

A BLIND APPROACH TO THE HAMMERSTEIN-WIENER MODEL IDENTIFICATION

Er-Wei Bai *

** Dept. of Electrical and Computer Engineering University
of Iowa, Iowa City, Iowa 52242 erwei@icaen.uiowa.edu*

Abstract: In this paper, we propose a blind approach to the sampled Hammerstein-Wiener model identification. By using the blind approach, it is shown that all internal variables can be recovered solely based on the output measurements. Then, identification of linear and nonlinear parts can be carried out. No a priori structural knowledge about the input nonlinearity is assumed and no white noise assumption is imposed on the input.

Keywords: system identification, parameter estimation, nonlinear systems, Wiener system, Hammerstein system

1. INTRODUCTION

The Hammerstein and Wiener models are special kinds of nonlinear systems where the nonlinear block is static and follows or is followed by a linear system. These models have applications in many engineering problems and therefore, identification of Hammerstein and Wiener models has been an active research area for many years. There exists a large number of research papers in the literature on the topics of Hammerstein and Wiener model identifications. There exists only scattered work reported in the literature on the Hammerstein-Wiener model identification.

Existing methods for Hammerstein model identification can be roughly divided into four categories (Bai): the iterative method, the over-parameterization method, the correlation method and the separable least squares method. In most cases, the structure of the nonlinearity is assumed to be known. Otherwise, identification becomes a structural estimation problem. Identification of Wiener models is more difficult. The reason is the lack of a good representation of the output nonlinearity for identification purpose. Unlike in

the Hammerstein model case, the commonly used polynomial representation for the output nonlinearities makes identification in terms of unknown parameters very hard. The main technique used for Wiener model identification is the correlation analysis. Suppose the input is a zero mean white noise, identification of linear part and identification of nonlinear part are shown to be separable. Moreover, it was shown that even the structure of the nonlinearity may be estimated with white noise inputs under certain conditions. The difficulty with the correlation method is that it is limited to white inputs. To overcome this difficulty, inverse representation method was proposed to model the inverse of the output nonlinearity allowing non-white inputs.

In this paper, we continue our work on the Hammerstein-Wiener model identification. Unlike previous works where a very special structure is assumed, however, the blind approach in this paper allows a very general structure on the nonlinearities. In particular, the input nonlinearity structure can be arbitrary and is not assumed to be known. By using the blind identification approach, all the unknown internal variables can

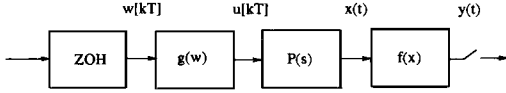


Fig. 1. The sampled Hammerstein-Wiener model. be recovered solely based on the output measurements. Once all interval variables are recovered, linear part and nonlinear parts including the structure can be identified. Our scheme applies to either white or non-white inputs. The blind techniques adopted in this paper use our previous results presented for blind channel equalizations of IIR systems and are also based on blind techniques developed for Hammerstein models. Although the algorithm proposed in this paper is for the Hammerstein-Wiener model, it applies directly to the identification of either Wiener models or Hammerstein models with trivial modifications.

Because of page limit, we only list the technical report version of this paper (Bai). Interested readers can find all the references in (Bai).

2. PROBLEM STATEMENT AND PRELIMINARIES

Consider the sampled Hammerstein-Wiener model shown in Figure 1, which consists of a Zero-Order-Hold, an input nonlinearity, a scalar linear stable continuous time system and an output nonlinearity. It is assumed that

Assumption 1:

- The unknown continuous time system $P(s)$ possesses a standard continuous and discrete state space representations (A, b, c) and (Φ, Γ, c) along with its discrete transfer function

$$G(z) = \frac{\beta(z)}{\alpha(z)} = \quad (2.1)$$

$$\frac{\beta_1 z^{-1} + \beta_2 z^{-2} + \dots + \beta_n z^{-n}}{1 - \alpha_1 z^{-1} - \alpha_2 z^{-2} - \dots - \alpha_n z^{-n}}$$

form some α_i 's and β_j 's at the given sampling interval T .

