Nonlinear System Identification of Discrete Systems
using GLO-MAP

Monika Marwaha,* John Valasek† and Puneet Singla‡

A Global-Local Mapping Approximation method is presented in this paper for identifying discrete systems using input-output data. The method is based on the idea that any nonlinear system can be represented as a sum of a discrete linear model and unmodeled nonlinearities. Linear system is then perturbed by a nonlinear term which represents the system nonlinearities that are not captured by the linear model. To identify the discrete systems, discrete learning laws are derived using Lyapunov stability analysis. Numerical examples show the successful application of this technique for identification of nonlinear discrete models using data obtained from the simulations.

I. Introduction

The modeling and identification of a nonlinear mathematical model of the system is of great importance in analysis and designing control of the nonlinear systems. Generally experimental data on the input-output behavior of the observable systems is readily available. However, the development of accurate physics-based dynamical model for complex systems using only input-output data is not trivial. In some cases Newton’s second law is used to derive the physics based dynamical model for various systems. Numerous system-identification algorithms1, 2 have been discussed in detail in literature. The theory behind these algorithms deals with linear systems using well established techniques of linear algebra and the theory of ordinary differential or difference equations. A number of system identification methods can be found in References 3–8. Most of these algorithms are based either on time-domain or frequency-domain methods. Observer Kalman filter identification (OKID),9–11 which makes use of the Eigensystem Realization Algorithm (ERA),1,12 is a time domain technique which identifies a discrete input-output mapping from known input-output data.

* Graduate Research Assistant, Aerospace Engineering Department, Texas A&M University, College Station, TX 77843-3141
† Associate Professor and Director, Vehicle Systems & Control Laboratory, Aerospace Engineering Department, Texas A&M University, College Station, TX 77843-3141, Associate Fellow AIAA, valasek@tamu.edu, http://jungfrau.tamu.edu/valasek
‡ Assistant Professor, Department of Mechanical & Aerospace Engineering, University at Buffalo, The State University of New York, Buffalo, NY 14260
records. Since first being developed by Juang in the early 1990’s, the method has been successfully employed to identify linear system models of flexible spacecraft structures\textsuperscript{13} and aircrafts.\textsuperscript{9} Modified Maximum Likelihood Estimation (MMLE)\textsuperscript{14} is also often used for linear and near-linear systems. However, most of the well developed and efficient identification algorithms exist only for linear system identification. While the linear system identification techniques provide the mechanism to identify the physically motivated state-space model to capture the input-output behavior, they might not be able to capture the unknown dynamics accurately.

Due to the fact that systems are linear only in a small region about the operating point, identified linear system holds only for that region and sometimes leads to misrepresentation of the systems accurate behavior. This leads to the necessity of developing nonlinear mathematical models. Many methods have been developed in the past to identify nonlinear dynamics of systems. Recently orthogonal least squares method for basis hunting has been developed to construct sparse radial basis functions (RBF) models for nonlinear systems.\textsuperscript{15} Two-layered neural networks with sigmoid functions are used as activation functions for system identification.\textsuperscript{16–18} In References 19 and 20, a two-layer neural network with RBF has been used for nonlinear aircraft dynamics modeling. Other methods rely upon order reduction to deal with the dimensionality issue.\textsuperscript{21, 22} Reference 23 uses a neural network based method where the theoretical model, assumed to exist, is used to reduce the error between the identified and theoretical model. Authors in Reference 24 use neural networks for system identification of discrete systems. In another study, a gradient based method is used to identify multivariable discrete systems and it is concluded that parameter estimation error converges to zero under persistent excitation.\textsuperscript{25} But as is seen in most of the continuous cases, the selection of weights and number of layers need a lot of tuning in discrete system identification also.

