Local Jacobian based Galerkin Order Reduction for the Approximation of Large-Scale Nonlinear Dynamical Systems

Georg Fuchs, Alois Steindl, Stefan Jakubek

***Abstract*—**In automotive applications large-scale nonlinear dynamical models are utilized for hardware-in-the-loop simulations and model-based controller design. A projection-based order reduction of these models, on the one hand, yields substantial advantages in computational speed and on the other hand, simplifies the controller design procedure. In this work a mathematical-empirical approach is chosen for the order reduction of a real-time diesel engine model. It is based on recorded time-snapshots for typical system excitations. Flat and nonlinear Galerkin approximations are obtained by projection onto a lower-dimensional sub-space. In the nonlinear Galerkin approach a novel scheme for the reconstruction of the omitted states is introduced. It makes use of the local model parameters in the local Jacobian matrix, obtained from a linearization of the complete nonlinear model for various points of a local model network. The results from the application of the reduction methods to the engine model are presented and discussed for different reduced model orders and the benefits of the iteration scheme are demonstrated.

***Keywords*—**Diesel engine modeling, Model order reduction, Singular value decomposition, Snapshot method, Galerkin methods, Local model network

I. Introduction

Modern internal combustion engines are very complex systems with a variety of different calibration parameters and actuators. Especially regarding today's rigorous legal emissions regulations, conflicting goals between fuel efficiency, driveability, performance and emissions must be handled (see Fig. 1), which is a very challenging task for control engineers [1],[2].

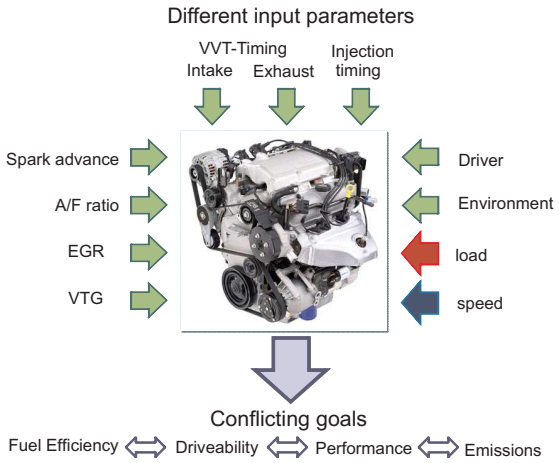

Fig. 1. Modern diesel engine's conflicting goals

The mathematical models used in the automotive industry are often deployed for offline and online simulation of the real engine behavior. They mostly comprise nonlinear ordinary differential equations and are typically utilized in practice using commercially available modeling and simulation tools. Online simulations require real-time capable models which are computationally very demanding due to their extensive, nonlinear structure. These models are often run in hardware-in-the-loop applications, using the engine model on a test bench for the validation of the engine control unit (ECU) performance, and frequently serve for model based controller design [3],[4]. They require very powerful and fast numerical integration algorithms. The main requirements for such real-time engine models are: the ability to sufficiently reproduce the system's main dynamic behavior, computational speed, and numerical accuracy and stability. From these standpoints a reduction of the complex nonlinear model to a simplified version only containing the main dynamic characteristics seems to be a very promising approach for subsequent online hardware-in-the-loop operation as well as controller design procedures.

The main idea behind model order reduction techniques is to simplify the original system to its dominating dynamic modes by significantly reducing the model order. There exists a multitude of different model order reduction techniques in literature, some of which are specified in section II. In this work a model order reduction will be applied to a real-time heavy-duty diesel engine model. The engine features a single-stage turbocharger and exhaust gas recirculation (EGR), and measurements can directly be acquired at the test bench for the parameter identification. It offers highly nonlinear dynamics through a wide operating range in engine speed and torque and constitutes a well-suited application example for the performance evaluation of model order reduction approaches. Two powerful methods are applied: the linear (flat) and the nonlinear Galerkin approximation methods. In the application of the nonlinear Galerkin method, a novel approach for the iterative solution of the nonlinear invariant manifold is adopted. It is based on the idea of decomposing the operating range into single subdomains (see section III) where the local model parameters in the local Jacobians are obtained by numerical linearization of the nonlinear model. The local Jacobians are assumed to be constant within each subdomain and then utilized for the iterative solution in the nonlinear Galerkin manifold (see section V). This approach provides substantial advantages concerning computation time, since the local Jacobians can be calculated from the original model in an offline a priori linearization. Direct simulation

outputs from the original nonlinear model for all states are used in the snapshot method for the assembly of the projection matrices (see section IV-D).

The paper is structured in the following way: in section II the basics of model order reduction with an overview of state-of-the-art techniques are presented. Then, in section III, the fundamentals of local model structures of nonlinear model domains are introduced. Section IV describes the principles of the flat and nonlinear Galerkin methods and shows the procedure of the snapshot methodology for the composition of the projection matrices calculated from a proper orthogonal decomposition. Section V deals with the iteration on the nonlinear invariant Galerkin manifold using the local Jacobian. In section VI the results of the application of the Galerkin methods to the engine model are given and the performances compared to each other. Section VII gives a conclusion and an outlook to forthcoming work.

II. Model Reduction Schemes

When dealing with large, complex nonlinear models, consisting of systems of nonlinear ordinary differential equation systems, there exist different schemes of model order reduction. One main group of such methods are singular value decomposition (SVD)-based approximation methods. They contain balanced approximations (first encountered in [5]), Hankel-norm approximations (see [6], [7]), proper orthogonal decomposition (POD), which comprises the Galerkin projection, and modal approximation methods. A very well-arranged compilation of POD and other model order reduction methods can be found in [8] and [9].

The basic idea of POD is the assumption that a state trajectory in the original state manifold of dimension d can be approximated by a projection of the trajectory onto a lower-dimensional state space (a sub-manifold) of dimension $m < d$. In this work the projection is obtained using a Galerkin projection which will be discussed in this paper in section IV. Because of the fact that there is no linear system behavior in this application, the eigenfunctions used for the projection are obtained from an empirical approach, called the snapshot method (see section IV-D), which takes advantage of system outputs collected in measurements or simulation [10].

The application of the aforementioned model order reduction methods has already been focus of several publications. In [11], the dimension reduction of the dynamics of a fluid conveying tube is presented, using linear and nonlinear Galerkin methods and center manifold reduction. In [12] an approach for the identification of the temporal coefficients of an empirical approximator of a process is shown. It uses experimental data gained from the process for a POD. [13] introduces a model order reduction using a nonlinear Galerkin projection for a finite element model of a horizontal axis wind turbine which serves for material fatigue assessment in long-time simulations. The application of the method of model reduction is controlled by error estimation. In [14], state aggregation techniques are applied in order to obtain a reduced-order model of tokamak devices. [15] uses Krylov subspaces for the model order reduction. Further approaches can be found in [16], [17], [18] and [19].

III. Local Model Network Approach

A. General

Modern diesel engines, may it be in the automotive area or for heavy-duty applications, offer a variety of inputs used for control purposes (Fig. 1). As mentioned before, the models consist of generally extensive systems of coupled nonlinear ordinary differential equations.

