[3] H. Deng, M. Krstić, and R. J. Williams, "Stabilization of stochastic nonlinear systems driven by noise of unknown covariance," *IEEE Trans. Autom. Control*, vol. 46, no. 8, pp. 1237–1253, Aug. 2001.

[4] M. K. Ghosh, A. Arapostathis, and S. I. Marcus, "Optimal control of switching diffusions with application to flexible manufacturing systems," *SIAM J. Control and Optimization*, vol. 31, pp. 1183–1204, 1993.

[5] R. Z. Has'minskii, *Stochastic Stability of Differential Equations*. Alphen aan den Rijn, The Netherlands: Sithoff & Noordhoff, 1980.

[6] L. Huang and X. Mao, "On input-to-state stability of stochastic retarded systems with Markovian switching," *IEEE Trans. Autom. Control*, vol. 54, no. 8, pp. 1898–1902, Aug. 2009.

[7] Y. Ji and H. J. Chizeck, "Controllability, stabilizability and continuous-time Markovian jump linear quadratic control," *IEEE Trans. Autom. Control*, vol. 35, no. 7, pp. 777–788, Jul. 1990.

[8] R. S. Lipster and A. N. Shiryayev, *Theory of Martingales*. Dordrecht, The Netherlands: Kluwer Academic, 1989.

[9] Q. Luo and X. Mao, "Stochastic population dynamics under regime switching II," *J. Math. Anal. Appl.*, vol. 355, pp. 577–593, 2009.

[10] X. Mao, *Stochastic Differential Equations & Applications*. Chichester, U.K.: Horwood Publishing, 1997.

[11] X. Mao, "A note on the LaSalle-type theorems for stochastic differential delay equations," *J. Math. Anal. Appl.*, vol. 268, pp. 125–142, 2002.

[12] X. Mao, "Exponential stability of stochastic delay interval systems with Markovian switching," *IEEE Trans. Autom. Control*, vol. 47, no. 10, pp. 1604–1612, Oct. 2002.

[13] X. Mao, J. Lam, S. Xu, and H. Gao, "Razumikhin method and exponential stability of hybrid stochastic delay interval systems," *J. Math. Anal. Appl.*, vol. 314, pp. 45–66, 2006.

[14] X. Mao, J. Lam, and L. Huang, "Stabilisation of hybrid stochastic differential equations by delay feedback control," *Syst. Control Lett.*, vol. 57, pp. 927–935, 2008.

[15] X. Mao, Y. Shen, and C. Yuan, "Almost surely asymptotic stability of neutral stochastic differential delay equations with Markovian switching," *Stoch. Processes Appl.*, vol. 118, pp. 1385–1406, 2008.

[16] M. Mariton, *Jump Linear Systems in Automatic Control*. New York: Marcel Dekker, 1990.

[17] H. Shao, "Delay-range-dependent robust H^∞ filtering for uncertain stochastic systems with mode-dependent time delays and Markovian jump parameters," *J. Math. Anal. Appl.*, vol. 342, pp. 1084–1095, 2008.

[18] A. V. Skorohod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*. Providence, RI: American Mathematical Society, 1989.

[19] D. D. Sworder and R. O. Rogers, "An LQ-solution to a control problem associated with a solar thermal central receiver," *IEEE Trans. Autom. Control*, vol. AC-28, no. 10, pp. 971–978, Oct. 1983.

[20] E. I. Verriest and W. Michiels, "Stability analysis of systems with stochastically varying delays," *Syst. Control Lett.*, vol. 58, pp. 783–791, 2009.

[21] Z. Yang, X. Mao, and C. Yuan, "Comparison theorem of one-dimensional stochastic hybrid delay systems," *Syst. Control Lett.*, vol. 57, pp. 56–63, 2008.

[22] C. Yuan and X. Mao, "Robust stability and controllability of stochastic differential delay equations with Markovian switching," *Automatica*, vol. 40, pp. 343–354, 2004.

[23] D. Yue and Q.-L. Han, "Delay-dependent exponential stability of stochastic systems with time-varying delay, nonlinearity, and Markovian switching," *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 217–222, Feb. 2005.

An Interactive Term Approach to Non-Parametric FIR Nonlinear System Identification

Er-Wei Bai and Manfred Deistler

***Abstract*—In this technical note, a framework for designing specially structured input sequences for non-parametric nonlinear system identification is presented so that interaction terms which describe interactions among variables can be identified separately. In a sense, the approach decomposes a general difficult nonlinear identification problem into a number of problems that are of lower orders. Corresponding identification algorithms are proposed.**

***Index Terms*—Additive systems, kernel estimation, nonlinear system identification.**

I. Introduction

Nonlinear system identification is an important problem. Unfortunately, non-parametric nonlinear system identification without *a priori* structural information remains a very tough task. This is partially because the model class is extremely rich and wide. In a sense, nonlinear system identification is a world like "non-elephant zoology." Several classical papers [3], [9] and two recent special issues [4], [10] provide valuable insight into this problem.

