Multibody Syst Dyn (2011) 25: 313–334
DOI 10.1007/s11044-010-9238-3

Simulation process of flexible multibody systems with non-modal model order reduction techniques

Jörg Fehr · Peter Eberhard

Received: 18 November 2009 / Accepted: 12 November 2010 / Published online: 4 December 2010

Abstract One important issue for the simulation of flexible multibody systems is the reduction of the flexible body's degrees of freedom. In this work, nonmodal model reduction techniques for flexible multibody systems within the floating frame of reference framework are considered. While traditionally in the multibody community modal techniques in many different forms are used, here other methods from system dynamics and mathematics are in the focus. For the reduction process, finite element data and user inputs are necessary. Prior to the reduction process, the user first needs to choose boundary conditions fitting the chosen reference frame before defining the appropriate in- and outputs. In this work, four different possibilities of modeling appropriate interface points to reduce the number of inputs and outputs are presented.

The main model reduction techniques to be considered in this context are moment-matching by projection on Krylov-subspaces, singular vaule decomposition (SVD)-based reduction techniques and combinations of those which are also compared to traditional modal approaches. All these reduction techniques are implemented in the model order reduction code Morembs. In addition, an error estimator for Krylov-subspace methods exists and an a-priori error bound can be calculated if frequency weighted Gramian matrices are used for the reduction process. This allows a fully automated reduction process. We evaluate and compare these methods in the frequency as well as in the time domain by reducing the flexible degrees of freedom of a rack used for active vibration damping of a scanning tunneling microscope.

Keywords Flexible multibody systems · Model reduction · Krylov-subspace · Gramian matrices · Modal reduction · Floating frame of reference · Interface definition

J. Fehr · P. Eberhard (✉)
Institute of Engineering and Computational Mechanics, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany
e-mail: peter.eberhard@itm.uni-stuttgart.de

J. Fehr
e-mail: joerg.fehr@itm.uni-stuttgart.de

1 Introduction

Typical elements of classical multibody systems are rigid bodies which are interconnected with ideal joints and constraint elements between each other and the surrounding environment. Due to increased usage of lightweight structures and the increased working speed, the negligence of elastic effects often is no longer appropriate. If a system consists of rigid bodies as well as bodies where the deformations have to be considered, one has to deal with flexible multibody systems. They are used to examine the dynamic behavior of gear-boxes, robot-arms, crank-shafts, windmills, impact drills, etc. The concept of flexible multibody systems is described, e.g., in [1] and [2]. This method is especially well suited for problems with a large movement of the reference frame and only small elastic deformations. Furthermore, it is possible to reduce the computational burden by using model reduction techniques; see, e.g. [3–6]. In this work, nonmodal model reduction techniques for flexible multibody systems within the floating frame of reference framework are considered. The simulation of a flexible multibody system involves some preprocessing steps shown in Fig. 1.

On the one hand, there is a nonlinear simulation of the motion of rigid bodies consisting of p bodies with f degrees of freedom and q reaction forces. On the other hand, there are elastic bodies. The deformations of elastic bodies can be described by a set of partial differential equations (PDEs). For general bodies, spatial discretization techniques are used transforming the PDEs to a set of ordinary differential equations (ODEs). Mostly, the Finite Element Method (FEM) is used, leading to bodies described by often more than 100,000 elements. Within the framework of a floating frame of reference formulation, the motion $\boldsymbol{r}$ of points of an elastic body is separated into an usually nonlinear motion of the reference frame $\boldsymbol{K}_i$ and into a linear elastic deformation $\boldsymbol{u}$ with respect to the reference frame; see Fig. 2. Similarly, the transformation matrix of frame $\boldsymbol{K}_P$ with respect to frame $\boldsymbol{K}_I$, is spitted into two transformation matrices $\boldsymbol{A}_{Ik}(t) = \boldsymbol{A}_{Ii}(t) \cdot \boldsymbol{A}_{iP}(t)$, in which $\boldsymbol{A}_{Ii}(t)$ defines a coordinate transformation from frame $\boldsymbol{K}_i$ to frame $\boldsymbol{K}_I$. In this

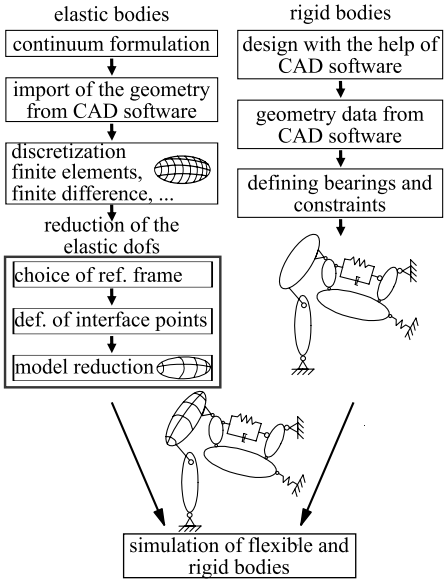


Fig. 1 Preprocessing for the simulation of a flexible multibody system

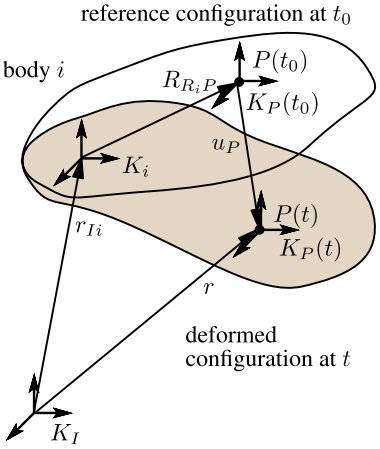


Fig. 2 Floating frame of reference formulation

work, only small deformations are considered and the transformation matrix $\boldsymbol{A}_{iP}(t)$ from frame $\boldsymbol{K}_i$ to frame $\boldsymbol{K}_P$ is separated into a constant part $\boldsymbol{\Gamma}_{ik}$ and a time dependent part $\boldsymbol{I} + \tilde{\boldsymbol{\vartheta}}_k(t)$

$$\boldsymbol{A}_{iP}(t) = \boldsymbol{\Gamma}_{ik} \cdot \left(\boldsymbol{I} + \tilde{\boldsymbol{\vartheta}}_k(t)\right), \tag{1}$$

where $\tilde{\boldsymbol{\vartheta}}_k(t)$ is the skew-symmetric matrix of the the rotational angles collected in the vector $\boldsymbol{\vartheta}_k(t)$. The linear elastic deformation $\boldsymbol{u} \in \mathbb{R}^3$ and the small rotations $\boldsymbol{\vartheta} \in \mathbb{R}^3$ are then approximated by $\boldsymbol{u} = \boldsymbol{\Phi} \cdot \boldsymbol{q}$ and $\boldsymbol{\vartheta} = \boldsymbol{\Phi} \cdot \boldsymbol{q}$, where $\boldsymbol{q} \in \mathbb{R}^N$ summarizes the nodal degrees of freedom of the finite element model and $\boldsymbol{\Phi}$ contains the shape functions.

Using Jourdain's principles of dynamics, the equation of motion for a single body can be derived, as shown, e.g., in [1], and read

$$\begin{bmatrix} \boldsymbol{M}_r & \boldsymbol{M}_{er}^T \\ \boldsymbol{M}_{er} & \boldsymbol{M}_e \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{a} \\ \ddot{\boldsymbol{q}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_r \\ \boldsymbol{h}_e \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ -\boldsymbol{K}_e \cdot \boldsymbol{q} - \boldsymbol{D}_e \cdot \dot{\boldsymbol{q}} \end{bmatrix}, \tag{2}$$

where the submatrix $\boldsymbol{M}_r \in \mathbb{R}^{6\times6}$ corresponds to the mass matrix known from rigid multibody dynamics, $\boldsymbol{M}_e \in \mathbb{R}^{N\times N}$, $\boldsymbol{D}_e \in \mathbb{R}^{N\times N}$ and $\boldsymbol{K}_e \in \mathbb{R}^{N\times N}$ are the flexible mass, damping and stiffness matrices, whereas $\boldsymbol{M}_{er} \in \mathbb{R}^{6xN}$ provides the coupling between the rigid body movement and the elastic deformation. The vector $\boldsymbol{a}$ contains the global accelerations of the floating frame of reference, the vectors $\boldsymbol{h}_r \in \mathbb{R}^6$ and $\boldsymbol{h}_e \in \mathbb{R}^N$ collect generalized inertia forces, gravitational forces, and forces acting on the body's surface. This approach leads to $6 + N$ degrees of freedom per elastic body. Three translational and three rotational degrees of freedom of the rigid body dynamics plus the N elastic degrees of freedom.