For a large class of these nonlinear dynamic systems, there exist methods that are based on the identification of subdomains of the system that can sufficiently accurately be described by local models [20]. The idea is to partition the utilized operating range of the model into smaller subdomains and approximate the nonlinear behavior by local models within each of these subdomains. The assembly of the subdomains is called local model network (LMN) which provides multilateral characterization of the overall system [21],[22]. A schematic example of such a local model network can be seen in Fig. 2. It shows the operating range in speed and load of a diesel engine with the single local model domains.

The local model networks are a very frequently adopted approach in automotive industry because they feature an appropriate structure for the representation of nonlinear static and dynamic systems where the incorporation of prior physical knowledge into the model structure is easily possible [23]. Additionally, in automotive applications at test benches usually numerous measurements from the engine are available for various dynamical test runs. Typically, the adopted local model network approaches make use of the input/output data of the system for the identification [24],[25].

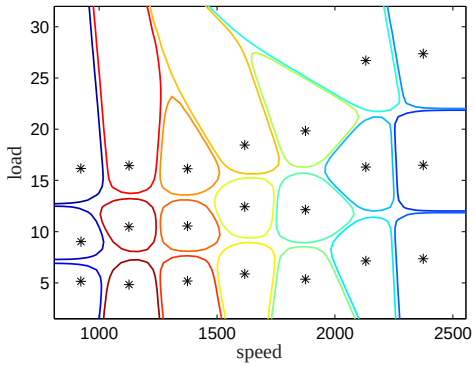

Fig. 2. Operating regime modeling using local models

In contrast to the identification purely from measurement data, the local model structure can also be composed from the known system behavior in the form of nonlinear ordinary differential equations [26]. In the present work the local models, which are derived from the nonlinear model, are used for a novel iterative scheme for the solution of the nonlinear invariant Galerkin manifold (see section V). The idea of this scheme is based on the notion that for typical operating modes of a heavy-duty engine the system dynamics in the vicinity of an operating point can sufficiently accurately be described by the local Jacobian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ (see section (V-A)). The choice of appropriate operating points for the linearization and the partitioning of the operating range into adequate

subdomains, for which the local model parameters are valid, is of particular importance. This procedure is to be specified in the following.

B. Local Model Network Architecture

A local model network approximates the behavior of the nonlinear model within its operating range as good as possible by using locally valid sub-models. In the following only linear local models will be considered.

The nonlinear model of the system is described by the state equation

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)), \tag{1}$$

with the state vector $\boldsymbol{x} \in \mathbb{R}^d$ and the input vector $\boldsymbol{u} \in \mathbb{R}^r$.

The operating range of the model is described by a so-called partition space. The partition space is a set of (mostly input) variables $\tilde{\boldsymbol{x}} = [\tilde{x}_1 \ \ldots \ \tilde{x}_p]$, with which the nonlinear behavior of the model can sufficiently be described. Within the partition space of the model a local linearization can be carried out for an arbitrary stationary operating point $\tilde{\boldsymbol{x}}$:

$$\boldsymbol{A}(\tilde{\boldsymbol{x}}) = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})}{\partial \boldsymbol{x}}\right|_{\tilde{\boldsymbol{x}}}. \tag{2}$$

$\boldsymbol{A}$ is the (local) Jacobian matrix, representing the local system dynamics around the operating point.

In the local model network approach the partition space is decomposed into n smaller subdomains. Every subdomain k possesses a local model which is specified by two components: its local Jacobian matrix

$$\boldsymbol{A}_k = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})}{\partial \boldsymbol{x}}\right|_{\tilde{\boldsymbol{x}}_k} \tag{3}$$

and a corresponding validity function

$$\phi_k = \phi_k(\tilde{x}_1, \tilde{x}_2) = \phi_k(\tilde{\boldsymbol{x}}), \tag{4}$$

which defines the area of validity of each local model. The task is now the determination of the locations of the particular operating points $\tilde{\boldsymbol{x}}_1 \ \ldots \ \tilde{\boldsymbol{x}}_n$ for the local Jacobians $\boldsymbol{A}_1 \ \ldots \ \boldsymbol{A}_n$ and the validity functions $\phi_1 \ \ldots \ \phi_n$ for each local model, which will be described below.

For an arbitrary point $\tilde{\boldsymbol{x}}$ of the partition space the local model network approach yields

$$\boldsymbol{A}^*(\tilde{\boldsymbol{x}}) = \sum_{k=0}^{n} \boldsymbol{A}_k \phi_k(\tilde{\boldsymbol{x}}), \tag{5}$$

with n being the number of local models.

Fig. 3 shows the example of a two-dimensional partition space spanned by the input variables $\tilde{x}_1$, $\tilde{x}_2$ and two local models ($n = 2$). In the application of the diesel engine the input variables could be e.g. the engine speed and the injection mass. In this case

$$\boldsymbol{A}^*(\tilde{x}_1, \tilde{x}_2) = \boldsymbol{A}_1 \phi_1(\tilde{x}_1, \tilde{x}_2) + \boldsymbol{A}_2 \phi_2(\tilde{x}_1, \tilde{x}_2). \tag{6}$$

The locations $\tilde{\boldsymbol{x}}_1$, $\tilde{\boldsymbol{x}}_2$ of the particular operating points for the local Jacobians $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ of the two local models are now determined in the following way: the partition space is discretized by a grid along the partition space axes, using the grid indices i_1 and i_2 for the description of every grid point in the partition space. Then, the following performance function is minimized over the complete partition space grid:

$$\sum_{i_1, i_2} \|\boldsymbol{A}^*(\tilde{x}_1(i_1), \tilde{x}_2(i_2)) - \boldsymbol{A}(\tilde{x}_1(i_1), \tilde{x}_2(i_2))\|_2^2 \rightarrow \min_{\phi_1, \phi_2, \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2}. \tag{7}$$

When the constraint

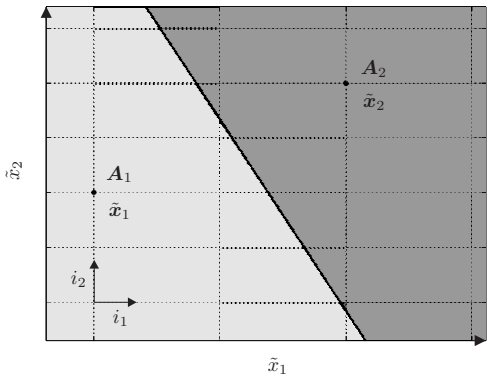

Fig. 3. Partition space for two inputs with two local models

$$\|\boldsymbol{A}^*(\tilde{x}_1(i_1), \tilde{x}_2(i_2)) - \boldsymbol{A}(\tilde{x}_1(i_1), \tilde{x}_2(i_2))\|_2 \leq q \tag{8}$$

for the error boundary q is not fulfilled for every grid point within the partition space, then the number of local models is increased and the optimization procedure is repeated.

Fig. 3 shows the two local models with the local Jacobians $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ obtained at the operating points $\tilde{\boldsymbol{x}}_1$ and $\tilde{\boldsymbol{x}}_2$, as well as the model transition boundary which is marked by the black bold line.

