Chapter 2
Refined Instrumental Variable Methods for Hammerstein Box-Jenkins Models

Vincent Laurain, Marion Gilson, and Hugues Garnier

2.1 Introduction

Hammerstein block diagram model is widely represented for modelling nonlinear systems [3, 6, 8, 26]. The nonlinear block can be represented as a piecewise linear function [2] or as a sum of basis functions [7, 21].

Among the very recent work on discrete-time (DT) Hammerstein models in the time domain, the most exposed methods are the extended least squares for Hammerstein ARMAX models [6] which were further extended to Hammerstein OE models [7]. E.R Bai exposed a two stage algorithm involving least squares and single value decomposition used in different configurations [3, 4, 19] and was very recently analysed for Hammerstein Box-Jenkins models [28]. Nonetheless, the convergence properties of the algorithm are studied but there was no study driven in case of noise modelling error. Suboptimal Hammerstein model estimation in case of a bounded noise was studied in [5]. A blind maximum likelihood method is derived in [27] but the output signal is considered to be errorless.

In the continuous-time (CT) case, an exhaustive survey by Rao and Unbehauen [25] shows that CT model identification methods applied to Hammerstein models are poorly studied in the literature. In [22], the authors focus on the time-derivative approximation problems while solving the optimization problem using least squares. A non-parametric method can be found in [14] while an approach

V. Laurain (✉) · M. Gilson · H. Garnier
CNRS, Nancy-Université, Vandoeuvre-lès-Nancy Cedex, France
e-mail: vincent.laurain@cran.uhp-nancy.fr

M. Gilson
e-mail: marion.gilson@cran.uhp-nancy.fr

H. Garnier
e-mail: hugues.garnier@cran.uhp-nancy.fr

L. Wang, H. Garnier (eds.), *System Identification, Environmental Modelling, and Control System Design*,
DOI 10.1007/978-0-85729-974-1_2,

dedicated to periodic input signals can be found in [33]. To the best of the authors' knowledge, the parametric estimation problem has not been addressed yet for CT Hammerstein models with colored added noise.

Section 2.2 shows how the refined instrumental variable (RIV) method introduced in [29] can be extended in order to deal with Hammerstein BJ models. Moreover, the development of instrumental variable techniques able to cope with the direct continuous-time model estimation in colored noise conditions are exposed in Sect. 2.3. All presented methods are statistically analyzed through relevant Monte Carlo simulations and the features of the proposed method are studied in the different pre-cited contexts.

2.2 Discrete-Time Hammerstein Model Identification

2.2.1 System Description

Consider the Hammerstein system represented in Fig. 2.1 and assume that both input and output signals, $u(t_k)$ and $y(t_k)$ are uniformly sampled at a constant sampling time $\mathtt{T}_\mathrm{s}$ over N samples. The Hammerstein system $\mathscr{S}_\mathrm{o}$, is described by the following input-output relationship:

$$\mathscr{S}_\mathrm{o} \quad \begin{cases} \bar{u}(t_k) = f(u(t_k)), \\ \chi_\mathrm{o}(t_k) = G_\mathrm{o}(q)\bar{u}(t_k), \\ y(t_k) = \chi_\mathrm{o}(t_k) + v_\mathrm{o}(t_k), \end{cases} \tag{2.1}$$

where u and y are the deterministic input and noisy output respectively, χ_o is the noise-free output and v_o the additive noise with bounded spectral density. $G_\mathrm{o}(q)$ is the linear transfer function which can be written as

$$G_\mathrm{o}(q) = \frac{B_\mathrm{o}(q^{-1})}{A_\mathrm{o}(q^{-1})}, \tag{2.2}$$

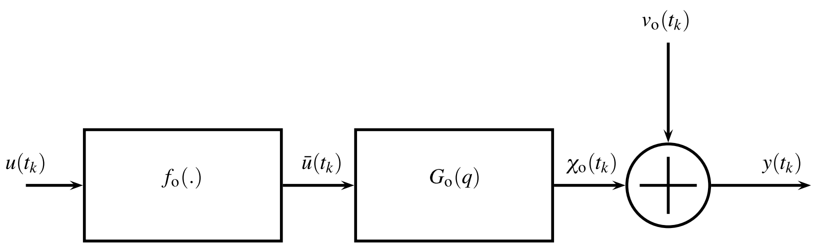

Fig. 2.1 Hammerstein block representation

where $B_o(q^{-1})$ and $A_o(q^{-1})$ are polynomial in q^{-1} of degree n_b and n_a respectively:

$$A_o(q^{-1}) = 1 + \sum_{i=1}^{n_a} a_i^o q^{-i}, \quad \text{and} \quad B_o(q^{-1}) = \sum_{j=0}^{n_b} b_j^o q^{-j}, \tag{2.3}$$

where the coefficients a_i^o and $b_j^o \in \mathbb{R}$. The most general case is considered where the colored noise associated with the sampled output measurement $y(t_k)$ is assumed to have a rational spectral density which might have no relation to the actual process dynamics of $\mathscr{S}_o$. Therefore, v_o is represented by a discrete-time *autoregressive moving average* (ARMA) model:

$$v_o(t_k) = H_o(q)e_o(t_k) = \frac{C_o(q^{-1})}{D_o(q^{-1})} e_o(t_k), \tag{2.4}$$

where $C_o(q^{-1})$ and $D_o(q^{-1})$ are monic polynomials with constant coefficients and with respective degree n_c and n_d. Furthermore, all roots of $z^{n_d} D_o(z^{-1})$ are inside the unit disc. It can be noticed that in case $C_o(q^{-1}) = D_o(q^{-1}) = 1$, (2.4) defines an OE noise model. It can be noticed that the same theory could be straightforwardly used if some pure delay was present on the input but this case is not exposed here for clarity's sake.

2.2.2 *Model Considered*

Next we introduce a discrete-time Hammerstein Box-Jenkins (BJ) type of model structure that we propose for the identification of the data-generating system (2.1) with noise model (2.4). In the chosen model structure, the noise model and the process model are parameterized separately.

2.2.2.1 Linear Part of the Hammerstein Model

The linear process model is denoted by $\mathscr{L}_{\rho_L}$ and is defined in a linear representation form as:

$$\mathscr{L}_{\rho_L} : \big(A(q^{-1}, \rho_L), B(q^{-1}, \rho_L)\big), \tag{2.5}$$

where the polynomials A and B are parameterized as

$$\mathscr{L}_{\rho_L} \quad \left\{ A(q^{-1}, \rho_L) = 1 + \sum_{i=1}^{n_a} a_i q^{-i}, \quad \text{and} \quad B(q^{-1}, \rho_L) = \sum_{j=0}^{n_b} b_j q^{-j}. \right.$$

The associated model parameters ρ_L are stacked columnwise:

$$\rho_L = \big[a_1 \dots a_{n_a}\ b_0 \dots b_{n_b}\big]^\top \in \mathbb{R}^{n_a+n_b+1}. \tag{2.6}$$

Introduce also $\mathcal{L} = \{\mathscr{L}_{\rho_\mathrm{L}} \mid \rho \in \mathbb{R}^{n_{\rho_\mathrm{L}}}\}$, as the collection of all process models in the form of (2.5).