Consider a rather general SISO nonlinear finite response system

$$\begin{aligned} y[k] &= f\left(u[k-1], u[k-2], \ldots, u[k-n]\right) + v[k] \\ &= \bar{c} + \sum_{j=1}^{n} \bar{f}_j\left(u[k-j]\right) + \sum_{1\le j_1<j_2\le n} \bar{f}_{j_1 j_2}\left(u[k-j_1], u[k-j_2]\right) + \ldots \\ &+ \sum_{1\le j_1<j_2<\ldots<j_n\le n} \bar{f}_{j_1 j_2 \ldots j_n}\left(u[k-j_1], u[k-j_2], \ldots, u[k-j_n]\right) \\ &+ v[k], k = 1, 2, \ldots \end{aligned} \quad (1)$$

where $y[k]$ and $u[k]$ are output and input measurements respectively, and the noise $v[k]$ is a sequence of independent random variables (not necessarily identically distributed) with zero mean and uniformly bounded variance. The functions $\bar{f}_j, \bar{f}_{j_1 j_2}, \ldots, \bar{f}_{j_1 j_2 \ldots j_n}$ are unknown and describe interactions of input variables. In the technical note, the input $u[\cdot]$ is assumed to be in $[\underline{a}, \bar{a}]$ for some $\underline{a} < \bar{a}$.

Almost all the works reported in the literature aim to find directly the function $f(u[k-1], \ldots, u[k-n])$ representing the input-output relationship of the system. This amounts to solve a high order nonlinear identification problem directly and can be difficult. What we are interested in this technical note is to ask the following question: Can this high order identification problem be decomposed into a number of lower order identification problems? What we are specially interested

Manuscript received November 30, 2006; revised June 09, 2007, and November 30, 2009; accepted December 08, 2009. First published May 17, 2010; current version published July 30, 2010. This work was supported in part by NSF ECS-0555394. Recommended by Associate Editor B. Ninness.

E.-W. Bai is with Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242 USA (e-mail: er-wei-bai@uiowa.edu).

M. Deistler is with Institute for Mathematical Methods in Economics, University of Technology, Vienna, Austria (e-mail: deistler@tuwien.ac.at).

Digital Object Identifier 10.1109/TAC.2010.2050355

in this technical note are nonlinear systems that contain only lower order interaction terms, for instance, up to third order terms

$$y[k] = \bar{c} + \sum_{j=1}^{n} \bar{f}_j \left(u[k-j]\right) + \sum_{1 \le j_1 < j_2 \le n} \bar{f}_{j_1 j_2} \left(u[k-j_1], u[k-j_2]\right) + \sum_{1 \le j_1 < j_2 < j_3 \le n} \bar{f}_{j_1 j_2 j_3} \left(u[k-j_1], u[k-j_2], u[k-j_3]\right) + v[k]. \quad (2)$$

Then, the question is how to identify the unknown interaction terms $\bar{f}_j$, $\bar{f}_{j_1 j_2}$ and $\bar{f}_{j_1 j_2 j_3}$ separately as well as the unknown constant, $\bar{c}$, based on the input and output measurements. In general, identification of these interaction terms separately is difficult because the output depends on all the terms. In the technical note, we show however that identification of each interaction term is possible by a specially designed input sequence. Moreover, identification of each interaction term is separable in some sense. Further, an identification algorithm is proposed which is consistent.

The work reported here was largely inspired by the study of additive systems [2] in statistics. The normalization procedure proposed is in the same spirit of the work [11] but in a different setting.

We make a comment by comparing the proposed method with the existing ones, in particular, orthogonal bases and Volterra series. The proposed method has to assume that (1) the system is FIR and the order n is not too high, and (2) the interactions among the variables are not too strong limited to 2 or 3 interactive terms. However, the method does allow highly nonlinear systems with no a priori information on the structure. Orthogonal bases method [10] is effective if the chosen base resembles the structure of the unknown system which is not trivial or a large number of terms is needed that has negative implications on identification. The proposed method does not need such a priori knowledge. The major advantage of Volterra series [7] is to make a non-parametric nonlinear system parametric which is effective if the unknown system is close to a low order polynomial. However, Volterra series is ineffective if the unknown nonlinearity is non-polynomial or the order is high. The proposed method can deal with non-polynomial or high order polynomial nonlinearities. In summary, the proposed method is not a substitute of the existing ones but a complement.

II. Input Design

Let the input range be given by m partitions

$$\underline{a} \le a_1 < a_2 < \ldots < a_m \le \bar{a}$$

where the choice of m balances how fine the resolution is with the length of the input sequence. The idea of input design is to have an input sequence $u[\cdot]$ that excites the system preferably with the minimal length and makes identification of each nonlinearity separable in some sense. We assume first that for every arbitrary combination of n-tuples $(a_{l_1}, \ldots, a_{l_n})$, $a_{l_i} \in \{a_1, a_2, \ldots, a_m\}$, there exists an integer q such that

$$\left(u[q-1], \ldots, u[q-n]\right) = \left(a_{l_1}, \ldots, a_{l_n}\right).$$

Denote by U_{m^n} the input matrix made of the input $u[\cdot]$

$$U_{m^n} = \begin{pmatrix} u[1] & \ldots & u[-n+2] \\ u[2] & \ldots & u[-n+3] \\ \vdots & \ddots & \vdots \\ u[m^n] & \ldots & u[m^n-n+1] \end{pmatrix} \quad (3)$$

which is applied to the system as one string

$$\left\{u[-n+2], \ldots, u[1], \ldots\ldots, u[m^n]\right\}.$$

Obviously, the minimal length of the input sequence that can generate all n-tuples as in (3) is $m^n + n - 1$ and in addition, each row of the input matrix (3) produces one combination of the n-tuple exactly once. Further, to average out the effect of noise, the input sequence should be repeated L times, i.e.

