signals which are described as stochastic processes with deterministic components. For a common framework for deterministic and stochastic signals, we employ the definition of quasi-stationary signals and the notation

$$\bar{E}[f(t)] \triangleq \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} E[f(t)] \tag{8}$$

which works on the deterministic components as well as the stochastic parts of quasi-stationary signal $f(t)$, where E denotes the mathematical expectation (Ljung 1999, p. 34). We implicitly assume that the limit in (8) exists when $\bar{E}$ is used.

We introduce the following assumptions.

A1. The structure of model, the non-linear degree L, the dynamic order M, and the number of inputs n are known *a priori*. If the system is dynamic, it is asymptotically stable.

A2. Measurements $v(t)$ and $y(t)$ are quasi-stationary and jointly quasi-stationary.

A3. Noises $\eta(t)$ and $\nu(t)$ are zero-mean, stationary and $E[\nu_j^L(t)] = 0, j = 1,\ldots,n$ for all odd L. There exists $\tau > 0$ for all $|k| \geq \tau$ such that

$$\bar{E}\big[\nu(t)\nu^T(t-k)\big] = 0, \tag{9}$$

$$\bar{E}[\nu(t)\eta(t-k)] = 0. \tag{10}$$

A4. For τ in A3, none of the elements of $v(t)$ is constant and $v(t)$ satisfies

$$\operatorname{rank}\big\{\bar{R}_{vv}(t, t-\tau, N) + \bar{R}_{vv}(t-\tau, t, N)\big\} = n, \tag{11}$$

where N denotes the number of samples and the empirical correlation $\bar{R}_{vv}(t_1, t_2, N)$ with $t_1 = t, t_2 = t-\tau$ or $t_1 = t-\tau, t_2 = t$ is defined by

$$\bar{R}_{vv}(t_1, t_2, N) \triangleq \frac{1}{N-\tau}\sum_{t=1+\tau}^{N} v(t_1)v^T(t_2),$$
$$\tau \triangleq |t_1 - t_2|. \tag{12}$$

Assumptions A3 and A4 express that correlations between signals are stronger than those between noises as well as those between signals and noises. Conditions (9) and (10) are equivalent to

$$\bar{E}\big[u(t)\nu^T(t-k)\big] = 0, \quad \bar{E}\big[\nu(t)\nu^T(t-k)\big] = 0, \tag{13}$$

$$\bar{E}[u(t)\eta(t-k)] = 0, \quad \bar{E}[\nu(t)\eta(t-k)] = 0, \tag{14}$$

for all $|k| \geq \tau > 0$, respectively.

3. Main results

Discussions in this section focus on static non-linear systems with multi-input single-output. The results can be extended with ease to multi-input multi-output systems.

3.1 *Second-order non-linear systems*

Consider a second-order non-linear model

$$z(t) = \sum_{j=1}^{n} a_j u_j^2(t) + \eta_1(t) \tag{15}$$

which is obtained from the Volterra model (1)–(2) by

$$z_0 = 0,\ M = 0,\ L = 2,$$
$$\alpha_1^j(0) = 0, \alpha_2^j(0,0) = a_j, j = 1,\ldots,n. \tag{16}$$

Using (3) and (4) with (15) or applying (16) to (5)–(7) yields

$$y(t) = \sum_{j=1}^{n} a_j(v_j(t) - \nu_j(t))^2(t) + \eta(t). \tag{17}$$

For $\nu_j(t)$ included in each term where the odd-order of $\nu_j(t)$ appears in the expansion of (17), substituting $v_j(t) = u_j(t) + \nu_j(t)$ gives

$$y(t) = \sum_{j=1}^{n} a_j\Big[v_j^2(t) - \nu_j^2(t) - 2u_j(t)\nu_j(t)\Big] + \eta(t). \tag{18}$$

Let (18) be a linear regression form

$$y(t) = \psi_a^T(t)\theta_a(t) + e(t) \tag{19}$$

with the error $e(t)$, the augmented regression vector $\psi_a(t) \in \mathbb{R}^{n+1}$ and the extended parameter vector $\theta_a(t) \in \mathbb{R}^{n+1}$ defined by

$$e(t) = \eta(t) - 2\sum_{j=1}^{n} a_j u_j(t)\nu_j(t), \tag{20}$$

$$\psi_a(t) \triangleq \big[v_1^2(t) \ \cdots \ v_n^2(t) - 1\big]^T, \tag{21}$$

$$\theta_a(t) \triangleq \big[\, a^T \quad \textstyle\sum_{j=1}^{n} a_j\nu_j^2(t)\,\big]^T, \tag{22}$$

where $a^T \triangleq [a_1 \cdots a_n]$. Given an estimate $\bar{\theta}_a$ and $N_{g(\tau)} \triangleq N - g(\tau)$, consider a criterion

$$J^2(\bar{\theta}_a, \tau, N) = \left(\frac{1}{N_\tau}\big(Y_0 - \Psi_0\bar{\theta}_a\big)^T\big(Y_\tau - \Psi_\tau\bar{\theta}_a\big)\right)^2, \quad (23)$$

where Y_0, Y_τ, Ψ_0 and Ψ_τ are defined by

$$Y_0 \triangleq \begin{bmatrix} y(N) \\ y(N_1) \\ \vdots \\ y(1+\tau) \end{bmatrix}, \quad Y_\tau \triangleq \begin{bmatrix} y(N_\tau) \\ y(N_{\tau+1}) \\ \vdots \\ y(1) \end{bmatrix}, \quad (24)$$

$$\Psi_0 \triangleq \begin{bmatrix} \psi_a^T(N) \\ \psi_a^T(N_1) \\ \vdots \\ \psi_a^T(1+\tau) \end{bmatrix}, \quad \Psi_\tau \triangleq \begin{bmatrix} \psi_a^T(N_\tau) \\ \psi_a^T(N_{\tau+1}) \\ \vdots \\ \psi_a^T(1) \end{bmatrix}. \quad (25)$$

Necessary and sufficient condition to minimize (23) yields the ELC estimate

$$\hat{\theta}_a(\tau, N) = \left(\Psi_{0/\tau}^T \Psi_{\tau/0}\right)^{-1} \Psi_{0/\tau}^T Y_{\tau/0}, \quad (26)$$

where the relevant matrices and vectors are composed of

$$\Psi_{0/\tau} \triangleq \begin{bmatrix} \Psi_0 \\ \Psi_\tau \end{bmatrix}, \quad \Psi_{\tau/0} \triangleq \begin{bmatrix} \Psi_\tau \\ \Psi_0 \end{bmatrix}, \quad Y_{\tau/0} \triangleq \begin{bmatrix} Y_\tau \\ Y_0 \end{bmatrix}. \quad (27)$$

The matrix $\Psi_{0/\tau}^T \Psi_{\tau/0}$ has a full rank since each component of $v(t)$ is independent of and is not constant. The ELC estimate (26) has the following property.