- The input nonlinearity is static $u = g(w)$ and its structure is not assumed to be known.
- The output nonlinearity is also static $y = f(x)$. However, it is assumed that $f(\cdot)$ is one-to-one so that the inverse $x = f^{-1}(y)$ exists and admits a polynomial representation $x = \sum_{i=1}^m r_i y^i$.

In identification, we will first estimate the inverse coefficients r_i 's and then, find the best forward function

$$y = \sum_{i=1}^q a_i x^i$$

of $x = \sum_{i=1}^m r_i y^i$ in the least squares sense using the observed data. Finally, we represent the output nonlinearity $y = f(x)$ by $y = \sum_{i=1}^q a_i x^i$. We remark that if $y = f(x)$ is one-to-one and continuous, its inverse $x = f^{-1}(y)$ exists and is also continuous. With a bounded input and a stable system, x and y are always bounded and this implies that the inverse $x = f^{-1}(y)$ can be approximated to any accuracy by a polynomial $x = \sum_{i=1}^m r_i y^i$. In particular, as the order m goes to infinite, $x = f^{-1}(y) = \sum_{i=1}^m r_i y^i$. This shows that the inverse representation of (??) is theoretically justified. Practically, of course, a high order m usually means a high sensitivity to noise and model uncertainty in the identification setting. Therefore, there is a limitation to the inverse approach. A discussion on this topic will be provided later.

For a given sampling interval T , the goal of the Hammerstein-Wiener model identification is to estimate the transfer function $G(z)$ in terms of its parameters α_i 's and β_j 's, the output nonlinearity $x = \sum_{i=1}^m r_i y^i$ and its inverse $y = \sum_{i=1}^q a_i x^i$, and the input nonlinearity $u = g(w)$ solely based on the measurements of w and y . No internal variables x and u are assumed available. Moreover, the structure of the input nonlinearity $u = g(w)$ is unknown.

Our idea of identification is the blind approach, i.e., to sample the output at a higher rate. Given the sampling interval T , let the output sampling interval be

$$h = T/l, \quad l \geq 1$$

for some positive integer l , referred to as the over-sampling ratio.

Minimality is important in identification. Without minimality assumption, the transfer function $G(z)$ has pole-zero cancellations and this makes the parameterization non-unique. In other words, the identifiability is lost. To this end, we make the following assumption.

Assumption 2: The sampled system at sampling interval T is assumed to be minimal at the sampling interval T .

Before closing this section, we observe that the parameterization of the Hammerstein-Wiener model is actually not unique. Suppose the system is represented by three blocks $u = g(w)$, $G(z)$ and $y = f(x)$. Then, any triple $ag(w)$, $bG(z)$ and $cf(x)$, for some constants a, b, c , would produce the identical input-output measurements, provided that $a \cdot b \cdot c = 1$. In other words, any identification setting can not distinguish between (g, G, f) and (ag, bG, cf) . To obtain a unique parameterization,

two blocks need to be normalized. Since the structure of the input nonlinearity is not assumed to be known, we normalize the linear block and the output nonlinearity.

Assumption 3: For the Hammerstein-Wiener model, it is assumed that $\beta_1 = 1$ and $r_1 = 1$.

With this normalization assumption and persistent exciting (PE) input, we will show that g , G and f can be uniquely identified. Note that there are other ways to normalize the system, Assumption 3.1 is the simplest one. The purpose is to avoid unnecessary complications so that our ideas can be presented clearly.

3. IDENTIFICATION OF THE HAMMERSTEIN-WIENER MODEL

If the structure of the input nonlinearity $u = g(w)$ is unknown, identification of the Hammerstein-Wiener model is no longer a parameter estimation problem. Obviously, the identification involves structural estimation. It is clear, however, that if u were available, we would be able to estimate the input nonlinearity structure. At least, the complete picture of $u = g(w)$ can be graphed by using the pairs of (w, u) and this graphical picture provides us accuracy information on the unknown input nonlinearity $u = g(w)$. Therefore, the key is to estimate $G(z)$, $x = \sum r_i y^i$, and then to recover u solely based on the output measurements. We accomplish this goal in several steps, estimating the output nonlinearity, finding the linear part and then recovering u .

