Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

A Bootstrap Term Selection Method for the Identification of Time-Varying Nonlinear Systems

Bashiru I. Ikharia and David T. Westwick

***Abstract*— Most physical systems are nonlinear and often time-varying. Constructing accurate models for nonlinear systems require specialized model structures that include their nonlinearities, whereas models of time-varying systems must include the time courses of the model's parameters. This contribution implements a technique in which the time dependence of the system's parameters are modeled by projecting them onto one or more expansion bases. However, unless an appropriate set of bases is chosen, it is likely that many of the basis functions used in the expansion will not contribute appreciably to the final model and may cause inaccuracy in the parameter estimates. Our study addresses the selection of a minimal number of parameters for optimization from a basis expansion of a system's time variations. The bootstrap method is used to select significant basis coefficients and hence basis functions in order to improve the overall accuracy of the model. The performance of the algorithm is demonstrated on simulated data from a system used to model the reflex contribution to joint stiffness.**

I. INTRODUCTION

System identification deals with building mathematical models of dynamic systems from observed data. This modeling approach has found many applications in biomedical fields where the system descriptions are often nonlinear and/or time-varying (TV). The TV and nonlinear nature of many physiological systems has led to the study and development of specialized procedures for their identification.

Identification methods for systems with rapidly changing dynamics, those that change too quickly to be tracked using methods based on adaptive filtering, have been developed [1], [2] but these algorithms generally have limits on their suitability for practical applications. For example, ensemble approaches assume that an ensemble of data, in which the system undergoes identical time variations in each trial, can be obtained [2]; basis expansion applies a set of basis functions to the time variations in order to make the parameters describing the system time-invariant (TIV), besides the need to select the basis *a priori*, it also requires the estimation of a relatively large number of parameters [1]. Recently, [3] combined the ensemble [2] and basis expansion [1] approaches in order to effectively deal with some of the lapses in each method.

In this study, a basis expansion technique for identifying rapidly changing TV nonlinear systems is derived. We employ a single data record, hence avoiding the alignment problems inherent in the ensemble approach, however the technique could readily be modified to use an ensemble of data records, by following the derivation presented in [3]. The dimension of the identification problem is reduced by applying a term selection algorithm to retain only the relevant expansion coefficients and hence basis functions for the final model description. Note that a dimension reduction step is commonly used in basis expansion techniques, for example Optimal Parameter Search [1], but since our system model is nonlinear in its parameters, the methods often employed (such as in [1]) cannot be used.

B. Ikharia and D. Westwick are with the Department of Electrical and Computer Engineering, Schulich School of Engineering, University of Calgary, 2500 University Dr. NW., Calgary, AB, Canada, T2N 1N4. {biikhari, dwestwic}@ucalgary.ca

The proposed method uses a separable least squares (SLS) optimization [4], [5] to estimate the parameters and a bootstrap [6] term selection algorithm to select significant basis functions from an initial set of basis functions. The algorithm seeks to obtain the sparsest representation of the final model that provides an accurate prediction of the system output.

An outline of the paper is as follows: In Section II we give a description of the class of TV nonlinear systems under study and equally apply basis expansion for transformation of the identification problem. Bootstrap methodology is discussed in Section III, where we also give a summary of the overall algorithm. A simulation study using a TV nonlinear Hammerstein model of the reflex contribution to joint stiffness is presented in Section IV and Section V concludes the paper.

II. NONLINEAR MODEL DESCRIPTION

Many biomedical systems can be efficiently modeled as a Hammerstein system [2], [5], [7], a two block cascade consisting of a static nonlinearity followed by a dynamic linearity. A TV Hammerstein cascade [2], [4] is shown in Fig. 1, where $u(t)$ and $y(t)$ are the system's input and output respectively at time t and are assumed to be available through measurements, $e(t)$ is the measurement noise, or innovation, to the system and $x(t)$ is an internal signal that is not directly available. The static nonlinearity, $m(\cdot, t)$, is modeled as a polynomial expansion of the input, $u(t)$. Other representations for the nonlinearity are also possible [7]. The

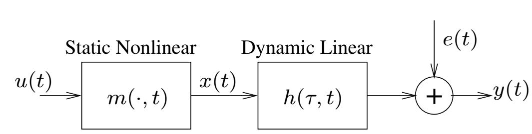

Fig. 1. Time-varying Hammerstein model structure

1-4244-0033-3/06/$20.00 ©2006 IEEE.