While existing system identification techniques are quite useful, there are no means of including any a-priori knowledge about the unknown system dynamics that one may have. For example, one may have a very good knowledge about the frequency and damping characteristics or some regions may be highly irregular and others may be smooth and linear. Hence, it is desirable to vary the order of approximation depending upon the operating point. In this paper, the recently developed Global-Local Orthogonal MAPping (GLO-MAP) algorithm which provides a mechanism to blend diverse local models into a statistically unbiased global representation\textsuperscript{26} is used to identify discrete systems. Marwaha and Valasek\textsuperscript{27} used this approach to successfully identify the nonlinear model for small acrobatic aircraft using flight data. The results are presented in continuous domain in their study. However there are many areas of interest where discrete models are required for implementation to reduce the computational time. For example, most of the control laws are implemented using the discrete models. This leads to the necessity of the identification of discrete nonlinear models using this approach.

The contribution of this paper is to derive and apply the GLO-MAP approach to discrete domain. GLO-MAP method of system identification has never been used to identify the discrete systems in the past. The
The uniqueness of this method gives the freedom to choose any local basis function and hence can capture the system properties better than the global methods. The objective of this paper is to validate the approach and algorithm using the sampled data for nonlinear systems. The state space based nonlinear discrete system modeled using this approach can further be used in control law formulation or model-based studies.

The paper is organized as follows: The basic idea of the proposed nonlinear system identification algorithm is presented in section II. The introduction to GLO-MAP algorithm is presented in Section III and the derivation of the update laws for GLO-MAP parameters in discrete domain is given in Section IV. Numerical examples to validate the algorithm are given in Section V. Conclusions are presented in Section VI.

II. Nonlinear System Identification

Figure 1 schematically represents the basic idea of system identification. System identification can be defined as a process to identify a mathematical model which when subjected to an actual input vector \( u(k) \), produces an output estimate \( \hat{y}(k) \) which approximates the actual output \( y(k) \) of the system, such that \( \| y(k) - \hat{y}(k) \| < \epsilon \) where \( \epsilon \) represents the desired accuracy in the identification and \( \| . \| \) is the suitable norm on the system output space. In other words, the aim is to find a mathematical model which produces the outputs as close as desired to the true system outputs when the same input is applied to both. The proposed nonlinear system identification algorithm divides the overall procedure into two sequential processes, the process of linear system identification followed by the nonlinear system identification process. The linear system identification captures the best linear representation of the system input-output behavior. Let the
best linear discrete time-invariant model be written as

\[ x_l(k + 1) = A_l x_l(k) + B_l u(k) \] (1)

\[ y(k) = C_l x_l(k) + D_l u(k) \] (2)

Here, \( x_l(k) \in \mathbb{R}^n \) is a hidden state vector corresponding to the best linear approximation of the given input-output data. It should be noted here that the matrices \( A_l, B_l, C_l, D_l \) are not unique but the underlying input-output mapping is unique. Any of the well known methods for system identification such as ERA, \( ^{12} \) OKID, \( ^{10} \) etc can be used to identify the state-space realization of the system \( (A_l, B_l, C_l, D_l) \). Equation (1) is then perturbed by a nonlinear term \( g(x(k)) \) which accounts for the nonlinear departure from the best known linear model where \( x(k) \) is nonlinear state vector.

\[ x(k + 1) = A_l x(k) + B u(k) + g(x(k)) \] (3)

Here, \( g(x(k)) \) represents a vector of unknown nonlinearities that are not captured by the linear model. The problem now reduces to the modeling of vector \( g(x(k)) \). If \( g \) is assumed to be a continuous function in \( x \) then according to Weierstrass approximation theorem, \( g \) can be approximated arbitrarily close by any set of complete functions, including a polynomial series. It can be written as

\[ g(x(k)) = C^T \Phi(x(k)) + \epsilon \] (4)

