Journal XXX manuscript No.
(will be inserted by the editor)

Heuristic Order Reduction of NARX-OBF models Applied to Nonlinear System Identification

Elder Oroski, Beatriz Pês, Adolfo Bauchspiess, and Marco Egito Coelho

Received: date / Accepted: date

Abstract Nonlinear system identification concerns the determination of the model structure and its parameters. Although the designers often seek the best model for each system, it can be tricky to determine, at the same time, the best structure and the parameters which optimize the model performance. This paper proposes the use of a Genetic Algorithm, GA, and the Levenberg-Marquardt, LM, method to obtain the model parameters, as well as perform the order reduction of the model. In order to validate the proposed methodology, the identification of a magnetic levitator, operating in closed loop, was performed. The class NARX-OBF, Nonlinear Auto Regressive with eXogenous input-Orthonormal Basis Function, was used. The use of OBF functions aims to reduce the number of terms in NARX models. Once the model is found, the order reduction is performed using GA and LM, in an hybrid application, capable of determining the model parameters and reducing the original model order, simultaneously. The results show, considering the inherent trade-off between accuracy and computational effort, the proposed methodology provided an implementation with good mean square error, when compared with the full NARX-OBF model.

Elder Oroski
Federal University of Technology – Paraná (UTFPR), Avenida Sete de Setembro 3165, 80230-901, Curitiba-PR, Brazil. E-mail: oroski@utfpr.edu.br

Beatriz do Santos Pês
Federal Institute of Paraná (IFPR),
Rua Engenheiro Tourinho, 829, , Campo Largo-PR, Brazil.
E-mail: beatriz.santos@ifpr.edu.br

Adolfo Bauchspiess · Marco Egito Coelho
University of Brasília (UnB), Campus Darcy Ribeiro, Asa Norte, Brasília-DF, Brazil. E-mail: adolfobs e egito@ene.unb.br

Keywords NARX-OBF Models · Genetic Algorithm · Levenberg Marquardt · System Identification.

1 Introduction

A system can be defined as a structure in which different variables interact and produce observable signals. In the same way, a model expresses the relation between observable quantities. Therefore, allowing the prediction of properties and behavior of an object [1].

In control engineering, modeling systems is a constant need. The model can provide a better understanding of the system operation, it can also be a powerful prediction tool which may prevent faults in the real system [2].

System Identification can be regarded as the field responsible for relating a purely mathematical model to a system. That is, in the development of the model, only data concerning the system input and output are needed [3]. That approach is known as black box modeling and it is very appealing, since no simplifying assumptions are requested when building the model [4].

There are several classes of models that can be used in the development of a black box model. The designer must consider representation capability and computational effort when choosing a class. That is not an easy choice and is often empirically made. As examples of classes of nonlinear models frequently mentioned in literature, one can cite Volterra Series [5] and NARX [6], Nonlinear Auto Regressive with eXogenous input. These classes may present a dimensionality problem, since the number of terms to be determined in the model is often high [3].

The dimensionality problem is less critical in NARX models, since past outputs are considered in the model, which reduces the number of terms required by the model. Additionally, it is possible to reduce the computational cost of NARX models even more, selecting the most relevant series' terms [7] (in other words, performing the order reduction of the model).

Order reduction techniques are specially interesting for the representation of highly nonlinear systems, since the number of terms in a polynomial model increases with the nonlinearity degree of the system [3]. NARX models are known for the high representation capability. In order to avoid the lost of this characteristic, the order reduction must be carefully parametrized [8].

It is worth mentioning that the use of OBF, Orthonormal Basis Functions, is widely known in literature. ARX-OBF [9], which are Infinite Impulse Response, IIR, models and Volterra-OBF [8,10,11], which are Finite Impulse Response, FIR, models have already been proposed. NARX-OBF models [11], which can be seen as a feedback version of the Volterra-OBF models, have also been treated in previous work.

It can be considered that a NARX-OBF model is a more compact implementation of the classical NARX models. That is, the use of OBF reduces the number of terms necessary to the model. In order to improve the performance of the NARX-OBF models, this paper proposes an order reduction methodology for this class of models.

A Genetic Algorithm, GA, is used to select the most representative terms of the full NARX-OBF model. Thus, simplifying the model realization and reducing the simulation time. The evaluation function applied in the GA is inspired by the Akaike Information Criterion [12], AIC. This criterion quantifies the impact in the Mean Square Error (MSE) caused by the insertion of a new term in the model. At the same time, the AIC penalizes the insertion of this new term, since it increases the model complexity.

With a fitness function which considers the AIC, the GA proposed in this work is capable of realize the joint minimization of the MSE and of the number of terms in the model. That is, the GA performs a multi-objective optimization, aiming the simplest model with the best representation.

It is important to mention that the use of GA in the order reduction of NARX polynomial models has already been investigated [13]. However, in this paper it is proposed the use of GA as a mechanism of search for the poles of the Kautz functions, in NARX-OBF models, and also to reduce the number of terms in the series that implement such a model. Furthermore, in the proposed methodology, the GA acts by interleaving its actions with the Levenberg-Marquardt method, which is the latter responsible for determining the coefficients of the NARX-OBF model.

The main contributions of this work can be defined as:

- the joint minimization of: (i) the MSE and; (ii) the complexity of the NARX-OBF models, (i.e., the joint search for the model structure and its parameters);
- the interleaving application of the GA method (searching for Kautz poles and the model structure), and the Levenberg Marquardt method (which search for models' coefficients).

This paper is divided as follows: section 2 presents the structure of NARX-OBF models, section 3 gives a summarized description of the methods for parameter selection, section 4 presents the identification of the magnetic levitation system, section 5 presents the main results obtained and, finally, section 6 is dedicated to conclusions and future work.

