ADCONIP 2014 Hiroshima

Development of Nonlinear Black Box Models using Orthonormal Basis Filters: A Review*

Sachin C. Patwardhan

***Abstract*— Over the last two decades, there has been a growing interest in the use of orthonormal basis filters (OBF) for developing dynamic black box models. When compared with the conventional approach, a substantial dimensionality reduction can be achieved through OBF parameterization. Moreover, the orthonormal filters, because of their similarity to the Padé approximation, can model systems that exhibit long time delays. Due to these advantages, several authors have recently resorted to OBF based parameterization of block oriented nonlinear black box models. This paper presents a review of nonlinear output error (NOE) and nonlinear ARX (NARX) model development using OBF. To begin with, the linear time series modelling using OBF parameterization is briefly reviewed. The methods available in the literature for the development of models with Wiener, Hammerstein and Wiener-Hammerstein structures are presented next. Features and properties of different model structures are examined in the light of their abilities to model the unmeasured disturbances and to capture complex nonlinear behaviour, such as input and output multiplicities.**

I. INTRODUCTION

With the availability of cheap and fast computing, nonlinear model based controllers are increasingly being used in variety of industrial applications [1]. Development of control relevant dynamic models can be singled out as the most important step in the process of controller synthesis. Over the last two decades, there has been growing interest in the use of orthonormal basis filters (OBF) for representing process dynamics ([3], [4], [5], [6]). The OBFs provide a simple and elegant method of parameterizing stable transfer functions. The conventional approach to parameterization of time series models is using operator $\left\{z^{-k} : k = 0, 1, 2, ..\right\}$ ([2], [3]), which can lead to models with large number of number parameters. The OBF based representation seeks to parameterize transfer functions using filters $F_k(z^{-1})$, which are orthogonal rational polynomials in z^{-1}. Signals filtered through OBF have much longer memory of the past. As a consequence, and due to orthogonality of these filters, a substantial dimensionality reduction can be achieved through OBF based parameterization of dynamic black box models. Moreover, because of their similarity to the Padé approximation, development of OBF based models can be carried out without *a priori* knowledge about system time delays. Also, *a priori* information about system dynamics can be explicitly used in the model development exercise. Heuberger at al. [7] provides an excellent exposure to various theoretical and practical aspects of linear dynamic model development using OBF parameterization.

[1]Sachin C. Patwardhan is with Faculty of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, India. sachinp@iitb.ac.in

For many industrial systems, however, nonlinear model based control schemes have to be employed for achieving uniformly satisfactory performance over the desired operating range. While a mechanistic / grey-box model has better extrapolation ability and portability, development of a nonlinear black box model, directly from perturbed plant data, can be a relatively easy and economically attractive alternative in many situations. Selection of a suitable model structure is a crucial step in the development of such black box models. Sjoberg *et al*. [13] and Pearson [15] provide excellent reviews of the variety of model structures available for black box modeling. From the viewpoint of structures used for modeling unmeasured disturbances, various black-box models employed in the literature can be broadly classified into two classes (a) Nonlinear Output Error (NOE) models and (b) Nonlinear ARX (NARX) Models. NOE models make no attempt to model the effect of unmeasured disturbances. The NARX models, on the other hand, explicitly capture the dynamics of unmeasured disturbances using the past output measurements.

The conventional parameterization of black box models often results in large number of unknown parameters and a large set of data is required to keep the variance errors in check. This translates to the long time required for conducting identification experiments and possible loss of production during this period. Due to the advantages OBF, several authors have recently resorted to OBF based parameterization of NOE and NARX models [14], [16], [17], [18], [20], [21], [24]. The resulting models are parsimonious in parameters. A brief review of nonlinear time series model development using OBF parameterization is presented here. Features and properties of different model structures are examined in the light of their abilities to model the unmeasured disturbances and to capture complex nonlinear behavior, such as input and output multiplicities.

This paper is organized in five sections. To begin with, a brief introduction to generalized orthonormal basis filters is presented. Linear time series modeling using OBF parameterization is then briefly reviewed . The development of nonlinear time-series models using OBF parameterization is presented next.

II. ORTHONORMAL BASIS FILTERS

Following [8], let $\mathcal{T}$ denote the unit circle $\{z : |z| = 1\}$ and $\mathcal{E}$ denote exterior of unit disc $\{z : |z| > 1\}$. We consider the Hardy space $\mathcal{H}_2$, of square integrable functions on $\mathcal{T}$ and analytic in $\mathcal{E}$. The corresponding inner product of $G_1(z), G_2(z) \in \mathcal{H}_2$ is denoted by

$$\langle G_1(z),\, G_2(z)\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} G_1(e^{i\omega})G_2(e^{i\omega})^*\, d\omega \qquad (1)$$

where * denotes the complex conjugate, $\omega = \omega' T$ denotes the normalized frequency and T denotes the sampling interval. Consider a SISO system represented by a strictly proper stable transfer function $G(z) \in \mathcal{H}_2$ and

$$\vartheta(z) = G(z)\, \upsilon(z) \qquad (2)$$

where $\upsilon(z)$ represents input and $\vartheta(z)$ represents the model output. Let $\{F_i(z)\}$ for $i = 1, 2, ...$ be an orthonormal basis for $\mathcal{H}_2$ such that

$$\langle F_i(z),\, F_j(z)\rangle = \left\{ \begin{array}{ll} 1 & (i = j) \\ 0 & (i \neq j) \end{array} \right\} \qquad (3)$$