2.2.2.2 Nonlinear Part of the Hammerstein Model

The static nonlinearity model is denoted by $\mathscr{F}_{\rho_\mathrm{NL}}$ and defined:

$$\mathscr{F}_{\rho_\mathrm{NL}} : (f(u, \rho_\mathrm{NL})) \tag{2.7}$$

where $f(u, \rho_\mathrm{NL})$ is parameterized as a sum of basis functions

$$f(u(t_k), \rho_\mathrm{NL}) = \sum_{i=1}^{l} \alpha_i(\rho_\mathrm{NL})\gamma_i(u(t_k)). \tag{2.8}$$

In this parametrization, $\{\gamma_i\}_{i=1}^{l}$ are meromorphic functions[1] of $u(t_k)$ which are assumed to be *a priori* known. Furthermore, they have a static dependence on u, and are chosen such that they allow the identifiability of the model (pairwise orthogonal functions on $\mathbb{R}$ for example). The associated model parameters ρ_NL are stacked columnwise:

$$\rho_\mathrm{NL} = \begin{bmatrix} \alpha_1 \dots \alpha_l \end{bmatrix}^\top \in \mathbb{R}^l, \tag{2.9}$$

Introduce also $\mathfrak{F} = \{\mathscr{F}_{\rho_\mathrm{NL}} \mid \rho_\mathrm{NL} \in \mathbb{R}^l\}$, as the collection of all process models in the form of (2.7).

Remark Note that the Hammerstein model $(\beta f(u, \rho_\mathrm{NL}), \frac{G(q, \rho_\mathrm{L})}{\beta})$ produces the same input-output data for any β. Therefore, to get a unique parametrization, the gain of $(\beta f(u, \rho_\mathrm{NL})$ or $G(q, \rho_\mathrm{L})/\beta$ has to be fixed [1, 6]. Hence, the first coefficient of the function $f(.)$ is fixed to 1, i.e. $\alpha_1 = 1$ in (2.9).

2.2.2.3 Noise Model

The noise model denoted by $\mathscr{H}$ is defined as a *linear time invariant* (LTI) transfer function:

$$\mathscr{H}_\eta : (H(q, \eta)), \tag{2.10}$$

where H is a monic rational function given in the form of

$$H(q, \eta) = \frac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)} = \frac{1 + c_1 q^{-1} + \dots + c_{n_\mathrm{c}} q^{-n_\mathrm{c}}}{1 + d_1 q^{-1} + \dots + d_{n_\mathrm{d}} q^{-n_\mathrm{d}}}. \tag{2.11}$$

[1] $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a real meromorphic function if $f = g/h$ with g, h analytic and $h \neq 0$.

The associated model parameters η are stacked columnwise in the parameter vector,

$$\eta = \begin{bmatrix} c_1 \ \dots \ c_{n_\mathrm{c}} \ d_1 \ \dots \ d_{n_\mathrm{d}} \end{bmatrix}^\top \in \mathbb{R}^{n_\eta}, \tag{2.12}$$

where $n_\eta = n_\mathrm{c} + n_\mathrm{d}$. Additionally, denote $\mathcal{H} = \{\mathscr{H}_\eta \mid \eta \in \mathbb{R}^{n_\eta}\}$, the collection of all noise models in the form of (2.10).

2.2.2.4 Whole Hammerstein Model

With respect to a given nonlinear, linear process and noise part $(\mathscr{F}_{\rho_\mathrm{NL}}, \mathscr{L}_{\rho_\mathrm{L}}, \mathscr{H}_\eta)$, the parameters can be collected as

$$\theta_H = [\, \rho_\mathrm{L}^\top \ \rho_\mathrm{NL}^\top \ \eta^\top \,], \tag{2.13}$$

and the signal relations of the Hammerstein BJ model, denoted in the sequel as $\mathscr{M}_\theta$, are defined as:

$$\mathscr{M}_{\theta_\mathrm{H}} \quad \begin{cases} \bar{u}(t_k) = \sum_{i=1}^{l} \alpha_i(\rho_\mathrm{NL}) \gamma_i(u(t_k)), \\ A(q^{-1}, \rho_\mathrm{L}) \chi(t_k) = B(q^{-1}, \rho_\mathrm{L}) \bar{u}(t_k), \\ v(t_k) = \dfrac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)} e(t_k), \\ y(t_k) = \chi(t_k) + v(t_k). \end{cases} \tag{2.14}$$

Based on this model structure, the model set, denoted as $\mathcal{M}$, with the linear process $(\mathscr{L}_{\rho_\mathrm{L}})$, the nonlinearity $(\mathscr{F}_{\rho_\mathrm{NL}})$ and noise $(\mathscr{H}_\eta)$ models parameterized independently, takes the form

$$\mathcal{M} = \left\{ (\mathscr{F}_{\rho_\mathrm{NL}}, \mathscr{L}_{\rho_\mathrm{L}}, \mathscr{H}_\eta) \ \mid \ \mathrm{col}(\rho_\mathrm{NL}, \rho_\mathrm{L}, \eta) = \theta_\mathrm{H} \in \mathbb{R}^{n_{\rho_\mathrm{NL}} + n_{\rho_\mathrm{L}} + n_\eta} \right\}. \tag{2.15}$$

2.2.2.5 Reformulation of the Model

The optimization problem is not convex in general. However, it can be clearly seen from the parametrization (2.8) that the model (2.14) can be rewritten in order to obtain a linear regression structure. By combining the first two equations in (2.14), the model can be rewritten as:

$$\mathscr{M}_{\theta_\mathrm{H}} \quad \begin{cases} A(q^{-1}, \rho_\mathrm{L}) \chi(t_k) = B(q^{-1}, \rho_\mathrm{L}) \sum_{i=1}^{l} \alpha_i(\rho_\mathrm{NL}) \gamma_i(u(t_k)), \\ v(t_k) = \dfrac{C(q^{-1}, \eta)}{D(q^{-1}, \eta)} e(t_k), \\ y(t_k) = \chi(t_k) + v(t_k), \end{cases} \tag{2.16}$$

which can be expanded as (note that for clarity's sake $\gamma_i(u(t_k))$ is denoted $u_i(t_k)$ in the sequel)

$$\mathscr{M}_{\theta_{\mathrm{H}}} \quad \begin{cases} A(q^{-1},\rho_{\mathrm{L}})\chi(t_k) = \sum_{i=1}^{l} \underbrace{\alpha_i(\rho_{\mathrm{NL}})B(q^{-1},\rho_{\mathrm{L}})}_{B_i(q^{-1},\rho_{\mathrm{NL}},\rho_{\mathrm{L}})}\underbrace{\gamma_i(u(t_k))}_{u_i(t_k)}, \\ v(t_k) = \dfrac{C(q^{-1},\eta)}{D(q^{-1},\eta)}e(t_k), \\ y(t_k) = \chi(t_k) + v(t_k). \end{cases} \tag{2.17}$$

Under these modelling settings, the nonlinearity model and the linear process model can be combined into the process model, denoted by $\mathscr{G}_\rho$ and defined in the form:

$$\mathscr{G}_\rho : \big(A(q^{-1},\rho), B_i(q^{-1},\rho)\big), \tag{2.18}$$

where the polynomials A and B_i are given by

$$\mathscr{G}_\rho \quad \begin{cases} A(q^{-1},\rho) = 1 + \sum_{i=1}^{n_{\mathrm{a}}} a_i q^{-i}, \\ B_i(q^{-1},\rho) = \alpha_i \sum_{j=0}^{n_{\mathrm{b}}} b_j q^{-j}, \quad i = 1 \dots l,\ \alpha_1 = 1. \end{cases}$$