3.1 Output nonlinearity estimation

Given the input sampling interval T , let the output sampling interval be $h = T/(n+1)$ where n is the order of $G(z)$. We remark that $h = T/(n+1)$ is not necessary but does make analysis and notation simple. In fact, $h = T/(\bar{n} + 1)$ for any $\bar{n} \geq n$ will work, see remarks in Discussion section for details. Now, consider the sampled system at the sampling interval $h = T/(n+1)$. It is clear that the transfer function of the sampled system at the sampling interval h is also an n th order strictly proper rational function

$$\tilde{G}(z) = \frac{\tilde{\beta}(z)}{\tilde{\alpha}(z)} = \quad (3.1)$$

$$\frac{\tilde{\beta}_1 z^{-1} + \tilde{\beta}_2 z^{-2} + \dots + \tilde{\beta}_n z^{-n}}{1 - \tilde{\alpha}_1 z^{-1} - \tilde{\alpha}_2 z^{-2} - \dots - \tilde{\alpha}_n z^{-n}}$$

for some unknown $\tilde{\alpha}_i$'s and $\tilde{\beta}_j$'s. Its time domain equation is accordingly given by

$$x[kh] = \sum_{i=1}^n \tilde{\alpha}_i x[kh - ih] + \sum_{i=1}^n \tilde{\beta}_i u[kh - ih].$$

Substituting the value $x[kh] = \sum_{j=1}^m r_j y^j[kh]$ into equation, it follows that

$$\sum_{j=1}^m r_j y^j[kh] = \sum_{i=1}^n \tilde{\alpha}_i \sum_{j=1}^m r_j y^j[kh - ih] + \quad (3.2)$$

$$\sum_{i=1}^n \tilde{\beta}_i u[kh - ih] + v_1[kh], \quad i = 1, \dots, N$$

where $v_1[kh]$ denotes any discrepancy not counted by the model, e.g., the contribution of noise, model uncertainty and approximation errors.

By observing that $r_1 = 1$ from the normalization assumption, the above equation can be re-written as

$$y[kh] = \phi_1'[kh]\theta_1 + \sum_{i=1}^n \tilde{\beta}_i u[kh - ih] + v_1[kh].$$

This is the basic equation for the estimation of the output nonlinearity $x = \sum r_i y^i$ in terms of its coefficients r_i 's. Now, consider two consecutive equations at $k = l(n+1)$ and $k = l(n+1) - 1$,

$$y[l(n+1)h] = \phi_1'[l(n+1)h]\theta_1 + \sum_{i=1}^n \tilde{\beta}_i u[l(n+1)h - ih] + v_1[l(n+1)h],$$

$$y[l(n+1)h - h] = \phi_1'[l(n+1)h - h]\theta_1 + \quad (3.3)$$

$$\sum_{i=1}^n \tilde{\beta}_i u[l(n+1)h - ih - h] + v_1[l(n+1)h - h].$$

Since the input sampling interval is fixed at $T = (n+1)h$ where h is the output sampling interval, we have

$$w[(l-1)T] = w[(l-1)T+h] = \dots = w[(l-1)T+nh],$$

and this implies

$$u[(l-1)T] = u[(l-1)T+h] = \dots = u[(l-1)T+nh],$$

$$\Delta y[l] = \Delta \phi_1'[l]\theta_1 + \Delta v[l] \quad (3.4)$$

with

$$\Delta y[l] = y[lT] - y[lT-h], \Delta v[l] = v_1[lT] - v_1[lT-h],$$

$$\Delta \phi_1[l] = \phi_1[lT] - \phi_1[lT-h].$$

In equation (3.4), $\Delta y[l]$ and $\Delta \phi_1[l]$ consist of output measurements $y[kh]$ only and thus are available. Moreover, this equation is linear in the unknown parameter vector θ_1 which can be estimated by many standard methods, e.g., the least squares method or the (normalized) LMS algorithm.

Note that the estimates $\hat{r}_2, \dots, \hat{r}_m, \hat{\alpha}_1, \dots, \hat{\alpha}_n$ of $r_2, \dots, r_m, \alpha_1, \dots, \alpha_n$ are the first $(m-1+n)$ entries of $\hat{\theta}_1$. Also note that r_1 is normalized to be

1. Therefore, once $\hat{\theta}_1$ is obtained, we have the estimate of the inverse output nonlinearity

$$\hat{f}^{-1}(y) = \sum_{i=1}^m \hat{r}_i y^i[kh] \quad (3.5)$$

with $\hat{r}_1 = 1$, as well as the estimate of the denominator $\hat{\alpha}(z)$ of $\hat{G}(z)$.

