15th Triennial World Congress, Barcelona, Spain

IDENTIFICATION OF LINEAR SYSTEMS WITH HARD INPUT NONLINEARITIES OF KNOWN STRUCTURE

Er-Wei Bai *

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

Abstract: This paper studies identification of systems with input nonlinearities of known structure. For input nonlinearities parameterized by one parameter, a deterministic approach is proposed based on the idea of separable least squares. The identification problem is shown to be equivalent to a one-dimensional minimization problem. The method is very effective for several common static and non-static input nonlinearities. For a general input nonlinearity, a correlation analysis based identification algorithm is presented which is shown to be convergent.

Keywords: system identification, parameterestimation, nonlinear systems

1. PROBLEM STATEMENT

Hard input nonlinearities are common in engineering practice. These nonlinearities severely limit the performance of control systems. Therefore, robust controls are often used (Gao et al) to cancel or reduce the effect of these harmful nonlinearities. Those control designs require values of the parameters that represent the hard nonlinearities. Clearly, system identification constitutes a crucial part in such control designs if the parameters are unknown. The difficulty of identification for the system with a hard input nonlinearity is that the unknown parameters of the nonlinearity and the linear system are coupled. Moreover, the output of the hard nonlinear block may not be written as an analytic function of the input. Surprisingly, there is only scattered work reported in the literature on identification of systems with hard nonlinearities (Gu & Voros), although robust control designs involving these hard nonlinearities become a very active research area in recent years.

This paper studies identification of a stable SISO discrete time linear system with a hard input nonlinearity as shown in Figure 1, where $y(k)$, $u(k)$ and $v(k)$ are system output, input and noise respectively. Note that the internal signal $x(k)$ is not measurable.

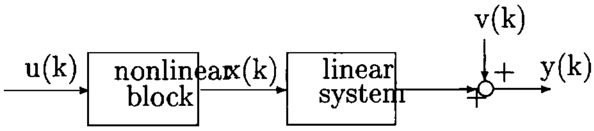

Figure 1: The nonlinear system

The linear system is assumed to be stable and is represented by the transfer function

$$H(z) = \frac{\beta_1 z^{-(n-1)} + \beta_2 z^{-(n-2)} + \ldots + \beta_n}{z^n - \alpha_1 z^{-(n-1)} - \ldots - \alpha_n}, \quad (1.1)$$

parameterized by the parameter vector

$$\theta^T = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n). \quad (1.2)$$

The nonlinear block represents a static or non-static nonlinearity in the form of

$$x(k) = \mathcal{N}(u(k), \ldots, u(0), a) \quad (1.3)$$

for some nonlinear functions $\mathcal{N}$ parameterized by the parameter vector $a \in R^l$. Common examples of input nonlinearities are the Saturation, Preload, Relay, Dead-zone, Hysteresis-relay and Hysteresis nonlinearities shown in Figure 2,

$$
\begin{aligned}
x_{sa}(k) &= \frac{1+sgn(a-|u(k)|)}{2}u(k)+ \\
&\quad \frac{1+sgn(|u(k)|-a)}{2}a\cdot sgn(u(k)) \\
x_{pr}(k) &= u(k)+a\cdot sgn(u(k)) \\
x_{de}(k) &= u(k)-a\cdot sgn(u(k))-\frac{[1+sgn(a-|u(k)|)]}{2}(u(k) \\
&\quad -a\cdot sgn(u(k))) \\
x_{re}(k) &= [sgn(u(k)-a)+sgn(u(k)+a)]/2 \\
x_{hr}(k) &= \begin{cases} 1 & (u(k)>a) \text{ or } (|u(k)|\le a \\ & \text{and } u(k)-u(k-1)<0) \\ & \text{or } (|u(k)|\le a \\ & \text{and } u(k)=u(k-1) \text{ and } x(k-1)=1) \\ -1 & (u(k)<-a) \text{ or } (|u(k)|\le a \\ & \text{and } u(k)-u(k-1)>0) \\ & \text{or } (|u(k)|\le a \\ & \text{and } u(k)=u(k-1) \text{ and } x(k-1)=-1) \end{cases} \\
x_{hy}(k) &= \begin{cases} u(k)-a & u(k)-u(k-1)>0 \\ u(k)+a & u(k)-u(k-1)<0 \\ x(k-1) & u(k)=u(k-1) \end{cases}
\end{aligned}
\tag{1.4}
$$

where sgn is the standard sign function.