In the following, all flexible bodies are considered simultaneously which yields the same structure of the equation with merely different dimensions. As a consequence, transient simulations, endurance tests or design problems of such huge systems are often hardly feasible. Because of this, it is necessary to reduce the flexible degrees of freedom (dof) of every elastic body, e.g., by a Petrov–Galerkin projection of the flexible coordinates $\boldsymbol{q}$ on an appropriate subspace $\text{span}(\boldsymbol{V}) \in \mathbb{R}^{N\times n}$ by $\boldsymbol{q} = \boldsymbol{V} \cdot \bar{\boldsymbol{q}}$, with $n = \dim(\bar{\boldsymbol{q}}) \ll \dim(\boldsymbol{q}) = N$ and requiring the residual to be orthogonal to the right projection space $\text{span}(\boldsymbol{W}) \in \mathbb{R}^{N\times n}$. A projection is called orthogonal if $\boldsymbol{V} = \boldsymbol{W}$ and otherwise oblique. This procedure leads to the reduced

equations of motion

$$\begin{bmatrix} \boldsymbol{M}_r & \boldsymbol{M}_{er}^T \cdot \boldsymbol{V} \\ \boldsymbol{W}^T \cdot \boldsymbol{M}_{er} & \boldsymbol{W}^T \cdot \boldsymbol{M}_e \cdot \boldsymbol{V} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{a} \\ \ddot{\boldsymbol{q}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_r \\ \boldsymbol{W}^T \cdot \boldsymbol{h}_e \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ -\boldsymbol{W}^T \cdot \boldsymbol{K}_e \cdot \boldsymbol{V} \cdot \overline{\boldsymbol{q}} - \boldsymbol{W}^T \cdot \boldsymbol{D}_e \cdot \boldsymbol{V} \cdot \dot{\overline{\boldsymbol{q}}} \end{bmatrix}. \tag{3}$$

For model reduction, the question of how to choose appropriate projection subspaces span($\boldsymbol{V}$) and span($\boldsymbol{W}$) is of interest. With the projection matrices $\boldsymbol{V}$ and $\boldsymbol{W}$ the standard data can be calculated as explained, e.g., in [1] and [2]. The elastic shape functions $\boldsymbol{\Phi}_{\mathrm{el}}$ are the global shape functions times the projection matrix $\boldsymbol{\Phi}_{\mathrm{el}} = \boldsymbol{\Phi} \cdot \boldsymbol{V}$. In state of the art reduction methods like modal reduction, the projection spaces span($\boldsymbol{V}$) = span($\boldsymbol{W}$) consist of the dominant eigenvectors plus some additional modes. The component mode method developed in [7] was utilized in [4] for the reduction of flexible multibody systems. A combination of fixed boundary modes plus some "constraint modes" which account for local effects at boundaries are utilized. However, the convergence of modal reduction can be slow, e.g., because the spatial distribution of loads is not considered. In addition, no information about the error introduced by model reduction can be gained if modal reduction techniques are used and a tuning of the reduced model for certain frequency ranges is not possible.

Concerning model reduction methods from system dynamics and mathematic on the contrary, like moment-matching or SVD-based reduction, they have rigorous error bounds and can be tuned in a certain frequency range. These methods are utilized in this work to find better elastic ansatz functions and gain information about the error introduced by model reduction. If all the reaction and applied forces acting on the elastic body are considered as inputs $\boldsymbol{B}_e \cdot \boldsymbol{u}(t)$ and the most important displacements of the elastic body as outputs $\boldsymbol{y}(t) = \boldsymbol{C}_e \cdot \boldsymbol{q}(t)$ to the elastic body, where $\boldsymbol{B}_e \in \mathbb{R}^{N \times p}$ and $\boldsymbol{C}_e \in \mathbb{R}^{r \times N}$, then the elastic part of the body can be considered as a linear time-invariant second-order multi input multi output (MIMO)-system $\boldsymbol{M}_e \cdot \ddot{\boldsymbol{q}}(t) + \boldsymbol{D}_e \cdot \dot{\boldsymbol{q}}(t) + \boldsymbol{K}_e \cdot \boldsymbol{q}(t) = \boldsymbol{B}_e \cdot \boldsymbol{u}(t)$, $\boldsymbol{y}(t) = \boldsymbol{C}_e \cdot \boldsymbol{q}(t)$. The projection spaces span($\boldsymbol{V}$) and span($\boldsymbol{W}$) are then yielding the reduction space of the second-order MIMO-system reduction and can be found with methods like moment-matching based on projection on Krylov-subspaces, SVD-based reduction techniques and combinations of those. It must be emphasized that now $\boldsymbol{V}$ and $\boldsymbol{W}$ no longer contain eigenmodes belonging to certain eigenfrequencies but mathematically determined ansatz functions with convincing properties. The reduced order system reads

$$\begin{aligned} \bar{\boldsymbol{M}}_e \cdot \ddot{\bar{\boldsymbol{q}}}(t) + \bar{\boldsymbol{D}}_e \cdot \dot{\bar{\boldsymbol{q}}}(t) + \bar{\boldsymbol{K}}_e \cdot \bar{\boldsymbol{q}}(t) &= \bar{\boldsymbol{B}}_e \cdot \boldsymbol{u}(t), \\ \boldsymbol{y}(t) &= \bar{\boldsymbol{C}}_e \cdot \bar{\boldsymbol{q}}(t) \end{aligned} \tag{4}$$

with the reduced order input matrix $\bar{\boldsymbol{B}}_e = \boldsymbol{W}^T \cdot \boldsymbol{B}_e$ and output matrix $\bar{\boldsymbol{C}}_e = \boldsymbol{C}_e \cdot \boldsymbol{V}$. Due to the physical fact that $\boldsymbol{M}_e$ is not singular, the second-order model can be written as an equivalent state-space model

$$\begin{aligned} \underbrace{\begin{bmatrix} \dot{\boldsymbol{q}}(t) \\ \ddot{\boldsymbol{q}}(t) \end{bmatrix}}_{\dot{\boldsymbol{x}}(t)} &= \underbrace{\begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\boldsymbol{M}_e^{-1} \cdot \boldsymbol{K}_e & -\boldsymbol{M}_e^{-1} \cdot \boldsymbol{D}_e \end{bmatrix}}_{\boldsymbol{A}} \cdot \underbrace{\begin{bmatrix} \boldsymbol{q}(t) \\ \dot{\boldsymbol{q}}(t) \end{bmatrix}}_{\boldsymbol{x}(t)} + \underbrace{\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{M}_e^{-1} \cdot \boldsymbol{B}_e \end{bmatrix}}_{\boldsymbol{B}} \cdot \boldsymbol{u}(t), \\ \boldsymbol{y}(t) &= \underbrace{\begin{bmatrix} \boldsymbol{C}_e & \boldsymbol{0} \end{bmatrix}}_{\boldsymbol{C}} \cdot \begin{bmatrix} \boldsymbol{q}(t) \\ \dot{\boldsymbol{q}}(t) \end{bmatrix}, \end{aligned} \tag{5}$$

where the state-space matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$, and the state vector $\boldsymbol{x}$ are introduced. Using the Laplace transformation, the transfer matrix of the system $\boldsymbol{H}(s) = \boldsymbol{C}_e \cdot (s^2\boldsymbol{M}_e + s\boldsymbol{D}_e + \boldsymbol{K}_e)^{-1} \cdot \boldsymbol{B}_e$ is obtained. Using the Laplace transformation, the transfer matrix of the reduced system $\bar{\boldsymbol{H}}(s) = \bar{\boldsymbol{C}}_e \cdot (s^2\bar{\boldsymbol{M}}_e + s\bar{\boldsymbol{D}}_e + \bar{\boldsymbol{K}}_e)^{-1} \cdot \bar{\boldsymbol{B}}_e$ is obtained.