$$U_{Lm^n} = \underbrace{\begin{pmatrix} U_{m^n} \\ U_{m^n} \\ \vdots \\ U_{m^n} \end{pmatrix}}_{repeat\ L\ times}. \quad (4)$$

In such a case, the input length is $Lm^n + n - 1$. We refer to any sequence having the above properties as a generating sequence. Generating sequences are closely related to finite fields or Galois fields in number theory [8]. It is well known that if m is a prime number, the Galois sequence $GF(m^n)$ [1], [5] is such a sequence which has the length $m^n + n - 1$, produces each combination of $(a_{l_1}, \ldots, a_{l_n})$ exactly once in U_{m^n} and repeats itself after m^n steps. We give two examples here, $m = 2$ & $n = 2$ and $m = 2$ & $n = 3$ respectively

$$U_{2^2} = \begin{pmatrix} u[1] & u[0] \\ u[2] & u[1] \\ u[3] & u[2] \\ u[4] & u[3] \end{pmatrix} = \begin{pmatrix} a_1 & a_1 \\ a_2 & a_1 \\ a_2 & a_2 \\ a_1 & a_2 \end{pmatrix},$$

$$U_{2^3} = \begin{pmatrix} u[1] & u[0] & u[-1] \\ u[2] & u[1] & u[0] \\ u[3] & u[2] & u[1] \\ u[4] & u[3] & u[2] \\ u[5] & u[4] & u[3] \\ u[6] & u[5] & u[4] \\ u[7] & u[6] & u[5] \\ u[8] & u[7] & u[6] \end{pmatrix} = \begin{pmatrix} a_1 & a_1 & a_1 \\ a_2 & a_1 & a_1 \\ a_2 & a_2 & a_1 \\ a_2 & a_2 & a_2 \\ a_1 & a_2 & a_2 \\ a_2 & a_1 & a_2 \\ a_1 & a_2 & a_1 \\ a_1 & a_1 & a_2 \end{pmatrix}.$$

For an arbitrary integer m, construction of a generating sequence with the minimal length is not trivial. One way to construct an input sequence is to list $u[\cdot]$ according to a_l in a natural way as

$$\begin{pmatrix} u[n-1] & \ldots & u[n-n] \\ u[2n-1] & \ldots & u[2n-n] \\ \vdots & \ddots & \vdots \\ u[m^n \cdot n-1] & \ldots & u[m^n \cdot n-n] \end{pmatrix} = \begin{pmatrix} a_1 & a_1 & \ldots & a_1 & a_1 \\ a_1 & a_1 & \ldots & a_1 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1 & a_1 & \ldots & a_1 & a_m \\ a_1 & a_1 & \ldots & a_2 & a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_m & a_m & \ldots & a_m & a_m \end{pmatrix}. \quad (5)$$

If only every knth output measurement, $y[kn]$, $k = 1, \ldots, m^n$, is considered, the corresponding input sequence $(u[kn-1], \ldots, u[kn-n])$ is the kth row of the matrix (5) that contains every possible combination of $(a_{l_1}, a_{l_2}, \ldots, a_{l_m})$ exactly once when k increases from 1 to m^n. Note, except the length of the input sequence in (5) is nm^n, the input matrix (3) made of the Galois sequences and the one made of the input (5) are the same modulo some row permutations that will not change any results derived in the technical note. With a little abuse of notation for the case when m is not a prime number, we define the input matrix through a fictitious input $u'[\cdot]$

$$U_{m^n} = \begin{pmatrix} u'[1] & \dots & u'[-n+2] \\ u'[2] & \dots & u'[-n+3] \\ \vdots & \ddots & \vdots \\ u'[m^n] & \dots & u'[m^n-n+1] \end{pmatrix} = \begin{pmatrix} u[n-1] & \dots & u[n-n] \\ u[2n-1] & \dots & u[2n-n] \\ \vdots & \ddots & \vdots \\ u[m^n n-1] & \dots & u[m^n n-n] \end{pmatrix}. \quad (6)$$

Therefore, notation of the input matrix U_{m^n} is the same independent if m is prime or not. Of course, in the case that m is not prime, $u[\cdot]$ is replaced by $u'[\cdot]$ as in (6).

We now make two comments.

- The main difficulty in identification of the interaction terms is the coupling because it is hard to determine, with respect to the output, which contribution is from which interaction term. The input sequences chosen including the Galois sequences and the one in (6) above have a nice property summarized in (V.1) that separates contributions of each term in some sense.
- Clearly, the input as in (5) is not minimal for the case when m is not a prime number in terms of the input length. Some more clever ways exist that could shorten the input length and achieve all possible combinations. On the other hand, however, those clever ways are unbearably complicated in terms of book-keeping. To convey the idea of the technical note clearly, we will use the input sequence of (5) when m is not a prime number.