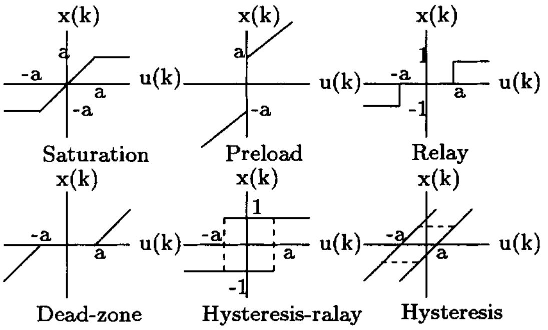

Figure 2: Examples of input nonlinearities.

Note that the gain of all nonlinearities in Figure 2 is assumed to be 1. This is to avoid the non-unique parameterization problem due to the product of the nonlinear block and the linear system. If the gain is not 1, say α, it can be absorbed by the linear system $\frac{\alpha\beta_1 z^{-(n-1)}+\alpha\beta_2 z^{-(n-2)}+\ldots+\alpha\beta_n}{z^n-\alpha_1 z^{-(n-1)}-\ldots-\alpha_n}$.

Our identification approach is based on the Hammerstein model (Bai, Billings & Stoica). There exists a large number of works in the literature on Hammerstein model identification. Most results require that the nonlinearity is static and analytic, usually a polynomial (Bai, Billings & Stoica) which is linear in the unknown parameter a. This is, however, not the case for hard nonlinearities. The hard nonlinearities may not be approximated by polynomials in stability analysis. Moreover, expressions of these hard nonlinearities are not linear in the unknown a. Determination of segments itself depends on the unknown a. To overcome these difficulties, some algorithms were proposed (Gu & Voros). For instance, an identification algorithm for a two-segment piecewise-linear nonlinearity was proposed in (Voros). This algorithm is based on alternative estimation of the parameters and some argument variables. Though simulations illustrate some good results, as pointed out in the paper, the convergence of the estimates is not analyzed and can be divergent in some applications (Stoica). Moreover, approaches of (Gu & Voros) do not apply to the non-static nonlinearity either. Two identification algorithms are proposed in the paper. For nonlinearities parameterized by a single unknown constant a as in Figure 2, a separable least squares approach is proposed. It is shown that the identification problem is equivalent to a one-dimensional minimization problem. This method makes full use of the low dimensionality of the nonlinearity and is found to be very effective. For a general input nonlinearity, a correlation analysis approach is presented. The novelty of this approach lies in the repeated applications of inputs.

2. DETERMINISTIC APPROACH

In this section, input nonlinearities parameterized by a lower dimensional parameter vector are considered. In particular, detailed analysis is given for input nonlinearities parameterized by an one-dimensional parameter a. Such nonlinearities are common in practice and examples are shown in Figure 2. The purpose is to develop an efficient method making full use of their low dimensionalities.

2.1 Identification algorithm

Re-write the equation as

$$y(k) = (y(k-1),...,y(k-n),x(k-1),... \tag{2.1}$$

$$,x(k-n))\theta + v(k),$$

with unknown parameters a and θ. Let

$$\hat{x}(k) = \mathcal{N}(u(k),...,u(0),\hat{a}) \tag{2.2}$$

denote the estimate of $x(k)$ using $\hat{a}$. Define

$$e_{\hat{\theta},\hat{a}}(k) = y(k) - (y(k-1),...,y(k-n), \tag{2.3}$$

$$\hat{x}(k-1),...\hat{x}(k-n))\hat{\theta},$$

$k = 1,2,...,N$, and the prediction error (Ljung)