where, \( \Phi(.) \in \mathbb{R}^N \) is a finite dimensional vector of polynomial functions and \( C \in \mathbb{R}^{n \times N} \) is a matrix of coefficients corresponding to polynomial basis functions. It is also known that as the number of basis functions increases, i.e., \( N \rightarrow \infty \), the approximation error goes to zero, i.e., \( \epsilon \rightarrow 0 \). It is although not possible practically as such a model has infinitely many parameters to optimize. One of the solutions is to use a multiresolution model for learning \( g(x(k)) \) which adequately minimizes the approximation error to the desired level. Multiresolution approximations e.g. conventional spline, \( ^{28} \) piecewise linear approximation \( ^{2} \) and wavelet approximation \( ^{29} \) involve hierarchically decomposing the input-output approximation. It helps in capturing both global and local features of the system dynamics. In Reference 30 learning algorithms are used with radial basis functions to tune the centers of Gaussian functions. Global-local separation of the region of concern leads to immensely improved approximation algorithms but for a specific multi-resolution algorithm one cannot use different basis functions to obtain different local approximations without introducing discontinuity across the boundary of different local regions. \( ^{26} \) For example, in the case of wavelet-based approximation, one is restricted to use the same wavelet function at different resolution levels. Recently, Singla and Junkins \( ^{26} \) have developed a multi-resolution approach known as the Global-Local Orthogonal Mapping (GLO-MAP) algorithm, which provides a mechanism to blend diverse local models into a statistically unbiased global representation. The adaptive nature of this approximation approach can essentially guarantee a small \( \epsilon \), if low noise measurement density in space and time is available. More details about this approach can be found in References 26 and 31.
III. Introduction to GLO-MAP

In this section, we briefly discuss the basic idea of the GLO-MAP-based approximation which allows us to blend locally independent approximations to a piecewise globally continuous function.

For illustration purposes, let us assume that $X = \{x_1, x_2, \ldots, x_N\}$ is a uniform grid with spacing $h$ and $F = \{f_1(x), \ldots, f_N(x)\}$ is a set of continuous functions that approximates the global function $f(x)$ at points $x_i \in X$. We define the weighted average approximation centered on $i$th vertex $x_i$ valid over $x \in [x_i, x_{i+1}]$:

$$\bar{f}_i(x) = w(\bar{x}_i) f_i(x) + w(\bar{x}_{i+1}) f_{i+1}(x), \quad \bar{x}_i = \frac{(x - x_i)}{h}$$

(5)

The weighting function $w(\bar{x}_i)$ is used to blend or average the two adjacent preliminary local approximations $f_i(x)$ and $f_{i+1}(x)$ and the global function is given by the expression:

$$f(x) = \sum_{i=1}^{N} w(\bar{x}_i) f_i(x), \quad \bar{x}_i \in [-1, 1].$$

(6)

Notice that the preliminary local approximations $f_i(x) \in F$ are completely arbitrary, as long as they are smooth and represents the local behavior of global unknown function $f(x)$ well. In References,[31], it is shown that if the weighting functions of Equation (5) satisfies the following boundary value problem (Equation 7), then the weighted average approximation in Equation (5) form an $m^{th}$-order continuous globally valid model with complete freedom in the choice of the local approximations in $F$.

$$w(0) = 1, \quad w(1) = 0,$$

$$\frac{d^k w}{dx^k} |_{x=0} = 0, \quad \frac{d^k w}{dx^k} |_{x=1} = 0, \quad k = 0, \cdots, m$$

$$w(\bar{x}) + w(\bar{x} - 1) = 1, \quad \forall \bar{x}, \quad -1 \leq \bar{x} \leq 1$$

(7)

Notice that the first two aforementioned conditions guarantee the continuity of the blended approximation, whereas the third condition guarantees that the blended approximation is un-biased. Assuming the weighting functions to be polynomial in nature, and adopting the procedure listed in Reference [26], the generic expression for $m^{th}$ order continuity weight function can be written as

$$w(\bar{x}) = 1 - \eta^{m+1} \left\{ \frac{(2m + 1)!(-1)^m}{(ml)^2} \sum_{k=0}^{m} \frac{(-1)^k}{2m - k + 1} \binom{m}{k} \eta^{m-k} \right\}, \quad \bar{x} \in [-1, 1], \quad \eta = |\bar{x}|$$

(8)