The associated model parameters are stacked columnwise in the parameter vector ρ,

$$\rho = \begin{bmatrix} \mathbf{a} \\ \alpha_1 \mathbf{b} \\ \vdots \\ \alpha_l \mathbf{b} \end{bmatrix} \in \mathbb{R}^{n_\rho}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_a} \end{bmatrix} \in \mathbb{R}^{n_a}, \ \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n_b} \end{bmatrix} \in \mathbb{R}^{n_b+1}, \tag{2.19}$$

with $n_\rho = n_a + l(n_b+1)$. Introduce also $\mathcal{G} = \{\mathscr{G}_\rho \mid \rho \in \mathbb{R}^{n_\rho}\}$, as the collection of all process models in the form of (2.18). Finally, with respect to the given process and noise part $(\mathscr{G}_\rho, \mathscr{H}_\eta)$, the parameters can be collected as $\theta = [\,\rho^\top\ \eta^\top\,]$ and the signal relations of the Hammerstein BJ model, denoted in the sequel as $\mathscr{M}_\theta$, are defined as:

$$\mathscr{M}_\theta : y(t_k) = \frac{\sum_{i=1}^{l} B_i(q^{-1},\rho)u_i(t_k)}{A(q^{-1},\rho)} + \frac{C(q^{-1},\eta)}{D(q^{-1},\eta)}e(t_k), \tag{2.20}$$

with $B_i(q^{-1},\rho) = \alpha_i B(q^{-1},\rho)$ and $u_i(t_k) = \gamma_i(u(t_k))$. Based on this model structure, the whole model set including the process $(\mathscr{G}_\rho)$ and noise $(\mathscr{H}_\eta)$ models parameterized independently, is denoted as $\mathcal{M}$ and takes finally the form

$$\mathcal{M} = \big\{(\mathscr{G}_\rho, \mathscr{H}_\eta) \ \mid \ \mathrm{col}(\rho,\eta) = \theta \in \mathbb{R}^{n_\rho + n_\eta}\big\}. \tag{2.21}$$

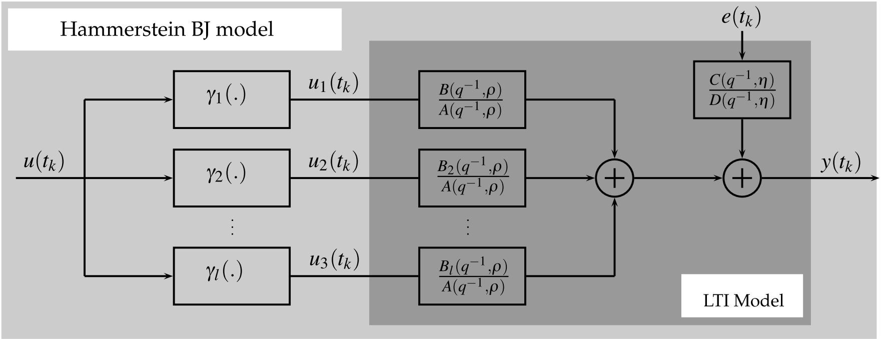

Fig. 2.2 Hammerstein augmented model

The set (2.21) corresponds to the set of candidate models in which we seek the best fitting model using data gathered from $\mathscr{S}_o$ under a given identification criterion (cost function).

Remarks It has to be noticed that this model transforms the Hammerstein structure into an augmented LTI *Multi Input Single Output* model structure such as presented in Fig. 2.2. Consequently, the number of parameters to be estimated is not minimal as $n_\rho = n_a + l(n_b + 1)$ which is in general greater than $n_{\rho_L} + n_{\rho_{NL}} = n_a + l + (n_b + 1)$. Therefore, as the model is not minimal, the optimal estimation of this augmented MISO model does not correspond to the optimal estimates of the true Hammerstein model. Nonetheless, the gain granted using this modelling is the possible linear regression form and therefore, the convexification of the optimization problem. In order to define the identification problem it is firstly necessary to define a minimization criterion. Nonetheless, the augmented model structure given in (2.20) is now an LTI structure, and therefore, the PEM framework from [20] can be directly used here.

2.2.3 Identification Problem Statement

Based on the previous considerations, the identification problem addressed can now be stated.

Problem 2.1 Given a discrete-time Hammerstein data generating system $\mathscr{S}_o$ defined as in (2.1) and a data set $\mathcal{D}_N$ collected from $\mathscr{S}_o$. Based on the Hammerstein BJ model structure $\mathscr{M}_\theta$ defined by (2.20), estimate the parameter vector θ using $\mathcal{D}_N$ under the following assumptions:

HA1 $\mathscr{S}_o \in \mathcal{M}$, i.e. there exits a θ_o defining a $\mathscr{G}_{\rho_o} \in \mathcal{G}$ and a $\mathscr{H}_{\eta_o} \in \mathcal{H}$ such that $(\mathscr{G}_{\rho_o}, \mathscr{H}_{\eta_o})$ is equal to $\mathscr{S}_o$.
HA2 $u(t_k)$ is not correlated to $e_o(t_k)$.

HA3 $\mathcal{D}_N$ is informative with respect to $\mathcal{M}$.
HA4 $\mathscr{S}_o$ is BIBO stable, i.e. for any bounded input signal u, the output of $\mathscr{S}_o$ is bounded.

2.2.4 Refined IV for Hammerstein Models

The *Hammerstein RIV* (HRIV) method derives from the RIV algorithm for DT linear systems. This was evolved by converting the maximum likelihood estimation equations to a pseudo-linear form involving optimal prefilters [29, 32]. A similar analysis can be utilised in the present situation since the problem is very similar, in both algebraic and statistical terms. The linear-in-the-parameters model (2.20) then takes the linear regression form [31]:

$$y(t_k) = \varphi^\top(t_k)\rho + \tilde{v}(t_k), \tag{2.22}$$

where ρ is as described in (2.19), $\tilde{v}(t_k) = A(q^{-1}, \rho)v(t_k)$ and

$$\varphi(t_k) = \begin{bmatrix} -\mathbf{y}(t_k) \\ \mathbf{u_1}(t_k) \\ \vdots \\ \mathbf{u_l}(t_k) \end{bmatrix}, \quad \mathbf{y}(t_k) = \begin{bmatrix} y(t_{k-1}) \\ \vdots \\ y(t_{k-n_a}) \end{bmatrix}, \quad \mathbf{u_i}(t_k) = \begin{bmatrix} u_i(t_k) \\ \vdots \\ u_i(t_{k-n_b}) \end{bmatrix}.$$

Using the conventional PEM approach on (2.22) leads to the prediction error $\varepsilon_\theta(t_k)$ given as:

$$\varepsilon_\theta(t_k) = \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)} \left\{ y(t_k) - \sum_{i=1}^{l} \frac{B_i(q^{-1}, \rho)}{A(q^{-1}, \rho)} u_i(t_k) \right\}, \tag{2.23}$$

which can be written as

$$\varepsilon_\theta(t_k) = \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)A(q^{-1}, \rho)} \left\{ A(q^{-1}, \rho)y(t_k) - \sum_{i=1}^{l} B_i(q^{-1}, \rho)u_i(t_k) \right\}, \tag{2.24}$$

where the prefilter $D(q^{-1}, \eta)/C(q^{-1}, \eta)$ will be recognised as the inverse of the ARMA(n_c, n_d) noise model. However, since the polynomial operators commute in this linear case, (2.24) can be considered in the alternative form:

$$\varepsilon_\theta(t_k) = A(q^{-1}, \rho)y_{\mathrm{f}}(t_k) - \sum_{i=1}^{l} B_i(q^{-1}, \rho)u_{i\mathrm{f}}(t_k) \tag{2.25}$$

where $y_{\mathrm{f}}(t_k)$ and $u_{i\mathrm{f}}(t_k)$ represent the outputs of the prefiltering operation using the filter:

$$Q(q, \theta) = \frac{D(q^{-1}, \eta)}{C(q^{-1}, \eta)A(q^{-1}, \rho)}. \tag{2.26}$$

Therefore, from (2.25), the associated linear-in-the-parameters model then takes the form:

$$y_f(t_k) = \varphi_f^\top(t_k)\rho + \tilde{v}_f(t_k), \tag{2.27}$$

where

$$\varphi_f(t_k) = \begin{bmatrix} -\mathbf{y}_f(t_k) \\ \mathbf{u}_{1f}(t_k) \\ \vdots \\ \mathbf{u}_{lf}(t_k) \end{bmatrix}, \quad \mathbf{y}_f(t_k) = \begin{bmatrix} y_f(t_{k-1}) \\ \vdots \\ y_f(t_{k-n_a}) \end{bmatrix}, \mathbf{u}_{if}(t_k) = \begin{bmatrix} u_{if}(t_k) \\ \vdots \\ u_{if}(t_{k-n_b}) \end{bmatrix}, \tag{2.28}$$

and $\tilde{v}_f(t_k) = Q(q, \theta)\tilde{v}(t_k) = e(t_k)$ which is a white noise.