III. Systems With Up to Second Order Interaction Terms

A key step in identification is a normalization which makes identification of interaction terms possible. Unfortunately, the normalization procedure involves some painfully tedious book-keeping for a general system. To convey the idea without unnecessary complications, we first study a system with interactions up to second order interaction terms

$$y[k] = \bar{c} + \sum_{j=1}^{n} \bar{f}_j\left(u[k-j]\right) + \sum_{1\le j_1<j_2\le n} \bar{f}_{j_1j_2}\left(u[k-j_1], u[k-j_2]\right) + v[k]. \quad (7)$$

In identification, we estimate $\bar{c}$, $\bar{f}_j$ and $\bar{f}_{j_1j_2}$ based on the input and output measurements. Immediately, we notice that the representation of (7) is actually not identifiable. For instance, $\bar{f}_1(\cdot)+c$ and $\bar{f}_2(\cdot)-c$, for any constant c, would produce identical input-output measurements. To this end, we propose a two step normalization procedure.

Step 1) Define, for each $1 \le j_1 < j_2 \le n$

$$f_{j_1j_2}(x_1,x_2) = \bar{f}_{j_1j_2}(x_1,x_2) - \underbrace{\frac{1}{m}\sum_{l=1}^{m}\bar{f}_{j_1j_2}(a_l,x_2)}_{g_{j_2,j_1j_2}(x_2)} - \underbrace{\frac{1}{m}\sum_{l=1}^{m}\bar{f}_{j_1j_2}(x_1,a_l)}_{g_{j_1,j_1j_2}(x_1)} + \underbrace{\frac{1}{m^2}\sum_{l_1=1}^{m}\sum_{l_2=1}^{m}\bar{f}_{j_1j_2}\left(a_{l_1},a_{l_2}\right)}_{c_{j_1j_2}}.$$

Step 2) Define, with $g_{i,jk}$'s being specified in Step 1

$$f_1(x) = \left[\bar{f}_1(x) + \sum_{i=2}^{n} g_{1,1i}(x)\right] - \underbrace{\frac{1}{m}\sum_{l=1}^{m}\left[\bar{f}_1(a_l) + \sum_{i=2}^{n} g_{1,1i}(a_l)\right]}_{c_1}$$

$$\vdots$$

$$f_{n-1}(x) = \left[\bar{f}_{n-1}(x) + \sum_{i=1}^{n-2} g_{(n-1),i(n-1)}(x) + g_{n,(n-1)n}(x)\right] - \underbrace{\frac{1}{m}\sum_{l=1}^{m}\left[\bar{f}_{n-1}(a_l) + \sum_{i=1}^{n-2} g_{(n-1),i(n-1)}(a_l) + g_{n,(n-1)n}(a_l)\right]}_{c_{n-1}}$$

$$f_n(x) = \left[\bar{f}_n(x) + \sum_{i=1}^{n-1} g_{n,in}(x)\right] - \underbrace{\frac{1}{m}\sum_{l=1}^{m}\left[\bar{f}_n(a_l) + \sum_{i=1}^{n-1} g_{n,in}(a_l)\right]}_{c_n}.$$

With $c = \bar{c} - \sum_{1\le j_1<j_2\le n} c_{j_1j_2} + \sum_{i=1}^{n} c_i$, it follows that:

$$y[k] = c + \sum_{j=1}^{n} f_j\left(u[k-j]\right) + \sum_{1\le j_1<j_2\le n} f_{j_1j_2}\left(u[k-j_1], u[k-j_2]\right) + v[k], \quad k = 1,2,\dots,Lm^{n+1}. \quad (8)$$

The above system (8) is well defined for identification purpose. One critical property, from lemmas (V.2) and (V.3) in Appendix, is that

$$\sum_{l=1}^{m} f_j(a_l) = 0, \quad \sum_{l=1}^{m} f_{j_1j_2}(a_l,x_2) = \sum_{l=1}^{m} f_{j_1j_2}(x_1,a_l) = 0$$

for every $1 \le j \le n$ and $1 \le j_1 < j_2 \le n$, which makes identification of each interaction terms separable.

A. Identification Algorithm

Denote Lm^n, Lm^{n-1} and Lm^{n-2} term averages of the noise, respectively, by

$$\bar{v}_{Lm^n} = \frac{1}{Lm^n}\sum_{k=1}^{Lm^n} v[k]$$

$$\bar{v}_{Lm^{n-1}} = \frac{1}{Lm^{n-1}} \sum_{\substack{k=1 \\ \underline{u[k-j]=a_l}}}^{Lm^n} v[k]$$

where the underline $u[k-j]=a_l$ stands for the fact the summation is with respect to all the terms $k \in [1, Lm^n]$ where the value of $u[k-j] = a_l$

$$\bar{v}_{Lm^{n-2}} = \frac{1}{Lm^{n-2}} \sum_{\substack{k=1 \\ u[k-j]a=a_{l_1},u[k-j]b=a_{l_2}}}^{Lm^n} v[k]$$

where the underline $u[k-j]a = a_{l_1}, u[k-j]b = a_{l_2}$ stands for the fact the summation is with respect to all $k \in [1, Lm^n]$ where the value of $u[k-j]a = a_{l_1}, u[k-j]b = a_{l_2}$, From lemmas (V.2) and (V.3) in Appendix, it is easily verified that

$$\begin{aligned}\frac{1}{Lm^n}\sum_{k=1}^{Lm^n} y[k] &= \frac{1}{Lm^n}\sum_{k=1}^{Lm^n} c + \sum_{j=1}^{n}\frac{1}{Lm^n}\sum_{k=1}^{Lm^n} f_j\left(u[k-j]\right)\\ &\quad + \sum_{1\le j_1<j_2\le n}\frac{1}{Lm^n}\\ &\quad \times \sum_{k=1}^{Lm^n} f_{j_1j_2}\left(u[k-j]a, u[k-j]b\right) + \bar{v}_{Lm^n}\\ &= c + \bar{v}_{Lm^n}.\end{aligned}$$