intermediate signal, $x(t)$, is thus given by

$$x(t) = m(u(t), t) = \sum_{q=0}^{Q} c(q,t) u^q(t) \qquad (1)$$

where Q is the polynomial order and $c(q,t),\ q = 1, 2, \cdots, Q$ are time-dependent polynomial coefficients. The output of the Hammerstein model is the sum of the convolution of the intermediate signal with the impulse response functions (IRFs) of the dynamic block, and the additive noise:

$$y(t) = \sum_{\tau=0}^{\infty} h(\tau, t) x(t-\tau) + e(t) \qquad (2)$$

where $h(\tau, t)$ is the TV IRFs. In practice, the IRF is truncated to a finite memory length T chosen to cover the range where the magnitude of the IRF is significant. Thus, the estimated output for the model with finite memory length T is

$$\begin{aligned} \hat{y}(t) &= \sum_{\tau=0}^{T-1} h(\tau, t) x(t-\tau) \\ &= \sum_{\tau=0}^{T-1} \sum_{q=0}^{Q} c(q,t) h(\tau, t) u^q(t-\tau) \end{aligned} \qquad (3)$$

The first step in implementing the algorithm, is to transform the TV parameters, $c(q,t)$ and $h(\tau,t)$ in (3), into TIV coefficients so that a standard optimization algorithm can be applied to the problem. This transformation is accomplished by projecting the TV parameters onto temporal expansion bases:

$$h(\tau, t) = \sum_{j=0}^{M_j} \alpha(\tau, j) \pi_j(t) \qquad (4)$$

$$c(q, t) = \sum_{i=0}^{M_i} \beta(q, i) \xi_i(t) \qquad (5)$$

where $\alpha(\tau, j)$ and $\beta(q, i)$ are the TIV expansion coefficients. Substituting (4) and (5) into (3) yields

$$\hat{y}(t) = \sum_{\tau=0}^{T-1} \sum_{j=0}^{M_j} \alpha(\tau, j) \sum_{q=0}^{Q} \sum_{i=0}^{M_i} \beta(q, i) u_{ij}^q(t-\tau) \qquad (6)$$

where $u_{ij}^q(t-\tau) = \pi_j(t)\xi_i(t)u^q(t-\tau)$. Clearly, (6) gives an expression that depends only on the TIV coefficients, $\alpha(\tau, j)$ and $\beta(q, i)$. Thus the identification problem reduces to estimating these expansion coefficients. The basis functions, $\pi_j(t)$ and $\xi_i(t)$, must be chosen *a priori* to represent the parameter variations. Some applicable basis functions include Walsh functions, Tchebyshev polynomials, wavelets and many more. All bases have different abilities for tracking a system's dynamics and so their choice depends on the TV characteristics of the system being modeled. Without *a priori* knowledge, they will have to be selected by trial and error.

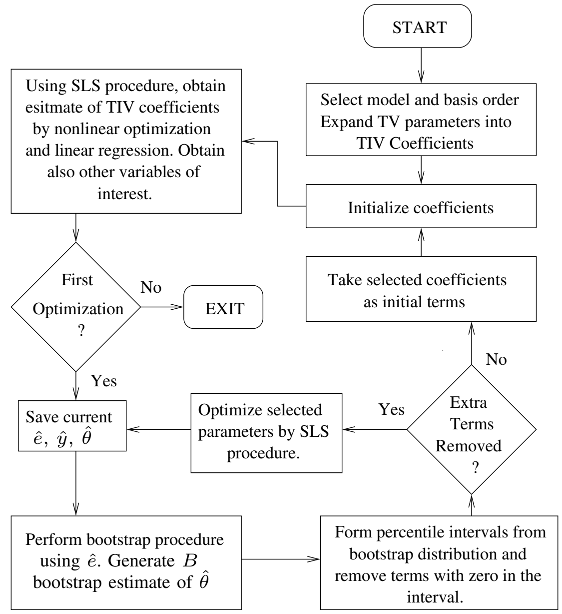

Fig. 2. Flowchart of the optimization algorithm

A. Parameter Estimation

The parameter vector, $\boldsymbol{\theta}$, is constructed from the TIV coefficients $\alpha(\tau, j)$ and $\beta(q, i)$. The model prediction error is $\epsilon(t, \boldsymbol{\theta}) = y(t) - \hat{y}(t, \boldsymbol{\theta})$, so that by defining the loss function, $V_N(\boldsymbol{\theta})$, as

$$V_N(\boldsymbol{\theta}) = \frac{1}{2N} \sum_{t=1}^{N} \epsilon^2(t, \boldsymbol{\theta}) \qquad (7)$$