$$J = \frac{1}{N}\sum_{k=1}^{N} e^2_{\hat{\theta},\hat{a}}(k). \tag{2.4}$$

The estimates $\hat{a}$ and $\hat{\theta}$ are the ones that minimize J. With

$$Y = (y(1),y(2),...,y(N))' \tag{2.5}$$

$$A(\hat{a}) = \begin{pmatrix} y(0) & \ldots & y(1-n) & \hat{x}(0) & \ldots & \hat{x}(1-n) \\ y(1) & \ldots & y(2-n) & \hat{x}(1) & \ldots & \hat{x}(2-n) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y(N-1) & \ldots & y(N-n) & \hat{x}(N-1) & \ldots & \hat{x}(N-n) \end{pmatrix},$$

the objective function J can be rewritten as

$$J = \frac{1}{N}\|Y - A(\hat{a})\hat{\theta}\|^2. \quad (2.6)$$

For a given data set $\{y(k), u(k)\}$, this minimization involves two variables $\hat{a}$ and $\hat{\theta}$. J may be a non-smooth function of $\hat{a}$, but is smooth in $\hat{\theta}$. Moreover,

$$0 = \frac{1}{2}\frac{\partial J}{\partial \hat{\theta}} = -A^T(\hat{a})Y + A^T(\hat{a})A(\hat{a})\hat{\theta}. \quad (2.7)$$

Clearly, if $A^T(\hat{a})A(\hat{a})$ is invertible, the necessary and sufficient condition for $\hat{\theta}$ to be optimal is that

$$\hat{\theta} = [A^T(\hat{a})A(\hat{a})]^{-1}A^T(\hat{a})Y \quad (2.8)$$

provided that $\hat{a}$ is optimal. Therefore, by substituting $\hat{\theta}$ in terms of $\hat{a}$ back into J, it follows that

$$J(\hat{a}) = \frac{1}{N}\|(I - A[A^TA]^{-1}A^T)Y\|^2. \quad (2.9)$$

By substitution, for all six nonlinearities shown in Figure 2, $\hat{\theta}$ is eliminated and the dimension of the search space is reduced from $(1 + 2n)$ to 1. This kind of elimination of variables in optimization literature is referred to as the separable nonlinear least squares problems (Rube).

Now, the original identification problem has been transformed into an one dimensional minimization problem (2.9) for all six nonlinearities in Figure 2. Once the optimal $\hat{a}$ is obtained, the optimal $\hat{\theta}$ follows from (2.8). It is important to remark that the minimization of (2.9) is always 1-dimensional for all six nonlinearities shown in Figure 2 independent of the linear part which could be parameterized by a high dimensional vector $\theta \in R^{2n}$.

The deterministic identification algorithm for systems with hard input nonlinearities parameterized by a single parameter a is now summarized.

Separable least squares identification algorithm for systems with hard input nonlinearities shown in Figure 2:

Step 1: Consider the system (2.1), collect the data set $\{u(k), y(k)\}$ and define Y and $A(\hat{a})$.

Step 2: Solve (2.9) for the optimal $\hat{a}$.

Step 3: Calculate the optimal $\hat{\theta}$ as in (2.8).

To illustrate the effectiveness of the proposed approach, the algorithm is tested with all six nonlinearities shown in Figure 2 with the following example.

Example 1: Let the linear system be

$$y(k) = \alpha_1 y(k-1) + \alpha_2 y(k-2) + \beta_1 x(k-1) + \beta_2 x(k-2) + v(k)$$

where $\theta^T = [\alpha_1, \alpha_2, \beta_1, \beta_2] = [-0.8333, -0.1667, 1, 1]$ which is unknown and $v(k)$ is an i.i.d. random sequence in $[-0.2, 0.2]$. For simulation, $N = 100$ and input is uniformly distributed in $[-4, 4]$. Now, consider the above linear system with the Preload nonlinearity of $a = 1$, Dead-zone nonlinearity of $a = 1$, Saturation nonlinearity of $a = 1$, Relay nonlinearity of $a = 1$, Hysteresis-relay nonlinearity of $a = 1$ and Hysteresis nonlinearity of $a = 1$, separately. The true values of a and θ, and the estimates $\hat{a}$ and $\hat{\theta}$ are, respectively, shown in Table 1.