Theorem 1: *Suppose that* A1–A4 *are satisfied. Then as N goes to infinity, the ELC estimate* $\hat{\theta}_a(\tau, N)$ *in* (26) *for* (19)–(22) *converges to the expectation of* $\theta_a(t)$ *in* (22), *that is,*

$$\lim_{N\to\infty} \hat{\theta}_a(\tau, N) = \bar{E}[\theta_a(t)] = \begin{bmatrix} a^T & \sum_{j=1}^n a_j\sigma_{v_j}^2 \end{bmatrix}^T, \quad (28)$$

where $\sigma_{v_j}^2$ *is the variance of* $v_j(t)$.

Proof: Refer to Appendix A. □

Theorem 1 addresses the consistency of ELC estimates in the sense that the first n elements of $\hat{\theta}_a(\tau, N)$ converge to the true parameters a as N goes to infinity. Above procedure to derive the ELC estimate is applied to multi-input multi-output systems according to the steps of least-squares estimate (Johansson 1993, pp. 97–98).

3.2 *Third-order non-linear systems*

For the third-order non-linear system

$$z(t) = \sum_{j=1}^n a_j u_j^3(t) + \eta_1(t), \quad (29)$$

the EIV model is written as

$$y(t) = \sum_{j=1}^n a_j\big(v_j(t) - \nu_j(t)\big)^3 + \eta(t). \quad (30)$$

Rearranging (30) for the even-order terms of $\nu_j(t)$ to be with $v_j(t)$ and for the odd-order terms of $\nu_j(t)$ to be with $u_j(t)$ yields (19) with

$$e(t) = \eta(t) - \sum_{j=1}^n a_j\Big(3u_j^2(t)\nu_j(t) - 2\nu_j^3(t)\Big), \quad (31)$$

$$\psi_a(t) = \left[v_1^3(t) \cdots v_n^3(t) - 3v^T(t)\right]^T, \quad (32)$$

$$\theta_a(t) = \left[a^T a_1\nu_1^2(t) \cdots a_n\nu_n^2(t)\right]^T. \quad (33)$$

Applying the method of least correlation to (19) with (31)–(33) gives an ELC estimator with the following property.

Theorem 2: *Suppose that* A1–A4 *are satisfied. Then as N goes to infinity,* $\hat{\theta}_a(\tau, N)$ *for* (19) *with* (31)–(33) *converges to the expectation of* $\theta_a(t)$ *in* (33), *that is,*

$$\lim_{N\to\infty} \hat{\theta}_a(\tau, N) = \bar{E}[\theta_a(t)] = \begin{bmatrix} a^T & a_1\sigma_{\nu_1}^2 & \cdots & a_n\sigma_{\nu_n}^2 \end{bmatrix}^T, \quad (34)$$

where $\sigma_{\nu_j}^2$ *is the variance of* $\nu_j(t)$.

Proof. Refer to Appendix B. □

3.3 *Higher-order non-linear systems*

Consider higher-order non-linear systems

$$z(t) = \sum_{j=1}^n a_j u_j^L(t) + \eta_1(t), \quad L \geq 4. \quad (35)$$

The EIV model of (35),

$$y(t) = \sum_{j=1}^n a_j\big(v_j(t) - \nu_j(t)\big)^L + \eta(t), \quad (36)$$

is rearranged for even L to

$$y(t) = \sum_{j=1}^{n} a_j \Big[\beta_0 v_j^L(t) + \beta_2 v_j^{L-2}(t) v_j^2(t) + \cdots + \beta_{L-2} v_j^2(t) v_j^{L-2}(t) + \beta_L v_j^L(t) \Big] + e(t), \tag{37}$$

$$e(t) = \sum_{j=1}^{n} a_j \Big[\beta_1 u_j^{L-1}(t) v_j(t) + \beta_3 u_j^{L-3}(t) v_j^3(t) + \cdots + \beta_{L-1} u_j(t) v_j^{L-1}(t) \Big] + \eta(t), \tag{38}$$

and for odd L to

$$y(t) = \sum_{j=1}^{n} a_j \Big[\beta_0 v_j^L(t) + \beta_2 v_j^{L-2}(t) v_j^2(t) + \cdots + \beta_{L-1} v_j(t) v_j^{L-1}(t) \Big] + e(t), \tag{39}$$

$$e(t) = \sum_{j=1}^{n} a_j \Big[\beta_1 u_j^{L-1}(t) v_j(t) + \beta_3 u_j^{L-3}(t) v_j^3(t) + \cdots + \beta_L v_j^L(t) \Big] + \eta(t). \tag{40}$$

Table 1 shows β_ℓ in (37)–(40) for systems with up to 8th-order non-linearity.

Each of (37) and (39) has an equivalent realization (19) with (38) and

$$\psi_a^T(t) = \begin{bmatrix} v_{(L)}'^T & v_{(L-2)}'^T & \cdots & v_{(2)}'^T & \beta_L \end{bmatrix}, \tag{41}$$

$$\theta_a^T(t) = \begin{bmatrix} v_{(0)}'^T & v_{(2)}'^T & \cdots & v_{(L-2)}'^T & \gamma_L(t) \end{bmatrix} \tag{42}$$

Table 1. Examples of the coefficients $\beta_0 = 1, \beta_\ell, \ell = 1, \ldots, 8$.

L	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
1	−1							
2	−2	−1						
3	−3	−3	2					
4	−4	−6	8	5				
5	−5	−10	20	25	−16			
6	−6	−15	40	75	−96	−61		
7	−7	−21	70	175	−336	−427	272	
8	−8	−28	112	350	−896	−1708	2176	1385

for even L, or with (40) and

$$\psi_a^T(t) = \begin{bmatrix} v_{(L)} & v_{(L-2)}'^T & \cdots & v_{(1)}'^T \end{bmatrix}, \tag{43}$$

$$\theta_a^T(t) = \begin{bmatrix} v_{(0)}'^T & v_{(2)}'^T & \cdots & v_{(L-1)}'^T \end{bmatrix} \tag{44}$$

for odd L, where $\gamma_L(t) \triangleq \sum_{j=1}^{n} a_j v_j^L(t)$, and $v'_{(\ell)}(t)$, $v'_{(\ell)}(t)$, $\ell = 0, 1, \ldots, L$ are defined by

$$v_{(\ell)}'^T(t) \triangleq \beta_{L-\ell} \begin{bmatrix} v_1^\ell(t) & \cdots & v_n^\ell(t) \end{bmatrix},$$
$$v_{(\ell)}'^T(t) \triangleq \begin{bmatrix} a_1 v_1^\ell(t) & \cdots & a_n v_n^\ell(t) \end{bmatrix}$$

For (19) with either (38) and (41)–(42) or (40) and (43)–(44), applying the method of least correlation yields $\hat{\theta}_a(\tau, N)$ in (26). The ELC estimate is evaluated as

$$\lim_{N \to \infty} \hat{\theta}_a(\tau, N) = \bar{E}[\theta_a(t)] + R_{\psi_a \psi_a}^{-1}(\tau) \mathcal{E}(t, t-\tau), \tag{45}$$

where each component of $\mathcal{E}(t, t-\tau)$ is given by (C20) and (C21) for even L and by (C25) for odd L, respectively, in Appendix C. The second term of (45) does not vanish unless the extended regressors are decoupled with the augmented parameters.