2 NARX-OBF Models

This section is dedicated to present basic concepts regarding orthonormal functions and NARX-OBF models [11].

2.1 Orthonormal Basis Functions

The main property of orthonormal functions is expressed by Eq. (1).

$$\langle \psi_m(k), \psi_n(k) \rangle = \begin{cases} 0 & m \neq n, \\ 1 & m = n, \end{cases} \tag{1}$$

in which k, m, $n \in \mathbb{Z}$, and $\psi_m(k)$ and $\psi_n(k)$ are orthonormal functions and $\langle . \rangle$ is an inner product, defined by Eq. (2).

$$\langle \psi_m(k), \psi_n(k) \rangle = \sum_{k=-\infty}^{\infty} \psi_m(k)\psi_n^*(k), \tag{2}$$

in which $\psi_n^*(k)$ represents the complex conjugate of $\psi_n(k)$. To be classified as orthonormal, a function must satisfy the following requirements:

- $\langle \psi_m, \psi_n \rangle = 0, \quad$ for $m \neq n$;
- $|\psi_n| = 1, \quad \forall\ n$;

in which $|\psi_n| = \sqrt{\langle \psi_n, \psi_n \rangle}$.

As examples of orthonormal functions, one might mention Hermite, Jacobi, Laguerre, Legendre, Kautz and the Generalized Orthonormal Basis Functions, GOBF. In this work, Kautz functions are employed. Therefore, next subsections present a summarized description of these functions.

2.2 Kautz Functions

Kautz functions are orthonormal functions parametrized by complex poles [8,9]. Equations (3) and (4) present the general form of Kautz functions:

$$K_{2m}(z) = \frac{\sqrt{(1-\tau^2)(1-\sigma^2)}}{z^2+\sigma(\tau-1)z-\tau}\left[\frac{-\tau z^2+\sigma(\tau-1)z+1}{z^2+\sigma(\tau-1)z-\tau}\right]^{m-1}, \quad (3)$$

$$K_{2m-1}(z) = \frac{z(z-\sigma)\sqrt{1-\tau^2}}{z^2+\sigma(\tau-1)z-\tau}\left[\frac{-\tau z^2+\sigma(\tau-1)z+1}{z^2+\sigma(\tau-1)z-\tau}\right]^{m-1}, \quad (4)$$

in which $m \in \mathbb{N}$, z stands for the complex variable associated with the Z transform, $K_{2m}(z)$ and $K_{2m-1}(z)$ are the even and odd Kautz functions, respectively. Considering that β and $\overline{\beta}$ are the complex conjugate poles that parametrize these functions, τ and σ can be expressed by:

$$\sigma = (\beta + \overline{\beta})/(1 + \beta\overline{\beta}), \quad (5)$$

$$\tau = -\beta\overline{\beta}. \quad (6)$$

The use of orthonormal functions in nonlinear models aims to reduce the number of terms required by the Volterra or NARX models [8,10]. In this context, FIR models described by Kautz basis (as Volterra-OBF models) can be implemented by concatenated filters, as depicted in Fig. 1.

The idea behind of NARX-OBF models came from the Volterra-OBF models. Thus, NARX-OBF can be seen as a feedback version of Volterra-OBF model [11]. Eq. (7) is the mathematical expression of a NARX-OBF model.

$$\hat{y}(k) = M_u(k) + M_y(k) + M_{uy}(k), \quad (7)$$

in which $M_u(k)$, $M_y(k)$ and $M_{uy}(k)$ stands for the input, output and hybrid model components, respectively. These components are expressed by equations (8), (9) and (10).

$$M_u = \sum_{m=1}^{n} c_m^u w_m^u + \sum_{p=1}^{n}\sum_{q=p}^{n} c_{p,q}^u w_p^u w_q^u, \quad (8)$$

$$M_y = \sum_{m=1}^{m} c_n^y w_m^y + \sum_{p=1}^{m}\sum_{q=p}^{m} c_{p,q}^y w_p^y w_q^y, \quad (9)$$

$$M_{uy} = \sum_{p=1}^{m}\sum_{q=1}^{n} c_{p,q}^{uy} w_q^u w_q^y, \quad (10)$$

in which the terms w_i^u are versions of the input, $u(k)$, filtered by a ith order Kautz function, being $i = 1, 2, \cdots, n$. The terms w_j^y are versions of the output, $y(k)$, filtered by a jth order Kautz function, being $j = 1, 2, \cdots, m$. Further, c_i^u (for $i = 1, 2, \cdots, n$), c_j^y (for $j = 1, 2, \cdots, m$) and c_{kl}^{uy} (for $k = 1, 2, \cdots, m$, and $l = 1, 2, \cdots, n$), are the coefficients of input, output and hybrid terms (nonlinear combination between the filtered input and output signals), respectively.

NARX-OBF models can also be expressed in the concatenated filters form, Fig. 2 shows the idea. It is worth mentioning that, in this work, the NARX-OBF model is truncated in the 2^{nd} order kernel.

Once the model is defined, it is necessary to determine its parameters, in order to capture the system's dynamic whose one desire to model. In this scenario, next section is dedicated to detailing the parameter selection in NARX-OBF models.

3 Parameter Selection in NARX-OBF Models

As stated in section 1, the goal of this work is to determine the structure and parameters of a NARX-OBF model for a nonlinear system. To that end, a methodology which combines a Genetic Algorithm, GA, and the Levenberg-Marquardt, LM, method is proposed.

The GA is used to find the model structure. It is worth pointing out that the GA searches for the terms that best represent the system dynamic. Thus, the algorithm finds a simplified structure for the model, aiming to reduce the computational effort. The GA is also used in the search for the pole that parametrizes the orthonormal functions. Further, LM is the method used to find the coefficients of the model.