	$a = 1$	$\theta^T = (-.83, -.17, 1, 1)$
Preload	$\hat{a} = 1$	$\hat{\theta}^T = (-.83, -.17, 1.0, 1.0)$
Dead-Zone	$\hat{a} = 1$	$\hat{\theta}^T = (-.83, -.17, 1.0, 1.0)$
Saturation	$\hat{a} = 1$	$\hat{\theta}^T = (-.82, -.16, 1.01, .98)$
Relay	$\hat{a} = 1$	$\hat{\theta}^T = (-.83, -.17, 1.01, 1.01)$
Hy-relay	$\hat{a} = 1.02$	$\hat{\theta}^T = (-.83, -.16, 1.0, .99)$
Hysteresis	$\hat{a} = 1$	$\hat{\theta}^T = (-.84, -.16, 1.0, 1.0)$

Table 1: True values and the estimates.

Note that only 100 data points are used to accurately estimate the unknown θ and a. This is because the dimension of the problem is reduced to one.

2.2 Consistency analysis and computational issues

Note that the estimates are derived from the minimization of $J(\hat{a})$. There are two questions that need to be answered: (1) how to find the global minimum of the nonlinear optimization problem $J(\hat{a})$, and (2) how the estimates perform in the noisy situation. At each N, the estimates $(\hat{\theta}, \hat{a})$ are derived from the prediction error. Therefore, with i.i.d. zero mean Gaussian noise, these estimates are actually the Maximum Likelihood estimates (Ljung) and are strong consistent. How to find the global minimum of $J(\hat{a})$ in general is a hard question that depends on the input nonlinearities. Recall, however, that $J(\hat{a})$ is one dimensional for all six input nonlinearities discussed in this paper. Thus, after collections of the data set $\{y(k), u(k)\}$, $\hat{x}(k)$ and consequently $A(\hat{a})$ can be constructed using $\hat{a}$, and therefore, the complete picture of $J(\hat{a})$ with respect $\hat{a}$ can be plotted. This graphical picture provides us accurate information on where the global minimum is. Then, local search algorithms, for instance, simplex method can be applied in that region to find the global minimum. In fact, the global minimum can also be obtained directly from the plot of $J(\hat{a})$ versus $\hat{a}$. Using the data generated in Example 1, the plots of $J(\hat{a})$ versus $\hat{a}$ for all six nonlinearities in Figure 2 are shown in Figure 3, where in each subplot the vertical axis is $J(\hat{a})$ and the horizontal axis is $\hat{a}$. In all six figures, the neighborhoods where the global minimum lies can be easily seen.

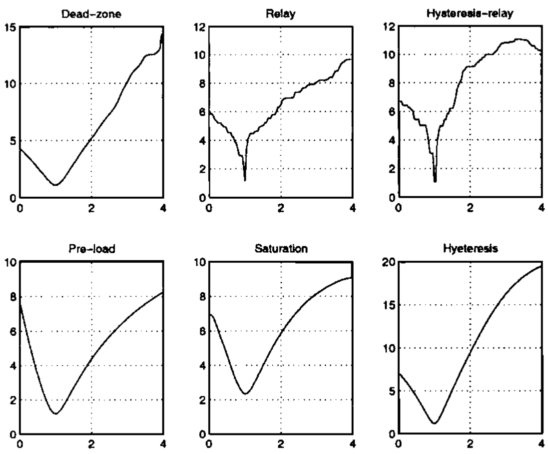

Fig. 3. $J(\hat{a})$ versus $\hat{a}$.