4. Applications

4.1 *Second-order Volterra models*

Consider the second-order Volterra model

$$z(t) = z_0 + \sum_{i=0}^{M} \alpha_1(i) u(t-i) + \sum_{i=0}^{M} \sum_{j=0}^{M} \alpha_2(i,j) u(t-i) u(t-j) + \eta_1(t) \tag{46}$$

which is obtained from (1) and (2) by setting $L = 2, n = 1$. Applying (3) and (4) to (46) yields

$$y(t) = \eta(t) + y_0 + \sum_{i=0}^{M} \alpha_1(i)[v(t-i) - v(t-i)] + \sum_{i=0}^{M} \sum_{j=0}^{M} \alpha_2(i,j)[v(t-i) - v(t-i)] \times [v(t-j) - v(t-j)], \tag{47}$$

where $\eta(t) \triangleq \eta_1(t) + \eta_2(t)$. Letting the even-order terms of v be with v and the odd-order terms of v be with u yields

$$
\begin{aligned}
y(t) &= y_0 - \sum_{i=0}^{M}\sum_{j=0}^{M}\alpha_2(i,j)v(t-i)v(t-j)\\
&\quad + \sum_{i=0}^{M}\alpha_1(i)v(t-i)\\
&\quad + \sum_{i=0}^{M}\sum_{j=0}^{M}\alpha_2(i,j)v(t-i)v(t-j) + e(t), \qquad (48)\\
e(t) &= \eta(t) - \sum_{i=0}^{M}\sum_{j=0}^{M}\alpha_2(i,j)\\
&\quad \times \Big[u(t-i)v(t-j) - v(t-i)u(t-j)\Big]. \qquad (49)
\end{aligned}
$$

The EIV Volterra model (48) is equivalent to (19) with

$$
\psi_a(t) = \begin{bmatrix} 1 \\ v(t) \\ \vdots \\ v(t-M) \\ v^2(t) \\ \vdots \\ v^2(t-M) \\ v(t)v(t-1) \\ \vdots \\ v(t)v(t-M) \\ \vdots \\ v(t-M+1)v(t-M) \end{bmatrix},
$$

$$
\theta_a(t) = \begin{bmatrix} \gamma_0(t) \\ \alpha_1(0) \\ \vdots \\ \alpha_1(M) \\ \alpha_2(0,0) \\ \vdots \\ \alpha_2(M,M) \\ \alpha_2(0,1)+\alpha_2(1,0) \\ \vdots \\ \alpha_2(0,M)+\alpha_2(M,0) \\ \vdots \\ \alpha_2(M-1,M)+\alpha_2(M,M-1) \end{bmatrix}, \qquad (50)
$$

where $\gamma_0(t) \triangleq y_0 - \sum_{i=0}^{M}\sum_{j=0}^{M}\alpha_2(i,j)v(t-i)v(t-j)$.

Corollary 1: *Suppose that $A1$–$A4$ are satisfied. Then as N goes to infinity,* $\hat{\theta}_a(\tau, N)$ for (19) *with* (49) and (50) *converges to the expectation of* $\theta_a(t,N)$ *in* (50), *that is,*

$$\lim_{N\to\infty}\hat{\theta}_a(\tau,N) = \bar{E}[\theta_a(t)].$$

Proof: Using the steps in Appendix A for (49)–(50) instead of (20)–(22) yields Corollary 1. □

4.2 *Hammerstein models*

Consider Hammerstein non-linear models (Doyle *et al.* 2002, p. 22). Figure 2 shows an EIV Hammerstein non-linear model. In this model, we restrict the non-linear part to a polynomial with finite degree L and the linear part to a class of FIR models with finite length M.

To the Hammerstein model written as

$$z(t) = \sum_{i=0}^{M} g(i)\sum_{\ell=0}^{L}\beta_\ell u^\ell(t-i) + \eta_1(t), \qquad (51)$$

applying (3) and (4) gives the EIV Hammerstein model

$$y(t) = \sum_{i=0}^{M} g(i)\sum_{\ell=0}^{L}\beta_\ell[v(t-i) - v(t-i)]^\ell + \eta(t). \qquad (52)$$

We consider two cases $L=2,3$. For $L=2$, (52) is rearranged as

$$
\begin{aligned}
y(t) &= e(t) + \sum_{i=0}^{M} g(i)\big[\beta_0 - \beta_2 v^2(t-i)\\
&\quad + \beta_1 v(t-i) + \beta_2 v^2(t-i)\big], \qquad (53)\\
e(t) &= \eta(t) - \sum_{i=0}^{M} g(i)\big[\beta_1 v(t-i) + 2\beta_2 u(t-i)v(t-i)\big] \quad (54)
\end{aligned}
$$

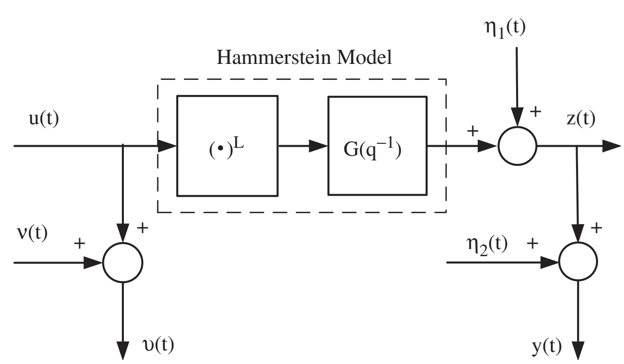

Figure 2. Description of EIV Hammerstein non-linear dynamic models.

and (53) is represented by (19) with

$$\psi_a(t) = \begin{bmatrix} 1 \\ v(t) \\ \vdots \\ v(t-M) \\ v^2(t) \\ \vdots \\ v^2(t-M) \end{bmatrix}, \quad \theta_a(t) = \begin{bmatrix} \gamma_0(t) \\ \beta_1 g(0) \\ \vdots \\ \beta_1 g(M) \\ \beta_2 g(0) \\ \cdots \\ \beta_2 g(M) \end{bmatrix}, \tag{55}$$

where $\gamma_0(t) \triangleq \sum_{i=0}^{M} g(i)\big[\beta_0 - \beta_2 v^2(t-i)\big]$.