In Fig. 1, the necessary steps involved in the simulation of a flexible multibody system are shown. It is neither intended to change the preprocessing methods for obtaining the elastic description of a flexible body nor to write a new flexible multibody code. Instead, we concentrate on the reduction process which is one of the integral parts of the simulation. The reduction process for every flexible body involves three crucial steps. First, choosing of boundary conditions fitting to the chosen reference frame. Secondly, the definition of appropriate inputs and outputs. Thirdly, the reduction of the linear time-invariant second-order system. Our aim is, e.g., to calculate the so-called standard input data of an elastic body (SID); see [8]. Based on the standard input data the equation of motion of a single elastic body (4) is assembled and solved by the commercial elastic multibody system code. With the previously calculated projection matrices $\boldsymbol{V}$ and $\boldsymbol{W}$ the standard data can be calculated as explained, e.g., in [1] and [2]. The improvement is represented by the application of new and more sophisticated reduction techniques to find the best approximation of the full elastic body with only a few ansatz functions and by the use of new metrics for the evaluation of the state of the art reduction method.

In Sect. 2, we introduce an example on the bases of which we compare and explain the necessary steps involved in the reduction process. We then shortly recapitulate the two essential reference frames used in the floating frame of references formulation before continuing with explanations about appropriate definitions of interface nodes. We finish this section with some notes about implementation issues. In Sect. 3, we give a short introduction of the different model reduction techniques implemented by the authors in recent years. Afterward, the used programs and the data flow is explained in Sect. 4. These different reduction techniques are evaluated with the example in the frequency as well as in the time domain in Sect. 5. This work is completed by some conclusions in Sect. 6.

2 Example

One example used in former work (see [9, 10]) was the finite element model of a rack consisting of two rigid plates and beam elements; see Fig. 3. The model used in this context has about 4000 elastic degrees of freedom. The two rigid plates are modeled as rigid bodies connected via constraint equations to the rods; see Fig. 3. The beam model served as an initial example for testing the different possible interface definitions, different reduction techniques and the efficient implementation of these methods. One goal of this work is the application of the new reduction methods to industrial size problems. In this paper, the full model, serves as an example from which the simplified academic beam model was derived. The rack is used for active vibration damping of a scanning tunneling microscope (STM) as shown in Fig. 4 and explained in [11]. The geometric information about the rack is given as CAD data. As a further step, the finite element model is obtained by meshing the CAD data with an appropriate FE Program, leading to a model with about 35,000 degrees of freedom.

The volume body is meshed with a 20 node structural solid element. All the nodes belonging to the structural solid element have three degrees of freedom. From the Finite Element data, the standard-data of an elastic body can be calculated as described in [1]. The interesting frequency range for this problem is between $f_{\min} = 0$ Hz and $f_{\max} = 1000$ Hz. The upper part of the system weights about 1.2 kg; the lower plate weights about 0.2 kg.

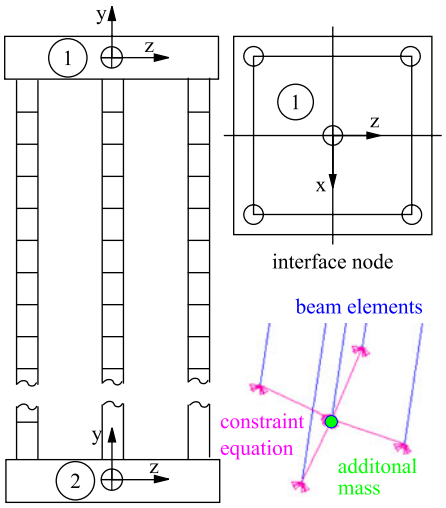


Fig. 3 Model of the academic rack consisting of two rigid bodies and beams

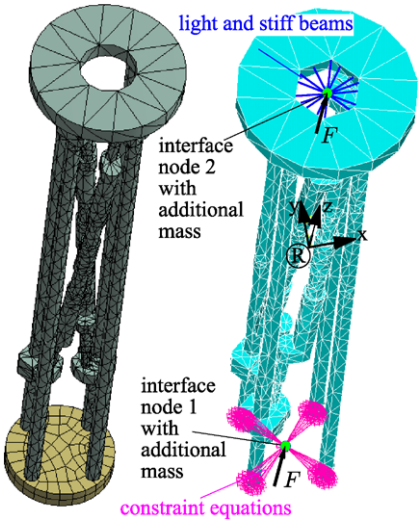


Fig. 4 Model of the rack

Forces are acting on the lower plate and at the hole in the upper plate in all three translational directions. The upper plate is the interface to the remaining part of the system. The lower plate is assumed to be a rigid body rigidly connected to the beams. The exact method of how the upper interface is appropriately modeled is investigated in Sect. 2.2. The following points need to be considered when calculating the standard data: Location and appropriate boundary conditions for the reference frame, modeling of interface nodes and master slave relations between the nodes.

2.1 Choice of reference frame

In the floating frame of reference approach twelve variables are used for describing the location $\boldsymbol{r}_P$ and orientation $\boldsymbol{A}_P$ of an arbitrary frame. Six of the twelve variables are redundant. This redundancy is eliminated by using special reference conditions. The use of principle

axes as a reference frame is described in [12]. The origin of the reference frame must remain at the instantaneous mass center and the three products of inertia must remain zero as the body deforms. This condition is imposed by adding constraint equations at the position, velocity, and acceleration level. In our work, commercial multibody system codes should be utilized for solving the dynamic system. Because the use of principle axes is not implemented in the commercial multibody system codes, different reference frames need to be utilized. Three additional methods are describe in [5]. The user's responsibility is the appropriate choice of the reference frame. Mostly, a tangent frame or a Buckens frame is chosen. The tangent frame imposes a kinematic constraint on the body by identifying the reference system $\boldsymbol{K}_i$ with the frame of the body having the material coordinate $\boldsymbol{R_0} = \boldsymbol{0}$. The Buckens frame, however, imposes a dynamical constraint on the body and is located in the center of gravity of the undeformed body. The dynamical constraint is defined as

$$\int_V \boldsymbol{u}(\boldsymbol{R}, t) \cdot \boldsymbol{u}(\boldsymbol{R}, t)\, dm \stackrel{!}{=} \min. \tag{6}$$

The choice of the Buckens frame leads to the smallest elastic deformation possible. For arbitrary shape functions, condition (6) must be enforced by imposing an algebraic constraint while solving the system's equations of motion. However, it was shown, e.g., in [1] that by using eigenfunctions of unsupported structures, so called free-free modes, and deleting the six rigid body modes, constraint (6) is automatically met. Although the Buckens frame leads to the smallest elastic deformation possible, a situation can arise that causes a reference frame with kinematic constraints to be favored. Comparison of the different approaches can be found in [5]. For general model reduction approaches, the following points need to be considered for fulfilling the floating frame of reference formulation. The shape function of the elastic body needs to fulfill the kinematic boundary conditions, e.g., $\boldsymbol{\Phi}(\boldsymbol{R_0}) = \boldsymbol{0}$ for the tangent frame. This implies implicitly that the projection spaces also fulfill these boundary conditions. This means that the user first has to translate the global coordinate system of the FE body into the node where the reference frame is located. Secondly, the elastic degrees of freedom of the reference nodes get bounded by setting them to zero. The columns of the projection matrices corresponding to this dof contain zeros only. For a Buckens frame, the global coordinate system of the FE body is translated into the center of gravity of the undeformed body. In addition, it is necessary to project the system with projection spaces span($\boldsymbol{V}$) and span($\boldsymbol{W}$), which are orthogonal to the rigid body modes of the body.