Then, there exists a unique generalized Fourier series expansion of $G(z)$ such that

$$G(z) = \sum_{i=1}^{\infty} c_i F_i(z) \qquad (4)$$

where $\{c_i\}$ represent Fourier coefficients defined as $c_i = \langle G(z),\, F_i(z)\rangle$. Thus, given a linear time invariant system $G(z)$, the n'th order finite expansion model that approximates $G(z)$ best in a $\mathcal{H}_2$ sense is given by

$$G_n(z) \approx \sum_{i=1}^{n} c_i F_i(z) \qquad (5)$$

Orthogonal filters can be constructed using a single real pole inside the unit circle (*Laguerre filters [3]*) *or* using a pair of complex conjugate poles inside the unit circle (*Kautz filters [4]*). In fact, the *classical impulse response model* can be viewed as a special case of GOBF with $F_j(z) = z^{-j}$. Van den Hof et. al. [5] showed that an orthonormal basis for $\mathcal{H}_2$ can be generated with repeated and fixed set of poles via balanced realization of all pass filters. Ninness and Gustafsson [6] unified the construction of orthonormal basis filters by proving that

$$F_j(z, \boldsymbol{\eta}) = \frac{\sqrt{\left(1 - \left|\eta_j\right|^2\right)}}{\left(z - \eta_j\right)} \prod_{i=1}^{j-1} \frac{(1 - \eta_i^* z)}{(z - \eta_i)} \qquad (6)$$

forms a complete orthogonal set in $\mathcal{H}_2$, where $\boldsymbol{\eta} \equiv \{\eta_i : i = 1, 2, ...\}$ is an arbitrary sequence of poles inside the unit circle appearing in complex conjugate pairs.

The OBF can also be constructed for a $r \times m$ MIMO transfer function matrices $\mathbf{G}(z) \in \mathcal{H}_2^{r\times m}$. Two approaches are available in the literature to define MIMO OBF expansions [7]. A MIMO transfer function, $\mathbf{G}(z)$, can be expressed as

$$\mathbf{G}(z) = \sum_{i=1}^{\infty} \mathbf{C}_i F_i(z)$$

where $\{F_i(z)\}$ represent scalar basis filters and $\mathbf{C}_i$ represent $r \times m$ coefficient matrices. An alternate approach of constructing vector basis functions using a multivariable, square and all pass function is discussed in [12].

While dealing with the parameter identification and control problems, it is often convenient to work with state space realizations of OBF models. For example, consider a SISO model of the form

$$\vartheta(z) = \left[\sum_{i=1}^{n} c_i F_i(z, \boldsymbol{\eta})\right] \upsilon(z) \qquad (7)$$

where $\boldsymbol{\eta}$ consists of only real poles. Defining a state vector $\mathbf{x}(k, \boldsymbol{\eta}) \in R^n$ such that

$$\mathbf{x}(z, \boldsymbol{\eta}) = \left[\begin{array}{ccc} F_1(z, \boldsymbol{\eta})\upsilon(z) & ... & F_n(z, \boldsymbol{\eta})\upsilon(z) \end{array}\right]^T \qquad (8)$$

a state realization of the form

$$\begin{aligned} \mathbf{x}(k+1) &= \Psi(\boldsymbol{\eta})\, \mathbf{x}(k) + \Lambda(\boldsymbol{\eta})\, \upsilon(k) \qquad (9) \\ \vartheta(k) &= \mathbf{C}^T \mathbf{x}(k) \qquad (10) \end{aligned}$$

can be developed for the model given by equation (7) as discussed in [10]. Here, the elements of the $\mathbf{C}$ vector are the Fourier coefficients $\{c_i\}$. A detailed discussion on state realizations for real as well as complex poles can be found in [8]. For a multiple input system, a non-minimal state realization can be created simply by stacking individual SISO state realizations [10]. In the subsequent development, transfer function representation and its state space realization of a GOBF model are used interchangeably.

III. LINEAR BLACK BOX MODELS

In this section, linear black-box model development using OBF parameterization is briefly discussed. The models are represented using time shift operator q and time domain signals. Consider a SISO linear, time invariant discrete time system modelled as

$$y(k) = G(q)u(k) + H(q)e(k) \qquad (11)$$

where $G(q)$ is a strictly proper stable transfer function, $H(q)$ is a stable rational monic transfer function and $e(k)$ represents a zero mean Gaussian white noise sequence. Identification of $G(q)$ and $H(q)$ can be carried out either sequentially or simultaneously [10], [11].