Corollary 2: *Suppose that A1–A4 are satisfied. Then $\hat{\theta}_a(\tau, N)$ for (19) with (54) and (55) asymptotically converges to $\bar{E}[\theta_a(t)]$.*

Proof: Applying the approach in Appendix A to (54) and (55) in place of (20)–(22) leads to Corollary 2. □

For $L = 3$, (52) is written as

$$\begin{aligned} y(t) = e(t) + \sum_{i=0}^{M} g(i)\big[\big\{\beta_0 - \beta_2 v^2(t-i)\big\} \\ + \big\{\beta_1 - 3\beta_3 v^2(t-i)\big\} v(t-i) \\ + \beta_2 v^2(t-i) + \beta_3 v^3(t-i)\big], \end{aligned} \tag{56}$$

$$\begin{aligned} e(t) = 2\beta_3 \sum_{i=0}^{M} g(i)\big[u^2(t-i) - v^2(t-i)\big] v(t-i) \\ - \sum_{i=0}^{M} g(i)[\beta_1 + 2\beta_2 u(t-i)] v(t-i) + \eta(t) \end{aligned} \tag{57}$$

and (56) is equivalent to (19) with

$$\psi_a(t) = \begin{bmatrix} 1 \\ v(t) \\ \vdots \\ v(t-M) \\ v^2(t) \\ \vdots \\ v^2(t-M) \\ v^3(t) \\ \vdots \\ v^3(t-M) \end{bmatrix}, \quad \theta_a(t) = \begin{bmatrix} \sum_{i=0}^{M} g(i)\big[\beta_0 - \beta_2 v^2(t-i)\big] \\ \gamma_0\big[\beta_1 - 3\beta_3 v^2(t)\big] \\ \vdots \\ \gamma_0\big[\beta_1 - 3\beta_3 v^2(t-M)\big] \\ \beta_2 g(0) \\ \vdots \\ \beta_2 g(M) \\ \beta_3 g(0) \\ \vdots \\ \beta_3 g(M) \end{bmatrix}. \tag{58}$$

Corollary 3: *Suppose that A1–A4 are satisfied. Then $\hat{\theta}_a(\tau, N)$ for (19) with (57) and (58) asymptotically converges to $\bar{E}[\theta_a(t)]$.*

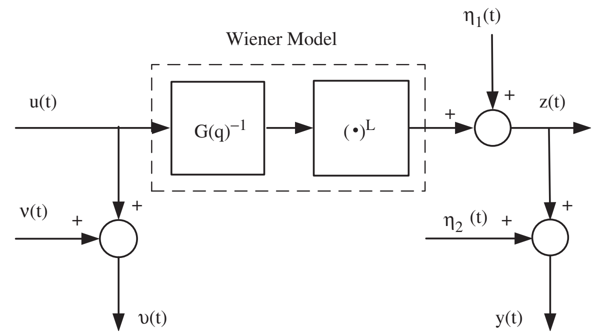

Figure 3. Errors-in-variables Wiener non-linear dynamic models.

Proof: Applying the steps in Appendix B to (57) and (58) instead of (31)–(33) proves Corollary 3. □

4.3 *Wiener models*

Consider the Wiener non-linear models (Doyle *et al.* 2002, p. 24). Figure 3 shows EIV Wiener models, where we restrict the linear part to a class of FIR models and the non-linear part to a polynomial of degree L.

To the Wiener model

$$z(t) = \sum_{\ell=0}^{L} \beta_\ell \left(\sum_{i=0}^{M} g(i) u(t-i) \right)^{\ell} + \eta_1(t). \tag{59}$$

applying (3) and (4) yields the EIV Wiener model

$$y(t) = \sum_{\ell=0}^{L} \beta_\ell \left(\sum_{i=0}^{M} g(i)[v(t-i) - \nu(t-i)] \right)^{\ell} + \eta(t). \tag{60}$$

Consider a case with $L = 2$. In this case, (60) is expressed as

$$\begin{aligned} y(t) = \beta_0 - \beta_2 \sum_{i=0}^{M} g^2(i) v^2(t-i) \\ - 2\beta_2 \times \sum_{i=0}^{M-1} \sum_{j=i+1}^{M} g(i)g(j) v(t-i) v(t-j) \\ + \beta_1 \sum_{i=0}^{M} g(i) v(t-i) + \beta_2 \sum_{i=0}^{M} g^2(i) v^2(t-i) \\ + 2\beta_2 \sum_{i=0}^{M-1} \sum_{j=i+1}^{M} g(i)g(j) v(t-i) v(t-j) + e(t), \end{aligned} \tag{61}$$

$$\begin{aligned} e(t) = \eta(t) - \beta_1 \sum_{i=0}^{M} g(i) v(t-i) \\ - 2\beta_2 \sum_{i=0}^{M-1} \sum_{j=i+1}^{M} g(i)g(j)[u(t-i) v(t-j) \\ - u(t-j) v(t-i)]. \end{aligned} \tag{62}$$

Equation (61) is represented by (19) with

$$\psi_a(t) = \begin{bmatrix} 1 \\ v(t) \\ \vdots \\ v(t-M) \\ v^2(t) \\ \vdots \\ v^2(t-M) \\ v(t)v(t-1) \\ \vdots \\ v(t)v(t-M) \\ \vdots \\ v(t-M+1)v(t-M) \end{bmatrix}, \quad \theta_a(t) = \begin{bmatrix} \gamma_0(t) \\ \beta_1 g(0) \\ \vdots \\ \beta_1 g(M) \\ \beta_2 g^2(0) \\ \vdots \\ \beta_2 g^2(M) \\ 2\beta_2 g(0)g(1) \\ \vdots \\ 2\beta_2 g(0)g(M) \\ \vdots \\ 2\beta_2 g(M-1)g(M) \end{bmatrix}. \tag{63}$$

where $\gamma_0(t) \triangleq \beta_0 - \beta_2 \sum_{i=0}^{M} g^2(i)v^2(t-i) - 2\beta_2 \times \sum_{j=i+1}^{M} g(i)g(j)v(t-i)v(t-j)$.