Similarly

$$\frac{1}{Lm^{n-1}}\sum_{\substack{k=1\\ u[k-j]=a_l}}^{Lm^n} y[k] = c + f_j(a_l) + \bar{v}_{Lm^{n-1}},$$

$$\begin{aligned}\frac{1}{Lm^{n-2}}\sum_{\substack{k=1\\ u[k-j]a=a_{l_1}, u[k-j]b=a_{l_2}}}^{Lm^n} y[k] &= + f_{j_1j_2}\left(a_{l_1}, a_{l_2}\right)\\ &\quad + f_{j_1}\left(a_{l_1}\right) + f_{j_2}\left(a_{l_2}\right)\\ &\quad + \bar{v}_{Lm^{n-2}}.\end{aligned}$$

Therefore, the estimates $\hat{c}$, $\hat{f}_j$ and $\hat{f}_{j_1j_2}$ of c, f_j and $f_{j_1j_2}$ can be defined as follows:

$$\begin{aligned}\hat{c} &= \frac{1}{Lm^n}\sum_{k=1}^{Lm^n} y[k]\\ \hat{f}_j(a_l) &= \frac{1}{Lm^{n-1}}\sum_{\substack{k=1\\ u[k-j]=a_l}}^{Lm^n} y[k] - \hat{c},\\ &\qquad 1\le j\le n, \quad 1\le l\le m\\ \hat{f}_{j_1j_2}\left(a_{l_1}, a_{l_2}\right) &= \frac{1}{Lm^{n-2}}\sum_{\substack{k=1\\ u[k-j]a=a_{l_1}, u[k-j]b=a_{l_2}}}^{Lm^n} y[k]\\ &\qquad - \hat{f}_{j_1}\left(a_{l_1}\right) - \hat{f}_{j_2}\left(a_{l_2}\right) - \hat{c}\\ &\qquad 1\le j_1<j_2\le n, \quad 1\le l_1, l_2\le m. \end{aligned} \tag{9}$$

Theorem III.1: Consider the system (8), the input (3) or (6) and the estimator (9). Then

$$\hat{c}\to c, \quad \hat{f}_j(a_l)\to f_j(a_l), \quad \hat{f}_{j_1j_2}\left(a_{l_1}, a_{l_2}\right)\to f_{j_1j_2}\left(a_{l_1}, a_{l_2}\right)$$

with probability one as $Lm^{n-2}\to\infty$ for every $1\le j\le n$, $1\le j_1 < j_2 \le n$ and $1\le l_1, l_2, l\le m$.

Proof: Since $\bar{v}_{Lm^n}$, $\bar{v}_{Lm^{n-1}}$ and $\bar{v}_{Lm^{n-2}}$ converge to zero with probability one as $Lm^{n-2}\to\infty$ by a suitable strong law of large numbers, the results follow directly.

In fact, the above results hold for filtered white noise disturbance, see [6], Section 14.9.

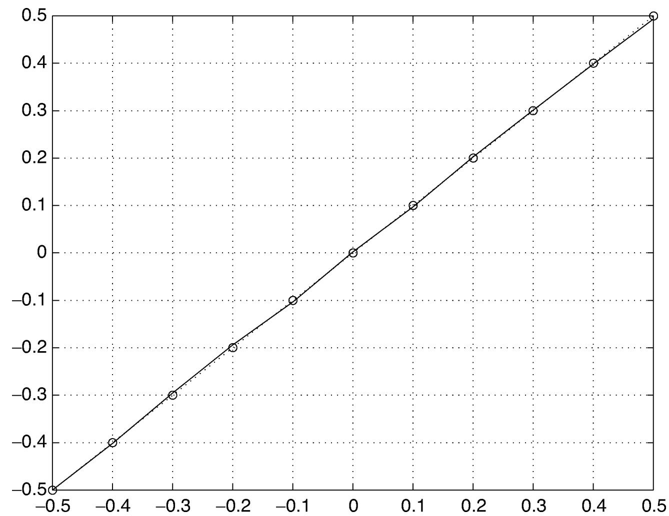

Fig. 1. $f_1(x)$(solid) and $\hat{f}_1(x)$(dashed and circles).