Therefore, according to the conditions for optimal IV estimates [30], the optimal instrument and filter for the augmented LTI MISO model structure (2.20) depicted in Fig. 2.2 are given as:

$$\begin{aligned} \zeta^{\text{opt}}(t_k) = \big[&-\chi_o(t_{k-1}) \dots -\chi_o(t_{k-n_a})\; u_1(t_k) \dots u_1(t_{k-n_b}) \\ &\dots u_l(t_k) \dots u_l(t_{k-n_b})\big]^\top, \end{aligned} \tag{2.29}$$

and

$$L^{\text{opt}}(q) = Q(q, \theta_o) = \frac{D_o(q^{-1})}{C_o(q^{-1})A_o(q^{-1})}. \tag{2.30}$$

2.2.5 The Hammerstein RIV (HRIV) Algorithm for BJ Models

Of course none of $A(q^{-1}, \rho_o)$, $B_i(q^{-1}, \rho_o)$, $C(q^{-1}, \eta_o)$ or $D(q^{-1}, \eta_o)$ is known and only their estimates are available. Therefore, neither the optimal prefilter nor the optimal instrument can be accessed and they can only be estimated. The 'auxiliary model' used to generate the noise-free output as well as the computation of the associated prefilter (2.26), are updated based on the parameter estimates obtained at the previous iteration to overcome this problem.

Algorithm 2.1 (HRIV)

Step 1 Generate an initial estimate of the process model parameter $\hat{\rho}^{(0)}$(*e.g.* using the LS method). Set $C(q^{-1}, \hat{\eta}^{(0)}) = D(q^{-1}, \hat{\eta}^{(0)}) = 1$. Set $\tau = 0$.

Step 2 Compute an estimate of $\chi(t_k)$ via

$$\hat{\chi}(t_k) = \frac{\sum_{i=1}^{l} B_i(q^{-1}, \hat{\rho}^{(\tau)})u_i(t_k)}{A(q^{-1}, \hat{\rho}^{(\tau)})},$$

where $\hat{\rho}^{(\tau)}$ is the estimate obtained at the previous iteration. According to assumption HA4 each $\hat{\chi}$ is bounded.

Step 3 Compute the filter as in (2.26):

$$L(q,\hat{\theta}^{(\tau)}) = \frac{D(q^{-1},\hat{\eta}^{(\tau)})}{C(q^{-1},\hat{\eta}^{(\tau)})A(q^{-1},\hat{\rho}^{(\tau)})}$$

and the associated filtered signals $\{u_{i\mathrm{f}} = \gamma_i(u)_\mathrm{f}\}_{i=1}^{l}$, y_f and $\{\hat{\chi}_\mathrm{f}\}_{i=1,l=0}^{n_\mathrm{a},n_\alpha}$.

Step 4 Build the filtered regressor $\varphi_\mathrm{f}(t_k)$ and the filtered instrument $\hat{\zeta}_\mathrm{f}(t_k)$ which equal in the given context:

$$\begin{aligned}\varphi_\mathrm{f}(t_k) &= \big[-y_\mathrm{f}(t_{k-1}) \ldots -y_\mathrm{f}(t_{k-n_\mathrm{a}})\; u_{1\mathrm{f}}(t_k) \ldots u_{1\mathrm{f}}(t_{k-n_b}) \\ &\qquad \ldots u_{l\mathrm{f}}(t_k) \ldots u_{l\mathrm{f}}(t_{k-n_b})\big]^\top, \\ \hat{\zeta}_\mathrm{f}(t_k) &= \big[-\hat{\chi}_\mathrm{f}(t_{k-1}) \ldots -\hat{\chi}_\mathrm{f}(t_{k-n_\mathrm{a}})\; u_{1\mathrm{f}}(t_k) \ldots u_{1\mathrm{f}}(t_{k-n_b}) \\ &\qquad \ldots u_{l\mathrm{f}}(t_k) \ldots u_{l\mathrm{f}}(t_{k-n_b})\big]^\top. \end{aligned} \tag{2.31}$$

Step 5 The IV optimization problem can be stated in the form

$$\hat{\rho}^{(\tau+1)}(N) = \arg\min_{\rho\in\mathbb{R}^{n_\rho}} \left\| \left[\frac{1}{N}\sum_{k=1}^{N}\hat{\zeta}_\mathrm{f}(t_k)\varphi_\mathrm{f}^\top(t_k)\right]\rho - \left[\frac{1}{N}\sum_{k=1}^{N}\hat{\zeta}_\mathrm{f}(t_k)y_\mathrm{f}(t_k)\right]\right\|^2, \tag{2.32}$$

where the solution is obtained as

$$\hat{\rho}^{(\tau+1)}(N) = \left[\sum_{k=1}^{N}\hat{\zeta}_\mathrm{f}(t_k)\varphi_\mathrm{f}^\top(t_k)\right]^{-1}\sum_{k=1}^{N}\hat{\zeta}_\mathrm{f}(t_k)y_\mathrm{f}(t_k).$$

The resulting $\hat{\rho}^{(\tau+1)}(N)$ is the IV estimate of the process model associated parameter vector at iteration $\tau+1$ based on the prefiltered input/output data.

Step 6 An estimate of the noise signal v is obtained as

$$\hat{v}(t_k) = y(t_k) - \hat{\chi}(t_k,\hat{\rho}^{(\tau)}). \tag{2.33}$$

Based on $\hat{v}$, the estimation of the noise model parameter vector $\hat{\eta}^{(\tau+1)}$ follows, using in this case the ARMA estimation algorithm of the MATLAB identification toolbox (an IV approach can also be used for this purpose, see [30]).

Step 7 If $\theta^{(\tau+1)}$ has converged or the maximum number of iterations is reached, then stop, else increase τ by 1 and go to Step 2.

At the end of the iterative process, coefficients $\hat{\alpha}_i$ are not directly accessible. They are however deduced from polynomial $\hat{B}_i(q^{-1})$ as $B_i(q^{-1},\rho) = \alpha_i B(q^{-1},\rho)$. The hypothesis $\alpha_1 = 1$ guarantees that $B_1(q^{-1},\rho) = B(q^{-1},\rho)$ and $\hat{\alpha}_i$ can be computed from:

$$\hat{\alpha}_i = \frac{1}{n_b+1}\sum_{j=0}^{n_b}\frac{\hat{b}_{i,j}}{\hat{b}_{1,j}}, \tag{2.34}$$

where $\hat{b}_{i,j}$ is the jth coefficient of polynomial term $B_i(q^{-1}, \rho)$ for $i = 2 \dots l$.

Moreover, after the convergence is complete, it is possible to compute the estimated parametric error covariance matrix $\hat{\mathbf{P}}_\rho$ from the expression:

$$\hat{\mathbf{P}}_\rho = \hat{\sigma}_e^2 \left(\sum_{k=1}^{N} \hat{\zeta}_{\mathrm{f}}(t_k) \hat{\zeta}_{\mathrm{f}}^\top(t_k) \right)^{-1}, \tag{2.35}$$

where $\hat{\zeta}$ is the IV vector obtained at convergence and $\hat{\sigma}_e^2$ is the estimated residual variance.