In the current work the local Jacobian matrix is obtained from a numerical linearization about the respective network subdomain operating point. It is then adopted in the iteration scheme of the nonlinear invariant Galerkin manifold (see section IV-C) for the computation of the vector of the inessential states. The realization of this approach is shown in section V.

IV. Model Reduction by Galerkin Methods

The Galerkin reduction methods were originally introduced for the approximation of dissipative partial differential equation problems. In a geometrical interpretation, they can be regarded as an approximation of the system dynamics on the phase manifold by projection onto a sub-space, which is able to capture the essential dynamical behavior of the original system. The two methods shown in this paper are the linear (also called flat) and the nonlinear Galerkin methods (see sections IV-B and IV-C).

In the general case a nonlinear model of the real process is obtained from a mathematical-physical modeling approach. Such a model can be written in state space representation as

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)), \tag{9}$$

with $\boldsymbol{x} \in \mathbb{R}^d$ the state vector of dimension d and $\boldsymbol{u} \in \mathbb{R}^r$ the r-dimensional input vector.

A. System Projection

The assumption is made that only certain modes of the original system play a relevant role in the representation of the system dynamics. In order to reduce the system order, a projection of the complete differential equation from the original state space onto a sub-space is performed. Galerkin's method now assumes that only the essential modes on this sub-space of the underlying model are important for the main dynamical behavior, whereas the remaining modes can either be neglected completely (flat Galerkin method) or are governed by the "main" modes in some algebraic relation (nonlinear Galerkin method).

For the solution of equation (9) the following approach is chosen:

$$\boldsymbol{x} = \boldsymbol{\Phi}_1 \boldsymbol{\xi} + \boldsymbol{\Phi}_2 \boldsymbol{\eta}, \tag{10}$$

where $\boldsymbol{\xi} \in \mathbb{R}^m$ is the vector containing the essential modes and $\boldsymbol{\eta} \in \mathbb{R}^{d-m}$ contains the less important remaining modes in the reduced system. The matrices $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2$ span m- and $d-m$-dimensional sub-spaces $\mathscr{X}$ and $\mathscr{Y}$. The composition of these matrices will be the topic of chapter IV-D.

Substituting (10) in (9) yields

$$\boldsymbol{\Phi}_1 \dot{\boldsymbol{\xi}} + \boldsymbol{\Phi}_2 \dot{\boldsymbol{\eta}} = \boldsymbol{f}(\boldsymbol{\Phi}_1 \boldsymbol{\xi} + \boldsymbol{\Phi}_2 \boldsymbol{\eta}, \boldsymbol{u}). \tag{11}$$

A projection onto the sub-spaces $\mathscr{X}$ and $\mathscr{Y}$ is accomplished by multiplication of (11) with the matrices $\boldsymbol{\Phi}_1^T$ and $\boldsymbol{\Phi}_2^T$:

$$\boldsymbol{\Phi}_1^T \boldsymbol{\Phi}_1 \dot{\boldsymbol{\xi}} + \boldsymbol{\Phi}_1^T \boldsymbol{\Phi}_2 \dot{\boldsymbol{\eta}} = \boldsymbol{\Phi}_1^T \boldsymbol{f}(\boldsymbol{\Phi}_1 \boldsymbol{\xi} + \boldsymbol{\Phi}_2 \boldsymbol{\eta}, \boldsymbol{u}) \tag{12}$$

$$\boldsymbol{\Phi}_2^T \boldsymbol{\Phi}_1 \dot{\boldsymbol{\xi}} + \boldsymbol{\Phi}_2^T \boldsymbol{\Phi}_2 \dot{\boldsymbol{\eta}} = \boldsymbol{\Phi}_2^T \boldsymbol{f}(\boldsymbol{\Phi}_1 \boldsymbol{\xi} + \boldsymbol{\Phi}_2 \boldsymbol{\eta}, \boldsymbol{u}) \tag{13}$$

Assuming $\boldsymbol{\Phi}_1^T \boldsymbol{\Phi}_1 = \boldsymbol{I}$, $\boldsymbol{\Phi}_1^T \boldsymbol{\Phi}_2 = \boldsymbol{0}$ and $\boldsymbol{\Phi}_2^T \boldsymbol{\Phi}_1 = \boldsymbol{0}$, $\boldsymbol{\Phi}_2^T \boldsymbol{\Phi}_2 = \boldsymbol{I}$, equations (12) and (13) lead to

$$\dot{\boldsymbol{\xi}} = \boldsymbol{\Phi}_1^T \boldsymbol{f}(\boldsymbol{\Phi}_1 \boldsymbol{\xi} + \boldsymbol{\Phi}_2 \boldsymbol{\eta}, \boldsymbol{u}) \tag{14}$$

$$\dot{\boldsymbol{\eta}} = \boldsymbol{\Phi}_2^T \boldsymbol{f}(\boldsymbol{\Phi}_1 \boldsymbol{\xi} + \boldsymbol{\Phi}_2 \boldsymbol{\eta}, \boldsymbol{u}). \tag{15}$$

After the projection of the original system two coupled systems of differential equations are thus obtained.

B. Linear Galerkin Method

The linear, or also called flat Galerkin method finds an approximation to the system (14),(15) by neglecting the states $\boldsymbol{\eta}$: $\boldsymbol{\eta} = \boldsymbol{0}$. For the approximate solution only equation (14) is considered:

$$\dot{\boldsymbol{\xi}} = \boldsymbol{\Phi}_1^T \boldsymbol{f}(\boldsymbol{\Phi}_1 \boldsymbol{\xi}, \boldsymbol{u}). \tag{16}$$

Since the remaining states are neglected in this approach, the flat Galerkin method only yields the approximation

$$\boldsymbol{x} \approx \boldsymbol{\Phi}_1 \boldsymbol{\xi}. \tag{17}$$

Geometrically, the linear Galerkin approach respresents a projection of the original differential equation from the state space of order d onto a sub-space of order m without accounting for the remaining $d-m$ states.

C. Nonlinear Galerkin Method

The nonlinear Galerkin method assumes that an algebraic relationship

$$\boldsymbol{\eta} = \boldsymbol{\Theta}(\boldsymbol{\xi}) \tag{18}$$

can be found which means that the behavior of the modes $\boldsymbol{\eta}$ is directly determined by the dynamic behavior of the $\boldsymbol{\xi}$ modes. Such a relation is called invariant manifold. The problem in this case is that there is no a priori information about an invariant manifold and the calculation can be very complex. Instead, one can come up with an approximate invariant manifold (AIM), which can be found without too much computational effort. The AIM can be described as the approximate solution of the equation

$$\dot{\boldsymbol{\eta}} = \boldsymbol{\Phi}_2^T \boldsymbol{f}(\boldsymbol{\Phi}_1 \boldsymbol{\xi} + \boldsymbol{\Phi}_2 \boldsymbol{\eta}, \boldsymbol{u}). \tag{19}$$

There exist different methods of finding these approximations (see [27], [28] and [9] for more detailed information about the AIM calculation). In this work the approach of Titi [29] was used. Here, the dynamics of $\boldsymbol{\eta}$ are disregarded in order to obtain a quasi-stationary AIM. (19) then becomes an algebraic relation:

$$\boldsymbol{\Phi}_2^T \boldsymbol{f}(\boldsymbol{\Phi}_1 \boldsymbol{\xi} + \boldsymbol{\Phi}_2 \boldsymbol{\eta}, \boldsymbol{u}) = \boldsymbol{0}. \tag{20}$$

This approach leads to a coupled system of differential-algebraic equations (14) and (20). The algebraic part can be solved by a fixed-point iteration, which is shown in section V.