Observe that by choosing the weighting functions given by Equation (8), we are guaranteed global piecewise continuity for all possible continuous local approximations in $F$. One retains the freedom to choose the local approximations as needed, to fit the local behavior of $f(x)$, and to rely upon $w(\bar{x})$ to enforce continuity across knot points, $x_i$. In References,[26, 31, 32], these ideas are developed systematically and extended rigorously to approximation with arbitrary order continuity in an $n$–dimensional space. In general, the final approximation in any hypercube is obtained by averaging $2^n$ overlapping approximations centered at the vertices of that local hypercube.
Now making use of the GLO-MAP averaging procedure, the unknown nonlinear term \( g(x) \) can be approximated as

\[
g(x) = \sum_{i=1}^{N_l} w_i(x) g_i(x)
\]  

(9)

where \( N_l \) is the total number of knot points in general \( n \)-dimensional space, \( \mathbb{R}^n \) and \( g_i(.) \) represent \( i^{th} \) the local approximation of the global unknown function \( g(.) \) centered at \( i^{th} \) knot point. Furthermore, we can approximate each \( g_i(.) \) using a set of independent basis functions, \( \phi_j(x) \).

\[
g_i(x) = C_i \phi_j(x), \quad i = 1, 2, \ldots, N_l
\]  

(10)

where, \( \phi_j(.) \) is a \( N \times 1 \) vector of independent basis functions set, and \( C_i \in \mathbb{R}^{n \times N} \) is a matrix of corresponding basis function coefficients. Notice that one can choose a basis function set \( \phi_j \) for each local approximation independently to each other. After substituting Equation (10) in Equation (9), we can re-write the unknown nonlinear term \( g(x) \) as

\[
g(x) = \sum_{i=1}^{N_l} w_i(x) C_i \phi_j(x) = C \Phi(x) W(x) = C \Psi(x)
\]  

(11)

where \( C \in \mathbb{R}^{n \times N_l} \) is a matrix of basis function coefficients for all local approximations. Note that due to the compact support of GLO-MAP weighting functions, only \( 2^n \) neighboring approximations need to be considered, depending upon the value of \( x \).

### IV. Adaptation Laws for GLO-MAP Parameters

Discrete Lyapunov stability analysis\(^{26,33}\) is used to find the adaptation laws for the unknown coefficient matrix \( C \) and the control distribution matrix \( B \). The nonlinear perturbed dynamical system can be written as

\[
x(k+1) = A_1 x(k) + Bu(k) + g(x(k))
\]  

(12)

Substituting the value of \( g(x(k)) \) from Equation 11 we get following equation

\[
x(k+1) = A_1 x(k) + Bu(k) + C^T \Psi(x(k)) + \epsilon
\]  

(13)

As Fourier coefficient matrix \( C \) and control effectiveness matrix \( B \) are unknown, we can write the Equation 13 in terms of estimated parameters.

\[
\hat{x}(k+1) = A_1 \hat{x}(k) + \hat{B}u(k) + \hat{C}^T \Psi(x(k))
\]  

(14)

Assuming the error between the true state and the approximated state as \( e(k+1) = x(k+1) - \hat{x}(k+1) \), the difference equation for error can be written as

\[
e(k+1) = A_1 e(k) + \hat{B}u(k) + \hat{C}^T \Psi(x(k)) + \epsilon
\]  

(15)
where $\hat{B} = B - \tilde{B}$ and $\hat{C} = C - \tilde{C}$.

Following Lyapunov function is selected to derive the adaptation laws:

$$ V(k + 1) = e^T(k + 1)Pe(k + 1) + \frac{1}{2}Tr(\hat{B}(k + 1)\Gamma_1\hat{B}^T(k + 1)) $$

$$ + \frac{1}{2}Tr(\hat{C}(k + 1)\Gamma_2\hat{C}^T(k + 1)) \tag{16} $$

where $P$ is a positive definite symmetric matrix, $e(k)$ represents error between the true state and the estimated state and $\hat{B}$ and $\hat{C}$ are errors between true and estimated matrices respectively. Taking the difference equation of $V$ leads to the following equation

$$ V(k + 1) - V(k) = e^T(k + 1)Pe(k + 1) - e^T(k)Pe(k) + \frac{1}{2}Tr(\hat{B}(k + 1)\Gamma_1\hat{B}^T(k + 1)) - \frac{1}{2}Tr(\hat{B}(k)\Gamma_1\hat{B}^T(k)) $$

$$ + \frac{1}{2}Tr(\hat{C}(k + 1)\Gamma_2\hat{C}^T(k)) - \frac{1}{2}Tr(\hat{C}(k)\Gamma_2\hat{C}^T(k)) \tag{17} $$