Comments By using the described algorithm, if convergence occurs, the HRIV estimates might be statistically optimal for the augmented model proposed, but the minimal number of parameters needed for representing the MISO structure and the Hammerstein structure are not equal. Consequently, the HRIV estimates cannot be statistically optimal for the Hammerstein model structure. Nonetheless, even if not optimal, the HRIV estimates are unbiased with a low variance as it will be seen in the result Sect. 2.2.7.

2.2.6 HSRIV Algorithm for OE Models

A simplified version of HRIV algorithm named HSRIV follows the exact same theory for estimation of Hammerstein output error models. It is mathematically described by, $C(q^{-1}, \eta^j) = C_{\mathrm{o}}(q^{-1}) = 1$ and $D(q^{-1}, \eta^j) = D_{\mathrm{o}}(q^{-1}) = 1$. All previous given equations remain true, and it suffices to estimate ρ^j as $\theta^j = \rho^j$. The implementation of HSRIV is much simpler than HRIV as there is no noise model estimation in the algorithm.

2.2.7 Performance Evaluation of the Proposed HRIV and HSRIV Algorithms

This section presents numerical evaluation of both suggested HRIV and HSRIV methods. For the presented example, the nonlinear block has a polynomial form, i.e. $\gamma_i(u(t_k)) = u^i(t_k), \forall i$ and the system to identify is given by

$$\mathscr{S}_{\mathrm{o}} \quad \begin{cases} \bar{u}(t_k) = u(t_k) + 0.5u^2(t_k) + 0.25u^3(t_k), \\ G_{\mathrm{o}}(q) = \dfrac{0.5q^{-1} + 0.2q^{-2}}{1 + q^{-1} + 0.5q^{-2}}, \\ H_{\mathrm{o}}(q) = \dfrac{1}{1 - q^{-1} + 0.2q^{-2}}, \end{cases}$$

where $u(t_k)$ follows a uniform distribution with values between -2 and 2.

The models considered for estimation are:

$$\mathscr{M}_{\texttt{HRIV}} \quad \begin{cases} G(q,\rho) = \dfrac{b_1 q^{-1} + b_2 q^{-2}}{1 + a_1 q^{-1} + a_2 q^{-2}}, \\ H(q,\eta) = \dfrac{1}{1 + d_1 q^{-1} + d_2 q^{-1}}, \\ f(u(t_k)) = u(t_k) + \alpha_1 u^2(t_k) + \alpha_2 u^3(t_k) \end{cases} \tag{2.36}$$

for the HRIV method which fulfills [HA1] and

$$\mathscr{M}_{\texttt{HSRIV}} \quad \begin{cases} G(q,\rho) = \dfrac{b_1 q^{-1} + b_2 q^{-2}}{1 + a_1 q^{-1} + a_2 q^{-2}}, \\ H(q,\eta) = 1, \\ f(u(t_k)) = u(t_k) + \alpha_1 u^2(t_k) + \alpha_2 u^3(t_k) \end{cases} \tag{2.37}$$

for the HSRIV method which only fulfills $G_o \in \mathscr{G}$ ($H_o \notin \mathscr{H}$).

The result of a Monte Carlo simulation (MCs) analysis is shown in Table 2.1 and the algorithms considered are: HRIV, HSRIV and LSQNONLIN. The LSQNONLIN is a nonlinear optimization algorithm from the MATLAB® optimization toolbox. It assumes the same model as the HRIV method ($S_o \in \mathscr{M}$) and hands out the statistically optimal estimates if the method is properly initialized. In order to place the LSQNONLIN method at its advantage, it is initialized with the true parameter values and therefore this method can be considered as the ground truth.

The MCs results are based on $N_{\text{run}} = 100$ random realization, with the Gaussian white noise input to the ARMA noise model being selected randomly for each realization. In order to compare the statistical performance of the different approaches, the computed mean and standard deviation of the estimated parameters are presented. The noise added at the output is adjusted such that it corresponds to a *Signal-to-Noise-Ratio* (SNR) of 5dB using:

$$\text{SNR} = 10 \log\left(\frac{P_\chi}{P_{v_o}}\right), \tag{2.38}$$

where P_g is the average power of signal g. The number of samples is chosen as $N = 2000$.

As expected, Table 2.1 shows that the proposed algorithms produce unbiased estimates of the Hammerstein model parameters. It can be further noticed that the standard deviation of the estimates remains low even under the unrealistic noise level of 5 dB. Even though, the ratio between the HSRIV and HRIV estimate standard deviation equals to 2. This can be logically explained by the fact that the HSRIV algorithm assumes a wrong noise model and such result remains acceptable in practical applications. Finally it can be depicted that the HRIV provides the statistical optimal estimates for the parameters which are not replicated inside the parameter vector ρ, that is a_1, a_2, d_1 and d_2. Concerning the other coefficients the standard

Table 2.1 Estimation results of the proposed algorithm

Method		b_0	b_1	a_1	a_2	α_1	α_2	d_1	d_2
	True value	0.5	0.2	1	0.5	0.5	0.25	−1	0.2
LSQNONLIN	*mean*($\hat{\theta}$)	0.4991	0.1983	0.9984	0.4992	0.5011	0.2512	−1.0004	0.2001
	std($\hat{\theta}$)	0.0159	0.0109	0.0114	0.0059	0.0187	0.0194	0.0224	0.0219
HSRIV	*mean*($\hat{\theta}$)	0.4992	0.1975	0.9944	0.4976	0.4956	0.2657	X	X
	std($\hat{\theta}$)	0.0402	0.0471	0.0186	0.0071	0.1107	0.0999	X	X
HRIV	*mean*($\hat{\theta}$)	0.5004	0.2009	0.9984	0.4992	0.5006	0.2487	−1.0011	0.2007
	std($\hat{\theta}$)	0.0193	0.0208	0.0114	0.0059	0.0384	0.0397	0.0224	0.0220

deviation is approximately multiplied by 2 but the absolute value remains acceptable considering the level of noise added. It can be concluded that the presented algorithms, even if not optimal in the Hammerstein case, constitute good candidates for practical applications where the noise is unknown, and can be a strong help for initializing optimal methods such as LSQNONLIN.

2.3 Continuous-Time Hammerstein Model Identification

Even if measured data are sampled, the underlying dynamic of a real system is continuous and direct continuous-time model identification methods regained interest in the recent years [15]. The advantage of using direct continuous-time model identification has been pointed out in many different contexts in the LTI framework [9–13, 15, 24]. Nonetheless, a survey by Rao and Unbehauen [25] shows that CT model identification methods applied to Hammerstein models are poorly represented in literature and only a few methods can be found. In [22], the authors focus on the time-derivative approximation problems while solving the optimization problem using least squares. A non-parametric method can be found in [14] while an approach dedicated to periodic input signals can be found in [33]. To the best of the authors' knowledge, the parametric estimation problem has not been addressed yet for CT Hammerstein models which focus with some colored added noise. Consequently, this section presents an RIV algorithm for direct CT model identification for CT Hammerstein models.

2.3.1 System Description

Consider the CT Hammerstein data generating system depicted in Fig. 2.3 corresponding to the following input-output relationship:

$$\mathscr{S}_o \quad \begin{cases} \bar{u}(t) = f_o(u(t)), \\ \chi_o(t) = G_o(d)\bar{u}(t), \\ y(t) = \chi(t) + v_o(t), \end{cases} \tag{2.39}$$

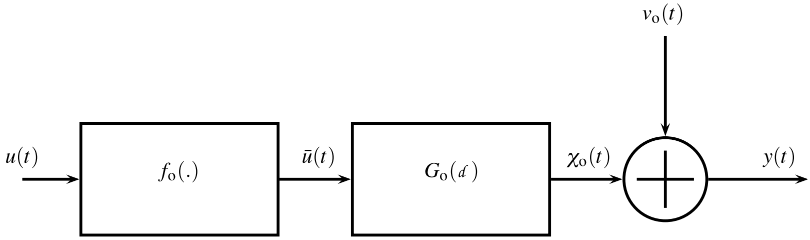

Fig. 2.3 CT Hammerstein block representation

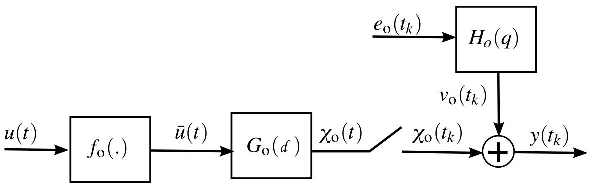

Fig. 2.4 Hybrid Hammerstein block representation

where

$$G_o(d) = \frac{B_o(d)}{A_o(d)} \tag{2.40}$$

and $B_o(d)$ and $A_o(d)$ are polynomials in the differential operator d ($d^i x(t) = \frac{d^i x(t)}{dt^i}$) of respective degree n_b and n_a ($n_a \geq n_b$).