D. Snapshot methodology for POD

In section IV-A the projection of the state space equation onto the sub-spaces $\mathscr{X}$ and $\mathscr{Y}$ was brought up. In this context the matrices $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2$ were utilized, computed from a snapshot analysis, which is the topic of this section.

Measurement results and analysis of the complete nonlinear model show that for typical operating conditions of an engine only certain modes show significant dynamic behavior. The idea behind the snapshot method is to excite the system with the inputs $\boldsymbol{u}$ in a way that is typical for the real engine operation. It is based on an empirical concept using the outputs generated from measurements or directly from the model simulation. It is a well-acknowledged technique, which has proved to be very efficient in several previous works, see e.g. [18] and [12]. For a defined input signal the states of the system are recorded and analyzed according to the dominant dynamic behavior. The $(n \times d)$ matrix $\boldsymbol{X}$ is called the snapshot record matrix, where d is the number of states in the reduced system and n the number of recorded snapshots, according to the simulation time and the time step width. For the assembly of $\mathscr{X}$ and $\mathscr{Y}$ the principal eigenmodes are needed, extracted from the $(d \times d)$ covariance matrix

$$\boldsymbol{C} = \frac{1}{d}((\boldsymbol{X} - \boldsymbol{\mu})^T (\boldsymbol{X} - \boldsymbol{\mu})), \tag{21}$$

whith $\boldsymbol{\mu}$ containing the mean values over all samples for each state. The eigenvalues λ_j and the eigenvectors $\boldsymbol{s}_j$ of the covariance matrix are computed according to

$$\boldsymbol{C}\boldsymbol{s}_j = \lambda_j \boldsymbol{s}_j. \tag{22}$$

For the POD, the eigenvectors of $\boldsymbol{C}$ are assembled in the matrices $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2$ according to the relative magnitude of the eigenvalues compared to each other:

$$\boldsymbol{\Phi}_1 = [\boldsymbol{s}_1 \ \dots \ \boldsymbol{s}_m], \ \boldsymbol{\Phi}_2 = [\boldsymbol{s}_{m+1} \ \dots \ \boldsymbol{s}_d]. \tag{23}$$

For the application to the diesel engine system, appropriate input signal excitation sequences have to be found. As already mentioned in this section, the aim is to keep only the modes which show distinct dynamic behavior and disregard the remaining ones. Modern internal combustion engines undergo strongly dynamic exposures, as for example abrupt load alternations or rapid acceleration maneuvers. These cycles are accompanied by partly rough actuator position variations. For the composition of the snapshot matrix in order to obtain preferably much information from the original system, the input signals should contain as much dynamical portions as possible. Fig. 4 shows one example of the input excitation of the throttle actuator signal u_{thr}, actuated around its halfway opened position ($u_{thr} = 50\%$) in a rectangular distribution with an amplitude of $\pm 20\%$ and superposed with additional noise. Corresponding signal sequences have been chosen for the other system inputs.

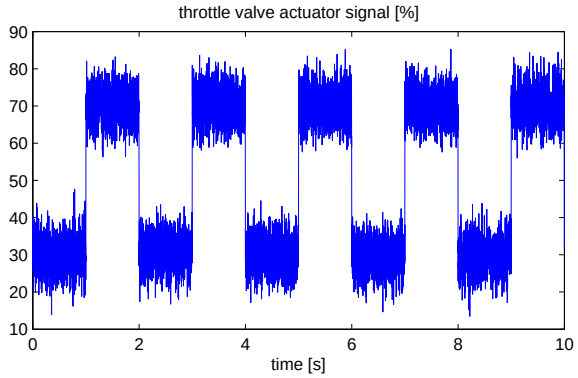

Fig. 4. Input signal sequence of the throttle actuator position u_{thr}

Of course the range of validity for the snapshot method is limited. For a certain excitation of the system with certain input variables, very sufficient results can be obtained in the reduced model using the same input variables. The method reaches its limits when additional inputs are actuated that are not accounted for in the snapshot acquisition process. Then the reduced model is not able to correctly map the dynamic response of the original system to these additional inputs.

V. Iterative nonlinear Galerkin Projection

In section IV-C the nonlinear equation (20) emerged. For reasons of computational speed and accuracy, an efficient numerical procedure has to be introduced to find approximate solutions on the invariant Galerkin manifold. On this account a local linear model structure, described in section III, is adopted, using the local Jacobian for the iteration scheme.

A. Linearized System

The nonlinear system equation (9) can be linearized in the current operating point (compare to section III-B), yielding the Jacobian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ according to

$$\boldsymbol{A} = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x},\boldsymbol{u})}{\partial \boldsymbol{x}}\right|_0, \ \boldsymbol{B} = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x},\boldsymbol{u})}{\partial \boldsymbol{u}}\right|_0. \tag{24}$$

Using an LMN approach, the operating region of the nonlinear model is separated into single subdomains (compare Fig. 2). Within these subdomains, the local Jacobians are computed and assumed to be constant for each of these partitions. The local Jacobians are now applied for the AIM iteration of (20) (see the following subsection).

B. Equation Error Minimization Problem

The solution of (20) is obtained recursively in the following way: the equation error is calculated and minimized in order to compute the optimal step width in the $\boldsymbol{\eta}$ direction. According to a Taylor series expansion approximation, the change of $\boldsymbol{f}$ in the $\boldsymbol{\eta}$ direction is

$$\Delta \boldsymbol{f} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\eta}}(\boldsymbol{x},\boldsymbol{u})\Delta\boldsymbol{\eta}. \tag{25}$$

$\Delta \boldsymbol{f}$ is used for the calculation of the change of $\boldsymbol{f}$ in the $\boldsymbol{\eta}$ direction:

$$\boldsymbol{f}_{k+1} = \boldsymbol{f}_k + \Delta \boldsymbol{f} = \boldsymbol{f}_k + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\eta}}(\boldsymbol{x},\boldsymbol{u})\Delta\boldsymbol{\eta} \tag{26}$$

Using the chain rule of differentiation and eq. (10), $\partial \boldsymbol{f}/\partial \boldsymbol{\eta}$ yields

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\eta}}(\boldsymbol{x}(\boldsymbol{\xi},\boldsymbol{\eta}),\boldsymbol{u}) = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\eta}} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\boldsymbol{\Phi}_2 \tag{27}$$