Substituting $e(k + 1)$ from Equation 15 results in

$$ V(k + 1) - V(k) = (A_ke(k) + \hat{B}u(k) + \hat{C}^T\Psi(x(k)) + e)P(A_ke(k) + \hat{B}u(k) + \hat{C}^T\Psi(x(k)) + e) $$

$$ - e^T(k)Pe(k) + \frac{1}{2}Tr(\hat{B}(k + 1)\Gamma_1\hat{B}^T(k + 1)) - \frac{1}{2}Tr(\hat{B}(k)\Gamma_1\hat{B}^T(k)) $$

$$ + \frac{1}{2}Tr(\hat{C}(k + 1)\Gamma_2\hat{C}^T(k)) - \frac{1}{2}Tr(\hat{C}(k)\Gamma_2\hat{C}^T(k)) \tag{18} $$

$$ \Delta V(k) = e^T(k)(A^T_kPA_k - P)e(k) + (\hat{B}u(k) + \hat{C}^T\Psi(x(k)) + e)P(e(k + 1)) + e(k)^T A^T_kP(\hat{B}u(k) + \hat{C}^T\Psi(x(k)) + e) $$

$$ + \frac{1}{2}Tr(\hat{B}(k + 1)\Gamma_1\hat{B}^T(k + 1)) - \frac{1}{2}Tr(\hat{B}(k)\Gamma_1\hat{B}^T(k)) + \frac{1}{2}Tr(\hat{C}(k + 1)\Gamma_2\hat{C}^T(k + 1)) - \frac{1}{2}Tr(\hat{C}(k)\Gamma_2\hat{C}^T(k)) \tag{19} $$

Let $Q \in R^{n \times n}$ be a positive definite matrix which satisfies the following discrete Lyapunov equation

$$ A^T_kPA_k - P = -Q \tag{19} $$

$$ \Delta V(k) = -e^T(k)(Q)e(k) + u^T(k)\hat{B}^T(k)Pe(k + 1) + \Psi^T(x(k))\hat{C}(k)Pe(k + 1) + e^T(k)Pe(k + 1) $$

$$ + e(k)^T A^T_kP(\hat{B}u(k) + \hat{C}^T\Psi(x(k)) + e) + \frac{1}{2}Tr((\hat{B}(k))\Gamma_1\hat{B}^T(k + 1)) - \frac{1}{2}Tr(\hat{B}(k)\Gamma_1\hat{B}^T(k)) $$

$$ + \frac{1}{2}Tr(\hat{C}(k + 1)\Gamma_2\hat{C}^T(k + 1)) - \frac{1}{2}Tr(\hat{C}(k)\Gamma_2\hat{C}^T(k)) \tag{20} $$

Using the relation between $C(k + 1)$ and $C(k)$, Equation 20 can be written as

$$ \Delta V(k) = -e^T(k)(Q)e(k) + u^T(k)\hat{B}^T(k)Pe(k + 1) + \Psi^T(x(k))\hat{C}(k)Pe(k + 1) + e^T(k)Pe(k + 1) $$

$$ + e(k)^T A^T_kP(\hat{B}u(k) + \hat{C}^T\Psi(x(k)) + e) + \frac{1}{2}Tr((\hat{B}(k))\Gamma_1\hat{B}^T(k + 1)) - \frac{1}{2}Tr(\hat{B}(k)\Gamma_1\hat{B}^T(k)) $$

$$ + \frac{1}{2}Tr((\hat{C}(k) + \Delta\hat{C}(k))\Gamma_2(\hat{C}(k) + \Delta\hat{C}(k))^T) - \frac{1}{2}Tr(\hat{C}(k)\Gamma_2\hat{C}^T(k)) \tag{21} $$