Remark 2.1. To compare with the existing results of deterministic identification algorithms for the hard input nonlinearities, for instance in (Voros), our method is very efficient for nonlinearities parameterize by one parameter. First, the global minimum can always be obtained at least numerically based the plot of $J(\hat{a})$ versus $\hat{a}$. Secondly, the estimates have strong consistency results and are well behaved in noisy situations. There is no consistency analysis for the estimates of (Voros) and it is not clear how they perform in a noisy situation. Moreover, the method of (Voros) is alternative estimate and can be divergent (Stoica) though rarely. Finally, the proposed method applies to non-static nonlinearities. The disadvantage of the proposed method is that it does not extend to the case that the nonlinearity is parameterized by a high dimensional vector due to nonlinear minimization of $J(\hat{a})$. We remark that this problem is not created by our formulation but is inherent in nonlinear system identification. Our approach makes full use of the low dimensionality of those nonlinearities parameterized by a one-dimensional parameter so that the global minimum is obtainable.

3. CORRELATION ANALYSIS METHOD

As discussed before, the separable least squares method can be easily extended to the case where the nonlinearity is parameterized by a two-dimensional vector. However, it seems hard to extend the method to the case where the nonlinearity is parameterized by a higher dimensional a. In this section, a general input nonlinearity parameterized by some $a \in R^l$ is considered and an algorithm for identification of systems with static hard input nonlinearities is proposed by using correlation analysis. Throughout this section, it is assumed that the input $u(k)$ is at our disposal and is a zero mean i.i.d. random sequence with finite variance. The noise $v(k)$ is assumed to be independent of the input.

Recall that all signals are ergodic and the input nonlinearity is assumed to be static, it follows that

$$\mathbf{E}v(k)u(k-j) = 0, \qquad (3.1)$$

$$\mathbf{E}x(k-i)u(k-j) = \mathbf{E}x(k)u(k+i-j) = q\delta(i-j)$$

where $q = \mathbf{E}x(k)u(k)$. Based on the system model, the following equations hold for $m \geq 2n$,

$$\mathbf{E}y(k)u(k-1) = \beta_1 \mathbf{E}x(k-1)u(k-1) \quad (3.2)$$

$$\mathbf{E}y(k)u(k-2) = \alpha_1 \mathbf{E}y(k-1)u(k-2) + \beta_2 \mathbf{E}x(k-2)u(k-2)$$

$$\mathbf{E}y(k)u(k-n) = \alpha_1 \mathbf{E}y(k-1)u(k-n) + \ldots + \alpha_{n-1} \mathbf{E}y(k-n+1) \cdot u(k-n) + \beta_n \mathbf{E}x(k-n)u(k-n)$$

$$\mathbf{E}y(k)u(k-n-1) = \alpha_1 \mathbf{E}y(k-1)u(k-n-1) + \ldots + \alpha_n \mathbf{E}y(k-n)u(k-n-1)$$

$$\mathbf{E}y(k)u(k-2n) = \alpha_1 \mathbf{E}y(k-1)u(k-2n) + \ldots + \alpha_n \mathbf{E}y(k-n)u(k-2n) \mathbf{E}y(k)u(k-m)$$

$$= \alpha_1 \mathbf{E}y(k-1)u(k-m) + \ldots + \alpha_n \mathbf{E}y(k-n)u(k-m)$$

Let $w(i) = \mathbf{E}y(k)u(k-i)$, then

$$\begin{pmatrix} w(1) \\ w(2) \\ \vdots \\ w(n) \\ w(n+1) \\ \vdots \\ w(2n) \\ \vdots \\ w(m) \end{pmatrix} = \Sigma \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ q\beta_1 \\ \vdots \\ q\beta_n \end{pmatrix} \qquad (3.3)$$

where

$$\Sigma = \begin{pmatrix} 0 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ w(1) & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ w(n-1) & w(n-2) & \ldots & 0 & 0 & 0 & \ldots & 1 \\ w(n) & w(n-1) & \ldots & w(1) & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ldots & 0 \\ w(2n-1) & w(2n-2) & \ldots & w(n) & 0 & \ldots & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ldots & 0 \\ w(m-1) & w(m-2) & \ldots & w(m-n) & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