2.2 Definition of interface nodes

Another important step previous to the model reduction step is the definition of appropriate interfaces which are necessary to reduce the number of inputs and for efficient coupling between bodies. Usually, forces or boundary conditions are not acting on a single node. They are acting within a specific area or on a specified user interface. In order to prepare the FE model for model reduction usually additional nodes so-called interface nodes are introduced; compare, e.g. [13]. In this example, one additional interface node is in the center of the upper hole; see Fig. 4 node 2. In a next step, a relation between the degrees of freedom of the interface nodes $\boldsymbol{q}_I$ and the degrees of freedom $\boldsymbol{q}_S$ of the nodes at the interface surface, in this example, the plane of the inner cylinder, are introduced by FE constraint equations; see Fig. 5. Such constraints are standard FE modeling capability and offer the option to define rigid regions that are associated to the interface node. The experienced user utilizes the correct interface definitions for correct simulation results. The user manual of ANSYS [14] suggests an improvement by adding a small mass to the node and then coupling the various

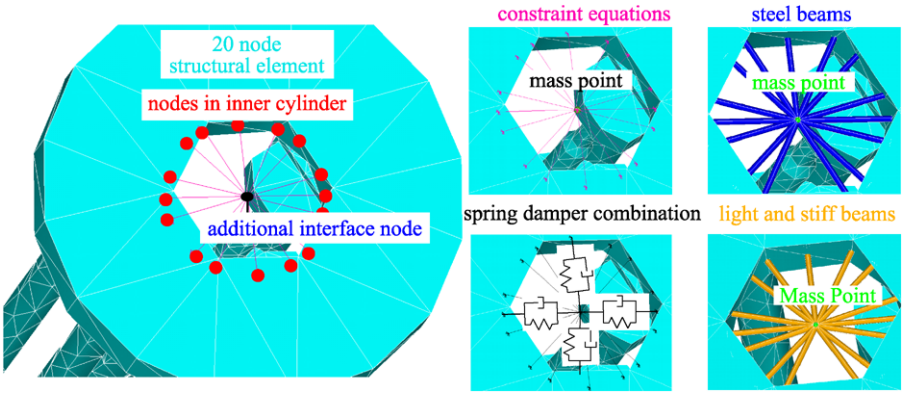
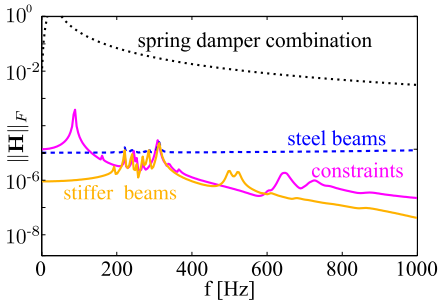


Fig. 5 Different possibilities to define an appropriate interface node

Fig. 6 Influence of different interface definitions



FE constraints to the additional point mass. Other possibilities for interface definitions are additionally shown in Fig. 5. One alternative, is the usage of beams instead of constraints. As a second example, 15 round steel beams, with a diameter of 1 mm are used for coupling the interface node with the hole in the upper plate. It is necessary that the eigenfrequencies of the additional beams are remote to the interesting frequency range. For a fixed-fixed Bernoulli beam, the first eigenfrequencies are calculated from

$$f_i = \frac{\Lambda_i}{2\pi}\sqrt{\frac{EI}{\rho A l^4}} \tag{7}$$

with E being Young's modulus, I the second moment of inertia, ρ the density, A the area and l the length of the beam. The Λ_i are resulting from the solution of the characteristic equation in the eigenproblem of fixed-fixed beam and can be found in tables; see, e.g. [15]. For the steel beam, the first two eigenfrequencies are at 4326 Hz and at 7182 Hz. The additional steel beams add a mass of $m_{\text{beams}} = 1.38$ mg. If the diameters of the beams are reduced a lower mass is introduced. However, the eigenfrequencies are proportional to the diameter of the steel beams. If beams with $d = 0.1$ mm diameter are used, the first eigenfrequency of a single beam is at at 433 Hz, which is in the interesting frequency range. In Fig. 6, the Frobenius norm of the frequency response matrix $\|\boldsymbol{H}\|_F$ is plotted depending on the different interface definitions. We see that if steel beams are used, the frequency response is different from the response with constraints coupling. The frequency response is constant for frequencies

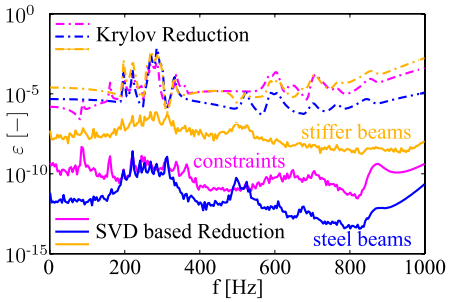


Fig. 7 Influence of different interface definitions

higher than 350 Hz which is not fitting the experimental data. Better results are achieved if non physical beams are used instead of steel beams. The nonphysical beams are stiffer, using Young's modulus $E = 10^{18}$ N/m^2 instead of $E_{\text{steel}} \approx 2.1 \times 10^{11}$ N/m^2, and lighter using density $\rho = 100$ kg/m^3 instead of $\rho_{\text{steel}} \approx 8000$ kg/m^3. This coupling is better than coupling with steel beams and the coupling with constraint equations. Another possibility for coupling the interface node is is the usage of a spring damper element. At the moment, an appropriate parameterization of the spring damper elements for this problem is not yet done and the authors are working on defining appropriate interface definitions in a way that the FE results fit the measurement results. In this example, the usage of a spring damper combination is not a good choice because the spring damper combination cannot distribute forces in z-direction because no forces act in this direction due to the orientation of the spring damper element. The wrong interface behavior can be seen in the frequency response plot; see Fig. 6. However, the developed model reduction tools can handle without additional user input all the different interface definitions in a mathematically correct way. The best interface definition depends on the actual system and needs verification by measurement. Our goal is to give the user the freedom of defining different interfaces and then using the different reduction techniques, but we want to emphasize that correct interface definition is crucial for good reduction results. In Fig. 7 a Krylov-subspace based reduction technique and a SVD-based reduction technique were applied to the different interface models. Model reduction for the spring-damper interface does not make sense because the parameterization is not yet satisfying. Because of that, we do not show these results in Fig. 6. For all the reduced models, the relative error $\epsilon(\omega) = \|\boldsymbol{H}(\omega) - \tilde{\boldsymbol{H}}(\omega)\|_F / \|\boldsymbol{H}(\omega)\|_F$ in the Frobenius norm is small meaning Morembs can handle the different interfaces definitions.

2.3 Implementation issues for model reduction

In the FE model different degrees of "freedom" families exist. Some degrees of freedom are dependent $\boldsymbol{q}_d$, some are constraint $\boldsymbol{q}_c$ and some are independent $\boldsymbol{q}_i$ dofs. The system dynamics of the elastic body is partitioned into the following form:

$$\begin{bmatrix} \boldsymbol{M}_e^{ii} & \boldsymbol{M}_e^{id} & \boldsymbol{M}_e^{ic} \\ \boldsymbol{M}_e^{di} & \boldsymbol{M}_e^{dd} & \boldsymbol{M}_e^{dc} \\ \boldsymbol{M}_e^{ci} & \boldsymbol{M}_e^{cd} & \boldsymbol{M}_e^{cc} \end{bmatrix} \cdot \begin{bmatrix} \ddot{\boldsymbol{q}}_i \\ \ddot{\boldsymbol{q}}_d \\ \ddot{\boldsymbol{q}}_e \end{bmatrix} + \begin{bmatrix} \boldsymbol{D}_e^{ii} & \boldsymbol{D}_e^{id} & \boldsymbol{D}_e^{ic} \\ \boldsymbol{D}_e^{di} & \boldsymbol{D}_e^{dd} & \boldsymbol{D}_e^{dc} \\ \boldsymbol{D}_e^{ci} & \boldsymbol{D}_e^{cd} & \boldsymbol{D}_e^{cc} \end{bmatrix} \cdot \begin{bmatrix} \dot{\boldsymbol{q}}_i \\ \dot{\boldsymbol{q}}_d \\ \dot{\boldsymbol{q}}_e \end{bmatrix}$$
$$+ \begin{bmatrix} \boldsymbol{K}_e^{ii} & \boldsymbol{K}_e^{id} & \boldsymbol{K}_e^{ic} \\ \boldsymbol{K}_e^{di} & \boldsymbol{K}_e^{dd} & \boldsymbol{K}_e^{dc} \\ \boldsymbol{K}_e^{ci} & \boldsymbol{K}_e^{cd} & \boldsymbol{K}_e^{cc} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{q}_i \\ \boldsymbol{q}_d \\ \boldsymbol{q}_e \end{bmatrix} = \boldsymbol{B}_e \cdot \boldsymbol{u}(t), \tag{8}$$
$$\boldsymbol{y} = \boldsymbol{C}_e \cdot \boldsymbol{q}.$$

The motion of the dependent degrees of freedom can be expressed with the movement of the independent degrees of freedom. The relation between the dependent degrees $\boldsymbol{q}_d(t)$ and the independent degrees of freedom $\boldsymbol{q}_s(t)$ are expressed by a constraint equation

$$\boldsymbol{q}_d(t) = \boldsymbol{G}_M \cdot \boldsymbol{q}_i(t) + \boldsymbol{g}. \tag{9}$$