**American Institute of Aeronautics and Astronautics**
Equation 22 can be written as

\[
\Delta V(k) = -e^T(k)(Q)e(k) + u^T(k)\hat{B}^T(k)Pe(k + 1) + \Psi^T(x(k))\hat{C}(k)Pe(k + 1) + e^TPe(k + 1)
\]

\[+ e(k)^T A^T P(\hat{B}u(k) + \check{C}^T \Psi(x(k)) + e) + Tr(\Delta \hat{B}(k)^T)\Gamma_1(\hat{B}(k)) + Tr(\Delta \hat{C}^T (k))\Gamma_2(\check{C}(k))\]  

Using the following Equations 23 and 24

\[
u(k)^T \hat{B}^T(k)Pe(k + 1) = tr(Pe(k + 1)u^T(k)\hat{B}(k))
\]

\[e(k)^T A^T P\hat{B}u(k) = tr(PA_te(k)u^T(k)\hat{B})\] 

Equation 22 can be written as

\[
\Delta V(k) = -e^T(k)(Q)e(k) + e^TPe(k + 1) + e(k)^T A^T Pe + Tr[(Pe(k + 1)u^T(k) + PA_te(k)u^T(k) + \Delta \hat{B}^T(k)\Gamma_1(\hat{B}(k))] + Tr[(Pe(k + 1)\Psi^T(x(k)) + PA_te(k)\Psi^T(x(k)) + \Delta \hat{C}^T (k)\Gamma_1)\check{C}(k)]
\]

Following adaptive laws are selected

\[
\Delta \hat{B}^T(k) = -P(e(k + 1) + A_te(k))u^T(k)\Gamma^{-1}
\]

\[\Delta \hat{B}(k) = -\Gamma^{-T}u(k)(e(k + 1) + A_te(k))^T P\]

\[\Delta \check{C}(k) = -\Gamma^{-T} \Psi^T(x(k))(e(k + 1) + A_te(k))^T P\]

so that Equation 25 reduces to

\[
\Delta V(k) = -e^T(k)(Q)e(k) + e^TPe(k + 1) + A_te(k))
\]

Coefficient matrix and control effectiveness matrix are updated as

\[
\hat{B}(k + 1) = \hat{B}(k) + \Delta \hat{B}(k)
\]

\[
\check{B}(k + 1) = \check{B}(k) + (-\Gamma^{-T}u(k)(e(k + 1) + A_te(k))^T P)
\]

\[
\hat{C}(k + 1) = \hat{C}(k) + \Delta \check{C}(k)
\]

\[
\check{C}(k + 1) = \check{C}(k) + (-\Gamma^{-T} \Psi^T(x(k))(e(k + 1) + A_te(k))^T P)
\]

Generalizing the above derived adaptive laws

\[
\hat{B}(k) = \hat{B}(k - 1) + (-\Gamma^{-T}u(k - 1)(e(k) + A_te(k - 1))^T P)
\]

\[
\check{C}(k) = \check{C}(k - 1) + (-\Gamma^{-T} \Psi^T(x(k - 1))(e(k) + A_te(k - 1))^T P)
\]
A. Stability Analysis

Using adaptive laws for $\hat{B}(k)$ and $\hat{C}(k)$ Equation 25 reduces to 26 which can be further written as

$$\Delta V(k) = -e^T(k)(Q)e(k) + e^T Pe(k) + e^T PAe(k)$$

(34)

using the fact that $\epsilon \Delta e$ is very small and hence can be neglected.

$$\Delta V(k) \leq -|\lambda_{min}Q| \|e(k)\|^2 + \|e\| \|P\| \|e(k)\| + \|e\| \|P\| \|Ae(k)\|$$

(35)

This shows that $\Delta V(k)$ is semi-negative definite if

$$\|e(k)\| \geq \frac{\|e\| \|P\| \|I + Ae(k)\|}{|\lambda_{min}Q|} = \epsilon_{lb}$$

(36)

This means that system residual error should always be greater than $\epsilon_{lb}$ so that the system parameters converge. As $\Delta V(k)$ is a function of $e(k)$ and hence $e(k) \in L_\infty$. This implies that $x(k)$ and $\hat{x}(k)$ are bounded and $\in L_\infty$. $V(k)$ is bounded and hence has a limit, i.e.