In terms of identification we can assume that sampled measurements of (y, u) are available at a sampling time $k\mathrm{T_s} > 0$. Hence, we will denote the discrete time samples of these signals as $u(t_k) = u(k\mathrm{T_s})$, where $k \in \mathbb{Z}$. The basic idea to solve the noisy *continuous-time* (CT) modelling problem is to assume that the CT noise process $v_o(t)$ can be written at the sampling instances as a *discrete-time* (DT) white noise process filtered by a DT transfer function [16, 23]. The practically general case is considered where the colored noise associated with the sampled output measurement $y(t_k)$ is assumed to have a rational spectral density which might has no relation to the actual process dynamics. Therefore, v_o is represented by a discrete-time *autoregressive moving average* (ARMA) model:

$$v_o(t_k) = H_o(q)e_o(t_k) = \frac{C_o(q^{-1})}{D_o(q^{-1})} e_o(t_k), \tag{2.41}$$

where $e_o(t_k)$ is a DT zero mean white noise process, q^{-1} is the backward time shift operator, i.e. $q^{-i}u(t_k) = u(t_{k-i})$, and C_o with D_o are monic polynomials with constant coefficients. This avoids the rather difficult mathematical problem of treating sampled CT random process [9] and their equivalent in terms of a filtered piecewise constant CT noise source (see [23]). Therefore, we will consider the Hammerstein system represented in Fig. 2.4 where it is assumed that both input and output signals, $u(t)$ and $y(t)$ are uniformly sampled at a constant sampling time $\mathrm{T_s}$ over N samples.

Consequently, in terms of (2.41), the Hammerstein system $\mathscr{S}_o$ (2.39), is described by the following input-output relationship:

$$\mathscr{S}_o \quad \begin{cases} \bar{u}(t) = f_o(u(t)), \\ \chi_o(t) = G_o(d)\bar{u}(t), \\ v_o(t_k) = H_o(q)e(t_k), \\ y(t_k) = \chi_o(t_k) + v_o(t_k). \end{cases} \tag{2.42}$$

This corresponds to a so-called Hammerstein hybrid Box-Jenkins system concept already used in CT identification of LTI systems (see [16, 23, 31]). Furthermore, in terms of (2.4), exactly the same noise assumption is made as in the classical DT Box-Jenkins models [20].

2.3.2 *Model Considered*

2.3.2.1 Process Modelling

Similarly to the discrete-time case, by aiming at the convexification of the optimization problem, the static nonlinearity model is modelled as the linear sum of basis functions:

$$f(u(t), \rho) = \sum_{i=1}^{l} \alpha_i(\rho)\gamma_i(u(t)), \quad \alpha_1 = 1 \tag{2.43}$$

while the CT linear part can be parameterized such that:

$$\chi(t) = G(d, \rho)\bar{u}(t) = \frac{B(d, \rho)}{A(d, \rho)} f(u(t), \rho), \tag{2.44}$$

with

$$A(d, \rho) = d^{n_a} + \sum_{i=1}^{n_\mathrm{a}} a_i d^{n_a - i} \quad \text{and} \quad B(d, \rho) = \sum_{j=0}^{n_\mathrm{b}} b_j d^{n_b - j}. \tag{2.45}$$

Just as in the DT case, both equations (2.43) and (2.44) can be combined such that:

$$\chi(t) = \frac{B(d, \rho)}{A(d, \rho)} \sum_{i=1}^{l} \alpha_i(\rho)\gamma_i(u(t)) = \frac{1}{A(d, \rho)} \sum_{i=1}^{l} \underbrace{\alpha_i(\rho)B(d, \rho)}_{B_i(d,\rho)}\underbrace{\gamma_i(u(t))}_{u_i(t)}. \tag{2.46}$$

Under these modelling settings, the nonlinearity model and the linear process model can be combined into a process model, denoted by $\mathscr{G}_\rho$ and defined in the form:

$$\mathscr{G}_\rho : (A(d, \rho), B_i(d, \rho)) \tag{2.47}$$

where the polynomials A and B_i are parameterized as

$$\mathcal{G}_\rho \quad \begin{cases} A(d,\rho) = 1 + \sum_{i=1}^{n_\mathrm{a}} a_i d^{n_a - i}, \\ B_i(d,\rho) = \alpha_i \sum_{j=0}^{n_\mathrm{b}} b_j d^{n_b - j}, \quad i = 1 \ldots l. \end{cases}$$

The associated model parameters are stacked columnwise in the parameter vector ρ,

$$\rho = \begin{bmatrix} \mathbf{a} \\ \alpha_1 \mathbf{b} \\ \vdots \\ \alpha_l \mathbf{b} \end{bmatrix} \in \mathbb{R}^{n_\rho}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_a} \end{bmatrix} \in \mathbb{R}^{n_a}, \ \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n_b} \end{bmatrix} \in \mathbb{R}^{n_b+1}, \tag{2.48}$$

with $n_\rho = n_a + l(n_b + 1)$. Introduce also $\mathcal{G} = \{\mathcal{G}_\rho \mid \rho \in \mathbb{R}^{n_\rho}\}$, as the collection of all process models in the form of (2.47).

2.3.2.2 Noise Model

The noise model being expressed in discrete-time, it is denoted by $\mathcal{H}$ and defined as in the DT case (see Sect. 2.2.2.3) Additionally, denote $\mathcal{H} = \{\mathcal{H}_\eta \mid \eta \in \mathbb{R}^{n_\eta}\}$, the collection of all noise models in the form of (2.10).

2.3.2.3 Whole Model

With respect to the given process and noise parts $(\mathcal{G}_\rho, \mathcal{H}_\eta)$, the parameters can be collected as $\theta = [\,\rho^\top \ \eta^\top\,]$ and the signal relations of the CT Hammerstein BJ model, denoted in the sequel as $\mathcal{M}_\theta$, are defined as:

$$\mathcal{M}_\theta \quad \begin{cases} \chi(t) = \dfrac{\sum_{i=1}^{l} B_i(d,\rho) u_i(t)}{A(d,\rho)}, \\ v(t_k) = \dfrac{C(q^{-1},\eta)}{D(q^{-1},\eta)} e(t_k), \\ y(t_k) = \chi(t_k) + v(t_k), \end{cases} \tag{2.49}$$

with $B_i(d,\rho) = \alpha_i(\rho) B(d,\rho)$ and $u_i(t) = \gamma_i(u(t))$. Based on this model structure, the model set, denoted as $\mathcal{M}$, with the linear process $(\mathcal{G}_\rho)$ and noise $(\mathcal{H}_\eta)$ models parameterized independently, takes the form

$$\mathcal{M} = \left\{ (\mathcal{G}_\rho, \mathcal{H}_\eta) \ \mid \ \mathrm{col}(\rho,\eta) = \theta \in \mathbb{R}^{n_\rho + n_\eta} \right\}. \tag{2.50}$$

Again, this set corresponds to the set of candidate models in which we seek the best fitting model using data gathered from $\mathscr{S}_{\mathrm{o}}$ under a given identification criterion (cost function). The identification problem can be stated in the exact same way as in the DT case (see Sect. 2.2.3).