In the sequential approach, initially output error (OE) model structure assumed, i.e.

$$y(k) = G(q)u(k) + v(k) = \sum_{i=1}^{n_u} c_i F_{u,i}(q, \boldsymbol{\eta}_u)u(k) + v(k)$$

and an estimate $\widehat{G}(z)$ of $G(z)$ is generated using OBF parameterization. In the next step, a filter that whitens the estimated residuals

$$\widehat{v}(k) = y(k) - \widehat{G}(q)u(k) = \widehat{H}(q)\, e(k) \qquad (12)$$

is identified by rearranging the model as an AR model follows

$$\widehat{v}(k) = \widehat{W_v}(q)\, \widehat{v}(k) +\, e(k) = \sum_{i=1}^{n_v} c_i F_{v,i}(q, \boldsymbol{\eta}_v)\widehat{v}(k) + e(k)$$

and parameterizing $\widehat{W_v}(q) = I - \widehat{H}(q)^{-1}$ using OBF [10].

Alternatively, for the purpose of parameter estimation, this model is rearranged in one step predictor form as follows

$$\begin{aligned} \hat{y}(k|k-1) &= H(q)^{-1}G(q)u(k) + (1 - H(q)^{-1})y(k) \\ &= W_u(q)u(k) + W_y(q)y(k) \qquad (13) \end{aligned}$$

which is similar to an ARX model. The R.H.S. of ARX model (13) is parameterized using OBF as follows [7], [11]

$$y(k)=\sum_{j=1}^{n_u}F_{u,j}(q,\boldsymbol{\eta}_u)u(k)+\sum_{j=1}^{n_y}F_{y,j}(q,\boldsymbol{\eta}_y)y(k)+e(k)$$

where $\{e(k)\}$ represents a zero mean white noise sequence. This model is referred to as OBF-ARX model. A state realization of the OBF-ARX model can be constructed as follows

$$\begin{aligned}\mathbf{x}(k+1) &= \boldsymbol{\Psi}\mathbf{x}(k)+\boldsymbol{\Gamma}u(k)+\mathbf{L}y(k) &(14)\\ y(k) &= \mathbf{C}\mathbf{x}(k)+e(k) &(15)\end{aligned}$$

where

$$\mathbf{x}(k)=\begin{bmatrix}\mathbf{x}_u(k)\\ \mathbf{x}_y(k)\end{bmatrix}\ ;\ \Psi=\begin{bmatrix}\Psi(\boldsymbol{\eta}_u) & [0]\\ [0] & \Psi(\boldsymbol{\eta}_y)\end{bmatrix} \tag{16}$$

$$\boldsymbol{\Gamma}=\begin{bmatrix}\boldsymbol{\Gamma}_u(\boldsymbol{\eta}_u)\\ [0]\end{bmatrix}\ ;\ \ \mathbf{L}=\begin{bmatrix}[0]\\ \boldsymbol{\Gamma}_y(\boldsymbol{\eta}_y)\end{bmatrix} \tag{17}$$

It is convenient to use model form (14-15) while carrying out parameter identification. After estimating the model parameters, the model (14-15) can be rearranged in standard innovation form of state space model as follows

$$\mathbf{x}(k+1)=\boldsymbol{\Phi}\mathbf{x}(k)+\boldsymbol{\Gamma}u(k)+\mathbf{L}e(k) \tag{18}$$

where $\boldsymbol{\Phi}=\boldsymbol{\Psi}+\mathbf{LC}$ and used to recover $G(q)=\mathbf{C}\left[q\mathbf{I}-\boldsymbol{\Phi}\right]^{-1}\boldsymbol{\Gamma}$ and $H(q)=\mathbf{C}\left[q\mathbf{I}-\boldsymbol{\Phi}\right]^{-1}\mathbf{L}+1$.

A key step in the development of the OBF models is the selection of filter poles and number of basis filters or *truncation order(s)*, (n_u,n_y), necessary to develop a reasonably good approximation of the system dynamics. While, an orthonormal basis for $\mathcal{H}_2$ can be constructed by selecting poles in an arbitrary manner, such a basis may lead to a large truncation order and a high dimensional model. Selecting a basis with poles that closely match the dominant poles of the system to be approximated significantly reduces the truncation order [3], [4], [9]. In many practical situations there is some *a priori* knowledge about the dominant system time constants, which can be used to reduce the truncation order. A significant advantage of choosing filter poles based on the *a priori* knowledge is that the resulting parameter estimation can be solved analytically. For example, given pole vectors $(\boldsymbol{\eta}_u,\boldsymbol{\eta}_y)$ of a OBF-ARX model, state sequence $\{\mathbf{x}(k):k=1,2...N_s\}$ can be generated using (14) and the Fourier coefficients $(\mathbf{C})$ can be estimated by minimizing the prediction error, i.e.

$$\arg\min_{\mathbf{C}}\frac{1}{N_s}\sum_{k=1}^{N_s}e(k)^2=\arg\min_{\mathbf{C}}\frac{1}{N_s}\sum_{k=1}^{N_s}\left(y(k)-\mathbf{C}\mathbf{x}(k)\right)^2$$

The optimal solution can be computed analytically as [9], [10]

$$\widehat{\mathbf{C}}\left(\boldsymbol{\eta}_u,\boldsymbol{\eta}_y\right)=\left(\overline{E}\left[\mathbf{x}(k)\mathbf{x}(k)^T\right]\right)^{-1}\overline{E}\left[\mathbf{x}(k)y(k)\right] \tag{19}$$

where $\overline{E}\left[.\right]$ represents the expectation operator. When *a priori* knowledge about the dominant system time constants is not available, the parameter estimation problem can be formulated as two nested optimization problems and optimal $(\boldsymbol{\eta}_u,\boldsymbol{\eta}_y)$ can be estimated simultaneously with the Fourier coefficients [10], [11].