An appealing advantage of heuristic methods concerns stability. NARX-OBF models are feedback models and, therefore, might be unstable. Instability is a problem for conventional parameter determination methods, which may not be able to solve the estimation problem. Heuristic methods, however, are based on a population of solutions. These solutions are categorized regarding the fitness function. A solution resulting in an unstable model is poorly evaluated. Hence, the natural dynamic of the GA is able to neglect unstable models.

Next sections are dedicated to present the main ideas concerning GAs and LM method.

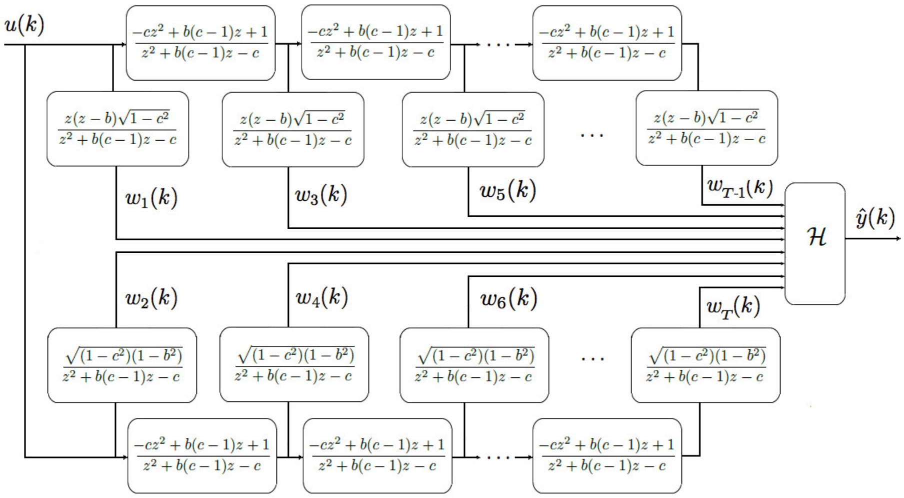

Fig. 1 OBF model with Kautz dynamics [8]. $u(k)$ is the input and $\hat{y}(k)$ is the estimated output of the system to be identified, [11].

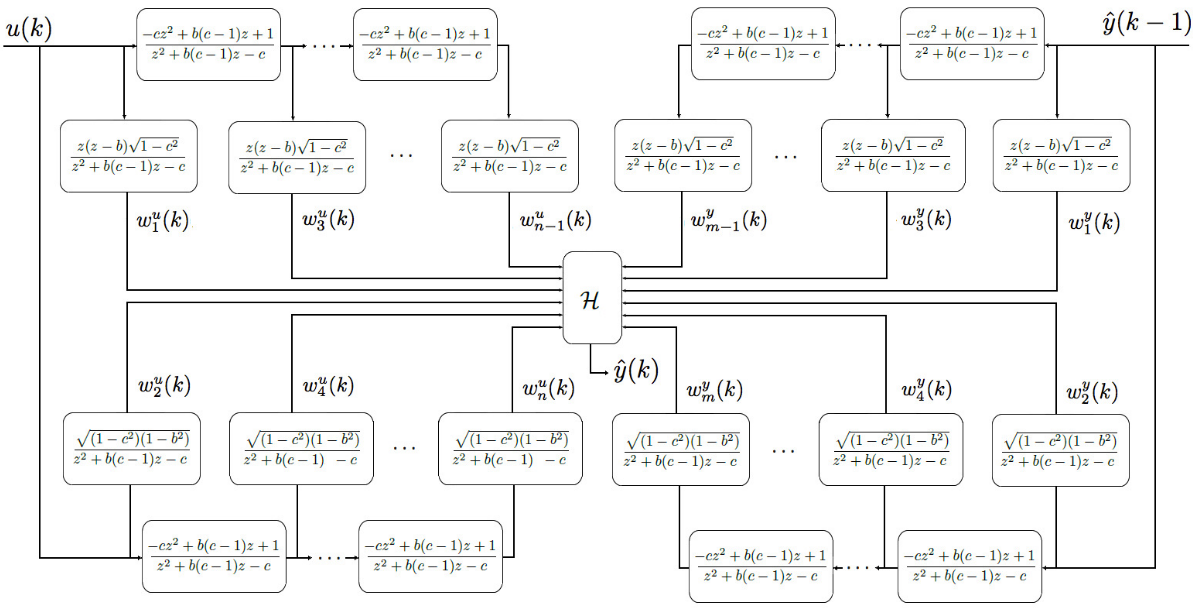

Fig. 2 NARX-OBF model in the form of concatenated filters. Operations expressed by equations (8), (9) and (10) are represented by the $\mathcal{H}$ operator, [11].

3.1 Genetic Algorithms

In the last decades, optimization problems have motivated great improvements in mathematics and engineering. Methods like Newton, steepest descendent and Levenberg-Marquardt have made possible the solution of a series of design optimization problems [14]. However, these methods require strong conditions to have their convergence proved, such as availability of gradients and convexity [15]. In several industrial applications the designer has to deal with some peculiarities such as nonlinearity, non-convexity, existence of several local minima, presence of discrete and continuous design variables, among others [16].

Optimization methods that can potentially circumvent the problems mentioned above are the heuristic algorithms. Some advantages of these algorithms include: (i) these methods do not require gradient information and can be applied to problems in which the gradient is not defined; (ii) these algorithms are not “trapped” in local minima, if correctly tuned; (iii) these methods can be applied to discontinuous functions; (iv) these algorithms provide a set of sub-optimal solutions instead of a single solution.