The estimate of α_i and $q\beta_j$ can be obtained by solving the above equation. To further find q and a, notice that $q = \mathbf{E}x(k)u(k)$ depends on the distribution f of $u(k)$ as well as the unknown a, i.e., $q = q(f, a)$. If identification is carried out $(l+1)$ times with different distribution $f_1, f_2, \ldots, f_{l+1}$, the ratios

$$c_i = \frac{q(f_1, a)\beta_j}{q(f_{i+1}, a)\beta_j} = \frac{q(f_1, a)}{q(f_{i+1}, a)}, \quad i = 1, 2, \ldots, l \quad (3.4)$$

are numerically obtained and this provides l equations for the unknown $a \in R^l$

$$c_i q(f_{i+1}, a) = q(f_1, a), \quad i = 1, 2, ..., l. \quad (3.5)$$

All variables c_i's and $q(f_i, a)$'s are computable and thus, a can be solved. Several examples are provided below.

Non-symmetric Relay nonlinearity: Consider a non-symmetric Relay

$$x(k) = \begin{cases} 1 & u(k) > a_2 \\ 0 & -a_1 < u(k) < a_2 \\ -1 & u(k) < -a_1 \end{cases} \quad (3.6)$$

where $a = (a_1, a_2)'$ is two dimensional. Let f_1 and f_2 be uniform distributions in $[-b_1, b_1]$ and $[-b_2, b_2]$ respectively. Let the third distribution be

$$f_3(u) = \begin{cases} 0 & u < -d_1 \text{ or } u > d_2 \\ \dfrac{d_2^2}{d_1 d_2^2 + d_2 d_1^2} & -d_1 \le u \le 0 \\ \dfrac{d_1^2}{d_1 d_2^2 + d_2 d_1^2} & 0 < u \le d_2 \end{cases} . (3.7)$$

For $d_2, d_1, b_1, b_2 > max(a_1, a_2)$, it follows that

$$q(f_i, a) = \frac{(2b_i^2 - a_1^2 - a_2^2)}{4b_i}, \; i = 1, 2 \quad (3.8)$$

$$q(f_3, a) = \frac{d_1^2 d_2^2}{d_1 d_2^2 + d_2 d_1^2} - \frac{1}{2} \frac{d_2^2 a_1^2 + d_1^2 a_2^2}{d_1 d_2^2 + d_2 d_1^2}.$$

From the definitions of $c_i = \frac{q(f_1, a)}{q(f_{i+1}, a)}$, it follows that

$$\begin{pmatrix} b_2 - c_1 b_1, & b_2 - c_1 b_1 \\ 1 - \frac{2b_1 c_2 d_2^2}{d_1 d_2^2 + d_2 d_1^2}, & 1 - \frac{2b_1 c_2 d_1^2}{d_1 d_2^2 + d_2 d_1^2} \end{pmatrix} \begin{pmatrix} a_1^2 \\ a_2^2 \end{pmatrix}$$

$$= \begin{pmatrix} 2b_1^2 b_2 - 2c_1 b_1 b_2^2 \\ 2b_1^2 - \frac{4b_1 c_2 d_1^2 d_2^2}{d_1 d_2^2 + d_2 d_1^2} \end{pmatrix}.$$

Hence, $a = (a_1, a_2)'$ are uniquely obtained by solving the above equation.

Non-symmetric Preload nonlinearity: In this case,

$$x(k) = \begin{cases} u(k) + a_2 & u(k) > 0 \\ 0 & u(k) = 0 \\ u(k) - a_1 & u(k) < 0 \end{cases} . \quad (3.9)$$