The motion of the constraint degrees of freedom are bounded by a boundary equation

$$\boldsymbol{q}_c = \boldsymbol{0}. \tag{10}$$

By inserting (9) and (10) into (8) the dynamics of the elastic body is described with the equation of motion of the independent degrees of freedom

$$\boldsymbol{M}_e^* \cdot \ddot{\boldsymbol{q}}_i(t) + \boldsymbol{D}_e^* \cdot \dot{\boldsymbol{q}}_i(t) \boldsymbol{M}_e^* \cdot \boldsymbol{q}_i(t) = \boldsymbol{B}_e^* \cdot \boldsymbol{u}(t) + \boldsymbol{f}^*, \tag{11}$$

$$\boldsymbol{y} = \boldsymbol{C}_e^* \cdot \boldsymbol{q}(t), \tag{12}$$

with

$$\boldsymbol{M}_e^* = \boldsymbol{M}_e^{ii} + \boldsymbol{G}_M^T \cdot \boldsymbol{M}_e^{di} + \boldsymbol{G}_M^T \cdot \boldsymbol{M}_e^{dd} \cdot \boldsymbol{G}_M + \boldsymbol{M}_e^{id} \cdot \boldsymbol{G}_M, \tag{13}$$

$$\boldsymbol{D}_e^* = \boldsymbol{D}_e^{ii} + \boldsymbol{G}_M^T \cdot \boldsymbol{D}_e^{di} + \boldsymbol{G}_M^T \cdot \boldsymbol{D}_e^{dd} \cdot \boldsymbol{G}_M + \boldsymbol{D}_e^{id} \cdot \boldsymbol{G}_M, \tag{14}$$

$$\boldsymbol{K}_e^* = \boldsymbol{K}_e^{ii} + \boldsymbol{G}_M^T \cdot \boldsymbol{K}_e^{di} + \boldsymbol{G}_M^T \cdot \boldsymbol{K}_e^{dd} \cdot \boldsymbol{G}_M + \boldsymbol{K}_e^{id} \cdot \boldsymbol{G}_M, \tag{15}$$

$$\boldsymbol{f}^* = -\boldsymbol{G}_M^T \cdot \boldsymbol{K}_e^{dd} \cdot \boldsymbol{g} - \boldsymbol{K}_e^{id} \cdot \boldsymbol{g}. \tag{16}$$

From the FE programs, the assembled elastic mass, stiffness and damping matrix can be extracted. Usually $\boldsymbol{M}_e^*, \boldsymbol{D}_e^*, \boldsymbol{K}_e^*$, and not $\boldsymbol{M}_e^{ii}, \boldsymbol{D}_e^{ii}, \boldsymbol{K}_e^{ii}$ are extracted. In addition, also the force vector $\boldsymbol{f}^*$ is extracted.

As the input, respectively, the output matrices $\boldsymbol{B}_e^*$ and $\boldsymbol{C}_e^*$, are assembled in the model reduction step. For the reduction process, the user has to define which node i is an input or output, respectively, and in which degrees of freedom directions l_i forces respective moments are acting on input node i. The FE program provides the user the node number of the input node. However, the node number in the FE program n_i is usually not the same number as the internally used node number n_{Si}. Whenever a user defines input degrees of freedom at a specific node i the global degrees of freedom $\boldsymbol{q}_{\text{global\_inputs}}$ has to be looked up ex ante in the dof table *dof_table*. After the global dof of the inputs $q_{\text{global\_inputs}}$ are known, ones are inserted at the corresponding places at the input matrix, i.e., $\boldsymbol{B}_e(\boldsymbol{q}_{\text{global\_inputs}(j)}, j)) = 1$. However, due to the usage of constraints in the FE Model it is likely that an input acts on a dependent degree of freedom. Then with the help of (9) it is ensured that all inputs act at the independent dof by distribution of the input forces from dependent dofs to independent dofs with

$$\boldsymbol{B}_e^* = \left((\boldsymbol{Q}_d \cdot \boldsymbol{G}_M)^T + \boldsymbol{Q}_i^T \cdot \boldsymbol{Q}_i \right) \cdot \boldsymbol{B}_e. \tag{17}$$

Here, $\boldsymbol{Q}_d$ is a localization matrix which localizes the dependent dof $\boldsymbol{q}_d(t)$ in their position in the global dof vector $\boldsymbol{q}(t)$ and $\boldsymbol{q}_i(t)$ is a localization matrix which localizes the independent dof $\boldsymbol{q}_i(t)$ to their position in the global dof vector $\boldsymbol{q}(t)$. The same procedure also needs to be done for the outputs. All this information *dof_table*, $\boldsymbol{G}_M, \ldots$ needs to be extracted from the different FE packages in order to reduce the FE model with the developed methods in Morembs. If some information is missing, it is necessary to calculate the missing data with the available data or determine it by additional user definitions.

The projection spaces are then calculated with the system of the independent dof (12); see Sect. 3 for the different reduction approaches.

After the reduction, the projection spaces need to be enlarged in a way that the dependent and constraint dof are also contained in the projection space. First, the projection spaces $\boldsymbol{V}_i$ and $\boldsymbol{W}_i$ are enlarged by the number of dependent plus constraint dofs by adding zero columns

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{V}_i \\ \boldsymbol{V}_d \\ \boldsymbol{V}_c \end{bmatrix}, \qquad \boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_i \\ \boldsymbol{W}_d \\ \boldsymbol{W}_c \end{bmatrix}. \tag{18}$$

Now the constraint dof are set according to the boundary conditions $\boldsymbol{V}_c = \boldsymbol{W}_c = \boldsymbol{0}$. Second, because the back projection for dependent dof would still be zero, the projection space for the depend dof is built with the constraint equation (9) to $\boldsymbol{V}_d = \boldsymbol{G}_M \cdot \boldsymbol{V}_i - \boldsymbol{g}$ and $\boldsymbol{W}_d = \boldsymbol{G}_M \cdot \boldsymbol{V}_i - \boldsymbol{g}$. With this procedure, it is ensured that the dependent dof move as a linear combination of the independent dof. For oblique projection techniques, i.e., $\boldsymbol{V} \neq \boldsymbol{W}$, a block diagonalization, as described in [16] of the reduced system is done which ensures that the final reduced system simulated in the commercial multibody system codes is diagonal with an unit mass matrix and the square root of the eigenvalues $\bar{\omega}_i$ of the reduced system are on the diagonal of the reduced stiffness matrix. Such a block-diagonal structure is important for the performance of the simulation process of the reduced system. After the projection matrices are calculated the standard data of an elastic body (see [8]) can be reconstructed as explained in [1, 17, 18]. Afterward, the elastic body can be simulated with a standard multibody system simulation tool. In the multibody system simulation tool, the user has to choose joint conditions fitting the chosen reference frame.

3 Model reduction techniques

During the last decades, a number of model order reduction techniques have been proposed. An excellent overview of the different reduction techniques is given in [19]. The main model reduction techniques are modal reduction, moment-matching based on projection on Krylov-subspaces, SVD-based reduction techniques and combinations of those. Whereas modal reduction can easily be extended to a second-order system by solving the quadratic eigenproblem, moment-matching and SVD-based model reduction have originally been developed for first-order ODEs. However, for the model order reduction process in flexible multibody systems the preservation of the second-order structure is important. Moment-matching for second-order systems can be achieved by projection on a second-order Krylov-subspace [20]. Also, for SVD-based reduction techniques extensions to second-order systems exist [21]. Each of these methods is characterized by certain advantages and disadvantages. The following points are of special interest for the user:

- computability for large scale systems,
- stability preservation,
- quality of the reduced order model,
- knowledge about the error induced through reduction methods,
- preservation of the second-order structure during the reduction, and
- emphasizing a certain frequency range.

In this paper, modal reduction is only used for comparing the results.