The Jacobian matrix $\partial \boldsymbol{f}/\partial \boldsymbol{x} = \boldsymbol{A}$ is computed offline for the different operating regimes and is chosen online according to the regime the model is currently running in. Combining the equations from above, the current equation error for the next step in $\boldsymbol{\eta}$ can be computed:

$$\boldsymbol{r}_k = \boldsymbol{\Phi}_2^T \boldsymbol{f}_k \tag{28}$$

$$\boldsymbol{r}_{k+1} = \boldsymbol{\Phi}_2^T \left(\boldsymbol{f}_k + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\eta}}\Delta\boldsymbol{\eta}\right) = \boldsymbol{\Phi}_2^T \left(\boldsymbol{f}_k + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\boldsymbol{\Phi_2}\Delta\boldsymbol{\eta}\right) \tag{29}$$

The error is now used for optimizing the performance function

$$J = \frac{1}{2}\boldsymbol{r}_{k+1}^T \boldsymbol{r}_{k+1} \rightarrow \min_{\Delta\boldsymbol{\eta}}, \tag{30}$$

from which the optimal $\Delta\boldsymbol{\eta}$ step size follows to

$$\Delta\boldsymbol{\eta} = -\boldsymbol{\Phi}_2^T \left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right)^{-1} \boldsymbol{f}_k. \tag{31}$$

The described iteration scheme is performed in parallel to the numerical integration of the reduced system states. The advantage of this method is that due to the fact that the local Jacobians, which are calculated offline for the subdomains of the LMN, can be switched online during simulation, a constitutive increase in performance with a low additional computing expense is allowed for. The results obtained are discussed in the following section.

VI. Results

The goal of this work was to adapt the model order reduction methods described in sections IV and V to a real-time heavy-duty diesel engine model. The model description and the results of the reduction methods are the topics of this section.

A. System description

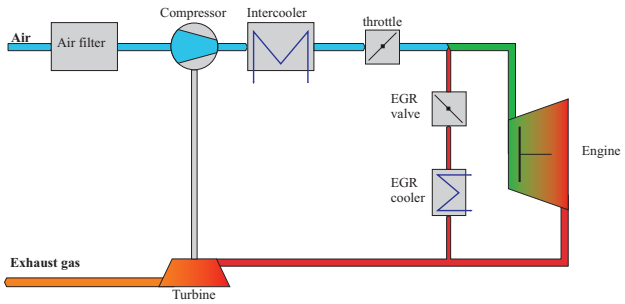

Fig. 5. Diesel engine system overview

The configuration of the diesel engine is schematically illustrated in Fig. 5. The model considered was obtained from a physical-mathematical modeling approach using the concept of mean value modeling (MVM) [30],[31]. MVM means the replacement of the discontinuous operation of the pistons by continuous processes for mass transportation through the cylinders and production of power. The thermodynamic and chemical processes inside the cylinders during the combustion cycles are considered as mean values over a cycle. Such a simplifying modeling approach is well-suited for realtime test stand applications because of low computational demand and thus higher simulation speed. Additionally, it achieves sufficient accuracy. The model is built up by modularly connectable zero-dimensional tank components, which reproduce the thermodynamic characteristics inside pipe connections, coolers, etc. They assume perfect mixing of the gas components inside so that a homogeneous distribution of pressure and temperature is sufficiently justified. These storage elements are connected to each other through coupling elements (valves, throttles). The exhaust gas turbocharger is modeled by a quasi-stationary, parameter based approach. The complete system has a state order of 16. The parameters of the model were identified using measurements from the test bench.

B. Results

The linear (flat) and nonlinear Galerkin methods were now applied to the engine model and the results were compared to the solution of the original system. The solution of the original system was computed using a fixed-step Runge-Kutta numerical integration method with an integration step size of 0.001s. The original system was also used for the offline calculation of the local Jacobians $\partial \boldsymbol{f}/\partial \boldsymbol{x}$. The main focus of this work lies in the examination of the general functionality of the two presented Galerkin methods and the sensitivity of the performance with respect to the reduced model order.

Different input signal and operating condition test cases were designed, with three input variables: the throttle actuator position u_{thr}, the EGR valve actuator position u_{egr} and the engine speed n_{eng}. For the snapshot recording procedure and the subsequent proper orthogonal decomposition, these input variables were varied in the range deployed for the respective use case (compare section IV-D). Consequently, the model order was reduced and the differences in performance of the model order reduction methods are discussed.

Case 1: $n_{eng} = 600\ U/min$, $u_{egr} = 100\ \%$, u_{thr} varied. The throttle actuator input signal was varied around its halfway opened position and superposed with additional noise, see also Fig. 4. In Fig. 6 the result for the exhaust turbocharger speed is shown for a reduced model order of 10. In the behavior of the solutions the following can be noticed: the system dynamics are reproduced sufficiently by both the flat and the nonlinear Galerkin methods. Obviously, the flat Galerkin method shows a clear offset from the original solution, which is also increasing with time.

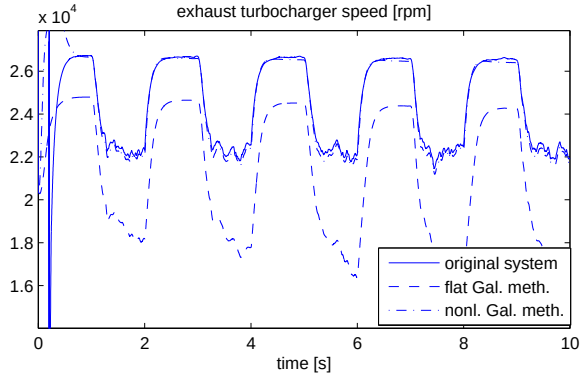

Fig. 6. Exhaust turbocharger speed, reduced system order 10

Next, the order of the reduced system was lowered to nine. For the same input signal the results were compared, see Fig. 7. It can be recognized that for a model order reduction by one the results of the flat and the nonlinear Galerkin methods show even more significant differences. The nonlinear Galerkin method traces the solution of the original system very well, whereas the flat method already indicates strong errors in the amplitude. The performance discrepancy is even more obvious, when the reduced order is dramatically decreased to 4 states, see Fig. 8. Here, the solution of the flat Galerkin method can not reproduce the dynamics of the original system any more (outside the plot range), but the nonlinear Galerkin method can still perform well, showing decent accord with the reference solution.

Case 2: $n_{eng} = 1200\ U/min$, $u_{egr} = 100\ \%$, u_{thr} varied. Here, the throttle actuator input signal was varied according to the sequence given in Fig. 9 for a higher engine speed of $1200\ U/min$. The results for the exhaust turbocharger speed, the intake manifold pressure and the intake throttle mass flow are presented in Fig. 10 for a reduced model order of 9.