Corollary 4: *Suppose that A1–A4 are satisfied. Then $\hat{\theta}_a(\tau, N)$ for (19) with (62) and (63) asymptotically converges and to $\bar{E}[\theta_a(t)]$.*

Proof: Using the steps in Appendix A to (62) and (63) instead of (20)–(22) gives Corollary 4. □

5. Numerical example

Consider the simple non-linear model

$$z(t) = \theta u^L(t), \quad L = 2, 3, 4, \ \theta = 1, \tag{64}$$

$$u(t) = \sqrt{2}\sin 2\pi t, \tag{65}$$

and the measurements (3) and (4). For simplicity, we assume that $\eta(t) = 0$ and $v(t)$ is white Gaussian with variance σ_v^2. Each simulation chooses σ_v^2 such that the signal-to-noise ratio (SNR),

$$\mathrm{SNR_i} = 10\log_{10}\left(\frac{\bar{E}[u^2(t))]}{E[v^2(t)]}\right) \ [\mathrm{dB}], \tag{66}$$

is realized.

$$\hat{\hat{\theta}} = \bar{\hat{\theta}} \pm 3\bar{\sigma}(\hat{\theta}),$$

where $\bar{\hat{\theta}}$ and $\bar{\sigma}(\hat{\theta})$ denote the empirical mean and standard deviation of $\hat{\theta}$, respectively. Figure 4 shows a clear trend that the ELC estimators decrease $\hat{\hat{\theta}}$ as the number of data increases, which supports Theorem 1. Observations on Figures 5–7 give a confidence that the

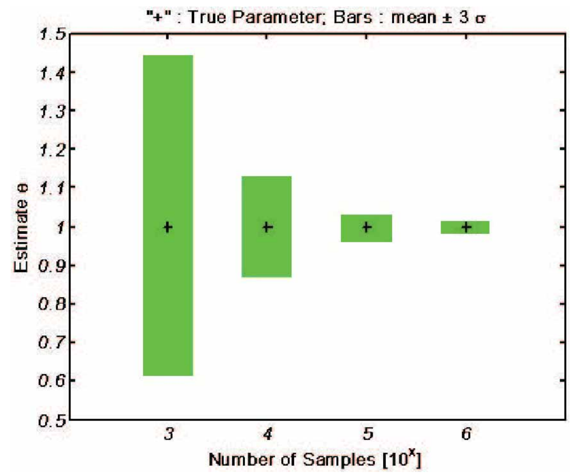

Figure 4. Effect of number of samples: $L = 2$, $\tau = 1$, $\mathrm{SNR_i} = 5\,\mathrm{dB}$.

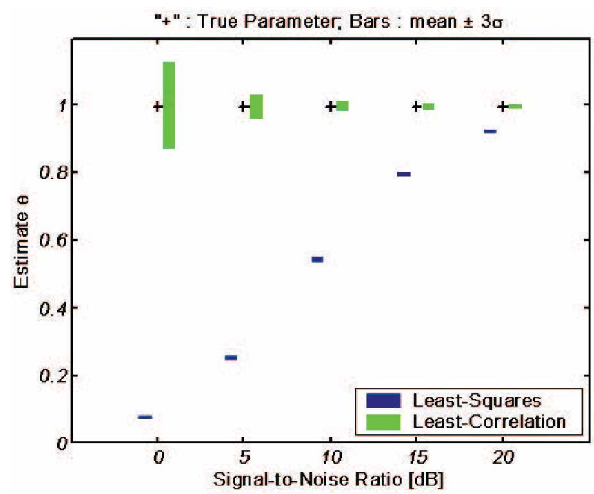

Figure 5. Input noise effect: $L = 2, \tau = 1, 10^5$ samples.

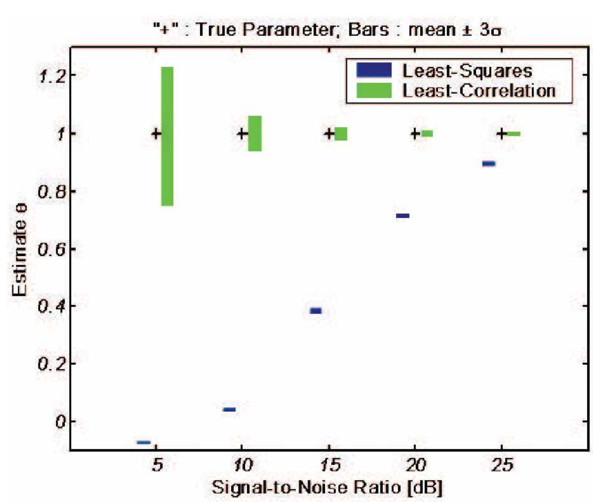

Figure 6. Input noise effect: $L = 3, \tau = 1, 10^5$ samples.

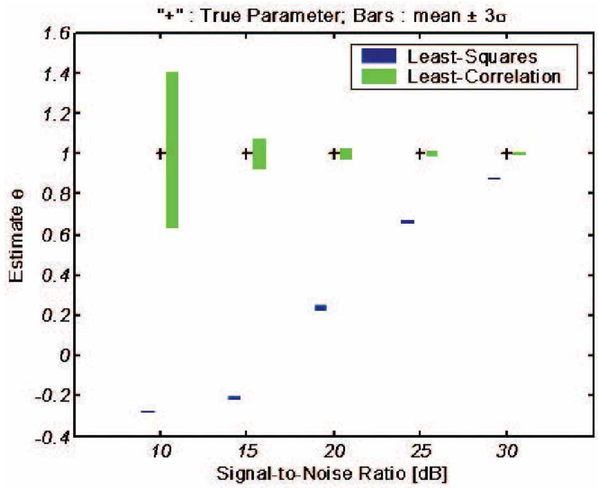

Figure 7. Input noise effect: $L = 4, \tau = 1, 10^5$ samples.

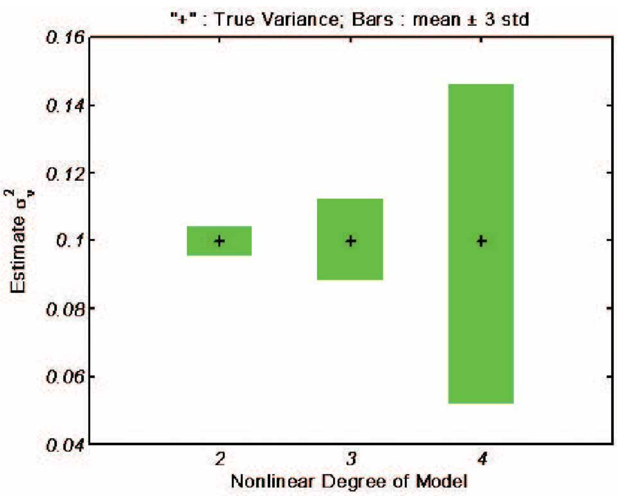

Figure 8. Input noise effect: $\mathrm{SNR_i} = 5\mathrm{dB}, \tau = 1$, 10^6 samples.

ELC estimators work well for the EIV non-linear models. Even though the ELC estimator is employed for a system with fourth-order non-linearity, figure 7 shows quite good estimates.