$$\lim_{k \to \infty} V(k) = V_\infty$$

$$|\lambda_{min}Q| \|e(k)\|^2 + \|e\| \|P\| \|e(k)\| + \|e\| \|P\| \|Ae(k)\| \|e(k)\| \leq$$

$$V(0) - V_\infty$$

(37)

Using this it can be shown that $e(k) \in L_2 \cap L_\infty$ and therefore from Barbalat’s Lemma $e(k) \to 0$ as $t \to \infty$. The adaptive parameters i.e. $\hat{B}$ and $\hat{C}$ matrices converge to true values of $B$ and $C$ under persistence of excitation. In case $u$ is not persistently excited these matrices converge to constant values but not true $B$ and $C$ matrices.

V. Numerical Examples

Two examples are considered to validate the algorithm. All the examples are selected from nonlinear systems widely used to model the nonlinearities in different areas of research. Data obtained from numerical simulation are considered for both the examples. Once model is identified it is simulated and compared with the true data to analyze the error between the two. Euclidean norm of the error is used as a standard to quantify the error between the two models.

A. Example 1 : Duffing oscillator

First example is to identify the discrete model for duffing oscillator which is widely used as a benchmark problem in different areas of research. It is useful to have nonlinear model for this oscillator to test control laws or any other algorithms. Simulation data are generated to use for identification purposes. Duffing
oscillator is a periodically forced oscillator with a nonlinear elasticity. Equation for duffing oscillator is given as

\[
\ddot{x} + \delta \dot{x} + \beta x + \alpha x^3 = F
\]

\[
F = \gamma \cos(\omega t)
\]

where the damping constant \( \delta \geq 0 \). Nonlinear time histories of states of duffing oscillator are generated by simulating the above system. The force term is kept zero and the initial condition response of the duffing oscillator is simulated. Nonlinear simulation data is collected at a sample period of 0.01 \text{ seconds}. Following constants are used for the simulation.

| Table 1. Constants for Duffing Oscillator Simulation |
|---|---|
| \textit{Constants} | \textit{Numerical Value} |
| \( \alpha \) | 3 |
| \( \beta \) | -1 |
| \( \delta \) | 0.2 |

OKID is used to get the best linear discrete model using the simulation data. Table 2 gives the information of order of weight function and polynomials used in this example for identification purpose.

| Table 2. Parameters for System Identification of Duffing Oscillator |
|---|---|
| \textit{Number of coefficients} | 108 |
| \textit{Order of polynomials} | 2 |
| \textit{Order of piecewise continuity}\(^3\) | 1 |

Figure 2 shows the norm of the error between identified models and simulation data. The maximum error between the linear model and simulation data is of the order 0.8 while the maximum error for GLO-MAP model is of the order 0.01. It can be observed from the figure that the error between simulation data and nonlinear model is less as compared to the error between simulation data and linear model and hence it verifies the need of nonlinear model over a linear model.

Time histories of each state is also observed and GLO-MAP data shows the same nonlinear behavior as the simulation data. Time history of position level state of duffing oscillator is shown in Figure 3. It can be observed that the error between linear and simulation data is very large as compared to error between the GLO-MAP and simulation data. Linear model damps in first 50 seconds and settles at a different equilibrium point than true nonlinear system. GLO-MAP exhibits the similar behavior and settles down at the same
Figure 2. Norm of Error (Duffing Oscillator)

Figure 3. Time histories of First State of Duffing Oscillator
equilibrium point as the true system. Figure 4 shows the error between the linear model and simulation data and error between the GLO-MAP model and simulation data. From Figure 5 behavior of second state is observed. It can be noted here that the nonlinearity is captured very well in the second state also with very small error. Figure 5 shows the error plot and the error in second state is greater than the error observed in the first state. Time histories of second state are shown in Figure 6.

Figure 7 shows that all the coefficients settle down at constants as the error between the states approaches to zero.