3.1 Krylov-subspace based reduction techniques

Krylov-subspace based reduction is an efficient method to impose that certain moments between the original transfer function $\boldsymbol{H}(s)$ and the reduced transfer function $\bar{\boldsymbol{H}}(s)$ match. Overview articles about Krylov-subspace reduction are, e.g. [22] or [23]. These methods are widely used as a tool for the reduction of large scale systems, e.g., in simulation of Micro-Electro-Mechanical Systems (MEMS) [24]. The fact that Krylov-subspace reduction methods are iterative methods and can be applied to large scale models represents their decisive advantage whereas they become inflexible for systems with many inputs if a small model is required. Recently, Krylov-subspace based reduction techniques have also been used for the reduction of mechanical systems [25]. In this work, moment matching for second-order systems is used. The projection with second-order Krylov-subspaces keeps the second-order structure of the system. The conservation of the second-order structure is necessary for the calculation of the standard data of an elastic body because only with the standard data a simulation of the elastic body in a flexible multibody system simulation environment, e.g., Simpack, is possible. Second-order Krylov-subspace methods are necessary if the system is not proportionally damped [17]. In [26], a second-order Krylov-subspace is defined as

$$\mathcal{G}_r(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{G}) = \operatorname{colspan}\{\boldsymbol{P}_0, \boldsymbol{P}_1, \ldots, \boldsymbol{P}_{r-1}\} \quad \text{with} \tag{19}$$

$$\begin{cases} \boldsymbol{P}_o = \boldsymbol{G}, \ \boldsymbol{P}_1 = \boldsymbol{A}_1, \\ \boldsymbol{P}_i = \boldsymbol{A}_1 \cdot \boldsymbol{P}_{i-1} + \boldsymbol{A}_2 \cdot \boldsymbol{P}_{i-2}, \quad i = 2, 3, \ldots, r-1. \end{cases} \tag{20}$$

The input and output second-order Krylov-subspace of a second-order system around the zero expansion point $s_k = 0$ are $\mathcal{G}_{rI}(-\boldsymbol{K}_e^{-1} \cdot \boldsymbol{D}_e, -\boldsymbol{K}_e^{-1} \cdot \boldsymbol{M}_e, -\boldsymbol{K}_e^{-1} \cdot \boldsymbol{B}_e)$ and $\mathcal{G}_{rO}(-\boldsymbol{K}_e^{-T} \cdot \boldsymbol{D}_e^T, -\boldsymbol{K}_e^{-T} \cdot \boldsymbol{M}_e^T, -\boldsymbol{K}_e^{-T} \cdot \boldsymbol{C}_e^T)$, respectively. Moment matching can be achieved at arbitrary expansion points s_k by projection on the Krylov-subspaces $\mathcal{G}_{r^k I}^k(-\check{\boldsymbol{K}}_k^{-1} \cdot \check{\boldsymbol{D}}_k, -\check{\boldsymbol{K}}_k^{-1} \cdot \boldsymbol{M}_e, -\check{\boldsymbol{K}}_k^{-1} \cdot \boldsymbol{B}_e)$ and $\mathcal{G}_{r^k O}^k(-\check{\boldsymbol{K}}_k^{-T} \cdot \check{\boldsymbol{D}}_k^T, -\check{\boldsymbol{K}}_k^{-T} \cdot \boldsymbol{M}_e^T, -\check{\boldsymbol{K}}_k^{-T} \cdot \boldsymbol{C}_e^T)$ where $\check{\boldsymbol{D}}_k = 2s_k\boldsymbol{M}_e + \boldsymbol{D}_e$ and $\check{\boldsymbol{K}}_k = s_k^2\boldsymbol{M}_e + s_k\boldsymbol{D}_e + \boldsymbol{K}_e$ are used instead of $\boldsymbol{K}_e$ and $\boldsymbol{D}_e$, respectively. If the projection spaces $\operatorname{span}(\boldsymbol{V})$ and $\operatorname{span}(\boldsymbol{W})$ span the union of the different second-order Krylov-subspaces at different expansion points s_k, i.e.,

$$\operatorname{span}(\boldsymbol{V}) = \bigcup_{k=1}^{\hat{k}} \mathcal{G}_{rI}^k\left(-\check{\boldsymbol{K}}_k^{-1} \cdot \check{\boldsymbol{D}}_k, -\check{\boldsymbol{K}}_k^{-1} \cdot \boldsymbol{M}_e, -\check{\boldsymbol{K}}_k^{-1} \cdot \boldsymbol{B}_e\right), \tag{21}$$

$$\operatorname{span}(\boldsymbol{W}) = \bigcup_{k=1}^{\hat{k}} \mathcal{G}_{rO}^k\left(-\check{\boldsymbol{K}}_k^{-T} \cdot \check{\boldsymbol{D}}_k^T, -\check{\boldsymbol{K}}_k^{-T} \cdot \boldsymbol{M}_e^T, -\check{\boldsymbol{K}}_k^{-T} \cdot \boldsymbol{C}_e^T\right), \tag{22}$$

it is ensured that $2r^k$ moments are matched at every expansion point, meaning $2\sum_{i=0}^{k} r^k$ moments are matched in total. For proportionally damped systems, the calculation of the projection spaces $\operatorname{span}(\boldsymbol{V})$ and $\operatorname{span}(\boldsymbol{W})$ can be done with first-order Krylov-subspaces [20]. If only the first moments are matched at $s_i = 0 + i\omega_{\text{frm}}$ with first-order techniques, a Krylov-subspace based reduction is equivalent to the usage of frequency-response modes. How many moments are considered and at which expansion points the moments match, is in the responsibility of the user and needs to be chosen as input.

Error estimation is one of the important features for model order reduction techniques. An error bound for reduced systems obtained by moment matching was introduced in [27] for a Padé approximation via the Lanczos process. The performance of Krylov-subspace

based reduction methods clearly depends on the choice of expansion points s_k. Recently in [28], a method was proposed where the $\mathcal{H}_2$-norm of the error $\boldsymbol{H}_e$ is minimized if the moment matching conditions are chosen in an optimal way. Also, by using Laguerre functions, it is possible to find a single optimal expansion point [29, 30]. These two applications are under current research for their applicability to big industrial problems. Another error estimator as well as an expansion point selection strategy were shown in [31] and were extended to second-order systems by the authors, see [32]. In this method, called Second-Order Adaptive Global Arnoldi (SOAGA), the error estimator and automated selection of appropriate expansion points is based on the global Arnoldi process; see [33]. The calculation starts with a given set $\hat{S} = \{s_1, s_2, \ldots, s_i\}$ of expansion points. By minimizing the error in every iteration step, the algorithm builds a new basis at that expansion point with the maximum output error.

The development of general a-priori error bounds independent from the choice of expansion points is still an active field of research. However, following [34] error estimation can be achieved within a given frequency range $[\omega_{\min},\ \omega_{\max}]$ by using two different sets of expansion points in compliance with the frequency range condition. Another error estimator was introduced by [35] but its applicability to EMBS requires additional research.

3.2 SVD-based reduction techniques, reduction with Gramian matrices

For first-order systems, the well-known balanced truncation as explained in [36] has a-priori error bounds [37] and asymptotic stability is preserved in the reduced-order system. The controllability and observability Gramian matrices, $\boldsymbol{P}$ and $\boldsymbol{Q}$, of a system in the state-space form (5) are strongly related to balanced truncation techniques and are important matrices for model reduction. For second-order systems, Gramian matrices also play the decisive role in the development of a-priori error bounds. According to [38], second-order Gramian matrices identify the important positions and the important velocities in the input/output (I/O) map of a second-order system. For first-order systems, the Gramian matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ tell what are the most controllable and observable states of the system. For second-order systems, more than two Gramian matrices exist. The position controllability Gramian matrix $\boldsymbol{P}_p$ identifies the most easily controllable positions whereas the velocity observability Gramian matrix $\boldsymbol{Q}_p$ tells which are the most easily observable velocities. The most easily controllable position coordinates can be calculated by solving the minimization problem

$$\mathcal{J}^* = \min_{\dot{\mathbf{q}}_0 \in \mathbb{R}^N} \min_{\mathbf{u}(t) \in \mathcal{L}_2[-\infty,0]} \left\{ \mathcal{J} \triangleq \int_{-\infty}^{0} \boldsymbol{u}^T(t) \cdot \boldsymbol{u}(t) dt \right\} \tag{23}$$

$$\text{subject to} \quad \begin{cases} \boldsymbol{M}_e \cdot \ddot{\boldsymbol{q}}(t) + \boldsymbol{D}_e \cdot \dot{\boldsymbol{q}}(t) + \boldsymbol{K}_e \cdot \boldsymbol{q}(t) = \boldsymbol{B}_e \cdot \boldsymbol{u}(t), \\ \boldsymbol{q}(-\infty) = \boldsymbol{q}_{-\infty}. \end{cases}$$