Let f_1 and f_2 be two uniform distributions in $[-b_1, b_1]$ and $[-b_2, b_2]$ respectively, and

$$f_3(u) = \begin{cases} 0.5\delta(u + b_3) & u = -b_3 \\ \dfrac{1}{2b_3} & 0 \le u \le b_3 \\ 0 & \text{otherwise} \end{cases} \quad (3.10)$$

where $\delta(t)$ is the δ function. It is easily calculated that

$$q(f_i, a) = \frac{b_i^2}{3} + \frac{b_i(a_1 + a_2)}{4}, \; i = 1, 2 \quad (3.11)$$

$$q(f_3, a) = \frac{2b_3^2}{3} + \frac{a_1 b_3}{2} + \frac{a_2 b_3}{4}.$$

Therefore, $a = (a_1, a_2)$ can be uniquely calculated from

$$\begin{pmatrix} 1, & 1 \\ \frac{b_3 c_2}{2} - \frac{b_1}{4}, & \frac{b_3 c_2}{4} - \frac{b_1}{4} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (3.12)$$

$$= \begin{pmatrix} \frac{4(b_1^2 - b_2^2 c_1)}{3(b_2 c_1 - b_1)} \\ \frac{b_1^2}{3} - \frac{2b_3^2 c_2}{3} \end{pmatrix}.$$

Saturation nonlinearity in Figure 2: Let f_1 and f_2 be two the uniform distributions in $[-b_i, b_i]$, $i = 1, 2$ with $b_2, b_1 > a$. It is a routine to calculate

$$q(f_i, a) = \frac{a}{b_i}\left(\frac{b_i^2}{2} - \frac{1}{6}a^2\right), \quad i = 1, 2 \quad (3.13)$$

and this implies, from the definition of c_1,

$$a = \sqrt{\left|\frac{3(c_1 b_2 - b_1) b_1 b_2}{c_1 b_1 - b_2}\right|}. \quad (3.14)$$

Dead-Zone nonlinearity in Figure 2. Let f_1 and f_2 be two uniform distributions in $[-b_1, b_1]$ and $[-b_2, b_2]$ respectively with $b_2 > b_1 > a$. It is easily calculated that

$$q(f_i, a) = \frac{1}{6b_i}a^3 - \frac{ab_i}{2} + \frac{1}{3}b_i^2, \; i = 1, 2 \; (3.15)$$

$$c_1 = \frac{\frac{1}{6b_1}a^3 - \frac{ab_1}{2} + \frac{1}{3}b_1^2}{\frac{1}{6b_2}a^3 - \frac{ab_2}{2} + \frac{1}{3}b_2^2}.$$

This implies

$$(c_1 b_1 - b_2)a^3 + 3b_1 b_2 (b_1 - c_1 b_2)a + \quad (3.16)$$

$$2b_1 b_2 (c_1 b_2^2 - b_1^2) = 0.$$

It can be shown that this equation always has three real roots. One lies in the interval $(-\infty, 0)$, the second one in $(0, b_1)$ and the last one in (b_1, ∞), provided $b_2 > b_1 > a > 0$. Since $0 < a < b_1$ and b_1 is known, a can be uniquely determined.

The identification algorithm using correlation analysis for static input nonlinearities is now summarized.

Identification algorithm using correlation analysis:

Step 1: Apply input $u(k)$ with the distributions f_1 and define

$$w(i) = \frac{1}{N} \sum_{k=1}^{N} y(k) u(k - i), \quad i = 1, 2, ..., m (3.17)$$

for some large N and $m \ge 2n$.

Step 2: Construct equation (3.3). Solve the equation and denote the solution by $\hat{\alpha}_i$ and $\widehat{q(f_1, a)}\beta_j$.

Step 3: Repeat Steps 1 and 2 by applying the input with different distributions f_i, $i = 2, ..., l+1$ to obtain $q(\widehat{f_i, a})\beta_j$.

Step 4: Calculate $q(f_i, a)$ and find

$$c_i = \frac{q(\widehat{f_1, a})\beta_j}{q(\widehat{f_i, a})\beta_j} = \frac{q(\widehat{f_1, a})}{q(\widehat{f_i, a})}, \quad i = 2, ..., l+1. \tag{3.18}$$

Denote the solution by $\hat{a}$. Compute $q(\widehat{f_1, a})$ using $\hat{a}$.