As a byproduct the method of ELC gives an estimate of noise variance. Figure 8 shows the expectation of the estimates. In the figure, '+' denotes the true variance and the bar means the expected range of variance estimates defined by

$$\hat{\hat{\sigma}}_v^2 = \bar{\hat{\hat{\sigma}}}_v^2 \pm 3\bar{\sigma}(\hat{\sigma}_v^2),$$

where $\bar{\hat{\hat{\sigma}}}_v^2$ and $\bar{\sigma}(\hat{\sigma}_v^2)$ denote the empirical mean and standard deviation of $\hat{\sigma}_v^2$, respectively. One thing observed in figure 8 is that the larger the non-linear degree L is the more data the estimate needs. It is thought that this trend is related to the condition number of $\bar{R}_{\psi_a\psi_a}$ which tends to increase as L increases.

6. Concluding remarks

In the present paper, we are interested in non-linear systems which contain polynomials and have equivalent realizations known as the linear regression model, and in which all measurements are contaminated by noise. The conventional method of least squares applied to the systems tends to give error-prone estimates. On the other hand, recently proposed estimators (Vandersteen *et al.* 1996, Fazekas and Kukush 1997, Hermey and Watson 1999, Baran 2000, Li 2002, Vajk and Hetthessy 2003) need *a priori* information about the moments of noise.

A main idea is to propose a way of augmenting regressors and extending parameters of the linear regression models so that the new models are equivalent to the original systems. Applying the method of least correlation to the equivalent model with extended parameters and augmented regressors yields the ELC estimator which has many good properties as follows. For third or lower order polynomials the ELC estimator gives consistent estimates of system parameters without *a priori* knowledge about noise covariance. It is easy to apply the method to Hammerstein models, Wiener models as well as Volterra models. The ELC estimate, moreover, conserves many good properties of the least-correlation estimator such as simple structure, clear understanding of minimization policy and recursive realization (Jun and Bernstein 2006).

The ELC estimator is designed for post processing or off-line application. For online or real-time application, the ELC estimator is realized to a recursive procedure, recursive least-correlation (RLC) algorithm (Jun and Bernstein 2006), expressed as

$$\begin{aligned}
\hat{\theta}_a(\tau, t) &= \hat{\theta}_a(\tau, t-1) + K(t)\Big(y_{t-\tau/t} - \psi_{t-\tau/t}^T \hat{\theta}_a(\tau, t-1)\Big)\\
K(t) &= P(t-1)\psi_{t/t-\tau}\Big(I + \psi_{t-\tau/t}^T P(t-1)\psi_{t/t-\tau}\Big)^{-1}\\
P(t) &= P(t-1) - K(t)\psi_{t-\tau/t}^T P(t-1),
\end{aligned}$$

where $\psi_{i/i-\tau} \triangleq [\,\psi_a(i) \;\; \psi_a(i-\tau)\,]$, $\psi_{i-\tau/i} \triangleq [\,\psi_a(i-\tau) \times \psi_a(i)]$, $y_{i-\tau/i} \triangleq [\,y(i-\tau) \;\; y(i)\,]^T$ for $t > \tau$ provided that $P(\tau)$ and $\hat{\theta}_a(\tau, \tau)$ are given. Computational burden of the RLC algorithm is approximately double the corresponding RLS (recursive least-squares) algorithm since $\psi_{i/i-\tau} \in \mathbb{R}^{n\times 2}$ or $\psi_{i-\tau/i} \in \mathbb{R}^{n\times 2}$ in the RLC plays the role of the regression vector with order n in the RLS.

For fourth or higher order non-linearity, unfortunately, the ELC estimator tends to give estimates containing a bit of error which is due to the correlation between the augmented regressors $\psi_a(t)$ and the extended parameters $\theta_a(t)$. Actually the correlation keeps the bicorrelation $\mathbf{t}_{\psi_a\psi_a\theta_a}(\tau,\tau)$ from being divided into $R_{\psi_a\psi_a}(\tau)$ and $E[\theta_a(t)]$ unless $\psi_a(t)$ is decoupled from $\theta_a(t)$. The rank condition of correlation matrices in A.4, which is related to the sufficient excitation or identifiability of the ELC estimate, is another weak point of the ELC estimate. When input signals are periodic, for instance, there exists a time-intervals τ which makes $R_{\psi_a\psi_a}(\tau)$ singular. Designing system inputs, if possible, will be an approach to avoiding the singularity.

Appendix

A. *Proof of Theorem* 1

With the empirical correlations defined by

$$\bar{R}_{\psi_a\psi_a}(t_1,t_2,N) \triangleq \frac{1}{N_\tau}\sum_{t=1+\tau}^{N}\psi_a(t_1)\psi_a^T(t_2), \tag{A1}$$

$$\bar{r}_{\psi_a y}(t_1,t_2,N) \triangleq \frac{1}{N_\tau}\sum_{t=1+\tau}^{N}\psi_a(t_1)y(t_2), \tag{A2}$$

where $\tau \triangleq |t_1 - t_2|$ and $t_1 = t,\ t_2 = t-\tau$ or $t_1 = t-\tau,\ t_2 = t$, (26) is rewritten as

$$\begin{aligned}\hat{\theta}_a(\tau,N) = {}&\left\{\bar{R}_{\psi_a\psi_a}(t,t-\tau,N) + \bar{R}_{\psi_a\psi_a}(t-\tau,t,N)\right\}^{-1}\\ &\times\left\{\bar{r}_{\psi_a y}(t,t-\tau,N) + \bar{r}_{\psi_a y}(t-\tau,t,N)\right\}.\end{aligned} \tag{A3}$$

Using (19) to (A2) gives

$$\bar{r}_{\psi_a y}(t_1,t_2,N) = \bar{\mathbf{t}}_{\psi_a\psi_a\theta_a}(t_1,t_2,t_2,N) + \bar{r}_{\psi_a e}(t_1,t_2,N), \tag{A4}$$

where the empirical bicorrelation $\bar{\mathbf{t}}_{\psi_a\psi_a\theta_a}(t_1,t_2,t_2,N)$ (Koh and Powers 1985) and the empirical correlation $\bar{r}_{\psi_a e}(t_1,t_2,N)$ are defined by

$$\bar{\mathbf{t}}_{\psi_a\psi_a\theta_a}(t_1,t_2,t_2,N) \triangleq \frac{1}{N_\tau}\sum_{t=1+\tau}^{N}\psi_a(t_1)\psi_a^T(t_2)\theta_a(t_2), \tag{A5}$$

$$\bar{r}_{\psi_a e}(t_1,t_2,N) \triangleq \frac{1}{N_\tau}\sum_{t=1+\tau}^{N}\psi_a(t_1)e(t_2), \tag{A6}$$

respectively.