the optimal parameter estimate is therefore the solution to

$$\hat{\boldsymbol{\theta}} = \arg \min_{\theta \in D_m} V_N(\boldsymbol{\theta}) \qquad (8)$$

where D_m is an open set which corresponds to useful models. Note that while the coefficients in (6) are TIV, the output is a nonlinear function of those TIV coefficients. As a result, (6) cannot be solved in closed form, so an iterative optimization will be required. We used the separable least squares optimization technique described in [4] to minimize the cost function in (7). The TV parameters, $h(\tau, t)$ and $c(q, t)$, were computed using (4) and (5) once optimal values had been obtained for $\alpha(\tau, j)$ and $\beta(q, i)$.

III. Bootstrap

Expanding the TV parameters onto temporal bases results in a large number of coefficients when compared to the size and complexity of the original model. Generally, several of the basis functions used in the expansion increase the parameter uncertainty without improving the model accuracy, so there is need to eliminate those insignificant terms from the model representation. This term selection problem is common in basis expansion approaches, however existing applications of TV basis expansions have used linear regression models [1], so term selection methods based on linear regression could be employed. Since the Hammerstein model

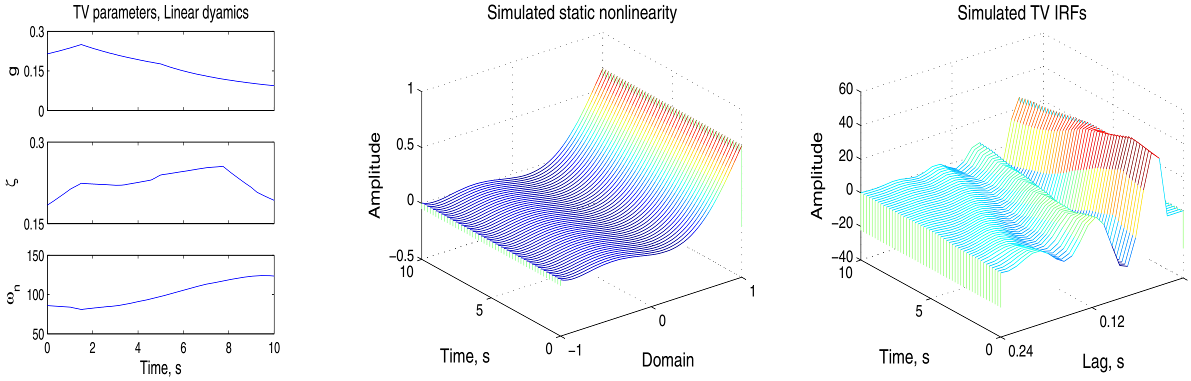

Fig. 3. Simulation description for the TV Hammerstein system. Left: Values of the static gain (g), damping factor (ζ) and natural frequency (ω_n). Center: Normalized TV static nonlinearity for the range of inputs. Right: Normalized TV impulse response functions for the linear dynamics

is nonlinear in its parameters, a different approach to term selection will be required.

The bootstrap technique has been well studied in the literature and has been shown to be a robust method for parameter and model selection even in applications that are ill-suited to traditional methods [6], [8]. It is a simple computer-based technique for estimating the statistical distribution of any variable of interest. The bootstrap is appealing due to its ease of implementation and minimal statistical requirements.

In our study, we use the bootstrap technique to estimate the probability density functions (PDFs) of the estimated TIV coefficients. Those coefficients whose values cannot be statistically distinguished from zero are taken to be insignificant and removed from the model.

A. Implementation Details

Bootstrap based techniques are Monte-Carlo repetitions of a simulation without a repeat of the experimental process since new noise realizations cannot be generated. The general bootstrap procedure is to assign an equal probability to the elements of the estimated residual, $\hat{\boldsymbol{\epsilon}} = [\hat{\epsilon}_1, \hat{\epsilon}_2, \ldots, \hat{\epsilon}_N]$, and then to sample with replacement from among the mean-detrended elements, $\hat{\boldsymbol{\epsilon}}_c^* = \hat{\boldsymbol{\epsilon}} - N^{-1}\sum_{i=1}^{N}\hat{\epsilon}_i$, to generate pseudo-noise sequence, $\hat{\boldsymbol{\epsilon}}^* = [\hat{\epsilon}_1^*, \hat{\epsilon}_2^*, \ldots, \hat{\epsilon}_N^*]$, called the bootstrap replicates, having same length as the original residual for the Monte-Carlo simulation. Note that by sampling with replacement, elements of the original data may be duplicated more than once while some elements may not appear at all in the bootstrap copy of the residual. Bootstrap estimates of the output are then generated as the sum of the estimated output, $\hat{\boldsymbol{y}}$, and the bootstrapped errors, $\hat{\boldsymbol{\epsilon}}^*$, $\boldsymbol{y}^* = \hat{\boldsymbol{y}} + \hat{\boldsymbol{\epsilon}}^*$. A number, B, of bootstrap estimates of the coefficients are obtained by repeating this procedure.