Among the most popular heuristic algorithms are the Genetic Algorithms, GA [17], the Ant Colony Optimization, ACO [18], and the Particle Swarm Optimization, PSO [19], all inspired by biological principles.

Genetic Algorithm, GA, was developed by John Holland in the 1960s. Inspired by Darwin's evolutionary ideas, Holland has created a method in order to solve optimization problems that dispenses Jacobian or Hessian matrices of the problem [17]. The GA extracts an emergent behavior of convergence. Emerging behaviors involve the application of simple rules, over and over again, which generates complex behaviors [17].

In a GA, each individual is modeled as a set of constants $[c_1, c_2, ..., c_n]$, called genes. These constants form a vector, $\mathfrak{C}$, known as a chromosome.

$$\mathfrak{C} : \boxed{c_1 \,|\, c_2 \,|\, c_3 \,|\, ... \,|\, c_n}$$

such constants are real for the treated problem. GA have a whole population, $\mathcal{P}$, of chromosomes:

$$\mathcal{P} = \{\mathfrak{C}_1, \mathfrak{C}_2, \mathfrak{C}_3, ..., \mathfrak{C}_m\}\,, \tag{11}$$

in which each chromosome is analyzed by an evaluation function, known as a fitness function, $\mathcal{F}_{it}(\mathfrak{C})$: $\mathbb{R}^n \to \mathbb{R}$, and the chromosomes with the best fitness will have greater reproducing likelihood in the next GA generations.

Genetic Algorithm is a highly parallel mathematical algorithm which transforms a population of mathematical objects, with well-defined evaluation function, $\mathcal{F}_{it}$, in a new population of mathematical objects, following the Darwinian principles of reproduction of the most adapted.

According to Algorithm 1, a classical GA creates a random population of solutions, expressed by $\mathcal{P}$. This population consists of possible solutions, $\mathfrak{C}$, to the problem addressed.

The solutions are, then, evaluated by the function $\mathcal{F}_{it}(\mathfrak{C})$. After these step, the best individuals (chromosomes with the lowest image in the evaluation function, $\mathcal{F}_{it}$) are (more likely) selected to be progenitors of the next generation.

The Crossover and Mutation operators are applied to the population and the generated individuals are evaluated. This cycle is repeated until the stop criterion is reached, i.e,. the fitness value of the best chromosome is smaller than a certain threshold: $\mathcal{F}_{it} < Stop_{criterion}$.

It should be emphasized that GAs do not guarantee convergence to the optimum of the problem, and may end up confined to a local region [20]. A point x_l is a local minimum if there is a neighborhood $\mathcal{V}$ (of x_l), such that $f(x_l) \leq f(x)$ for $x \in \mathcal{V}$ [14].

Algorithm 1: Genetic Algorithm

Initializes the population of solutions, $\mathcal{P}$;
Simulates the model generated by each chromosome;
Fitness is calculated for each chromosome, $\mathcal{F}_{it}(\mathfrak{C})$;
while $\mathcal{F}_{it} > Stop_{criterion}$ **do**
  The best individual is saved;
  Selecting parents;
  Apply Crossover Operator;
  Apply Mutation Operator;
  Simulate the model corresponding to each chromosome;
  Evaluate Fitness function for each chromosome;

3.1.1 Crossover

The Crossover operator was inspired by the biochemical process of Crossing-Over, in which parts of two chromosomes are exchanged in the process of sexual reproduction [20].

Fig. 3 shows the Crossover operator action under chromosomes $f = [f_1, f_2, \cdots, f_n]$ and $g = [g_1, g_2, \cdots, g_n]$. This operator performs the local search in the search space [17].

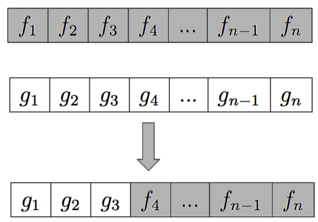

Fig. 3 Example of Crossover operator.

3.1.2 Mutation

The Mutation operator can perform an exploration by inserting a random constant into a random gene position, as expressed in Fig. 4.

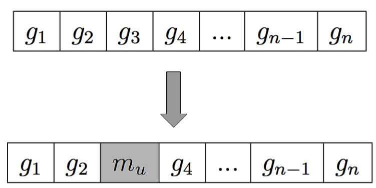

Fig. 4 Example of Mutation operator.

This operator has essentially the global search function, being complemented by the crossover operator, that performs local searches, composing the mechanism of a classical Genetic Algorithm.

3.2 Levenberg-Marquardt

The Levenberg-Marquardt (LM) algorithm is a second-order method, which shares with gradient-based methods the ability to converge rapidly from any starting point, even if this point is outside the region of convergence of other methods [21]. Further, Levenberg-Marquardt is one of the most efficient and most used optimization methods, especially in Systems Identification area [4].

The Levenberg-Marquardt method can be derived by substituting the exact line search strategy (see [15]), for the Confidence Regression strategy (see [14]). The use of the Confidence Regression strategy avoids the main problem of the Gauss Newton method, which occurs when the Jacobian, $J_\xi(x)$, stop being full rank, or near to it [14]. Generally, the Hessian of a generic function, $f(x)$, can be approximated as:

$$\nabla^2 f(x) \approx J_\xi(x)^T J_\xi(x), \tag{12}$$

In order to simplify (12), $\nabla^2 f(x)$ can be expressed by $B(x)$, as (13).

$$B(x) = J_\xi(x)^T J_\xi(x) \approx \nabla^2 f(x). \tag{13}$$

The idea of the LM method is to disturb the matrix $B(x)$, considering $B(x) + \rho I$, for $\rho > 0$. This method can be understood as the Gauss-Newton method with the following modification:

$$\begin{cases} x_{k+1} = x_k + \Delta x_k \\ \Delta x_k = -[J_\xi(x)^T J_\xi(x) + \lambda I]^{-1} [J_\xi(x)^T \xi(x)], \end{cases} \tag{14}$$

in which $\lambda \in \mathbb{R}$ and I is the identity matrix.