The forward output nonlinearity $\sum_{i=1}^q \hat{a}_i x^i[kh]$ can be constructed by minimizing

$$\hat{a} = \arg \min \sum_k \{y[kh] - \sum_{i=1}^q \hat{a}_i \hat{x}^i[kh]\}^2$$

where $y[kh]$'s are observed outputs and $\hat{x}[kh] = \sum_{i=1}^m \hat{r}_i y^i[kh]$'s are generated from (3.5). We comment that direct readings of \hat{r}_i and $\hat{\alpha}_k$ from $\hat{\theta}_1$ may not be a good policy in a noisy situation because it ignores a large number of identified parameters $\hat{\alpha}_k r_i$ without taking into account of their contribution. A more robust way should consider their contribution.

3.2 Linear transfer function estimation

In this section, we propose a blind method to estimate $G(z)$ without requiring $u[kT]$. Recall that the denominator $1 - \hat{\alpha}_1 z^{-1} - \dots - \hat{\alpha}_n z^{-n}$ of $\hat{G}(z)$ was obtained as a result of the estimation of $\hat{\theta}_1$ in the previous section. Note $\hat{G}(z)$ is the transfer function of the sampled system at the sampling interval $h = T/(n+1)$. Write

$$1 - \hat{\alpha}_1 z^{-1} - \dots - \hat{\alpha}_n z^{-n} = (1 - \tilde{s}_1 z^{-1}) \dots (1 - \tilde{s}_n z^{-1})$$

where \tilde{s}_i denotes the poles of $\hat{G}(z)$. The sampled system is assumed to be minimal at the sampling interval T and is minimal at any sampling interval $h = T/l$, $l \geq 1$. Clearly, s is a pole of the continuous time system if and only if $e^{sT} = e^{sh(n+1)}$ is a pole of $G(z)$ if and only if e^{sh} is a pole of $\hat{G}(z)$. In other words, if \tilde{s}_i 's are the poles of $\hat{G}(z)$, then \tilde{s}_i^{n+1} 's are the poles of $G(z)$. This implies that an estimate of $\alpha(z)$, the denominator of $G(z)$ is given by

$$\hat{\alpha}(z) = (1 - \tilde{s}_1^{n+1} z^{-1}) \dots (1 - \tilde{s}_n^{n+1} z^{-1}) \quad (3.6)$$

$$= 1 - \hat{\alpha}_1 z^{-1} - \dots - \hat{\alpha}_n z^{-n}.$$

Hence, an estimate of $\alpha(z)$ is already contained in $\hat{\theta}_1$ and what we have to estimate is only the numerator $\beta(z)$ of $G(z)$. To this end, consider two sequences

$$\{x[kT]\} \iff X(z) = \sum_{k=0}^{\infty} x[kT] z^{-k} \quad (3.7)$$

$$= G(z)U(z)$$

$$\begin{aligned} \{x[kT + T/2]\} &\iff \bar{X}(z) = \sum_{k=0}^{\infty} x[kT + T/2] z^{-k} \\ &= \bar{G}(z)U(z) \end{aligned}$$

where

$$U(z) = \sum_{k=0}^{\infty} u[kT] z^{-k}$$

is the Z-transform of the sequence $u[kT]$ at the sampling interval T , and $G(z)$ and $\bar{G}(z)$ represent the transfer functions from $U(z)$ to $X(z)$ and $\bar{X}(z)$ respectively. The transfer function $G(z)$ is derived in (2.1) and is strictly proper. $\bar{G}(z)$ needs a special attention. From the continuous time state space equation, we have

$$\bar{G}(z) = \frac{\bar{\beta}_0 + \bar{\beta}_1 z^{-1} + \dots + \bar{\beta}_n z^{-n}}{1 - \bar{\alpha}_1 z^{-1} - \bar{\alpha}_2 z^{-2} - \dots - \bar{\alpha}_n z^{-n}}.$$

It is interesting to note that $G(z)$ and $\bar{G}(z)$ share the same denominator but with different numerators and unlike $G(z)$, $\bar{G}(z)$ is proper but not strictly proper.