Remarks It has to be noticed that, just as in the DT case, this model transforms the Hammerstein structure into an augmented LTI *Multi Input Single Output* model similarly to the DT case. Consequently, the number of parameters to be estimated is not minimal as $n_\rho = n_a + l(n_b + 1)$ which is in general greater than $n_{\rho_{\mathrm{L}}} + n_{\rho_{\mathrm{NL}}} = n_a + l + (n_b + 1)$. Therefore, as the model is not minimal, and therefore the optimal estimation of this augmented MISO model does not correspond to the optimal estimates of the true Hammerstein model. Nonetheless, the gain granted using this modelling is the possible linear regression form and therefore, the convexification of the optimization problem.

2.3.3 Refined IV for CT Hammerstein BJ Models

Using the LTI model (2.49), $y(t_k)$ can be written in the regression form:

$$y^{(n_a)}(t_k) = \varphi^\top(t_k)\rho + \tilde{v}(t_k), \tag{2.51}$$

where

$$\varphi(t_k) = \left[-y^{(n_{\mathrm{a}}-1)}(t_k) \ldots -y(t_k)\; u_1^{(n_{\mathrm{b}})}(t_k) \ldots u_1(t_k) \ldots u_l^{(n_{\mathrm{b}})}(t_k) \ldots u_l(t_k)\right]^\top$$
$$\rho = \left[a_1 \ldots a_{n_{\mathrm{a}}}\; b_0 \ldots b_{n_{\mathrm{b}}} \ldots \alpha_l b_0 \ldots \alpha_l b_{n_{\mathrm{b}}}\right]^\top$$

and

$$\tilde{v}(t_k) = A(d, \rho)v(t_k),$$

where $x^{(n)}(t_k)$ denotes the sample of the nth derivative of the signal $x(t)$ sampled at time t_k.

By driving the exact same discussion as in Sect. 2.2.4 it can be shown that the optimal filtered instrument for the augmented LTI MISO model structure (2.49) is given as:

$$\zeta^{\mathrm{opt}}(t_k) = \left[-\chi_{\mathrm{o}}^{(n_{\mathrm{a}}-1)}(t_k) \ldots -\chi_{\mathrm{o}}(t_k)\; u_1^{(n_{\mathrm{b}})}(t_k) \ldots u_1(t_k) \ldots u_l^{(n_{\mathrm{b}})}(t_k) \ldots u_l(t_k)\right]^\top \tag{2.52}$$

while the optimal filter is given as the filter chain involving the continuous-time filtering operation using the filter (see [31]):

$$L_{\mathrm{c}}^{\mathrm{opt}} = Q_{\mathrm{c}}(d, \rho_{\mathrm{o}}) = \frac{1}{A_{\mathrm{o}}(d)}, \tag{2.53}$$

and the discrete-time filtering operation using the filter:

$$L_{\mathrm{d}}^{\mathrm{opt}} = Q_{\mathrm{d}}(q, \eta_{\mathrm{o}}) = \frac{D_{\mathrm{o}}(q^{-1})}{C_{\mathrm{o}}(q^{-1})}. \tag{2.54}$$

2.3.4 Hammerstein RIVC (HRIVC) Algorithm for BJ Models

For space and redundancy's sake the HRIVC algorithm is not described here, but the interested reader can find a detailed algorithm in [18]. By using the HRIVC algorithm, if convergence occurs, the HRIVC estimates might be statistically optimal for the augmented model proposed, but the minimal number of parameters needed for representing the MISO structure and the Hammerstein structure are not equal. Consequently, the HRIVC estimates cannot be statistically optimal for the CT Hammerstein model structure. Nonetheless, even if not optimal, the HRIVC estimates are unbiased with a low variance as it will be seen in the result Sect. 2.3.5. A simplified version of HRIVC algorithm named HSRIVC follows the exact same theory for estimation of Hammerstein output error models.

2.3.5 Performance Evaluation of the Proposed HRIVC and HSRIVC Algorithms

This section presents numerical evaluation of both suggested HRIVC and HSRIVC methods. For the presented example, the nonlinear block has a polynomial form, i.e. $\gamma_i(u(t)) = u^i(t), \forall i$ and

$$\bar{u}(t) = u(t) + 0.5u^2(t) + 0.25u^3(t), \tag{2.55}$$

where $u(t)$ follows a uniform distribution with values between -2 and 2. The system is simulated using a zero-order-hold on the input.

The system considered is a hybrid Hammerstein Box-Jenkins model in which the linear dynamic block is first a second-order system described by:

$$G_{\mathrm{o}}(d) = \frac{10d + 30}{d^2 + d + 5}, \tag{2.56}$$

and the noise is given by

$$H_{\mathrm{o}}(q) = \frac{1}{1 - q^{-1} + 0.2q^{-2}}.$$

Table 2.2 Estimation results for different noise models

SNR	Method		b_0	b_1	a_1	a_2	α_1	α_2	d_1	d_2
		True value	10	30	1	5	0.5	0.25	−1	0.2
15 dB	HSRIVC	$mean(\hat{\theta})$	9.9957	29.8760	1.0001	4.9991	0.5026	0.2523	X	X
		$std(\hat{\theta})$	0.3670	1.5660	0.0170	0.0436	0.0201	0.0180	X	X
		RMSE	0.0367	0.0523	0.0169	0.0087	0.0405	0.0723	X	X
	HRIVC	$mean(\hat{\theta})$	9.9906	30.0172	1.0006	5.0020	0.5008	0.2506	−1.0002	0.2005
		$std(\hat{\theta})$	0.2497	0.8954	0.0119	0.0265	0.0118	0.0115	0.0219	0.0223
		RMSE	0.0250	0.0298	0.0119	0.0053	0.0236	0.0460	0.0218	0.1112
5 dB	HSRIVC	$mean(\hat{\theta})$	10.0882	29.6146	1.0010	4.9814	0.5080	0.2604	X	X
		$std(\hat{\theta})$	1.0764	4.4585	0.0517	0.1291	0.0610	0.0542	X	X
		RMSE	0.1079	0.1490	0.0517	0.0261	0.1230	0.2208	X	X
	HRIVC	$mean(\hat{\theta})$	10.049	30.0277	0.9998	4.9980	0.5015	0.2522	−0.9997	0.1994
		$std(\hat{\theta})$	0.7861	2.8278	0.0379	0.0871	0.0369	0.0366	0.0227	0.0219
		RMSE	0.0787	0.0942	0.0378	0.0174	0.0738	0.1466	0.0227	0.1096

The models considered for estimation are:

$$\mathscr{M}_{\texttt{HRIVC}} \quad \begin{cases} G(d,\rho) = \dfrac{b_0 d + b_1}{d^2 + a_1 d + a_2}, \\ H(q,\eta) = \dfrac{1}{1 + d_1 q^{-1} + d_2 q^{-2}}, \\ f(u(t)) = u(t) + \alpha_1 u^2(t) + \alpha_2 u^3(t) \end{cases} \tag{2.57}$$

for the HRIVC method and

$$\mathscr{M}_{\texttt{HSRIVC}} \quad \begin{cases} G(d,\rho) = \dfrac{b_0 d + b_1}{d^2 + a_1 d + a_2}, \\ H(q,\eta) = 1, \\ f(u(t)) = u(t) + \alpha_1 u^2(t) + \alpha_2 u^3(t) \end{cases} \tag{2.58}$$

for the HSRIVC method.