It is reasonable to consider the use of hybrid algorithms. Such algorithms behave as LM, for small residues, and apply the Newton method, for larger residues [14].

As the Gauss-Newton methods, LM is based on an expansion into Taylor's series [21]. The search mechanism used by the Levenberg-Marquardt method can be observed in the algorithm 2.

Algorithm 2: Levenberg-Marquardt Algorithm.

Let be $x_0 \in \mathbb{R}^n$;
Calculate d_0, solution of:
$[B(x_0) + \lambda I]\, \Delta x_0 = -\nabla f(x_0)$;
in which: $B(x) = J_\xi(x)^T J_\xi(x)$;
$x_1 = x_0 + \Delta x_0$;
k=1;
while $\nabla f(x_k) \neq 0$ **do**
  Calculate Δx_k, solution of:
  $[B(x_k) + \lambda I]\, \Delta x_k = -\nabla f(x_k)$;
  Determine x_{k+1};
  $x_{k+1} = x_k + \Delta x_k$;
  k=k+1

3.3 Proposed Methodology

NARX-OBF models, in their complete form, have a high number of terms [9,8], therefore, reducing the order of the model is interesting from a computational perspective.

In order to reduce the model order, the GA fitness function take two aspects under consideration: (i) the number of terms, N_C; and (ii) the minimization of the MSE. Eq. (15) expresses the fitness function used in this work, which was inspired by the Akaike criterion [12].

$$f_{it}(MSE, NC) = N \times MSE + 0.1\ N_C \times log(N), \tag{15}$$

in which N is the number of samples in the input and output signals of the system, MSE is the mean square error, N_C stands for the number of coefficients involved in the model and the multiplier 0.1 is an empirical coefficient, used to balance the weight of terms.

The genes used in the GA are depicted in Fig. 5. The real and imaginary parts of the Kautz function pole and the presence of terms coefficients in the NARX-OBF model are genes in the GA chromosome. Thus, the evolutionary dynamics of the GA is responsible for selecting the terms of the model which are representative for the system to be identified, performing an order reduction of the NARX-OBF model.

$Real(\beta)$	$Imag(\beta)$

,

p_1	p_2	$\cdots$	p_n

,

Fig. 5 Structure of genes that compose a chromosome in the proposed GA. β is the pole of the Kautz functions and the genes of the vector $[p_1, p_2, p_3, \cdots, p_n]$ represent the presence or absence of each term of the NARX-OBF model, in its simplified version. If, e.g., p_1 is 0 the first term of the NARX-OBF model is disregarded. If p_2 is 1, the second term of the series is maintained.

In the proposed methodology the idea is interleaving an heuristic algorithm, GA, and a deterministic one, LM algorithm. Mixing these two algorithms, one can achieve the advantages of heuristic algorithms (do not get stuck in local minima) and the advantages of deterministic algorithms (the guaranty of finding the global minimum in a concave region of the search space). In this context, GA is responsible for finding: (i) the orthonormal functions pole; and (ii) the NARX-OBF model structure, whereas the LM is responsible for finding the model coefficients. The loop interaction between the two algorithms is illustrated in Fig 6.

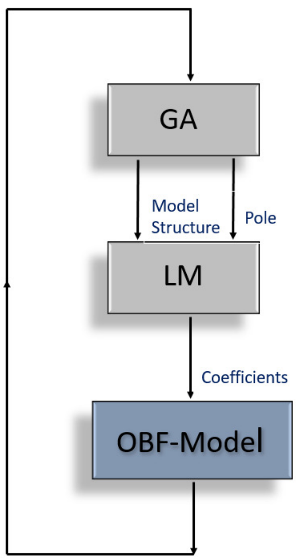

Fig. 6 Search mechanism of OBF-model parameters.

In order to testing this methodology, in the next section, it will be presented the identification process of a non-linear system.

4 Identification of a Magnetic Levitator

In order to validate the proposed method for reducing the order of a NARX-OBF model, a magnetic levitation system was chosen and identified. The system consists of 2 permanent magnets and a mobile magnetic disk. Four coils, operated two at a time, are able to control the disk movement. In this paper, the model is obtained concerning the x-axis. That is, the movements in the axes y and z are neglected [22]. A schematic of the system is depicted in Fig. 7.

The identification of a nonlinear system, such as the one in Fig. 7, starts by the application of a Persistently Exciting, PE, signal to its input. A signal is said to be

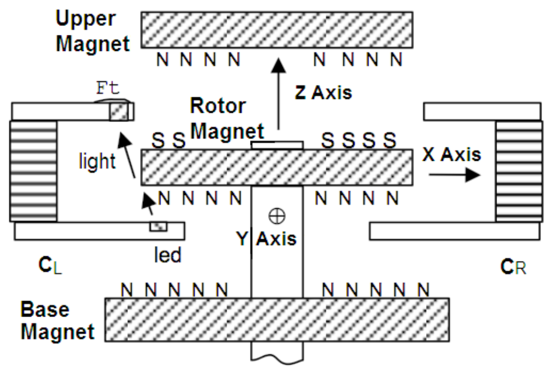

Fig. 7 Magnetic levitator schematic, adapted from [22].

persistently exciting if, considering the need to estimate N_p parameters, it has spectral power in N_p bands of frequency [4].

A widely used kind of PE signal is the Pseudo-Random Multi Level Signal, PRMLS. The variable range of the signal amplitude is desirable in the identification of nonlinear systems, since it provide the excitation of the system several dynamics [3].