B. Numerical Simulations

Let the unknown system contain up to second order interaction terms

$$\begin{aligned}y[k] &= \underbrace{0}_{\bar{c}} + \underbrace{u[k-1]}_{\bar{f}_1(u[k-1])} + \underbrace{7\cdot u[k-2]^3}_{\bar{f}_2(u[k-2])} + \underbrace{3\cdot u[k-3]^2}_{\bar{f}_3(u[k-3])}\\ &\quad + \underbrace{0.4\cdot u[k-1]\cdot u[k-2]}_{\bar{f}_{12}(u[k-1],u[k-2])} + \underbrace{0}_{\bar{f}_{13}(u[k-1],u[k-3])}\\ &\quad + \underbrace{0}_{\bar{f}_{23}(u[k-2],u[k-3])} + v[k]\\ &= \underbrace{0.3}_{c} + \underbrace{u[k-1]}_{f_1(u[k-1])} + \underbrace{7\cdot u[k-2]^3}_{f_2(u[k-2])} + \underbrace{3\cdot u[k-3]^2 - 0.3}_{f_3(u[k-3])}\\ &\quad + \underbrace{0.4\cdot u[k-1]\cdot u[k-2]}_{f_{12}(u[k-1],u[k-2])} + \underbrace{0}_{f_{13}(u[k-1],u[k-3])}\\ &\quad + \underbrace{0}_{f_{23}(u[k-2],u[k-3])} + v[k].\end{aligned}$$

In simulation, $n = 3$, $L = 3$, $[-a, a] = [-0.5, 0.5]$, $a_l = -0.5 + 0.1 * l$, $l = 1, 2, \ldots, 11 = m$ and iid Gaussian noises $v[k]$ is added with $\mathrm{SNR} = 20$ dB. Figs. 1–6 show the actual f_i and $f_{j_1j_2}$ (solid) super-imposed by their estimates $\hat{f}_j$ and $\hat{f}_{j_1j_2}$ (dashed). Obviously, a satisfactory result is achieved and the actual functions and their estimates are almost indistinguishable.

IV. Conclusion

We make a few final remarks in this section.

- The result applies to any nonlinear system with arbitrary lth order interaction terms. However, the number of interaction terms that need to be identified increases quickly when l increases. Also, the amount of data samples required is high, proportional to m^n. Thus, the method is effective only when n, m and l are small. It is interesting to investigate how to modify the approach when n, m or l is not small.

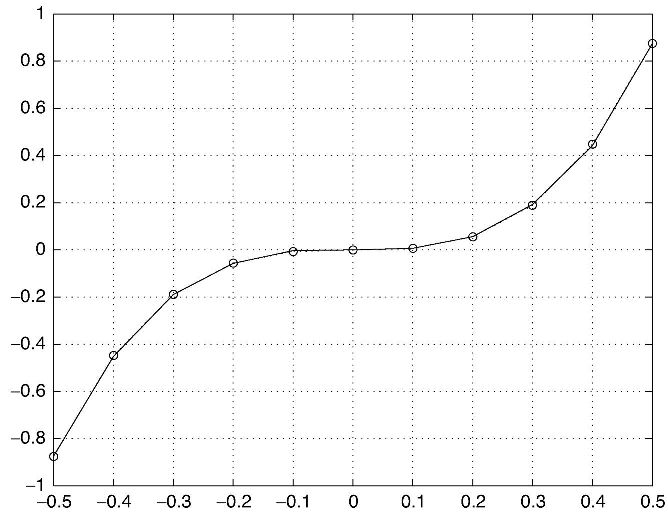

Fig. 2. $f_2(x)$(solid) and $\hat{f}_2(x)$(dashed and circles).

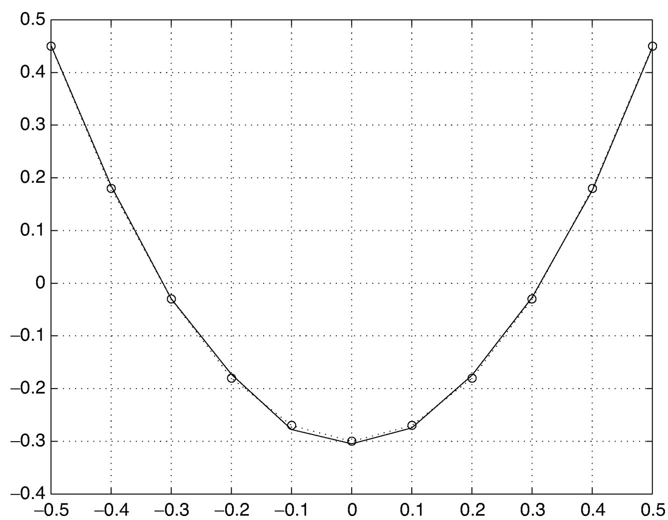

Fig. 3. $f_3(x)$(solid) and $\hat{f}_3(x)$(dashed and circles).

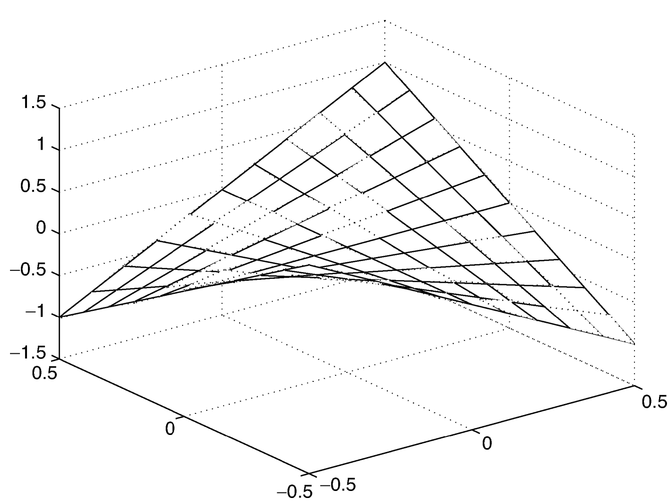

Fig. 4. $f_{12}(x_1, x_2)$(solid) and $\hat{f}_{12}(x_1, x_2)$(dashed).