Again, the dynamics of the flat and nonlinear Galerkin approximations achieve good dynamical accordance, with the flat solution showing a little offset. For a reduction of the reduced

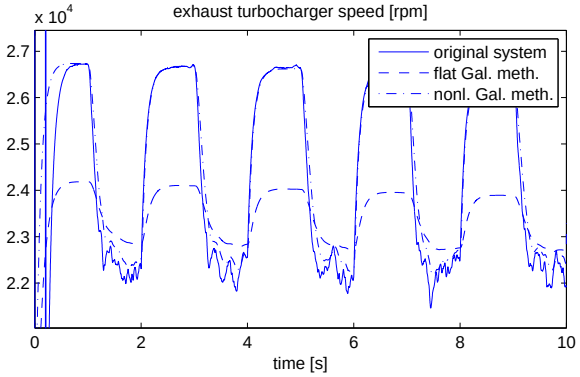

Fig. 7. Exhaust turbocharger speed, reduced system order 9

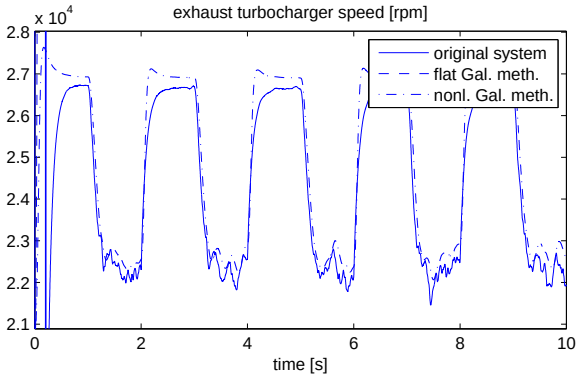

Fig. 8. Exhaust turbocharger speed, reduced system order 4

model order by one to 8 states (see Fig. 11), no dramatic decline of accuracy can be noticed for both approximation methods, but another interesting property can be recognized: the flat Galerkin solution shows signs of numerical instabilities for strong amplitudes in the range between 8 and 9 seconds of simulation time, whereas the nonlinear approximation features very robust numerical behavior. This is a crucial advantage of the nonlinear Galerkin method compared to the flat reduction when it comes to real-time hardware-in-the-loop applications where the limitation of the time step width and numerical stability are very important.

Case 3: $u_{egr} = 100\ \%$, u_{thr} and n_{eng} varied simultaneously. In this test case two input variables are varied at the same time: the throttle actuator position and the engine speed. The according pattern is shown in Fig. 12.

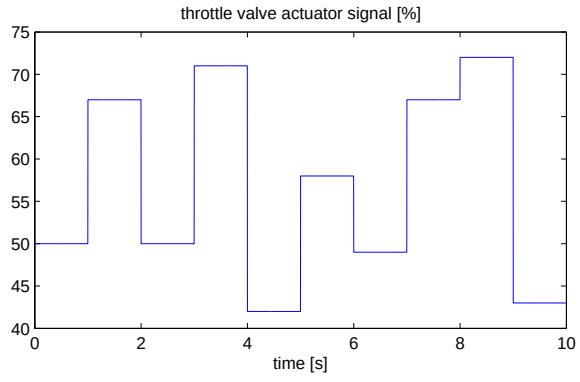

Fig. 9. Input signal sequences of the throttle actuator position u_{thr}

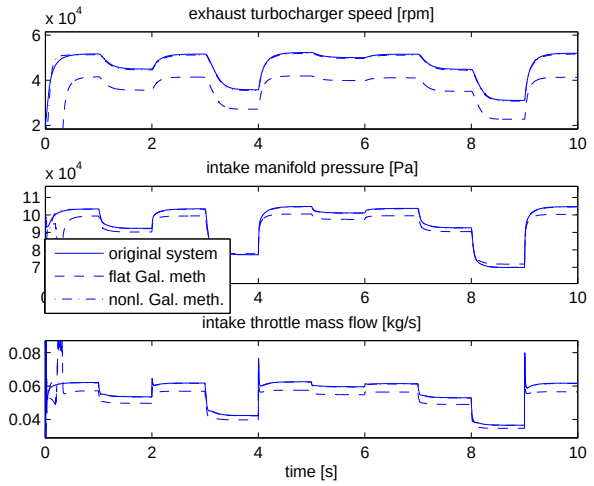

Fig. 10. Exhaust turbocharger speed, intake manifold pressure and throttle mass flow, reduced system order 9

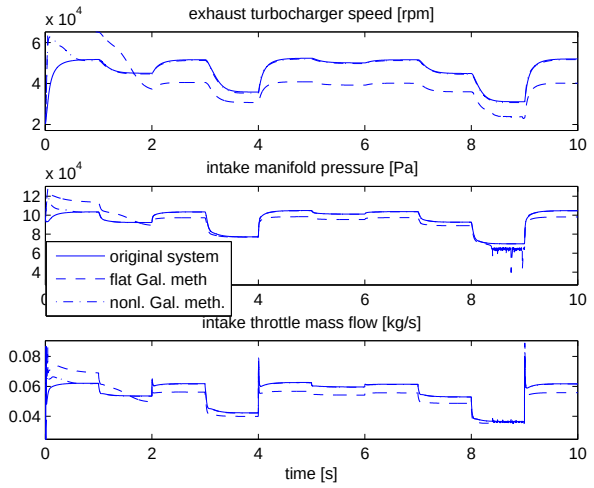

Fig. 11. Exhaust turbocharger speed, intake manifold pressure and throttle mass flow, reduced system order 8

Analyzing the results for the exhaust turbocharger speed, the intake and the exhaust manifold pressures (Fig. 13), it can be clearly seen that already for a system order of 9 the results of the nonlinear Galerkin method are much better, especially at the beginning of the simulation.

The difference becomes more distinct again when the system order is further reduced, in this case from 9 to 8 states (Fig. 14). The nonlinear Galerkin approximation fits the original solution very well, whereas the flat Galerkin method result shows poor behavior.

Case 4: $u_{thr} = 20\ \%$, u_{egr} and n_{eng} varied. Again, two input variables - the EGR valve actuator signal and the engine speed - are varied according to the pattern given in Fig. 15.

Fig. 16 shows the results for the intake manifold pressure,

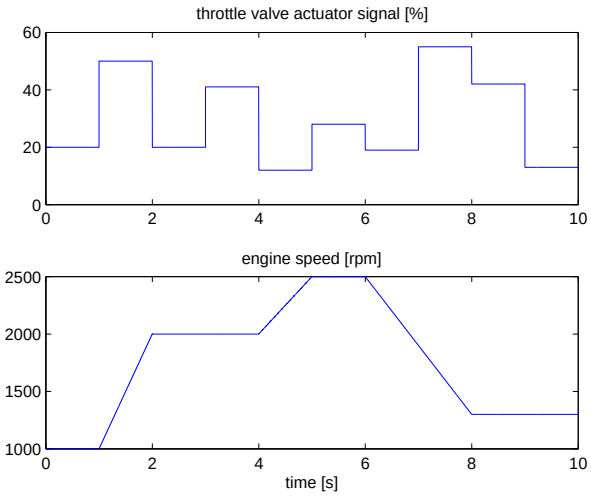

Fig. 12. Input signal sequences of the throttle actuator position u_{thr} and the engine speed n_{eng}