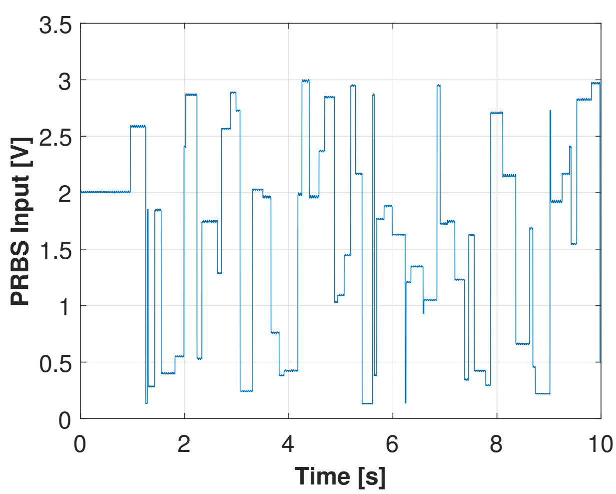

Fig. 8 Input data sampled in the magnetic levitator.

According to Earnshaw's theorem [23], the system in Fig. 7 is unstable. Thus, in order to circumvent the instability problem, the system operates in closed loop under the action of a Proportional Integral, PI, controller.

Fig. 9 depicts the system output response to the input signal (presented in Fig. 8). Both signals, input and output, are composed by 100,000 samples. The sampling period is of 1 ms. Moreover, in the validation of the model, 20,000 samples were used.

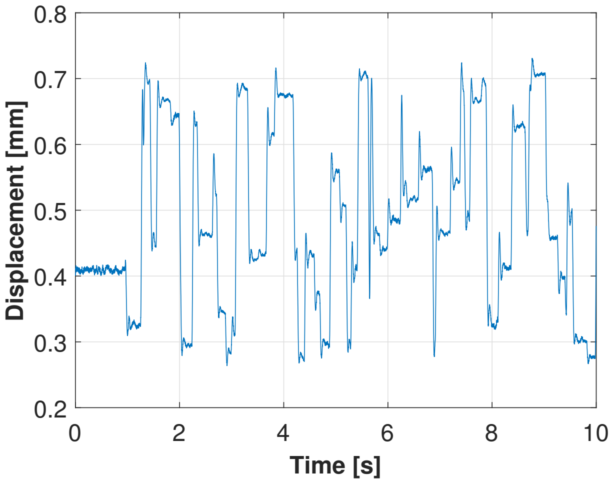

Fig. 9 Output data sampled in the magnetic levitator.

Table 1 Genetic Algorithm Parameters.

Parameter	Value
Population	200 Chromosomes
Selection Method	Tournement
Mutation Rate	varies linearly (5 to 20 %)
Crossover Rate	varies linearly (80 to 40 %)

Table 1 shows the specifications of the GA, which performed the parameter search for the NARX-OBF model. It is worth to emphasize that the metric applied in the identification process aims the minimization of the MSE as well as the reduction of the model complexity, as expressed in Eq. (15).

Genetic algorithms have a stochastic component [17]. Thus, there is no guarantee that the MSE has reached its global minimum. However, if the GA is well tuned, it is possible to find reasonable parameters. Further, in this work, the algorithm applies the GA to search the model structure and the OBF pole, and applies Levenberg-Marquardt method to find the model coefficients, interweaving the methods. It is important to mention that the GA loop keeps going on until the stop criterion is reached, i.e., the MSE of the best solution reach a value smaller than ε. The complete structure of this hybrid GA can be seen in Fig. 10.

Next section is dedicated to present the main results obtained in the identification of the magnetic levitator.

5 Results

After the closed-loop identification of the magnetic levitator, the MSE obtained in each complete NARX-OBF model, with different numbers of orthonormal functions, is shown in Table 2. In this table, N_F stands for the number of orthonormal functions used, and N_C corresponds to the number of coefficients for each NARX-OBF model. The larger N_C, the greater is the number of terms in the model. Therefore, models with larger N_C are more time consuming, computationally speaking.

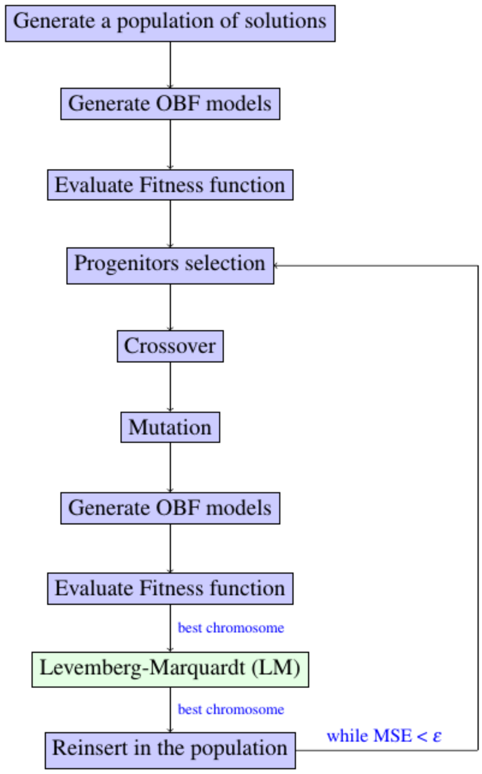

Fig. 10 The proposed Genetic Algorithm structure expressed as a flowchart. ε is a higher limit for the MSE.

Table 2 MSE for Complete NARX-OBF Models.

N_F	N_C	Pole	MSE
2	14	$0.6479 \pm 0.3812i$	1.9416×10^{-3}
4	44	$0.5521 \pm 0.4012i$	9.7212×10^{-4}
6	90	$0.6017 \pm 0.0686i$	5.4947×10^{-6}

The MSE for the NARX-OBF models with order reduction, and their fitness values, can be seen in Table 3.