It is straight forward to extend the sequential and the simultaneous modeling approaches discussed in this section to identification of MISO models using non-minimal state realizations of the MISO models [10], [11]. Two major concerns in system identification are bias and variance errors committed in a modeling exercise. A detailed discussion on this issue can be found in [9].

IV. NONLINEAR TIME SERIES MODELS

For the sake of brevity, development of SISO nonlinear time series models is discussed in this section. Extension to the MISO case is straight forward and will not be discussed separately. Let $\mathbf{Y}^k$ and $\mathbf{U}^k$ denote data sets

$$\begin{aligned}\mathbf{Y}^k &= \{y(0),y(1),...,y(k)\}\\ \mathbf{U}^k &= \{u(0),u(1),...,u(k)\}\end{aligned}$$

A nonlinear black box model can be represented as follows

$$\begin{aligned}y(k) &= \mathcal{F}\left[\boldsymbol{\varphi}(k),\boldsymbol{\theta}\right]+e(k) &(20)\\ \boldsymbol{\varphi}(k) &= \boldsymbol{\varphi}\left[\mathbf{U}^{k-1},\mathbf{Y}^{k-1},\boldsymbol{\eta}\right] &(21)\end{aligned}$$

where $\boldsymbol{\varphi}$ represents *regressor vector,* $\mathcal{F}[.]$ represents a *nonlinear mapping* from the regressor space to the output space and $e(k)$ represents the model residual [13]. Vectors $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$ represent parameters of the *regressor function,* $\boldsymbol{\varphi}[.]$, and the *nonlinear mapping,* $\mathcal{F}[.]$, respectively. Thus, the problem of nonlinear black box model development can be decomposed into (a) the regressor selection and (b) selection of the nonlinear mapping. The nonlinear mapping can further be expressed as a parameterized function family of the form

$$\mathcal{F}\left[\boldsymbol{\varphi},\boldsymbol{\theta}\right]=\sum_j\alpha_j\ f_j(\boldsymbol{\varphi},\boldsymbol{\beta})$$

where $f_j(\boldsymbol{\varphi})$ represent *basis functions* and $\boldsymbol{\theta}\equiv(\boldsymbol{\alpha},\boldsymbol{\beta})$. A variety of basis functions such as polynomials, splines, sigmoidal neural nets, wavelets, radial basis network etc. can be used to construct the nonlinear mapping. Sjoberg et al. [13] provide a comprehensive and unifying review of the basis functions used in nonlinear black box modeling.

The main difference between the nonlinear black box models parameterized using GOBF and the conventional black box models is the choice of the regressor vector. The basis functions used for constructing the nonlinear mapping, however, are not different from the conventional choices for the function space basis discussed in [13].

A. Regressor Construction

By the conventional approach, the regressor vectors are chosen as

$$\boldsymbol{\varphi}_u(k)=\left[\begin{array}{ccc}u(k-1) & ... & u(k-n)\end{array}\right]^T \tag{22}$$

$$\boldsymbol{\varphi}_y(k)=\left[\begin{array}{ccc}y(k-1) & ... & y(k-n)\end{array}\right]^T \tag{23}$$

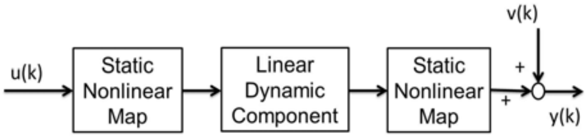

Fig. 1. Schematic Representation of Feedforward Block Oriented (Wiener Hammerstein) Model

i.e. to have a finite memory of n past samples [13]. In the GOBF based modeling, on the other hand, the regressors are typically chosen as

$$\boldsymbol{\varphi}_u(k) \equiv \left[\begin{array}{ccc} F_{u,1}(q,\boldsymbol{\eta}_u)u(k) & ... & F_{u,n}(q,\boldsymbol{\eta}_u)u(k) \end{array} \right]^T$$

$$\boldsymbol{\varphi}_y(k) \equiv \left[\begin{array}{ccc} F_{y,1}(q,\boldsymbol{\eta}_y)y(k) & ... & F_{y,n}(q,\boldsymbol{\eta}_y)y(k) \end{array} \right]^T$$

where $\boldsymbol{\eta}_u$ and $\boldsymbol{\eta}_y$ represent vectors of GOBF poles. Thus, given the identification data set ($\mathbf{Y}^N$, $\mathbf{U}^N$) and OBF poles, the regressor vectors can be constructed using state space realizations of the form (9). For example, $\boldsymbol{\varphi}_u(k)$ $(\equiv \mathbf{x}_u(k))$ can be generated recursively as follows

$$\mathbf{x}_u(k+1) = \boldsymbol{\Phi}(\boldsymbol{\eta}_u)\mathbf{x}_u(k) + \boldsymbol{\Gamma}_u(\boldsymbol{\eta}_u)u(k) \tag{24}$$