- The approach requires a design of input sequence. In some applications, this may or may not be possible.

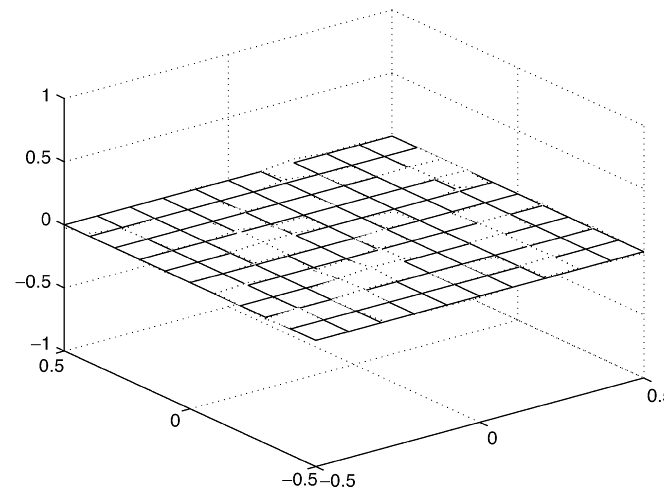

Fig. 5. $f_{13}(x_1, x_2)$(solid) and $\hat{f}_{13}(x_1, x_2)$(dashed).

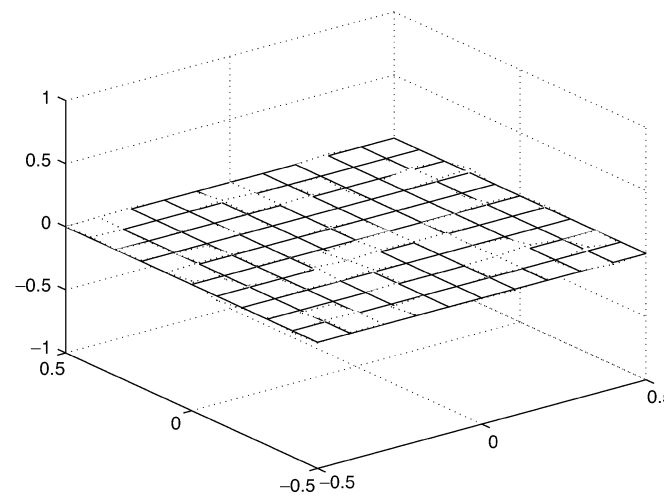

Fig. 6. $f_{23}(x_1, x_2)$(solid) and $\hat{f}_{23}(x_1, x_2)$(dashed).

APPENDIX

Consider the matrix U_{m^n} in (3) or (6) (in this case, replacing $u(k)$ by $u'(k) = u[kn]$) and the partition $a \le a_1 < a_2 < \ldots < a_m \le b$.

Then, we have

Lemma V.1:

1) For $1 \le j \le n$

$$\sum_{k=1}^{m^n} u[k-j] = m^{n-1} \sum_{l=1}^{m} a_l.$$

2) For $1 \le j \le n$ and $1 \le l \le m$

$$\sum_{\substack{k=1 \\ u[k-j]=a_l}}^{m^n} u[k-j] = m^{n-1} a_l.$$

3) For $1 \le j, j_1 \le n$, $j \ne j_1$ and $1 \le l \le m$,

$$\sum_{\substack{k=1 \\ u[k-j]a=a_l}}^{m^n} u[k-j] = m^{n-2} \sum_{l=1}^{m} a_l.$$

4) For $1 \le j, j_1 \le n$, $j \ne j_1$ and $1 \le l, p \le m$

$$\sum_{\substack{k=1 \\ u[k-j]=a_l, u[k-j]a=a_p}}^{m^n} u[k-j] = m^{n-2} a_l.$$

5) For $1 \le j \le n$, $1 \le j_1 < j_2 \le n$, $j \ne j_1, j_2$ and $1 \le l, p \le m$

$$\sum_{\substack{k=1 \\ u[k-j]a=a_l, u[k-j]b=a_p}}^{m^n} u[k-j] = m^{n-3} \sum_{l=1}^{m} a_l.$$

6) For $1 \le j_1 < j_2 \le n$

$$\sum_{k=1}^{m^n} (u[k-j]a, u[k-j]b) = m^{n-2} \sum_{l_1=1}^{m} \sum_{l_2=1}^{m} (a_{l_1}, a_{l_2}).$$

7) For $1 \le j_1 < j_2 \le n$, $j_3 \ne j_1, j_2$ and $1 \le l \le m$

$$\sum_{\substack{k=1 \\ u[k-j]c=a_l}}^{m^n} (u[k-j]a, u[k-j]b) = m^{n-3} \sum_{l_1=1}^{m} \sum_{l_2=1}^{m} (a_{l_1}, a_{l_2}).$$