The result $\mathcal{J}^*$ is the minimal energy needed to steer the system from initial position $\boldsymbol{q}_{-\infty}$ at time $t = -\infty$ to zero position $\boldsymbol{q}_0 = \boldsymbol{0}$ at time $t = 0$. This optimization problem does not depend on the velocities $\dot{\boldsymbol{q}}(t)$. The optimum solution is $\mathcal{J}^* = \boldsymbol{q}_0^T \cdot \boldsymbol{P}_p^{-1} \cdot \boldsymbol{q}_0$ where $\boldsymbol{P}_p = [\boldsymbol{I}\ \boldsymbol{0}] \cdot \boldsymbol{P} \cdot [\boldsymbol{I}\ \boldsymbol{0}]^T$ is the $N \times N$ upper left block of $\boldsymbol{P}$. A similar result can be obtained for the position observability Gramian matrix $\boldsymbol{Q}_p$ by using the dual system; see [17, 38]. In (23), the necessary energy to reach a given position $\boldsymbol{q}_0 = \boldsymbol{0}$ over all past inputs and initial velocities is minimized. If the system is balanced, i.e., $\boldsymbol{P}_p = \boldsymbol{Q}_p = \mathrm{diag}(\sigma_p)$, the previously defined Gramian matrices describe how the I/O energy is distributed among the positions. Analogously, the velocity Gramian matrices $\boldsymbol{P}_v$ and $\boldsymbol{Q}_v$ can be found by a slight

modification of the optimization problem; see [38]. These matrices describe how the I/O energy is distributed among the velocities and can be considered as the $N \times N$ lower right block of the controllability and observability Gramian matrices. Another important second-order Gramian matrix is $\boldsymbol{Q}_{pv} = \boldsymbol{M}_e^{-T} \cdot \boldsymbol{Q}_v \cdot \boldsymbol{M}_e^{-1}$, which is the position controllable Gramian matrix of the dual system if $\boldsymbol{C}_e^T$ is used as input matrix to the dual system. For the important special case, if input equals output, i.e., $\boldsymbol{B} = \boldsymbol{C}$, the matrix $\boldsymbol{Q}_{pv}$ is equal to $\boldsymbol{P}_p$. In [39] the different approaches for balanced truncation of second-order systems are described. However, not all of the proposed methods keep the reduced systems symmetric and stable. In addition, they only evaluate their methods in the frequency domain and the reverse transformation of their projector, which would allow an application to second-order systems, is still an unsolved problem; see [40]. In our approach, we do not explicitly balance the system but project it with the dominant eigenraum of the Gramian matrices.

Frequently, in mechanical systems, a certain frequency range is of special interest. The information about the interesting frequencies can be included in the reduction process by projecting the system with the dominant frequency-weighted version of the Gramian matrices $\boldsymbol{P}_p^{\omega}$ and $\boldsymbol{Q}_{pv}^{\omega}$. Frequency weighted Gramian matrices are defined in [17, 19]. In [17, 32] it is shown that if the projection space span$(\boldsymbol{V})$ consists of the dominant eigenvectors span$(\boldsymbol{V}_1)$ of $\boldsymbol{P}_p^{\omega} + \boldsymbol{Q}_{pv}^{w}$, which are associated with the largest eigenvalues σ_i^{PQ} of the sum of frequency-weighted position Gramian matrices then the frequency-weighted $\mathcal{H}_2$-error of the reduced system can be written as

$$\left\| \boldsymbol{H}_E(s) \cdot \boldsymbol{W}_i(s) \right\|_{\mathcal{H}_2}^2 \leq \frac{1}{2\pi} \max(\kappa, \kappa_T) \sum_{j=n+1}^{N} \sigma_j^{PQ}. \tag{24}$$

In (24), κ and κ_T are constants and $\boldsymbol{W}_i(s)$ is a frequency filter matrix of an ideal band pass filter (see [19]) which is used to emphasize a certain frequency range. The fact that the sum of neglected eigenvalues can be calculated as soon as the eigen decomposition of the Gramian matrices is known and causes the upper bound of the error to be known before the reduced system is obtained. The sum of neglected eigenvalues $\sum_j \sigma_j^{PQ}$ can be used to determine the size m of the reduced order model because the error is bounded below the sum of the neglected eigenvalues. Usually, the eigenvalues of the Gramian matrix decay rapidly in mechanical systems. As a consequence, the first neglected eigenvalue is the dominant share in the sum of neglected eigenvalues and allows a fully automated reduction process; see [10].

For small asymptotically stable systems, the Gramian matrices $\boldsymbol{P}_p^{w}$ and $\boldsymbol{Q}_{pv}^{w}$ can be calculated by evaluating a matrix logarithm in addition to the solution of a suitable Lyapunov equation; see [19]. Direct solution of the Lyapunov equation is only possible for small- to medium-scale models because the solution requires $\mathcal{O}(n^3)$ operations and the storage requirement is $\mathcal{O}(n^2)$. For large-scale models, somehow the subspace of dominant eigenvectors of the Gramian matrix has to be generated.

One approach uses the two-step approach explained in [9]. According to this approach, a medium-scale model is acquired in a first step with the help of modal or Krylov-subspace based methods. Subsequently, the Gramian matrix of the medium-scaled model is calculated and then used as an approximation for the Gramian matrix of the original system.

A second approach for approximating the Gramian matrix is introduced in [17]. According to this method, the Gramian matrix is numerically approximated by a POD based procedure. However, because the Gramian matrix is approximated some heuristics is included and the strict error bound is lost but the approximated error estimator usually is close to the

correct one. That is why in all our examples the estimated error and the real error coincide well.

A third approach is the calculation of the second-order Gramian matrices with iterative methods like the LR-ADI or Smith algorithm; see, e.g. [40, 41]. These methods are under current research for their applicability to the reduction of flexible multibody systems.

3.3 Modal reduction

Modal reduction is currently the widely used state of the art reduction method for flexible multibody systems. Here, we use modal reduction only to compare the other model reduction techniques and to discuss the difference to modal reduction. The projection space span($\boldsymbol{V}$) consists of selected eigenvectors of the system. For the calculation of the eigenmodes, the ansatz $\boldsymbol{q} = e^{\lambda_i t}\boldsymbol{\Phi_i}$ is used with $\lambda_i \in \mathbb{C}$ and $\boldsymbol{\Phi}_i \in \mathbb{C}^n$ leading to the quadratic eigenproblem $(\lambda_i^2 \boldsymbol{M}_e + \lambda_i \boldsymbol{D}_e + \boldsymbol{K}_e) \cdot \boldsymbol{\Phi}_i = \boldsymbol{0}$. If the space spanned by all eigenvectors is considered as projection space $\boldsymbol{\Phi} = [\Phi_1\ \Phi_2\ \ldots\ \Phi_n] = \boldsymbol{V}$, the flexible part of system (2), after the definition of appropriate interfaces can be written in modal coordinates $\tilde{\boldsymbol{q}}(t)$ as

$$\begin{aligned} \boldsymbol{I} \cdot \ddot{\tilde{\boldsymbol{q}}}(t) + \operatorname{diag}(2\omega_i \xi_i) \cdot \dot{\tilde{\boldsymbol{q}}}(t) + \operatorname{diag}\left(\omega_i^2\right) \cdot \tilde{\boldsymbol{q}}(t) &= \tilde{\boldsymbol{B}}_e \cdot \boldsymbol{u}(t), \\ \boldsymbol{y}(t) &= \tilde{\boldsymbol{C}}_e \cdot \tilde{\boldsymbol{q}}(t). \end{aligned} \tag{25}$$