Suppose that N increases to infinity. Then $\bar{R}_{\psi_a\psi_a}(t_1,t_2,N)$ and $\bar{r}_{\psi_a e}(t_1,t_2,N)$ converge to $R_{\psi_a\psi_a}(t_1-t_2)$ and $r_{\psi_a e}(t_1-t_2)$, respectively, due to the ergodic theory (Ljung 1999, Theorem 2.3 in p. 43). $R_{\psi_a\psi_a}(t_1-t_2)$ and $r_{\psi_a e}(t_1-t_2)$, moreover, depend on $\tau = |t_1 - t_2|$ owing to A2. Now (A1) and (A6) are expressed as

$$\lim_{N\to\infty}\bar{R}_{\psi_a\psi_a}(t_1,t_2,N) = R_{\psi_a\psi_a}(\tau) \tag{A7}$$

$$\lim_{N\to\infty}\bar{r}_{\psi_a e}(t_1,t_2,N) = r_{\psi_a e}(\tau), \tag{A8}$$

respectively, where $R_{\psi_a\psi_a}(\tau)$ and $r_{\psi_a e}(\tau)$ are evaluated as follows:

$$R_{\psi_a\psi_a}(\tau) = \bar{E}\begin{bmatrix} v_{(2)}(t)v_{(2)}^T(t-\tau) & -v_{(2)}(t)\\ -v_{(2)}^T(t-\tau) & 1\end{bmatrix}, \tag{A9}$$

$$\begin{aligned}r_{\psi_a e}(\tau) = {}&\bar{E}[\psi_a(t)\eta(t-\tau)]\\ &-\sum_{j=1}^{n}a_j\bar{E}[\psi_a(t)u_j(t-\tau)v_j(t-\tau)] = 0,\end{aligned} \tag{A10}$$

with $v_{(\ell)}^T(t) \triangleq [\, v_1^\ell(t) \quad \cdots \quad v_n^\ell(t)\,]$. Similarly the bicorrelation in (A5) is evaluated as

$$\begin{aligned}\mathbf{t}_{\psi_a\psi_a\theta_a}(\tau,\tau) &= \bar{E}\begin{bmatrix} v_1^2(t)\sum_{j=1}^{n}a_j\Big(v_j^2(t-\tau) - v_j^2(t-\tau)\Big)\\ \vdots\\ v_n^2(t)\sum_{j=1}^{n}a_j\Big(v_j^2(t-\tau) - v_j^2(t-\tau)\Big)\\ -\sum_{j=1}^{n}a_j\Big(v_j^2(t-\tau) - v_j^2(t-\tau)\Big)\end{bmatrix}\\ &= R_{\psi_a\psi_a}(\tau)E[\theta_a(t)].\end{aligned} \tag{A11}$$

Using (A10) and (A11) for (A4) at $N\to\infty$ yields

$$r_{\psi_a y}(\tau) = R_{\psi_a\psi_a}(\tau)E[\theta_a(t)]. \tag{A12}$$

Finally applying (A7) and (A12) to (A3) at $N\to\infty$ gives (28). □

B. *Proof of Theorem* 2

Equations (A1)–(A8) in Appendix A work on (19) with (31)–(33). For the third-order model, $R_{\psi_a\psi_a}(\tau)$ is evaluated as

$$R_{\psi_a\psi_a}(\tau) = \begin{bmatrix} R_{v^3v^3}(\tau) & -3R_{v^3v^1}(\tau)\\ -3R_{v^1v^3}(\tau) & 9R_{v^1v^1}(\tau)\end{bmatrix}, \tag{B13}$$

where $R_{v^\ell v^k}(\tau) \triangleq \bar{E}[v_{(\ell)}(t)v_{(k)}^T(t-\tau)]$ for $\ell, k = 1, 3$. Applying $\mathbf{t}_{\psi_a\psi_a\theta_a}(\tau,\tau)$ and $r_{\psi_a e}(\tau)$ evaluated as

$$\mathbf{t}_{\psi_a\psi_a\theta_a}(\tau,\tau) = \bar{E}\begin{bmatrix} v_1^3(t)\sum_{j=1}^n a_j\Big(v_j^3(t-\tau) - 3v_j(t-\tau)v_j^2(t-\tau)\Big) \\ \vdots \\ v_n^3(t)\sum_{j=1}^n a_j\Big(v_j^3(t-\tau) - 3v_j(t-\tau)v_j^2(t-\tau)\Big) \\ -3v_1(t)\sum_{j=1}^n a_j\Big(v_j^3(t-\tau) - 3v_j(t-\tau)v_j^2(t-\tau)\Big) \\ \vdots \\ -3v_n(t)\sum_{j=1}^n a_j\Big(v_j^3(t-\tau) - 3v_j(t-\tau)v_j^2(t-\tau)\Big) \end{bmatrix}$$

$$= R_{\psi_a\psi_a}(\tau)\bar{E}[\theta_a(t)], \tag{B14}$$

$$r_{\psi_a e}(\tau) = \bar{E}[\psi_a(t)\eta(t-\tau)] - 3\bar{E}\left[\psi_a(t)\sum_{j=1}^n a_j u_j^2(t-\tau)v_j(t-\tau)\right] + 2\bar{E}\left[\psi_a(t)\sum_{j=1}^n a_j v_j^3(t-\tau)\right] = 0, \tag{B15}$$

respectively, to (A4) at $N\to\infty$ yields (A12). Finally using (A7) and (A12) with (A3) at $N\to\infty$ gives (34). □

C. *Proof of equation* (45)

Refer to (A1)–(A8) in Appendix A. First suppose that L is even. Then $R_{\psi_a\psi_a}(\tau)$ is expressed as

$$R_{\psi_a\psi_a}(\tau) = \begin{bmatrix} R_{v'^L v'^L}(\tau) & \cdots & R_{v'^L v'^2}(\tau) & r_{v'^L} \\ \vdots & \cdots & \vdots & \vdots \\ R_{v'^2 v'^L}(\tau) & \cdots & R_{v'^2 v'^2}(\tau) & r_{v'^2} \\ r_{v'^L}^T & \cdots & r_{v'^2}^T & \beta_L^2 \end{bmatrix}, \tag{C16}$$

where each component of the matrix is given by

$$R_{v'^\ell v'^k}(\tau) \triangleq \bar{E}[v'_{(\ell)}(t)v'^T_{(k)}(t-\tau)], \tag{C17}$$

$$r_{v'^\ell} \triangleq \beta_L\beta_{L-\ell}\bar{E}[v_{(\ell)}(t)] = \beta_L\beta_{L-\ell}\bar{E}[v_{(\ell)}(t-\tau)], \tag{C18}$$

for $\ell, k \in \{L, L-2, \ldots, 2\}$. The bicorrelation $\mathbf{t}_{\psi_a\psi_a\theta_a}(\tau,\tau)$ is expressed as

$$\mathbf{t}_{\psi_a\psi_a\theta_a}(\tau,\tau) = R_{\psi_a\psi_a}(\tau)\bar{E}[\theta_a(t)] + \mathcal{E}(t, t-\tau) \tag{C19}$$