Since $F_{u,i}(q,\boldsymbol{\eta}_u) \in \mathcal{H}_2$, the regressor (state) sequence can be generated under the assumption $\mathbf{x}_u(0) = \overline{\mathbf{0}}$. Alternatively, while constructing block oriented nonlinear black-box models, regressors can also be constructed as follows

$$\mathbf{x}_u(k+1) = \boldsymbol{\Phi}(\boldsymbol{\eta}_u)\mathbf{x}_u(k) + \boldsymbol{\Gamma}_u(\boldsymbol{\eta}_u)g\left[u(k)\right] \tag{25}$$

where $g[.] : R \to R$ represent some preselected smooth nonlinear algebraic function. Moreover, if $g\left[u(k)\right]$ is a continuous function and $\{u(k)\}$ is restricted to a compact set, then it can be shown that resulting state sequence $\{\mathbf{x}_u(k)\}$ is also a bounded sequence. It may be noted that the resulting regressors have a memory over a growing time window in the past. A direct consequence of this feature is reduction in the dimensionality of the model.

B. Forward Block Oriented NOE Models

A general nonlinear output error (NOE) models can be represented as follows

$$\mathbf{y}(k) = \mathcal{F}\left[\boldsymbol{\varphi}(k), \boldsymbol{\theta}\right] + \mathbf{v}(k) \tag{26}$$

where the regressor vector, $\boldsymbol{\varphi}(k) = \boldsymbol{\varphi}\left[\mathbf{U}^{k-1}, \boldsymbol{\eta}_u\right]$, is as a function of past manipulated / known inputs alone and $\mathbf{v}(k)$ represents the model residual term. Forward block oriented nonlinear models (see Figure 1), such as Hammerstein, Wiener or Hammerstein - Wiener models, and Volterra series models are sub-classes of NOE models that have been extensively studied in the literature [14], [15]. In fact, majority of OBF based NOE models that have been proposed in the literature have forward block oriented structures. These OBF-NOE models can be represented by a general Wiener-Hammerstein form as follows

$$\begin{aligned} \mathbf{x}_u(k+1) &= \boldsymbol{\Phi}(\boldsymbol{\eta}_u)\mathbf{x}_u(k) + \boldsymbol{\Gamma}_u(\boldsymbol{\eta}_u)h\left[u(k), \boldsymbol{\beta}\right] && (27) \\ y(k) &= \mathcal{F}\left[\mathbf{x}_u(k), \boldsymbol{\theta}\right] + \mathbf{v}(k) && (28) \end{aligned}$$

where $\boldsymbol{\beta}$ represents parameter vector for the static map $h\left[.\right]$. Choosing $h\left[u(k)\right] = u(k)$ reduces it to the Wiener structure while letting $\mathcal{F}\left[.\right]$ to be a linear map reduces it to the Hammerstein structure. The OBF-NOE models proposed in the literature essentially differ in their choice of basis functions used for representing the static nonlinear blocks.

The conventional finite discrete Volterra series models can be represented as follows [15]

$$\mathcal{F}\left[\varphi_u(k)\right] = \alpha_o + \sum_{j=1}^{n} \alpha_1(j)u(k-j) + ...$$

$$+ \sum_{j_1=1}^{n} .. \sum_{j_M=1}^{n} \alpha_M(j_i, .., j_M)u(k-j_1)..u(k-j_M)$$

The finite Volterra series models result from the application of the Weierstrass approximation theorem to equation (26) under the assumption that (a) $\mathcal{F}[.]$ is a continuous function of regressor defined by equation (22) and (b) input $u(k)$ is restricted to a compact set for all k [15]. Boyd and Chua [22] have proved an important extension of Weierstrass approximation theorem to *fading memory systems*, i.e. systems that have weak dependence on the input signals from the remote past. They established that, if $\mathcal{F}\left[.\right]$ is a fading memory operator, then it can be approximated arbitrarily well by a finite Volterra model with some n and sufficiently large M. In fact, the Volterra series models can be viewed as nonlinear generalization of FIR models, which have been widely used for development of linear MPC schemes. However, these models are non parsimonious in parameters and, as a consequence, require a large data set to keep the variance errors low. This is an important issue, especially when they are used to model multi-variable systems. Choosing regressors parameterized using GOBF can alleviate this difficulty. Boyd and Chua [22] have also shown that a Wiener Laguerre model, constructed using Laguerre filters to parameterize (27), can approximate any time invariant causal operator with fading memory.

Several authors have subsequently developed and employed models that belong to this sub-class, i.e. OBF-Wiener models. Doyle at al. [14] have proposed idea of representing Volterra model coefficients in terms of discrete Laguerre functions, which can be viewed as the impulse responses of a family of Laguerre filters. Dumont et al. [16] develop SISO Laguerre-Wiener models with state to output map modelled as a quadratic polynomial

$$\mathcal{F}\left[\mathbf{x}_u(k)\right] = \mathbf{C}\mathbf{x}_u(k) + \mathbf{x}_u(k)^T\mathbf{D}\mathbf{x}_u(k)$$

and use it to formulate nonlinear generalized predictive control. Srinivasrao et al. [17] have developed a similar model for MISO systems and shown that it can be adopted easily to deal with irregularly sampled multi-rate measurement scenario. Deshpande et al. [29] have exploited structure of these models to develop closed form control laws and used it for controlling an experimental proton exchange membrane fuel cell (PEMFC) setup. Senotni et al. [18] initially parameterize the linear dynamics using OBF while the state to output map is constructed using a single hidden layer neural network.