Let λ be a chosen confidence level, $0 < \lambda < 0.5$, then the lower and upper percentile intervals for a coefficient are given by its $B\cdot\lambda$ and $B\cdot(1-\lambda)$ smallest values, respectively. If zero lies between the lower and upper confidence limits, the coefficient is taken to be statistically indistinguishable from zero, and is removed from the model.

An iteration of the optimization procedure and bootstrap term selection is carried out until the bootstrap does not eliminate any further terms, after which the final optimal values are obtained with the selected terms.

B. Implementation Procedure

A detailed procedure for the proposed algorithm is given below. The procedure is elucidated in Fig. 2.

1) First optimization:
   a) Choose model order, polynomial nonlinearity order and basis functions.
   b) Obtain TIV coefficients from TV parameters using basis expansion.
   c) Initialize nonlinear TIV coefficients.
   d) Estimate TIV expansion coefficients using SLS algorithm outlined in [4].
2) Bootstrap procedure: Using optimal TIV $\hat{\boldsymbol{\theta}}$, prediction errors $\hat{\boldsymbol{e}}$ and $\hat{\boldsymbol{y}}$ from first optimization:
   a) Obtain $\boldsymbol{\Theta} = [\hat{\boldsymbol{\theta}}_1^*, \hat{\boldsymbol{\theta}}_2^*, \cdots, \hat{\boldsymbol{\theta}}_B^*]$ bootstrap estimates of $\hat{\boldsymbol{\theta}}$.
   b) Form percentile intervals of $\boldsymbol{\Theta}$. Remove associated element(s) of $\hat{\boldsymbol{\theta}}$ that cannot be distinguished from zero and also associated basis functions.
   c) Remove corresponding terms from the model output (6).
   d) Go to $2(a)$ unless no zero term was removed in the current iteration
3) Final optimization: From selected TIV coefficients and basis functions
   a) Do $1(c)$ and $1(d)$.
   b) Evaluate TV parameters using final optimal TIV coefficients.

IV. Simulation

The technique developed was tested on a TV Hammerstein system simulation of the reflex contribution to joint stiffness [2], [9]. It gives the relationship between angular velocity of the ankle, and the torque arising from reflex mechanisms. The nonlinear component of the Hammerstein system in the reflex pathway is described by a half-wave rectifier, which we approximated using a fifth-order polynomial approximation with parameter values of: $p_0 = -0.11$; $p_1 = -0.45$; $p_2 = 0.79$; $p_3 = 2.59$; $p_4 = 0.34$; $p_5 = -1.14$. The Hammerstein

TABLE I

ESTIMATED VARIABLES OBTAINED FOR DIFFERENT OUTPUT SNR FOR INITIAL NUMBER OF PARAMETERS OF 162

SNR dB	Initial model % VAF	Initial model Avg. NMSE	Bootstrapped model # $\hat{\theta}$	Bootstrapped model % VAF	Bootstrapped model Avg. NMSE	% reduction in # $\hat{\theta}$
5	92.42	3.55	85	93.73	3.49	47.53
10	96.14	1.43	91	97.80	1.02	43.83
20	99.73	0.368	93	99.11	0.978	42.59

system has a one-degree freedom of redundancy which is removed by normalizing the TV parameters for each point in time for all the parameter values in order to obtain unique parameterization. A plot of the normalized static nonlinearity for the expected range of input values is given in Fig. 3. The impulse response describing the linear dynamics of the system was generated from a second-order low-pass system [2]

$$H(s) = \frac{g\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{9}$$

where g is the static gain, ω_n is the natural frequency and ζ is the damping parameter. Their values and IRFs obtained thereof for $0.24s$ are also shown in Fig. 3. Observe the short delay in the IRFs in figure, which accounts for the delay observed in the reflex pathway. The input and output data contains 1000 samples, corresponding to $10s$ of data. The input excitation signal is an independent sequence of uniform distribution between ± 1, while the noise was white Gaussian noise.