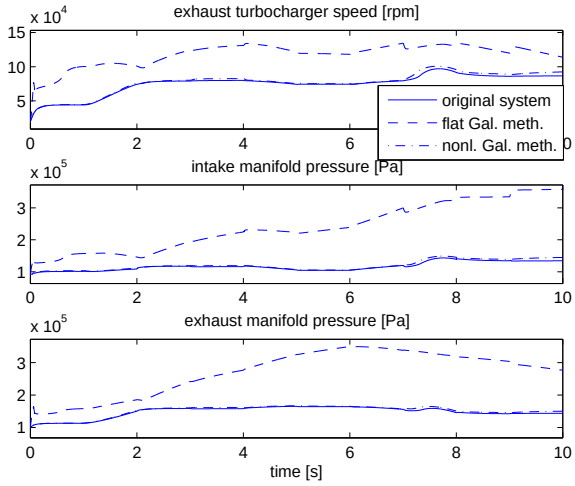

Fig. 14. Exhaust turbocharger speed, intake and exhaust manifold pressure, reduced system order 8

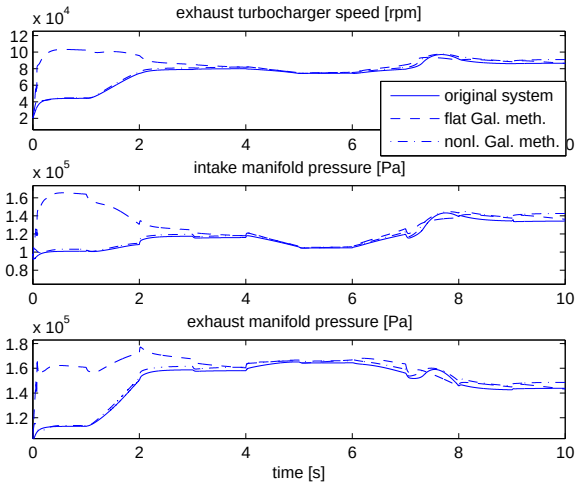

Fig. 13. Exhaust turbocharger speed, intake and exhaust manifold pressure, reduced system order 9

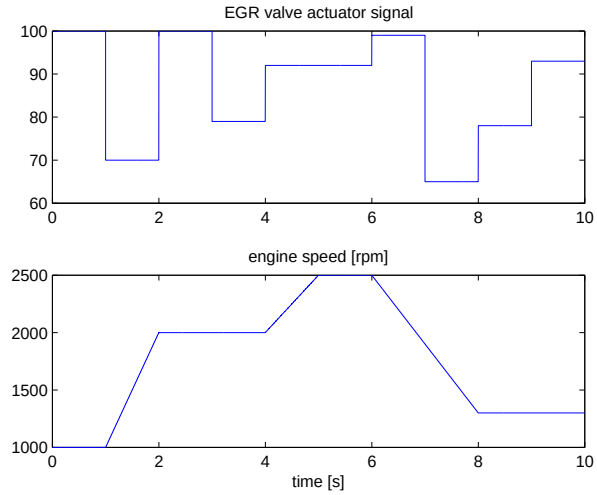

Fig. 15. Input signal sequences of the EGR valve actuator position and the engine speed

the EGR mass flow and the exhaust manifold pressure for the reduced system order 10. When the model order is further reduced - here from 10 to 8 states - again one can see the considerably better results of the nonlinear method (see Fig. 17). The results of the nonlinear Galerkin method seem even more remarkable when the wide range of passed nonlinear system behavior (1000 - 2500 rpm) is envisioned.

The results of this section allow for some important conclusions: The linear and nonlinear Galerkin methods are both able to reproduce the main system dynamics very well for high orders of the reduced model. In the process of decreasing the model order, the omitting of relevant dynamic states is clearly better compensated by the nonlinear Galerkin method due to the fact that it can account for these omitted states using the AIM. The use of the local Jacobian achieves impressive results even for running through strongly nonlinear operating regions. However, the linear method, completely neglecting the omitted states, shows considerable performance declines for reduced system orders.

VII. Conclusion

In this work the basic principles of flat and nonlinear Galerkin methods for the model reduction of computationally expensive dynamical systems, applied to a mean value diesel engine model, were presented. The existing approach of the nonlinear Galerkin method was combined with an iterative solution scheme using local Jacobians. The local Jacobians were obtained from a local model network approach and assumed constant within each operating subdomain. A proper orthogonal decomposition on the basis of the snapshot method for the determination of the principal eigenvalues was performed and used for the flat and nonlinear Galerkin reduction schemes. The achieved results proved that both the linear and nonlinear Galerkin method yield very viable reduced order

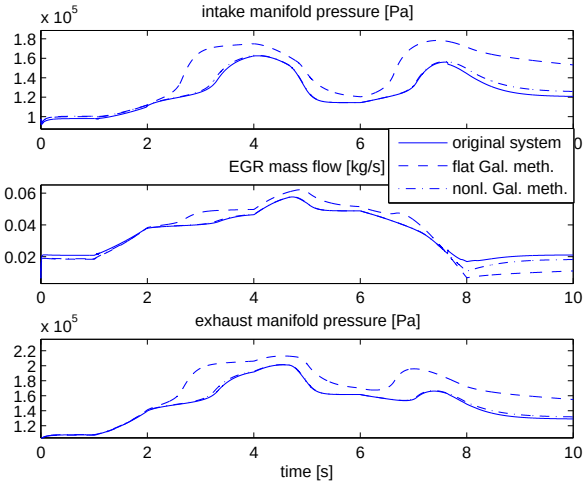

Fig. 16. Intake manifold pressure, EGR mass flow and exhaust manifold pressure, reduced system order 10

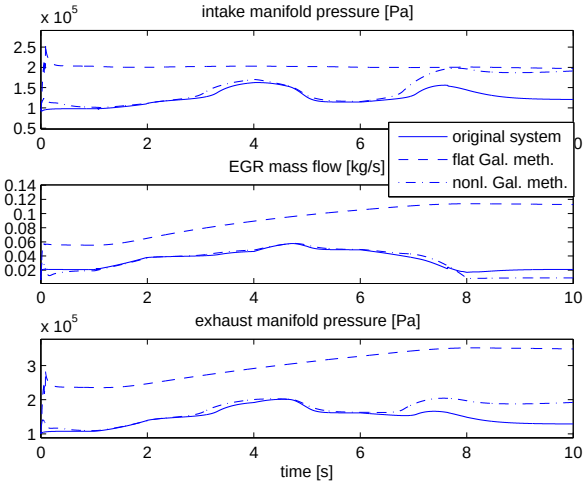

Fig. 17. Intake manifold pressure, EGR mass flow and exhaust manifold pressure, reduced system order 8

models. Subsequent reduction of the model order revealed significant performance advantages of the nonlinear Galerkin method using the local Jacobian-based iteration scheme. The low computational expense increase is outweighed by remarkable benefits in dynamic accuracy.

The obtained results provide an outlook to further applications in other technical disciplines. Forthcoming work is intended to deal with the comparison of the flat and nonlinear Galerkin methods with other model order reduction techniques, such as balanced reduction and Krylov-based approximation methods.