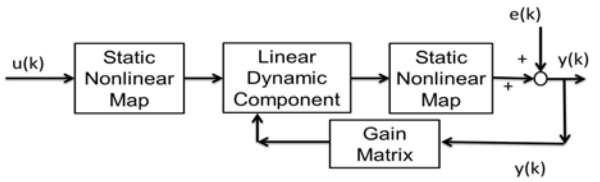

Fig. 2. Schematic Representation of Block Oriented Model with Feedback

To reduce the dimensionality of the inputs to the neural network, model order reduction is carried out using balanced truncation on the hidden nodes of the neural network. This approach was used in a commercial product (Aspen Target of Aspen Technology) [1], [19]. Kumar and Patwardhan [27] have proposed to develop a MIMO multi-model using OBF parameterization. Multiple local MIMO linear OBF-OE models, identified at different representative operating points, are combined using a neural network to arrive at a OBF-Wiener model.

Hammerstein and Wiener-Hammerstein models are other sub-classes of block oriented models that have been parameterized using OBF. Gomez and Baeyens [13] have proposed a multivariable Hammerstein model where linear dynamics is parameterized using MIMO GOBF. Thus, a MIMO Hammerstein modes of the form

$$\mathbf{y}(k) = \sum_{i=1}^{n_u} \mathbf{C}_i F_{u,i}(q) \left(\sum_{j=1}^{M} \mathbf{A}_j H_j \left[\mathbf{u}(k) \right] \right) + v(k) \qquad (29)$$

is proposed where $\mathbf{y} \in \mathbf{R}^r$, $\mathbf{u} \in \mathbf{R}^m$, $\mathbf{C}_i$ represents $(r \times m)$ matrices of Fourier coefficients, $\mathbf{A}_j$ represents $(m \times m)$ matrix parameters of the static nonlinear map and $H_j\left[.\right] : \mathbf{R}^m \rightarrow \mathbf{R}^m$ are chosen nonlinear basis functions. Dasgupta and Patwardhan [23] have developed OBF based Hammerstein and Wiener-Hammerstein model using irregularly sampled multi-rate data. Recently, McArthur [21] has proposed Hammerstein and Wiener-Hammerstein models parameterized using OBF. The nonlinear static maps in the block oriented models are constructed using ordinal splines. In particular, normalized cubic splines are used in modeling. A detailed discussion on selection of *instruments*, i.e. linear / nonlinear combinations of internal / intermediate variables, for model construction is provided. A high point of this work is successful application of Hammerstein and Wiener-Hammerstein modeling to an industrial air separation unit.

The main advantage of the OBF-NOE models is that they ensures good prediction capability over a large prediction horizon, which is of vital importance in any model predictive control formulation. In addition, under the assumption that the static nonlinear functions are smooth and $u(k)$ input is restricted to a compact set, these models can be shown to be BIBO stable [15], [14], [17].

C. Block Oriented Models with Feedback

The NARX models can be represented in generic form given by equation (20). This form explicitly captures the effect of unmeasured disturbances by including past the measurements in the regressor vector. When compared with the forward block orientated models with NOE structure, development of NARX type models using OBF parameterization has not attracted much attention. Srinivasrao et al. [24] and Dasgupta and Patwardhan [28] have proposed block oriented models with output feedback as shown in Figure 2. The proposed model can be expressed as follows

$$\begin{aligned} \mathbf{x}(k+1) &= \Psi(\boldsymbol{\eta}_u, \boldsymbol{\eta}_y)\mathbf{x}(k) + \mathbf{\Gamma}(\boldsymbol{\eta}_u) h\left[u(k)\right] + \mathbf{L}(\boldsymbol{\eta}_y) y(k) \\ y(k) &= \mathcal{F}\left[\mathbf{x}(k), \boldsymbol{\theta}\right] + e(k) \end{aligned} \qquad (30)$$

where matrices $\Psi(\boldsymbol{\eta}_u, \boldsymbol{\eta}_y)$, $\mathbf{\Gamma}(\boldsymbol{\eta}_u)$ and $\mathbf{L}(\boldsymbol{\eta}_y)$ are defined by equations (16-17) and innovations $\{e(k)\}$ represent a zero mean white noise sequence. The model represents a form of nonlinear observer. Combining the output map with the state dynamics, the model can be expressed in a *nonlinear innovation form* as follows

$$\begin{aligned} \mathbf{x}(k+1) &= \Psi(\boldsymbol{\eta}_u, \boldsymbol{\eta}_y)\mathbf{x}(k) + \mathbf{L}(\boldsymbol{\eta}_y)\mathcal{F}\left[\mathbf{x}(k), \boldsymbol{\theta}\right] \\ &\quad + \mathbf{\Gamma}(\boldsymbol{\eta}_u) h\left[u(k)\right] + \mathbf{L}(\boldsymbol{\eta}_y) e(k) \end{aligned} \qquad (31)$$

which reveals that the states are functions of the past innovations. The nonlinear innovation form has been used to develop and implement a nonlinear MPC formulation [25]. Similar to the linear case, modeling of unmeasured disturbances can also be carried out in a sequential manner. Srinivasrao et al. [17] model residuals generated from NOE model as nonlinear ARMA (or NARMA) process.