Table 3 MSE for simplified NARX-OBF models.

N_F	N_C	Pole	MSE	Fitness
2	8	$0.3833 \pm 0.7870i$	1.8370×10^{-3}	40.1808
4	11	$0.4770 \pm 0.5850i$	3.5540×10^{-4}	11.8331
6	54	$0.5487 \pm 0.0289i$	3.5256×10^{-4}	30.2768

Figures 11, 12 and 13 illustrate the time responses of the reduced order NARX-OBF models cited in Table 3.

In order to make a visual comparison, Fig. 14 depicts the performance of the complete NARX-OBF model, mentioned in Table 2.

Comparing tables 2 and 3 one can see that removing some terms of the complete NARX-OBF model does not represent a significant MSE increasing. Furthermore, analysing the first row of tables 2 and 3, it is possible to see that removing 6 terms of the complete model leads to a small reduction in the MSE. This result shows that not all terms of the complete model are in accordance with the system dynamic. Therefore, removing these terms has small impact in the MSE. Thus, some results obtained by order reduction of NARX-OBF models can approximate complete NARX-OBF models without loss of generality.

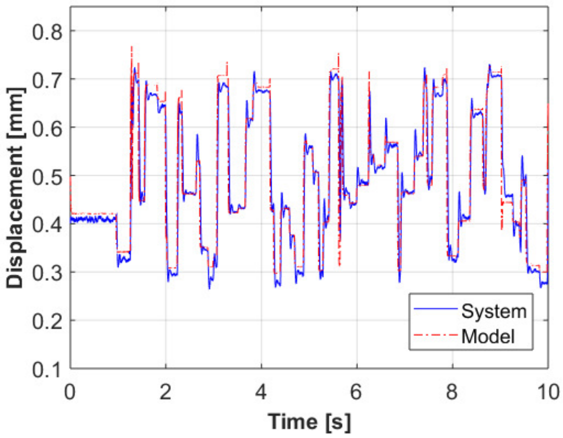

Fig. 11 Identification results for the magnetic levitator, with reduced order NARX-OBF, using 2 Kautz functions and 8 coefficients (terms).

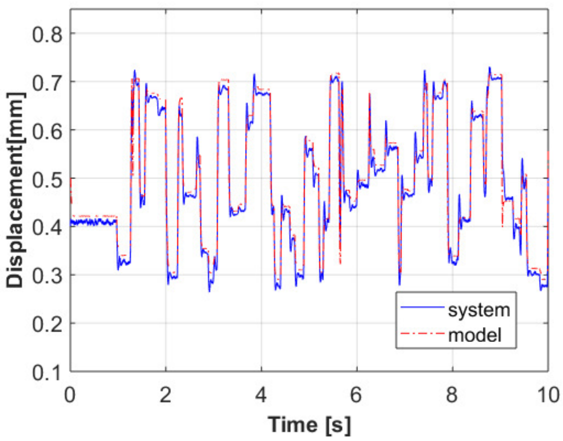

Fig. 12 Identification results for the magnetic levitator, with reduced order NARX-OBF, using 4 Kautz functions and 11 coefficients (terms).

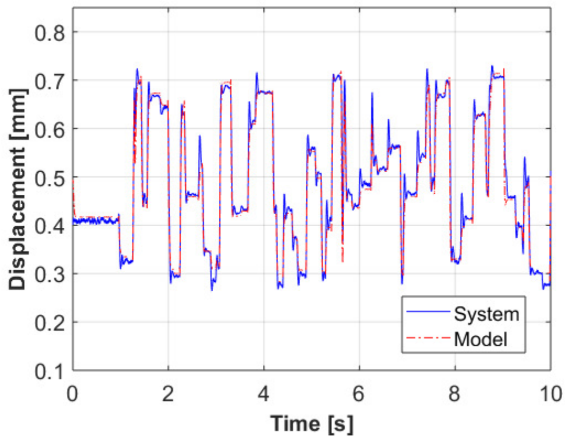

Fig. 13 Identification results for the magnetic levitator, with reduced order NARX-OBF, using 6 Kautz functions and 54 coefficients (terms).

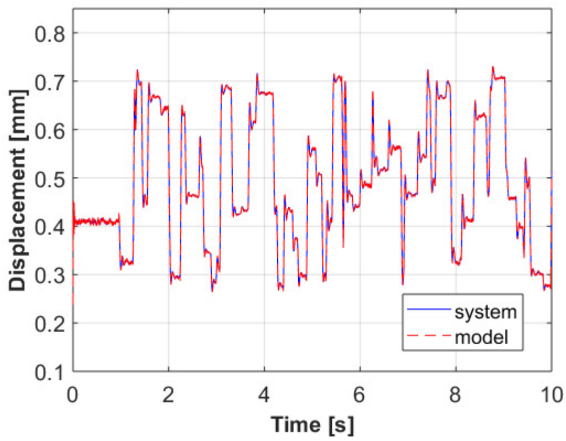

Fig. 14 Identification results for the magnetic levitator, with complete NARX-OBF, using 6 Kautz functions and 90 coefficients (terms).

The time that each simplified model, expressed in Table 3, took to be simulated is shown in Table 4. The simulation time was computed for 10 simulations of each model, the average of the simulation time is expressed in the final row of Table 4. In this table, K indicates the number of Kautz functions and C, the number of coefficients. Thus, the model $2K$ $8C$ stands for the model with 2 Kautz functions and 8 coefficients, which is in the first row of Table 3.

Next section presents the main conclusions of this work.

Table 4 Simulation time for simplified NARX-OBF models.