1. Number of Nodes

Due to highly nonlinear behavior of this system it is difficult to capture the nonlinearities in the whole region and hence nodes are placed and local behavior of each region is learned. Number of nodes play an important part in capturing the behavior of nonlinear systems. Two cases were simulated with different number of nodes to see the effect of nodes on the identified model. Figure 8 shows the error time histories using 4 nodes. Figure 9 shows the comparison of norm of the error between the two cases i.e. using four nodes and using 9 nodes. It can be seen in this figure that that by using just 4 nodes the nonlinearity of the system is not captured very well and hence nodes are increased to 9. Maximum error is 0.5 in four node case and reduces to less than 0.1 using nine nodes.
Figure 5. Time Histories of Second State Error (Duffing Oscillator)

Figure 6. Time histories of Second State of Duffing Oscillator
Figure 7. Adaptation of Coefficients for Duffing Oscillator

Figure 8. Norm of Error using Four Nodes for Duffing Oscillator
B. Example 2: Lorentz Systems

A chaotic system is a nonlinear deterministic system that displays nonlinear and unpredictable behavior. These systems are very sensitive to initial conditions and system’s parameters behavior. One of the ways to represent a chaotic system is using Lorentz system. It is a simplified model of several physical systems. In addition to the field of non-linear mathematics, the Lorenz model has important implications for climate and weather prediction. It also models planetary atmospheres that exhibits a variety of quasi-periodic regimes which are fully deterministic but are subject to abrupt and seemingly random changes. The Lorentz system is described by following equations

\[
\begin{align*}
\dot{x} &= \sigma(y - x) \\
\dot{y} &= rx - y - xz \\
\dot{z} &= xy - bz
\end{align*}
\]

where \(\sigma, r, b > 0\) are system parameters.\textsuperscript{35}

<table>
<thead>
<tr>
<th>Constants</th>
<th>Numerical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\sigma)</td>
<td>10</td>
</tr>
<tr>
<td>(r)</td>
<td>28</td>
</tr>
<tr>
<td>(b)</td>
<td>(8/3)</td>
</tr>
</tbody>
</table>
Table 3 shows all the parameters used in this example for simulating the nonlinear behavior of Lorentz system. Simulated data is used to identify the linear model and then nonlinear model using GLO-MAP.

Table 4. Parameters for System Identification of Lorentz System

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of coefficients</td>
<td>810</td>
</tr>
<tr>
<td>Order of polynomials</td>
<td>2</td>
</tr>
<tr>
<td>Order of piecewise continuity</td>
<td>1</td>
</tr>
</tbody>
</table>

![Figure 10](image1.png)

**Figure 10. Time Histories of First State of Lorentz System**

The GLOMAP parameters are tabulated in Table 4. All the three states are compared with simulation data to measure the performance of the algorithm. Figure 10 shows the time histories of first state of the Lorentz system. It shows that the linear system is very damped and does not show nonlinear behavior at all. GLO-MAP model follows the nonlinear behavior very well but it misses amplitude information at the peaks.

Figure 11 shows the second state of the Lorentz system and Figure 12 shows the time histories of the third state. Linear system shows the similar behavior in second and third state and damps out within first six seconds. GLO-MAP model captures the frequency information very well and shows small error in amplitude for the second state.
Figure 11. Time Histories of Second State of Lorentz System

Figure 12. Time Histories of Third State of Lorentz System
VI. Conclusions

In this paper a discrete Global-Local Mapping Approximation algorithm was presented. The method was applied to identify discrete nonlinear systems using two examples. The results demonstrate that the discrete Global-Local Mapping Approximation nonlinear identification algorithm is able to learn the unknown nonlinearities, using discrete adaptive laws developed by the Lyapunov analysis. The euclidean norm of the errors between estimated and true model was calculated to analyze the identified model. Effect of number of nodes for identification method was also studied and it is concluded that the number of nodes can be tuned to get a better model. The future work includes validation of this algorithm for nonlinear experimental data. The numerical results show that GLO-MAP is a promising candidate for nonlinear system identification of discrete systems.

References


31 Singla, P., Multi-resolution methods for high fidelity modeling and control allocation in large-scale dynamical systems, Ph.D. thesis, Texas A&M University, College Station, TX, 2006.