Now, consider again two sequences $\{x[kT]\}$, $\{x[kT + T/2]\}$ and their Z-transforms

$$X(z) = G(z)U(z), \quad \bar{X}(z) = \bar{G}(z)U(z).$$

Clearly,

$$\bar{G}(z)X(z) - G(z)\bar{X}(z) = 0, \quad \bar{\beta}(z)X(z) - \beta(z)\bar{X}(z) = 0$$

and this results in

$$\begin{aligned} (\bar{\beta}_0 + \bar{\beta}_1 z^{-1} + \dots + \bar{\beta}_n z^{-n})X(z) - \\ (\beta_1 z^{-1} + \dots + \beta_n z^{-n})\bar{X}(z) = 0. \end{aligned}$$

Its time domain equation is

$$\beta_1 x[kT + T/2] = \bar{\phi}'_2[k]\theta_2.$$

By noting that $\beta_1 = 1$, it follows that

$$x[kT + T/2] = \bar{\phi}'_2[k]\theta_2.$$

In this equation, $\bar{\phi}_2$ is a function of $x[kT]$ and $x[kT + T/2]$ which are not available. However, their estimates $\hat{x}[kT]$ and $\hat{x}[kT + T/2]$ are readily available by using the estimated output nonlinearity $\hat{x} = \sum_{i=1}^m \hat{r}_i y^i$ and the observed outputs $y[kT]$ and $y[kT + T/2]$. Let \hat{x} and ϕ_2 denote the estimates of x and $\bar{\phi}_2$ using \hat{x} instead of x respectively, we have

$$\hat{x}[kT + T/2] = \phi'_2[k]\theta_2 + v_2[k] \quad (3.8)$$

where $v_2[k]$ indicates the contribution due to the error between x and \hat{x} . This equation is again linear in the unknown parameter vector θ_2 and many standard estimation algorithms apply. $\hat{\theta}_2$ consists of the estimates $(\hat{\beta}_1, \dots, \hat{\beta}_n, \hat{\beta}_0, \dots, \hat{\beta}_n)$ with $\hat{\beta}_1 = 1$. Therefore, combining equation (3.6), we obtain the estimates $\hat{G}(z)$ and $\hat{\bar{G}}(z)$ of $G(z)$ and $\bar{G}(z)$ respectively.

3.3 Input nonlinearity estimation

Since the structure of the input nonlinearity is unknown, estimation of the input nonlinearity relies completely on the graph information determined by the pairs of $(w[kT], u[kT])$. The input $w[kT]$ is known, but not $u[kT]$. Therefore, estimation of $u[kT]$ becomes a key in determining the input nonlinearity.

Recall that the input sampling interval is T and thus $u[kT] = u[kT + T/2]$. Also recall

$$X(z) = G(z)U(z), \quad \bar{X}(z) = \bar{G}(z)U(z).$$

If either $G(z)$ or $\bar{G}(z)$ is minimum phase, $U(z)$ and consequently, $u[kT]$ can be recovered easily

$$U(z) = G^{-1}(z)X(z) \text{ or } U(z) = \bar{G}^{-1}(z)\bar{X}(z).$$

In time domain, these equations are

$$u[kT] = -\beta_2 u[kT - T] - \dots \quad (3.9)$$

$$\begin{aligned} & -\beta_n u[kT - (n-1)T] + x[kT + T] \\ & -\alpha_1 x[kT] - \dots - \alpha_n x[kT - (n-1)T] \end{aligned}$$

or

$$u[kT] = \frac{1}{\beta_0} (-\bar{\beta}_1 u[kT - T] - \dots \quad (3.10)$$

$$\begin{aligned} & -\bar{\beta}_n u[kT - (n-1)T] \\ & + x[kT + T/2] - \bar{\alpha}_1 x[kT + T/2 - T] \\ & - \dots - \bar{\alpha}_n x[kT + T/2 - nT]. \end{aligned}$$

In these equations, both the estimates $\hat{\beta}(z)$ and $\hat{\bar{\beta}}(z)$ have been obtained and so are the estimates of $\hat{x}[kT]$ and $\hat{x}[kT + T/2]$. Thus, $\hat{u}[kT]$ can be calculated.