Step 5: The estimates are $\hat{a}$, $\hat{\alpha}_i$ and

$$\hat{\beta}_j = \frac{q(\widehat{f_1, a})\beta_j}{q(\widehat{f_1, a})}, \quad j = 1, 2, ..., n. \tag{3.19}$$

Note that $w(i) = \frac{1}{N}\sum_{k=1}^{N} y(k)u(k-i) \rightarrow \mathbf{E}y(k)u(k-i)$ as $N \rightarrow \infty$ and therefore, the estimates derived by the correlation method converge to the true values. It is also noted that in calculating $c_i = \frac{q(\widehat{f_1, a})\beta_j}{q(\widehat{f_{i+1}, a})\beta_j}$ and the corresponding $\hat{a}$, any j can be used. It may be beneficial to use the average $\hat{c}_i = \frac{1}{n}\sum_{j=1}^{n} \frac{q(\widehat{f_1, a})\beta_j}{q(\widehat{f_{i+1}, a})\beta_j}$. Now, consider a numerical simulation example.

Example 2: Consider the same linear system as in Example 1 with the Saturation nonlinearity of $a = 1$. For simulation, $N = 2000$, and $v(k)$ is uniformly in $[-0.1, 0.1]$. Apply two inputs uniformly distributed in $[-2, 2]$ and $[-3, 3]$ respectively. The Identification Algorithm produces following estimates

$$[\hat{a}, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2] = [1.005, -0.8390, \tag{3.20}$$

$$-0.1555, 1.09, 1.08]$$

Remark 3.1. Comparing to the deterministic identification algorithm presented in the previous section, the correlation method needs to estimate the correlation between input and outputs and therefore, a long length of data is needed. In Example 2, the estimation results using 2000 data points are not as good as the one by the separable least squares approach using only 100 data points. However, the correlation method applies to nonlinearities parameterized by some $a \in R^l$ while the separable least squares method is limited to nonlinearities parameterized by a one or two dimensional a.

Remark 3.2. The choices of f_i's are arbitrary and the formula derived above are just examples. One may pick other input distributions for identification. Of course, formula would be different for different distributions. It is also interesting to note that if the nonlinearity is even, then any even distribution f_i gives rise to a zero $q(f_i, a)$. In this case, non-even distributions such as f_3 in the non-symmetric Preload example can be used. Similar discussion applies to odd nonlinearities.

4. CONCLUDING REMARKS

Two identification algorithms are proposed for the system with hard input nonlinearities. The first one notes the fact that a is one dimensional and thus, transforms a higher dimensional nonlinear identification problem into a one dimensional minimization problem. The method is particularly effective for many input nonlinearities which are parameterized by a single parameter a. The approach also applies to nonlinearities with memory. The second algorithm relies on repeated identifications with different random input sequences and is convergent.

REFERENCES

E.W. Bai (1998), An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems, *Automatica*, Vol. 34, No. 3. pp.333-338.

S.A. Billings and S.Y. Fakhouri (1978), Identification of a class of nonlinear systems using correlation analysis, *Proc. of IEE* Vol. 125, No. 7. pp.691-697.

X., Gu, Y. Bao and Z. Lang (1988), A parameter identification method for a class of discrete time nonlinear systems, *Proc. 12th IMACS World Congress, Paris*, Vol. 4, pp.627-629.

L. Ljung (1987) *System Identification*, Englewood Cliffs, NJ: Prentice -Hall.

A. Ruhe and P. Wedin (1980), Algorithms for separable nonlinear least squares problems, *SIAM Review*, Vol. 22, pp.318-337.

P. Stoica (1981), On the convergence of an iterative algorithm used for Hammerstein system identification, *IEEE Trans. on Automatic Control*, Vol. 26, pp.967-969.

G. Tao and C.A. Canudas de Wit (Eds) (1997), SPECIAL ISSUE ON ADAPTIVE SYSTEMS WITH NON-SMOOTH NONLINEARITIES, *Int. J. Adapt. Contr. Signal Process*, Vol.11, No.1.

J. Voros (1997), Parameter identification of discontinuous Hammerstein systems, *Automatica*, Vol. 33, No. 6, pp.1141-1146.