where $\mathcal{E}(t,t-\tau) = [\,\mathcal{E}_L^T \quad \mathcal{E}_{L-2}^T \quad \cdots \quad \mathcal{E}_2^T \quad \mathcal{E}_0\,]^T$ and $\mathcal{E}_0$, $\mathcal{E}_k$ are given as

$$\mathcal{E}_0 = \sum_{\ell=\mathcal{L}_e}\Big(\beta_L\bar{E}[v'^T_{(\ell)}(t-\tau)v'_{(L-\ell)}(t-\tau)] - r_{v'^\ell}^T E[v'_{(L-\ell)}(t)]\Big), \tag{C20}$$

$$\mathcal{E}_k = \sum_{\ell=\mathcal{L}_e}\Big(\bar{E}[v'_{(k)}(t)v'^T_{(\ell)}(t-\tau)v'_{(L-\ell)}(t-\tau)] - R_{v'^k v'^\ell}(\tau)E[v'_{(L-\ell)}(t)]\Big), \tag{C21}$$

respectively, for $\mathcal{L}_e = \{L-2,\ldots,2\}$ and $k \in \{L, L-2, \ldots, 2\}$. Then $r_{\psi_a e}(\tau)$ is evaluated as

$$r_{\psi_a e}(\tau) = \bar{E}[\psi_a(t)\eta(t-\tau)] + \beta_1\bar{E}\left[\psi_a(t)\sum_{j=1}^n a_j u_j^{L-1}(t-\tau)v_j(t-\tau)\right] + \cdots + \beta_{L-1}\bar{E}\left[\psi_a(t)\sum_{j=1}^n a_j u_j(t-\tau)v_j^{L-1}(t-\tau)\right] = 0. \tag{C22}$$

For odd L, $R_{\psi_a\psi_a}(\tau)$ and $r_{\psi_a e}(\tau)$ are written as

$$R_{\psi_a\psi_a}(\tau) = \begin{bmatrix} R_{v'^L v'^L}(\tau) & \cdots & R_{v'^L v'^1}(\tau) \\ \vdots & \cdots & \vdots \\ R_{v'^1 v'^L}(\tau) & \cdots & R_{v'^1 v'^1}(\tau) \end{bmatrix}, \tag{C23}$$

$$r_{\psi_a e}(\tau) = \bar{E}[\psi_a(t)\eta(t-\tau)] + \beta_1\bar{E}\left[\psi_a(t)\sum_{j=1}^n a_j u_j^{L-1}(t-\tau)v_j(t-\tau)\right] + \cdots + \beta_L\bar{E}\left[\psi_a(t)\sum_{j=1}^n a_j v_j^L(t-\tau)\right] = 0, \tag{C24}$$

respectively, and each component of $\mathcal{E}(t,t-\tau) = [\,\mathcal{E}_L^T \quad \mathcal{E}_{L-2}^T \quad \cdots \quad \mathcal{E}_3^T \quad \mathcal{E}_1^T\,]^T$ is given by

$$\mathcal{E}_k = \sum_{\ell=\mathcal{L}_o}\Big(\bar{E}[v'_{(k)}(t)v'^T_{(\ell)}(t-\tau)v'_{(L-\ell)}(t-\tau)] - R_{v'^k v'^\ell}(\tau)E[v'_{(L-\ell)}(t)]\Big), \tag{C25}$$

for $\mathcal{L}_o = \{L-2,\ldots,3\}$ and $k \in \{L, L-2, \ldots, 3, 1\}$. Employing either (C16)–(C22) or (C23)–(C25) to (A4) at $N\to\infty$, and applying the results to (A3) at $N\to\infty$ yields (45). □

References

S. Baran, "A consistent estimator in general functional errors-in-variables models", *Metrika*, 51, pp. 117–132, 2000.

F. J. Doyle, R. K. Pearson and B. A. Qgunnike, *Identification and Control Using Volterra Models*, London, UK: Springer, 2002.

I. Fazekas and A. G. Kukush, "Asymptotic properties of an estimator in nonlinear functional errors-in-variables models with dependent error terms", *Computers Math. Applic.*, 34, pp. 23–39, 1997.

D. Hermey, and G. A. Watson, "Fitting data with errors in variables using the Huber M-estimator", *SIAM J. Sci. Comput.*, 20, pp. 1276–1298, 1999.

R. Johansson, *System Modeling and Identification*, Englewood Cliffs, NJ: Prentice-Hall, 1993.

B. E. Jun and D. S. Bernstein, "Least-correlations estimates for errors-in-variables models", *Int. J. Adapt. Cont. Signal Processing*, 20, pp. 337–351, 2006.

T. Koh, and E. J. Powers, "Second-order volterra filtering and its application to nonlinear system identification", *IEEE Trans. Acoustics, Speech, and Signal Processing*, 33, pp. 1445–1455, 1985.

T. Li, "Robust and consistent estimation of nonlinear errors-in-variables models", *J. Economet.*, 110, pp. 1–26, 2002.

L. Ljung, *System Identification – Theory for the User* 2nd Edition, Englewood Cliffs, NJ: Prentice-Hall, 1999.

D. Mattera, "Identification of polyperiodic volterra systems by means of input–output noisy measurements", *Signal Processing*, 75, pp. 41–50, 1999.

D. Mattera and L. Paura, "Higher-order cyclostationary-based methods for identifying volterra systems by input-output noisy measurements", *Signal Processing*, 67, pp. 77–98, 1998.

W. J. Rugh, *Nonlinear System Theory – The Volterra/Wiener Approach*, Baltimore, MD: The Johns Hopkins University Press, 1981.

W. Scherrer and M. Deistler, "A structure theory for linear dynamic errors-in-variables models", *SIAM J. Cont. Optimi.*, 36, pp. 2148–2175, 1998.

T. Soderstrom, U. Soverini and K. Mahata, "Perspectives on errors-in-variables estimation for dynamic systems", *Signal Processing*, 82, pp. 1139–1154.

A. H. Tan and K. Godfrey, "Identification of wiener-hammerstein models using linear interpolation in the frequency domain (lifed)", *IEEE Trans. Instrumentation and Measurement*, 51, pp. 509–521, 2002.

I. Vajk, and J. Hetthessy, "Identification of nonlinear errors-in-variables models", *Automatica*, 39, pp. 2099–2107, 2003.

G. Vandersteen, Y. Rolain, J. Schoukens and P. Pintelon, "On the use of system identification for accurate parametric modeling of nonlinear systems using noisy measurements", *IEEE Trans. Instrumentation and Measurement*, 45, pp. 605–609, 1996.

S. van Huffel and P. Lemmerling, *Total Least Squares Technique and Errors-in-Variables Modeling: Analysis, Algorithms and Applications*, Dordrecht, Netherlands: Kluwer Academic Publishers, 2002.