In the case that both $G(z)$ and $\bar{G}(z)$ are non-minimum phase, recovery of $u[kT]$ becomes problematic. To overcome this difficulty, assume that $G(z)$ and $\bar{G}(z)$ do not share any common zeros. Then, from the Bezout identity, there exist two stable transfer functions $F(z)$ and $\bar{F}(z)$ such that

$$F(z)\bar{G}(z) + \bar{F}(z)G(z) = 1. \quad (3.11)$$

This implies that

$$F(z)\bar{X}(z) + \bar{F}(z)X(z) = \quad (3.12)$$

$$[F(z)\bar{G}(z) + \bar{F}(z)G(z)]U(z) = U(z)$$

and recovery of $U(z)$ or equivalently $u[kT]$ can be easily implemented by using $\hat{x}[kT]$ and $\hat{x}[kT + T/2]$. Once the estimate $\hat{u}[kT]$ of $u[kT]$ is obtained, the input nonlinearity $u[kT] = g(w[kT])$ can be graphed using the pairs $(w[kT], \hat{u}[kT])$ that provides complete information about the unknown input nonlinearity.

3.4 Algorithm and simulations

We are now in a position to summarize the identification algorithm for the Hammerstein-Wiener model with unknown structure of the input nonlinearity.

Identification algorithm:

Step 1: Consider the sampled Hammerstein-Wiener model in Figure 1. For a given sampling interval T , collect input and output measurements $w[kT]$, $y[kh]$, $y[kT]$ and $y[kT + T/2]$, where $h = T/(n+1)$.

Step 2: Construct $\Delta y[l]$ and $\Delta \phi_1[l]$ as in (3.4) and estimate θ_1 using, e.g., the LMS algorithm. From θ_1 , determine the inverse output nonlinearity $x = \sum_{i=1}^m \hat{r}_i y^i$. The estimate $y = \sum_{i=1}^q \hat{a}_i x^i$ of the forward output nonlinearity $y = f(x)$ is the best inverse of $x = \sum_{i=1}^m \hat{r}_i y^i$. Denote by $\hat{x}[kT] = \sum_{i=1}^m \hat{r}_i y^i[kT]$ and $\hat{x}[kT + T/2] = \sum_{i=1}^m \hat{r}_i y^i[kT + T/2]$, the estimates of $x[kT]$ and $x[kT + T/2]$ respectively.

Step 3: From $\hat{\theta}_1$, calculate the estimate $\hat{\alpha}(z)$ of $\alpha(z)$ using (3.6). Construct $\phi_2[k]$ as in (3.8) and the estimate $\hat{\theta}_2$. From $\hat{\theta}_2$, determine the estimates $\hat{G}(z)$, $\hat{\bar{G}}(z)$ of $G(z)$, $\bar{G}(z)$ respectively.

Step 4: If either $\hat{G}(z)$ or $\hat{\bar{G}}(z)$ is minimum phase, calculate $\hat{u}[kT]$ using either (3.9) or (3.10). If both $\hat{G}(z)$ and $\hat{\bar{G}}(z)$ are non-minimum phase, calculate $F(z)$ and $\bar{F}(z)$ as in (3.11) and calculate the estimate $\hat{u}[kT]$ using (3.12).

Step 5: Graph the input nonlinearity $u = g(w)$ by using the pairs $(w[kT], \hat{u}[kT])$. Estimate the nonlinearity based the information provided by the graph. If necessary, parameterize the input nonlinearity using some base functions.

3.5 Discussions

To avoid unnecessary complications so that the idea can be clearly conveyed, our attention was focused on presenting the basic algorithm. The algorithm can be in fact improved in several ways. Because of page limit, we only provide brief discussions here. Interested readers may find detail discussions in the full length paper.

(1) Parameterization of the input nonlinearity $u = g(w)$.

In the previous discussion, the structure of the input nonlinearity is assumed to be unknown and thus estimation relies on the graph given by the pairs $(w[kT], \hat{u}[kT])$. Once the picture of $u = g(w)$ is obtained, the structure of $u = g(w)$ can be determined. The next step is to parameterize this nonlinearity by using appropriate base functions, e.g.,

$$u = g(w) = \sum g_i(w, b_i)$$

for some known nonlinear functions g_i 's and unknown coefficients b_i 's. The choice of g_i 's of course depends on the structure shown in the graph.