REFERENCES

[1] J. Wahlström, L. Eriksson, L. Nielsen, and M. Pettersson, "PID controllers and their tuning for EGR and VGT control in diesel engines," ser. IFAC World Congress, Prague, Czech Republic, 2005.

[2] N. Buzbuchi and L. Stan, "Model simulation of high power diesel engine exhaust gas pollutants," *Proceedings of the 3rd WSEAS International Conference on Environmental and Geological Science and Engineering*, pp. 59–64, September 2010.

[3] L. Guzzella and C. Onder, *Introduction to Modeling and Control of Internal Combustion Engine Systems*. Springer-Verlag Berlin Heidelberg, 2004.

[4] Q. Song and K. Grigoriadis, "Diesel engine control using parameter varying methods," *WSEAS Trans. Systems*, vol. 3, no. 5, pp. 2335–2341, 2004.

[5] C. Mullis and R. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *Circuits and Systems, IEEE Transactions on*, vol. 23, no. 9, pp. 551 – 562, sep. 1976.

[6] V. Adamyan, D. Arov, and M. Krein, "Analytic properties of schmidt pairs for a hankel operator and the generalized schur-takagi problem," *Mat. Sb. (N.S.)*, vol. 86(128), pp. 34–75, 1971.

[7] ——, "Infinite block hankel matrices and related extension problems," *Amer. Math. Soc. Trans.*, vol. 111, pp. 133–156, 1978.

[8] A. C. Antoulas, *Approximation of Large Scale Dynamical Systems*, R. C. Smith, Ed. Society for Industrial and Applied Mathematics, 2005.

[9] A. Steindl and H. Troger, "Methods for dimension reduction and their application in nonlinear dynamics," *International Journal of Solids and Structures*, vol. 38, no. 10-13, pp. 2131 – 2147, 2001.

[10] L. Sirovich, "Method of snapshots," *Q. Appl. Math*, vol. XLV, pp. 561–571, 1987.

[11] A. Steindl, "Application of dimension reduction methods to the dynamics of a fluid conveying tube," *ENOC-2005, Eindhoven, Netherlands, 7-12 August 2005*, 2005.

[12] A. J. Enrique, "Estimation of the temporal coefficients for an empirical approximator. new approach based on the proper orthogonal decomposition modes," *Computers & Chemical Engineering*, vol. 34, no. 8, pp. 1220 – 1225, 2010.

[13] H. G. Matthies and M. Meyer, "Nonlinear galerkin methods for the model reduction of nonlinear dynamical systems," *Computers & Structures*, vol. 81, no. 12, pp. 1277 – 1286, 2003, advanced Computational Models and Techniques in Dynamics.

[14] A. Beghi, V. Coccorese, and A. Portone, "Model reduction for electromechanical systems with application to tokamak control," *Proceedings of the WSEAS Multiconference, Rethymno, Greece*, pp. 4831–4836, July 2001.

[15] Z. Yacine, A. Benfdila, and S. Djennoune, "Model order reduction of nonlinear systems via the trajectory piecewise linear approach application to transmission lines circuit," *Proceedings of the 8th WSEAS International Conference on Applied Informatics and Communications, Rhodes, Greece*, pp. 369–375, August 2008.

[16] K. Willcox and J. Peraire, "Balanced model reduction via the proper orthogonal decomposition," *AIAA Journal*, vol. 40, pp. 2323–2330, 2002.

[17] S. Prajna, "Pod model reduction with stability guarantee," vol. 5, dec. 2003, pp. 5254 – 5258 Vol.5.

[18] C. Frouzakis, Y. Kevrekidis, J. Lee, K. Boulouchos, and A. Alonso, "Proper orthogonal decomposition of direct numerical simulation data: Data reduction and observer construction," *Proceedings of the Combustion Institute*, vol. 28, no. 1, pp. 75 – 81, 2000.

[19] Y. Gheisari, "Toward structure preserving model reduction for analysis and simulation in micro-technology," *Proceedings of the 8th WSEAS International Conference on Microelectronics, Nanoelectronics, Optoelectronics*, pp. 32–36, May 2009.

[20] C. Hametner, "Nonlinear dynamic system identification using local model architectures," Ph.D. dissertation, Vienna University of Technology, 2007.

[21] R. Murray-Smith and T. A. Johansen, *Multiple Model Approaches to Modelling and Control*. Taylor & Francis, 1997.

[22] G. Gregorcic and G. Lightbody, "Local model network identification with gaussian processes," *Neural Networks, IEEE Transactions on*, vol. 18, no. 5, pp. 1404 –1423, sep. 2007.

[23] M. Hafner, M. Schüler, O. Nelles, and R. Isermann, "Fast neural networks for diesel engine control design," *Control Engineering Practice*, vol. 8, no. 11, pp. 1211 – 1221, 2000.

[24] O. Nelles, *Nonlinear System Identification*, 1st ed. Springer Verlag, 2002.

[25] C. Hametner and S. Jakubek, "Neuro-fuzzy modelling using a logistic discriminant tree," *American Control Conference, 2007. ACC '07*, pp. 864–869, July 2007.

[26] T. A. Johansen and B. A. Foss, "Operating Regime based Process Modeling and Identification," *Computers and Chemical Engineering*, vol. 21, pp. 159 – 176, 1997.

[27] C. Foias, M. Jolly, I. Kevrekidis, G. Sell, and E. Titi, "On the computation of inertial manifolds," *Physics Letters A*, vol. 131, no. 7-8, pp. 433 – 436, 1988.
[28] J. Carr, *Applications of Centre Manifold Theory*, R. C. Smith, Ed. Springer Verlag, 1981.
[29] E. S. Titi, "On approximate inertial manifolds to the navier-stokes equations," *Journal of Mathematical Analysis and Applications*, vol. 149, no. 2, pp. 540 – 557, 1990.
[30] L. Guzzella and C. Onder, *Introduction to Modeling and Control of Internal Combustion Engine Systems*. Springer-Verlag Berlin Heidelberg, 2004.
[31] J. Wahlström and L. Eriksson, "Modeling of a diesel engine with VGT and EGR including oxygen mass fraction," Department of Electrical Engineering, Linköpings Universitet, SE-581 83 Linköping, Sweden, Tech. Rep. LiTH-R-2747, 2006.

Georg Fuchs received the M.S. degree in Mechanical Engineering in 2008 from the Vienna University of Technology. He is presently research assistant at the Institute of Mechanics and Mechatronics at the Vienna University of Technology, pursuing his Ph.D. degree in automatic control engineering. His current research interests are internal combustion engine modeling, simulation and control, numerical linearization and integration methods, and real-time hardware applications.

Alois Steindl is Professor at the Vienna University of Technology, Institute of Mechanics and Mechatronics. His main research interests are nonlinear stability and bifurcation theory, optimization and dimension reduction techniques.

Stefan Jakubek received the M.S. degree in mechanical engineering in 1997, the PhD in 2000 and the Habilitation (professorial qualification) in early 2007 from Vienna University of Technology. He is head of the Division for Control and Process Automation of the Institute of Mechanics and Mechatronics at the Vienna University of Technology. His research interests include fault diagnosis, nonlinear system identification and simulation technology.