8) For $1 \le j_1 < j_2 \le n$ and $1 \le l \le m$

$$\sum_{\substack{k=1 \\ u[k-j]a=a_l}}^{m^n} (u[k-j]a, u[k-j]b) = m^{n-2} \sum_{i=1}^{m} (a_l, a_i).$$

Similarly

$$\sum_{\substack{k=1 \\ u[k-j]b=a_l}}^{m^n} (u[k-j]a, u[k-j]b) = m^{n-2} \sum_{i=1}^{m} (a_i, a_l).$$

9) For $1 \le j_1 < j_2 \le n$ and $1 \le l_1, l_2 \le m$

$$\sum_{\substack{k=1 \\ u[k-j]a=a_{l_1}, u[k-j]b=a_{l_2}}}^{m^n} (u[k-j]a, u[k-j]b) = m^{n-2} (a_{l_1}, a_{l_2}).$$

10) For $0 \le j_1 < j_2 \le n$, $j_3 \ne j_1, j_2$ and $0 \le l, l_1 \le m$

$$\sum_{\substack{k=1 \\ u[k-j]a=a_l, u[k-j]c=a_{l_1}}}^{m^n} (u[k-j]a, u[k-j]b) = m^{n-3} \sum_{i=1}^{m} (a_l, a_i).$$

Similarly

$$\sum_{\substack{k=1 \\ u[k-j]b=a_l, u[k-j]c=a_{l_1}}}^{m^n} (u[k-j]a, u[k-j]b) = m^{n-3} \sum_{i=1}^{m} (a_i, a_l).$$

11) For $1 \le j_1 < j_2 \le n$, $1 \le j_3 < j_4 \le n$, $j_3, j_4 \ne j_1, j_2$ and $1 \le p, q \le m$

$$\sum_{\substack{k=1 \\ u[k-j]c=a_p, u[k-j]d=a_q}}^{m^n} (u[k-j]a, u[k-j]b) = m^{n-4} \sum_{l_1=1}^{m} \sum_{l_2=1}^{m} (a_{l_1}, a_{l_2}).$$

Lemma V.2: Consider the input matrix U_{m^n} in (3) or (6) (in this case, replacing $u(k)$ by $u'(k) = u[kn]$). Let $\bar{g}(\cdot)$ be any function. Define

$$g(x) = \bar{g}(x) - \frac{1}{m} \sum_{l=1}^{m} \bar{g}(a_l).$$

Then, $\sum_{l=1}^{m} g(a_l) = 0$.

Lemma V.3: Consider the input matrix U_{m^n} in (3) or (6) (in this case, replacing $u(k)$ by $u'(k) = u[kn]$). Let $\bar{g}(\cdot, \cdot)$ be any function. Define

$$g(x_1, x_2) = \bar{g}(x_1, x_2) - \frac{1}{m} \sum_{l=1}^{m} \bar{g}(a_l, x_2) - \frac{1}{m} \sum_{l=1}^{m} \bar{g}(x_1, a_l) + \frac{1}{m^2} \sum_{l_1=1}^{m} \sum_{l_2=1}^{m} \bar{g}(a_{l_1}, a_{l_2})$$

Then, $\sum_{l=1}^{m} g(a_l, x_2) = \sum_{l=1}^{m} g(x_1, a_l) = 0$.

References

[1] K. Godfrey, *Perturbation Signals for Identification*. Hemel Hempstead, Hertfordshire, U.K.: Prentice-Hall, 1993.

[2] T. Hastie and R. Tibshirani, *Generalized Additive Models*. London, U.K.: Chapman & Hall, 1997.

[3] A. Juditsky, H. Hjalmarsson, A. Benveniste, B. Delyon, L. Ljung, J. Sjoberg, and Q. Zhang, "Nonlinear block-box models in system identification: Mathematical foundations," *Automatica*, vol. 31, pp. 1725–1750, 1995.

[4] L. Ljung and A. Vicino, "Special issue on identification," *IEEE Trans. Autom. Control*, vol. 50, no. 10, pp. 1477–1634, Oct. 2005.

[5] P. M. Makila, "Robust identification and Galois sequences," *Int. J. Control*, vol. 54, pp. 1189–1200, 1991.

[6] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach*. New York: IEEE Press, 2001.

[7] W. Rugh, *Nonlinear System Theory: A Volterra/Wiener Approach*. Baltimore, MD: Johns Hopkins Univ. Press, 1981.

[8] M. R. Schroeder, *Number Theory in Science and Communication*. Berlin, Germany: Springer-Verlag, 1986.

[9] J. Sjoberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky, "Nonlinear black-box modeling in system identification: A unified overview," *Automatica*, vol. 31, pp. 1691–1724, 1995.

[10] T. Soderstrom, P. Van den Hof, B. Wahlberg, and S. Weiland, Eds., "Special issue on data-based modeling and system identification," *Automatica*, vol. 41, no. 3, pp. 357–562, 2005.

[11] S. Sperlich, D. Tjostheim, and L. Yang, "Non-parametric estimation and testing of interaction in additive models," *Econ. Theory*, vol. 18, pp. 197–251, 2002.