(2) **Output nonlinearity order estimation.**

In the proposed algorithm, the order m of the inverse output nonlinearity $x = \sum_{i=1}^m r_i y^i$ is assumed to be known. In practice, m is unknown and needs to be estimated using the on-line data. A number of standard methods of order estimation for linear systems, e.g, rank test and the output error test, find their applications here. Interested readers can find details in (Bai).

(3) **Inverse parameterization of the output nonlinearity.**

The actual output nonlinearity $y = f(x)$ is unknown and we use the inverse parameterization $x = \sum r_i y^i$. This inverse approach has been used in the literature to model Wiener systems. Here, we adopt this approach to model Hammerstein-Wiener systems with unknown input nonlinearity structure. Because of the blind method, this inverse approach makes identification of the Hammerstein-Wiener model feasible even with the unknown structure of the input nonlinearity and non-white inputs. Clearly, the success of the proposed algorithm hinges on the accuracy between the approximation $x = \sum r_i y^i$ and the true $x = f^{-1}(y)$. Theoretically, as long as $y = f(x)$ is one-to-one and continuous, $x = \sum r_i y^i$ approximates $x = f^{-1}(y)$ to any accuracy as the order increases. Practically, however, a high order m introduces errors due to noise and model uncertainty and slows down the convergence rate. There is balance between the errors introduced by the approximation $x = \sum r_i y^i$ and the errors due to noise and model uncertainty.

(4) **Persistent excitation conditions.**

To have a robust identification algorithm in the presence of noise and model uncertainty, the regressors $\Delta\phi_1$ and ϕ_2 need to be persistently exciting (PE). The conditions that ϕ_2 is PE are derived (Bai), which basically say that ϕ_2 is PE if the spectral lines of $u[kT]$ is not concentrated on less than $2n$ points, a richness condition. This richness condition can also be translated into conditions in terms of the input $w[kT]$. Suppose that $w[kT]$ is i.i.d. zero mean random variable and $u[kT] = g(w[kT])$ assumes at least two distinctive values with nonzero probability. Then, $u[kT]$ is also i.i.d. and has infinitely many spectral lines which implies that ϕ_2 is PE. The second scenario is the polynomial

input nonlinearity $u = \sum b_i w^i$ and sinusoidal input $w[kT] = \sum c_i \sin(\Omega_i k)$. This is certainly the case if the input is periodic by the Fourier series representation. Then, $u[kT]$ has more than $2n$ spectral lines if $w[kT]$ has $2n$ spectral lines unless in a pathological case where either coefficients are zero or frequencies are the same module 2π .

(5) **Choice of the over-sampling ratio l .**

In the algorithm, the over-sampling ratio $l = (n + 1)$ is assumed, where n is the order of $G(z)$. This seems to imply that the order of $G(z)$ has to be known a priori. In fact, l does not have to be $(n + 1)$ and any $l > (n + 1)$ suffices (Bai).

(6) **Relation with the step response identification.**

In a way, the blind technique presented in this paper may be considered as repeatedly applying piece-wise constant inputs. Conceptually, a number of step responses could be used to give information first on the output nonlinearity and the linear part, and then on the input nonlinearity. However, the blind technique works fundamentally different from the traditional step response identification method (Bai). The traditional step response method relies heavily on the steady-state value $y(t)$, $t \rightarrow \infty$, of the step response and would suffer from large noises at the end of transient. This is specially true in the setting of parametric identification and therefore, it is suggested to apply the step response identification method several times to average out the effect of noises.

(7) **Identifiability**

With the PE inputs and assumptions 1, 2 and 3, the identifiability of the Hammerstein-Wiener model shown in Figure 1 can be easily established. Identifiability here means that the representation of the system is unique in the absence of noise and model uncertainties. This can be seen easily. With the PE regressors, the solutions of (3.4) and (3.8) are unique. Moreover, the true but unknown system parameters are solutions in the absence of noises and model uncertainties. This establishes the identifiability.

REFERENCES

- Bai, E.W. (2001), "A blind approach to the Hammerstein-Wiener model identification", Technical report, Dept. of ECE, University of Iowa, 2001