The result of a Monte Carlo simulation (MCs) analysis is shown in Table 2.2 for the algorithms considered. The MCs results are based on $N_{\text{run}} = 500$ random realization, with the Gaussian white noise input to the ARMA noise model being selected randomly for each realization. In order to compare the statistical performance of the different approaches, the computed mean, standard deviation and *Root Mean Squared Error* of the estimated parameters are presented. The noise added at the output is adjusted such that it corresponds to a SNR of 15 dB and 5 dB. The number of samples is $N = 2000$.

Table 2.2 shows that according to the theory, the HRIVC and HSRIVC methods provide similar, unbiased estimates of the model parameters. Both methods seem to be robust even at unrealistic noise level of 5 dB as the RMSE remain under 22% for both methods. Results obtained using the HRIVC algorithm, have standard deviations which are always smaller than the ones produced by HSRIVC. Even though, the HSRIVC algorithm based on an Output Error model is a reasonable alternative to the full HRIVC algorithm based on a Box-Jenkins model: in practice the noise model cannot be exactly known and therefore the use of the HRIVC algorithm would simply raise the number of parameters to be estimated. If the noise model is correctly assumed, it is as well correctly estimated as shown in Table 2.2.

2.4 Conclusion

In this chapter, some methods dedicated to Hammerstein CT and DT nonlinear models in open-loop were investigated. Extension to the closed-loop case has been published and can be found in [17]. Through a relevant set of examples, it was possible to show that the HRIV approach is robust to noise conditions and to noise error modelling. The presented methods are suboptimal as they estimate a larger number of parameters than the minimum needed for the system description. Nonetheless, the variance in the estimated parameters is acceptable in practical conditions, and if not satisfactory, the estimates can be used as initialisation values for some optimal method which usually are posed into some nonlinear optimization problems and often rely on some robust initialisation. The refined instrumental variable approach for Hammerstein models remains an interesting estimation method for practical applications where the noise condition are unknown.

References

1. Bai, E.-W.: A blind approach to the Hammerstein–Wiener model identification. Automatica **38**(6), 967–979 (2002)
2. Bai, E.-W.: Identification of linear systems with hard input nonlinearities of known structure. Automatica **38**(5), 853–860 (2002)
3. Bai, E.-W., Chan, K.-S.: Identification of an additive nonlinear system and its applications in generalized Hammerstein models. Automatica **44**(2), 430–436 (2008)
4. Bai, E.W.: An optimal two-stage identification algorithm for Hammerstein–Wiener nonlinear systems. Automatica **34**(3), 333–338 (1998)
5. Cerone, V., Regruto, D.: Parameter bounds for discrete-time Hammerstein models with bounded errors. IEEE Trans. Autom. Control **48**(10), 1855–1860 (2003)
6. Ding, F., Chen, T.: Identification of Hammerstein nonlinear ARMAX systems. Automatica **41**(9), 1479–1489 (2005)
7. Ding, F., Shi, Y., Chen, T.: Auxiliary model-based least-squares identification methods for Hammerstein output-error systems. Syst. Control Lett. **56**(5), 373–380 (2007)
8. Giri, F., Bai, E.-W. (eds.): Block-Oriented Nonlinear System Identification. Springer, London (2010)
9. Gillberg, J., Ljung, L.: Frequency domain identification of continuous-time ARMA models from sampled data. Automatica **45**(6), 1371–1378 (2009)
10. Gillberg, J., Ljung, L.: Frequency domain identification of continuous-time output error models, Part I: Uniformly sampled data and frequency function approximation. Automatica **46**(1), 1–10 (2010)

11. Gillberg, J., Ljung, L.: Frequency domain identification of continuous-time output error models, part II: Non-uniformly sampled data and B-spline output approximation. Automatica **46**(1), 11–18 (2010)
12. Gilson, M., Garnier, H., Young, P.C., Van den Hof, P.: Refined instrumental variable methods for closed-loop system identification methods. In: Proceedings of the 15th IFAC Symposium on System Identification, Saint-Malo, France (2009)
13. Gilson, M., Van den Hof, P.M.J.: Instrumental variable methods for closed-loop system identification. Automatica **41**, 241–249 (2005)
14. Greblicki, W.: Continuous-time Hammerstein system identification. IEEE Trans. Autom. Control **45**(6), 1232–1236 (2000)
15. Garnier, H., Wang, L. (eds.): Identification of Continuous-Time Models from Sampled Data. Springer, London (2008)
16. Johansson, R.: Identification of continuous-time models. IEEE Trans. Signal Process. **42** (1994)
17. Laurain, V., Gilson, M., Garnier, H.: Refined instrumental variable methods for identifying Hammerstein models operating in closed loop. In: Proceedings of the 48th IEEE Conference on Decision and Control, Shanghai, China (2009)
18. Laurain, V., Gilson, M., Garnier, H., Young, P.C.: Refined instrumental variable methods for identification of Hammerstein continuous time Box-Jenkins models. In: Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico (2008)
19. Liu, Y., Bai, E.W.: Iterative identification of Hammerstein systems. Automatica **43**(2), 346–354 (2007)
20. Ljung, L.: System Identification: Theory for the User, 2nd edn. Prentice-Hall, New York (1999)
21. Palanthandalam-Madapusi, H.J., Edamana, B., Bernstein, D.S., Manchester, W., Ridley, A.J.: NARMAX identification for space weather prediction using polynomial radial basis functions. In: 46th IEEE Conference on Decision and Control, New Orleans, LA, USA (2007)
22. Patra, A., Unbehauen, H.: Identification of a class of nonlinear continuous-time systems using Hartley modulating functions. Int. J. Control **62**(6), 1431–1451 (1995)
23. Pintelon, R., Schoukens, J., Rolain, Y.: Box-Jenkins continuous-time modeling. Automatica **36** (2000)
24. Rao, G.P., Garnier, H.: Identification of continuous-time systems: direct or indirect? Syst. Sci. **30**(3), 25–50 (2004)
25. Rao, G.P., Unbehauen, H.: Identification of continuous-time systems. IEE Proc. Control Theory Appl. **153**(2) (2006)
26. Schoukens, J., Widanage, W.D., Godfrey, K.R., Pintelon, R.: Initial estimates for the dynamics of a Hammerstein system. Automatica **43**(7), 1296–1301 (2007)
27. Vanbeylen, V., Pintelon, R., Schoukens, J.: Blind maximum likelihood identification of Hammerstein systems. Automatica **44**(11), 3139–3146 (2008)
28. Wang, J., Zhang, Q., Ljung, L.: Revisiting Hammerstein system identification through the two-stage algorithm for bilinear parameter estimation. Automatica **45**(9), 2627–2633 (2009)
29. Young, P.C.: Some observations on instrumental variable methods of time-series analysis. Int. J. Control **23**, 593–612 (1976)
30. Young, P.C.: The refined instrumental variable method: unified estimation of discrete and continuous-time transfer function models. J. Eur. Syst. Autom. **42**, 149–179 (2008)
31. Young, P.C., Garnier, H., Gilson, M.: Refined instrumental variable identification of continous-time hybrid Box-Jenkins models. In: Garnier, H., Wang, L. (eds.) Identification of Continuous-Time Models from Sampled Data, pp. 91–131. Springer, London (2008)
32. Young, P.C., Jakeman, A.: Refined instrumental variable methods of recursive time-series analysis—Part III. Extensions. Int. J. Control **31**(4), 741–764 (1980)
33. Zhai, D., Rollins, D.K., Bhandari, N., Wu, H.: Continuous-time Hammerstein and Wiener modeling under second-order static nonlinearity for periodic process signals. Comput. Chem. Eng. **31**, 1–12 (2006)

http://www.springer.com/978-0-85729-973-4

System Identification, Environmental Modelling, and Control System Design
(Eds.)L. Wang; H. Garnier
2012, XXVII, 645 p. 213 illus., 59 in color., Hardcover
ISBN: 978-0-85729-973-4