D. Model Parameter Estimation

Model parameters for OBF-NOE and well as OBF-NARX structure are estimated using the prediction error method. Thus, given a choice of OBF poles, the model parameter problem is posed as

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{N_s} \sum_{k=1}^{N_s} \left(y(k) - \mathcal{F}\left[\mathbf{x}(k), \boldsymbol{\theta}\right] \right)^2$$

For certain choices of mapping $\mathcal{F}\left[.\right]$, parameter estimation problem may have analytical solution [17], [20], [21], [24]. However, in most cases, the optimization problem has to be solved using a suitable numerical optimization approach. Choosing the GOBF poles can be a difficult task particularly for NARX structure. In such cases, a nested parameter estimation problem can be formulated as follows [24], [28], [17]

$$\left(\widehat{\boldsymbol{\eta}}_u, \widehat{\boldsymbol{\eta}}_y, \widehat{\boldsymbol{\theta}} \right) = \arg\min_{\boldsymbol{\eta}_u, \boldsymbol{\eta}_y} \left[\arg\min_{\boldsymbol{\theta}} \frac{1}{N_s} \sum_{k=1}^{N_s} e(k, \boldsymbol{\eta}_u, \boldsymbol{\eta}_y, \boldsymbol{\theta})^2 \right]$$

subject to $\left|\eta_{u,i}\right| < 1 \quad for\, i = 1, ...l_u$ and $\left|\eta_{y,i}\right| < 1 \quad for\, i = 1, ...l_y$.

E. Modeling Multiplicity Behavior

Many nonlinear dynamic systems exhibit two types of multiplicities viz. *input multiplicity* and *output multiplicity*. Input multiplicity is a phenomena in which different values of input variables produce identical value(s) of output variable(s) in the steady state solution, while output multiplicity is a phenomena in which different steady state solutions exist for the same value of inputs. A detailed discussion on various nonlinear black box model structures and their abilities to model input and output multiplicities can be found in [15]. If input is held constant at $u(k) = \overline{u}$, then, from (27), it follows that there is a unique steady state, $\overline{\mathbf{x}}_u = [\mathbf{I} - \mathbf{\Phi}]^{-1} \mathbf{\Gamma}_u h [\overline{u}]$, corresponding to $\overline{u}$. Thus, OBF-NOE models of the form (27-28) can only be used to capture input multiplicity behavior. Occurrence of an extremum and change in sign of the steady state gain around the extremum are often observed in systems exhibiting input multiplicity, which makes the task of controlling these systems extremely difficult. The OBF based nonlinear black-box models have been successfully used for extremum seeking control of system exhibiting input multiplicities such as wood-chip refiner motor [16], stirred tank reactor [25], continuous fermenter [17] and PEM fuel cell [26]. The OBF-NARX structure is capable of capturing input and well as output multiplicities. This follows from the fact that, at $u(k) = \overline{u}$ and $e(k) = 0$, equation (31) reduces to $[\mathbf{I} - \mathbf{\Phi}] \overline{\mathbf{x}} = \mathbf{L}\mathcal{F}[\overline{\mathbf{x}}] + \mathbf{\Gamma} h [\overline{u}]$, which can accommodate multiplicity of steady states ($\overline{\mathbf{x}}$) provided $\mathcal{F}[.]$ is chosen carefully. Srinivasrao et al [24] have shown that OBF-NARX model can be used to capture output multiplicity behavior in a CSTR system and control it in a highly nonlinear region.

V. CONCLUSIONS

While nonlinear black box models can facilitate nonlinear controller synthesis, the conventional NOE and NARX models often have large number of model parameters. This difficulty can be alleviated by constructing regressors using orthonormal basis filters. The use of OBF parameterization can significantly reduce the time required for identification tests and thereby accelerate the process of developing a nonlinear model based control scheme. Moreover, these models are capable of capturing input as well as output multiplicities exhibited by many nonlinear systems. A variety of OBF based forward block oriented nonlinear black box models with NOE structure have been proposed in the literature. While this modeling approach appears quite promising, its full potential still remains to been explored particularly for modeling of unmeasured disturbances (i.e. developing NARX models or nonlinear observers) and for multi-rate sampled data systems.

REFERENCES

[1] S. J. Qin and T. A. Badgwell, A survey of industrial model predictive control technology, *Control Engineering Practice*, 11, pp 733-764, 2003.

[2] L. Ljung, System Identification: Theory for the users. Prentice-Hall: New Jersey, 1999.

[3] B. Wahlberg, System identification using Laguerre filters, *IEEE Transactions on Automatic Control*, **36**(5), pp 551-562, 1991.

[4] B. Wahlberg, System identification using Kautz filters, *IEEE Transactions on Automatic Control*, **39**(6), 1276-1282, 1994.

[5] P. M. J. Van den Hof, P. S. C. Heuberger and J. Bokor**,** System identification with generalized orthonormal basis functions, *Automatica*, 31, 12, pp 1821-1834, 1995.

[6] B. M. Ninness and F. Gustafson, A Unifying Construction of Orthonormal Basis for System Identification. *IEEE Transactions on Automatic Control*, **42**(4), pp 515-521, 1997.