Time	$2K$ $8C$ [s]	$4K$ $11C$ [s]	$6K$ $54C$ [s]
1	0.1124	0.2796	0.3806
2	0.1039	0.2535	0.3599
3	0.1057	0.2300	0.3521
4	0.1284	0.3708	0.4583
5	0.0905	0.2021	0.3107
6	0.0885	0.1995	0.3089
7	0.8889	0.2032	0.3075
8	0.0903	0.2008	0.3040
9	0.0892	0.1949	0.3074
10	0.0888	0.2027	0.3059
Average Time	0.0906	0.2032	0.3107

6 Conclusion

This paper proposes a method to obtain the order reduction of a NARX-OBF model. A Genetic Algorithm combined with the Levenberg-Marquardt method is used to find the model structure and parameters. The validation of the proposed methodology was based on the closed-loop identification of a magnetic levitator.

The results shown in tables 2 and 3 allow the conclusion that the order reduction is not only possible, but can reduce the model complexity, and has little impact on the MSE value. Thus, one might say that some of the terms which compose the complete NARX-OBF model are irrelevant to modeling the system behavior. Moreover, in Table 3 it is possible to observe that the best fitness value is found in the intermediate situation, between the minimum MSE and the minimum number of terms, N_C. This case portrays the optimization of both criteria at the same time.

It is important to mention that GA is a probabilistic method and there is no guarantee in achieving the best MSE in the identification process. However, in average, it is possible to find reasonable parameters, as shown in tables 2 and 3.

The next steps of this work include the use of Genetic Programming to select the candidates to compose the simplified NARX-OBF model and the use of GOBFs (Generalized Orthonormal Basic Functions) instead of only Kautz functions.

Acknowledgment

The authors gratefully acknowledge the support provided by UTFPR and UnB.

References

1. L. Ljung, System Identification - Theory for the user, Prentice Hall - PTR (1999).
2. K. J. Aström and T. Hgglund, Advanced PID Control, ISA (2011).
3. S. A. Billings, Nonlinear System Identification: NARMAX Methods in the Time, Frequency and Spatio-Temporal Domains, John Wiley & Sons, 1st. ed., 2013.
4. R. Isermann and M. Münchhof, Identification of Dynamic Systems: An Introduction with Applications. Springer-Verlag Berlin Heidelberg, (2011).
5. S. Silva, Nonlinear Mechanical System Identification using Discrete-Time Volterra Models and Kautz Filter, 9th. Brazilian Conference on Dynamics, Control and their Applications, (2010).
6. A. Rahrooh and S. Shepard, Identification of Nonlinear Systems using NARMAX Models, Elsevier - Nonlinear Analisis, n. 71, pp. 1198-1202, (2009).
7. O. Akanyeti, I. Rañó, U. Nehmzow and S. A. Billings, An application of Lyapunov stability analysis to improve the performance of NARMAX models, Robotics and Autonomous Systems, v. 58, pp. 229-238, (2009).
8. P. S. C. Heuberger, P. M. J. Van der Hof and B. Wahlberg, Modelling and Identification with Rational Orthogonal Basis Functions. Springer Press, 1st. Edition, (2005).
9. D. T. Lemma, M. Ramasamy and M. Shuhaimi, Closed-loop Identification of Systems with uncertain Time Delays using ARX-OBF structure, Journal of Process Control, v. 21(8), pp. 1148-1154, (2011).
10. A. da Rosa, R. J. G. Campello and W. C. Amaral, Exact Search Directions for Optimizations of Linear and Nonlinear Models based on Generalized Orthonormal Functions, IEEE Trans. on Automatic Control, v. 54(12), pp. 2757-2772, (2009).
11. E. Oroski, A. Bauchspiess and R. H. Lopes, Identification of a Magnetic Levitator using NARX-OBF Models and Genetic Algorithm, International Journal of Modeling, Identification and Control, v. 28(4), pp. 307-317, (2017).
12. H. Akaike, A new Look at the Statistical Model Identification, IEEE Trans. on Automatic Control, v. 19(6), pp. 716-723, (1974).
13. C. M. Fonseca, E. M. Mendes, P. J. Fleming and S. A. Billings, Nonlinear Model Term selection with Genetic Algorithms, IEEE Workshop on Natural Algorithmsin Signal Processing, v. 2(27), pp. 1-8, (1993).
14. J. Nocedal and S. Wright, Numerical Optimization, Springer, (1999).
15. S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge Press, (2009).
16. J. Mockus, W. Eddy and G. Reklaitis, Bayesian Heuristic Approach to Discrete and Global Optimization: Algorithms, Visualization, Software and Applications, Kluwer Academic Publishers, 1^{st} ed., (1997).
17. J. Koza, Genetic Programming: On the Programming of Computers by Means of Natural Selection, MIT Press, 6^{th} ed, (1998).
18. C. Solnon, Ant Colony Optimization and Constraint Programming, Wiley Press, 1^{st} ed., (2010).
19. A. E. Olssom, Particle Swarm: Theory, Techniques and Applications, New Science Publishers, 1^{st} ed., (2011).
20. A. D. Coley, An Introduction to Genetic Algorithms for Scientists and Engineers, Would Scientific, (1999).
21. D. W. Marquardt, An Algorithm for Least Squares Estimation of Nonlinear Parameters, Journal of Society for Industrial and Appplied Mathematics, v. 11(2), pp. 431-441, (1963).
22. M. A. E. Coelho, Permanent Magnet based Magnetic Levitation Kit, International Journal of Applied Electromagnetics and Mechanics, v. 47, pp. 963-973, (2015).

23. L. Tonks, Note on Earnshaw's Theorem, Electrical Engineering, v. 59(3), pp. 118-119, (1940).