Without *a priori* knowledge on the type of variation being represented, the choice of basis functions is selected depending on which functions yield a minimum prediction error under the same estimation conditions. For this simulation, Tchebyshev and Laguerre polynomials were used as expansion bases for the nonlinear and linear block with initial order of 2 and 5 respectively. 100 Monte-Carlo bootstrap simulations were done for the term selection procedure. Model accuracy was assessed using percent variance accounted for $(\%VAF)$ [2] from cross validated data, which judges the goodness of fit of the actual output to the output estimates obtained from validation input signal. The average normalized mean squared error (NMSE) for the parameters is also provided (See Table I).

As the noise level was reduced, considerable improvement in the estimates were obtained with fewer terms removed from the model. Shown in Table I is a summary of results obtained with the proposed method for different noise level. A graphical representation of the result for a $5dB$ SNR is given in Fig. 4.

V. CONCLUSION

A bootstrap technique for the identification of a class of TV nonlinear system with a large number of estimation

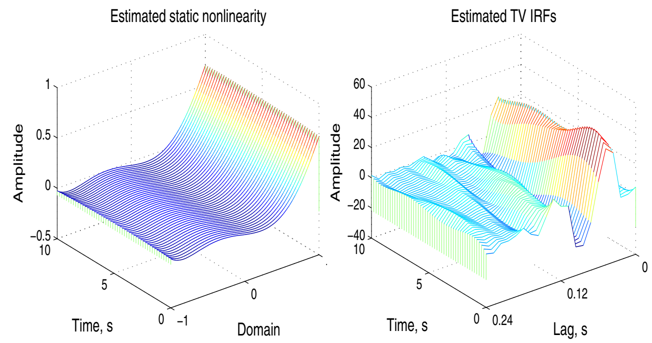

Fig. 4. Normalized model identified with $5dB$ SNR after the application of bootstrap. Left: Time-varying static nonlinearity. Right: Time varying impulse response functions.

parameters and severe noise corruption in the output signal has been developed. The performance of the algorithm was demonstrated using a simulated TV Hammerstein system with FIR linear dynamics. The transfer function (9) used to describe the impulse response of the linear block was used to model the reflex contribution to joint stiffness [9].

The percentage of terms eliminated from the model increases with increased noise level, so for a noise-free or low additive noise to the system, the algorithm gave good results even without term selection (See Table I). The results demonstrate that the proposed method is robust to noise and provide a suitable platform for identifying physiological systems when compared to the ensemble approach. Our interest is then to apply the proposed method to different realizations of experimental data obtained from a single subject and then use each realization to characterize the reflex component arising from changes in muscular activation due to sensory response to stretch. A correlation of the identified torques may provide better insight on reflex contribution to joint stiffness than the one obtained using a combined ensemble of data.

REFERENCES

[1] R. Zou, H. Wang, and K. H. Chon, "A robust time-varying identification algorithm using basis functions," *Annals Biom. Engg.*, vol. 31, pp. 840–853, March 2003.

[2] M. Lortie and R. E. Kearney, "Identification of time-varying Hammerstein systems from ensemble data," *Annals Biom. Engg.*, vol. 29, no. 8, pp. 619–635, August 2001.

[3] S. Sanyal and D. T. Westwick, "Identification of time-varying joint dynamics," *IEEE CCECE Conf.*, pp. 358–361, May 2005.

[4] B. I. Ikharia and D. T. Westwick, "Identitification of time-varying Hammerstein systems using a basis expansion approach," *IEEE CCECE conf.*, pp. 1827–1830, May 2006.

[5] D. T. Westwick and R. E. Kearney, "Separable least squares identification of nonlinear Hammerstein models: Application to stretch reflex dynamics," *Annals Biom. Engg.*, vol. 29, no. 8, pp. 707–718, August 2001.

[6] B. Efron and R. Tibshirani, *An introduction to the bootstrap*. Florida, USA: Chapman and Hall, 1993.

[7] E. J. Dempsey and D. T. Westwick, "Identification of Hammmerstein models with cubic spline nonlinearities," *IEEE Trans. Bio. Eng.*, vol. 51, no. 2, pp. 237–245, February 2004.

[8] A. Zoubir and D. Iskander, *Bootstrap techniques for signal processing*. New York, USA: Cambridge Universtiy Press, 2004.

[9] M. M. Mirgagheri, H. Barbeau, and R. E. Kearney, "Instrinsic and reflex contributions to human ankle stiffness: variation with activation level and position," *Exp. Brain Res.*, vol. 135, pp. 423–436, July 2000.