[7] P. S. C. Heuberger, P. M. J. Van den Hof and B. Walhberg (Eds.), *Modelling and Identification with Rational Orthogonal Basis Functions.* Springer-Verlag:London, 2005**.**

[8] B. Walhberg and T. O. Silva, Construction and analysis. In *Modelling and Identification with Rational Orthogonal Basis Functions.* P. S. C. Heuberger et al. (Eds.) Springer-Verlag:London, pp 15-39, 2005**.**

[9] P. M. J. Van den Hof and B. Ninness, System identification with generalized orthonormal basis functions. In *Modelling and Identification with Rational Orthogonal Basis Functions.* P. S. C. Heuberger et al. (Eds.) Springer-Verlag:London, pp 61-102, 2005**.**

[10] S. C. Patwardhan and S. L. Shah, From data to diagnosis and control using generalized orthonormal basis filters. Part I: Development of state observers, *Journal of Process control*,15,7, 819-835, 2005.

[11] M. Muddu, A. Narang, S. C. Patwardhan, Development of ARX models for predictive control using fractional order and orthonormal basis filter parameterization, *Ind. Eng. Chem. Res.*, 48, pp 8966–8979, 2009.

[12] P. S. C. Heuberger, Design Issues, In *Modelling and Identification with Rational Orthogonal Basis Functions*, P. S. C. Heuberger et al. (Eds.), Springer-Verlag:London, pp 269-296, 2005**.**

[13] J. Sjoberg, Q. Zhang, L. Ljung, A. Benveniste, B. Doyle, P. Glorennec, H. Hajmarsson and A. Judisky, Nonlinear black box modeling in system identification a unified approach, *Automatica*, 31(12), pp 1691-1724, 1995.

[14] F. Doyle, R.K. Pearson and B.A.Ogunnaike. Identification and control using Volterra models, Springler- Verlag London Ltd. Great Britain, 2002.

[15] R. K. Pearson, Selecting nonlinear model structures for computer control, *Journal of Process Control*, 13, 1-26, 2003.

[16] G. A. Dumont, Y. Fu, and G. Lu, Nonlinear adaptive and generalized predictive control and its applications. *In Advances in Model-Based Predictive Control*, ed. D. Clarke, 498 515, Oxford University Press, New York, 1994.

[17] M. Srinivasarao, S. C. Patwardhan, and R. D. Gudi, Nonlinear predictive control of irregularly sampled multirate systems using blackbox observers. *Journal of Process Control*, 17, 17–35, 2007.

[18] G. B. Sentoni, J. P. Guiver, H. Zhao and L. T. Biegler, A state space approach to nonlinear process modeling: Identification and universality. *AIChE J.,* **44, 10**, 2229, 1998.

[19] H. Zhao, J. Guiver, R. Neelakantan, L. T. Biegler (2001), A nonlinear industrial model predictive controller using integrated PLS and neural net state-space model. Control Engineering Practice, 9, 125-133.

[20] J. C. Gomez and E. Baeyens, Identification of block-oriented nonlinear systems using orthonormal bases" *Journal of Process Control*, 14, 685-697, 2004.

[21] W. McArthur, A new approach for nonlinear process identification using orthonormal bases and ordinal splines. *Journal of Process Control*, 22, 375– 389, 2012.

[22] Boyd, S. and L. Chua, Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Transactions on Circuits and Systems*, **CAS-32**, 11, 1150-1161, 1985.

[23] D. Dasgupta and S. C. Patwardhan, NMPC of a Continuous Fermenter Using Wiener-Hammerstein Model Developed from Irregularly Sampled Multi-Rate Data , *Proc. of IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS-2010)* , Leuven, Belgium, 5-7 July, 2010.

[24] M. Srinivasrao, S.C. Patwardhan and R. D. Gudi, From data to nonlinear predictive control. 1. Identification of multivariable nonlinear state observers. *Ind. Eng. Chem. Res.*, 45, 1989-2001, 2006.

[25] M. Srinivasrao, S. C. Patwardhan and R. D. Gudi (2005). From data to nonlinear predictive control. 2. Improving regulatory performance using identified observers. *Ind. Eng. Chem. Res.*, 45, 3593-3603.

[26] R. N. Methekar, S. C. Patwardhan, R. D. Gudi and V. Prasad, Adaptive peak seeking control of a proton exchange membrane fuel cell. *Journal of Process Control*, 20, 73–82, 2010.

[27] K. Kumar and S. C. Patwardhan, Nonlinear Predictive Control of Systems Exhibiting Input Multiplicities Using the Multimodel Approach, *Ind. Eng. Chem. Res.*, 41, 3186-3198, 2002.

[28] D. Dasgupta and S. C. Patwardhan, Development of block oriented nonlinear time-series models for a continuous fermenter, *Proc. of The Fifth International Symposium on Design, Operation and Control of Chemical Processes (PSE-ASIA 2010)*, Singapore, 25-28 July, 2010.

[29] S. Deshpande,† S. C. Patwardhan, R. Methekar and R. Rengaswamy, Development of a Closed Form Nonlinear Predictive Control Law Based on a Class of Wiener Models, *Ind. Eng. Chem. Res.*, 49, 148–165, 2010.