For the system (25), the transfer matrix is $\tilde{\boldsymbol{H}}(i\omega) = \sum_{i=0}^{N} [\tilde{\boldsymbol{C}}_e]_{*i} [\tilde{\boldsymbol{B}}_e]_{i*} / (-\omega^2 + \omega_i^2 + 2i\omega\omega_i\xi_i)$ with $[\tilde{\boldsymbol{C}}_e]_{*j} [\tilde{\boldsymbol{B}}_e]_{i*}$, being the dyadic product of the ith column of $\tilde{\boldsymbol{C}}_e$ and the i-th row of $\tilde{\boldsymbol{B}}_e$. In $\tilde{\boldsymbol{H}}(i\omega)$, we see that the denominator of every summand of $\tilde{\boldsymbol{H}}(i\omega)$ near ω_i is getting small and the summand can get high. The approximation idea is to truncate the series after the kth coefficient and to choose k with the aim of providing enough distance of the $k+1$ eigenfrequency from the largest excitation frequency ω_{ex}. Sometimes (see e.g. [42]) the recommendation is given to consider all the eigenmodes up to two times the maximum excitation frequency. As the transfer function indicates, this procedure leads to inadequate models taking into consideration that the value of a single summand also depends on the values of $[\tilde{\boldsymbol{C}}_e]_{*j}$ and $[\tilde{\boldsymbol{B}}_e]_{i*}$. As far as the limit case is concerned, i.e., one mode is neither controllable nor observable, the numerator is zero and the summand can be neglected without changing the transfer behavior of the system. It is well known from balanced truncation reduction techniques (see e.g. [19]) that the Hankel singular values (HSV) are strongly related to the error which is induced by reduction. In [43], it is shown that for proportionally damped systems in modal form it is possible to calculate approximate HSV by

$$\tilde{\sigma}_i = \frac{\|[\tilde{\boldsymbol{C}}_e]_{*i}\|_2 \, \|[\tilde{\boldsymbol{B}}_e]_{i*}\|_2}{4\,\omega_i\,\xi_i}, \tag{26}$$

which are a measurement to which extent a state is involved in the energy transfer from a given input to a certain output; see [17]. With the approximate HSV, the important modes can be selected. However, if the approximated HSV are sorted in descending order they have no steep decay as the HSV of a balanced mechanical system because the spatial distribution of loads is not considered; see [44]. Modal reduction can be improved by using only the important modes and extending the projection space span($\boldsymbol{V}$) with constraint modes or attachment modes similar to sub-structuring techniques in structural dynamics; see e.g. [45, 46]. The component mode method developed in [7] was utilized in [4] for the reduction of flexible multibody system. A combination of fixed boundary nodes plus some "constraint modes" which account for local effects at boundaries is used in their work. Frequency response

modes $\boldsymbol{\Phi}_{\text{frm}} := \boldsymbol{Q} = (-\omega_{\text{frm}}^2 \boldsymbol{M}_e + \boldsymbol{K}_e)^{-1} \cdot \boldsymbol{B}_e$ (see [44]) are a generalization of constraint modes and attachment modes. For fixed systems and an excitation frequency $\omega_{\text{frm}} = 0$, the inertia relief attachment modes are calculated which are related to the mode acceleration method [47] and ensure that the static solution is correct, i.e., $\tilde{\boldsymbol{H}}(0) = \boldsymbol{H}(0)$. One problem is the combination of eigenmodes and correction modes. Sometimes, the combined projection space do not have numerically full rank. That is why it is not known a-priori which and how many frequency response modes should be used in combination with the eigenmodes. In addition, error estimators are not available for modal reduction which makes modal reduction difficult to automate.

4 Used programs and data flow

During the last years, an extensive set of new model order reduction techniques has been developed, implemented, and tested by the authors. Each is characterized by certain advantages and disadvantages. By giving the user an easy-to-use tool for testing several model order reduction techniques and giving assistance how to choose an appropriate reduction space, the simulation process of EMBS can be improved. In addition, the development of automated reduction techniques is one essential topic in the development of parametric model reduction and represents also further goals. The methods are implemented in the C++ code Morembs and in a Matlab implementation MatMorembs; see Fig. 8. Morembs is able to handle data from several FE programs. Up to now data from the FE programs Ansys, Abaqus and Permas can be used for model reduction. Morembs provides either an SID-File for Simpack containing the reduced elastic body or the system matrices of the reduced system as output. In addition, there exist a direct converter to Neweul-M^2. The C++ implementation is

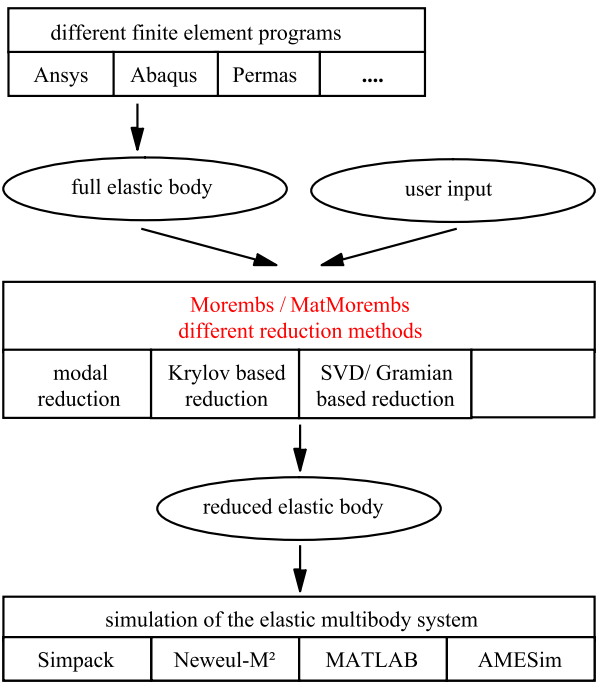


Fig. 8 Data flow and programs used in the simulation process

based on advanced numeric libraries leading to short computation times and can be used for the reduction of real industrial problems with more than 100,000 degrees of freedom. The MatMorembs implementation is appropriate for testing and debugging new algorithms before the methods are added to Morembs. Moreover, most of the new algorithms developed by mathematicians are written in Matlab, e.g. [40]. Furthermore, given the steeper learning curve, the MatMorembs version is easier accessible for students and beginners to become familiar with the concepts.

5 Example and results

In Sect. 2.2, we concluded that the spider web of beams with light stiff beams is an appropriate interface model. We now compare the different reduced models where the interface is always modeled with light stiff beams. In Fig. 9, the relative error

$$\epsilon(\omega) = \left\| \boldsymbol{H}(\omega) - \bar{\boldsymbol{H}}(\omega) \right\|_F / \left\| \boldsymbol{H}(\omega) \right\|_F \tag{27}$$

in the Frobenius norm for modal-based reduction techniques is plotted. By choosing eigenmodes based on approximated HSV (26), we get a small improvement in comparison to a reduction where the first 40 eigenmodes are used; see Fig. 9. However, by using the best 40 eigenmodes plus one frequency response mode at $\omega_{\text{frm}} = 1$ Hz a model of size 46 is obtained with which the results are improved, however, the projection space is also extended. In Fig. 9, the crosses are the results obtained with the C++-implementation of Morembs. The MatMorembs and the C++ version Morembs lead to the same reduction results which is a hint for the correct implementation.

For testing the Krylov-subspace based reduction method, the newly suggested SOAGA algorithm from [32] is used. Two medium-scaled systems of order 132 are calculated with two different expansion point sets $\bar{S}$ and $\check{S}$ equidistantly spaced between 1 and 1000 Hz. Both calculations stopped after 12 Frobenius orthonormal bases were calculated for input and output. Afterward, the Frobenius orthonormal bases were mass orthogonalized to satisfy the dynamic boundary condition. In Fig. 10, the relative error $\epsilon(\omega)$ is plotted for the SOAGA algorithm. Additionally, the error estimator mentioned in Sect. 3.1 is plotted, labeled $\check{\varepsilon}$ SOAGA. The good agreement between estimator for the approximated solution and the true error can be seen. Besides this, the method stands out due to its low error rate. Due to the low error of the SOAGA algorithm, it is admissible to use this medium scaled model for the calculation of a Gramian matrix based reduction. With the SVD based reduction method it is possible to reduce the medium scale system further to size $n = 35$. The error, labeled "Kry. + Gram," is still small in the whole interesting frequency range. This model

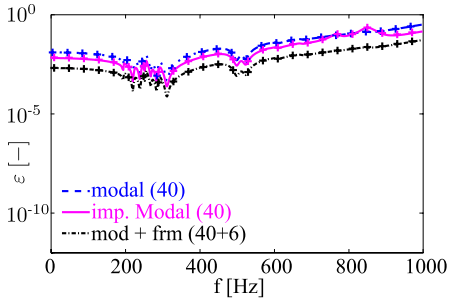


Fig. 9 Comparison between MatMorembs and C++ implementation of different reduction methods

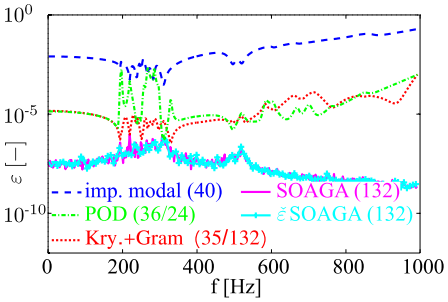


Fig. 10 Relative reduction error using different reduction approaches

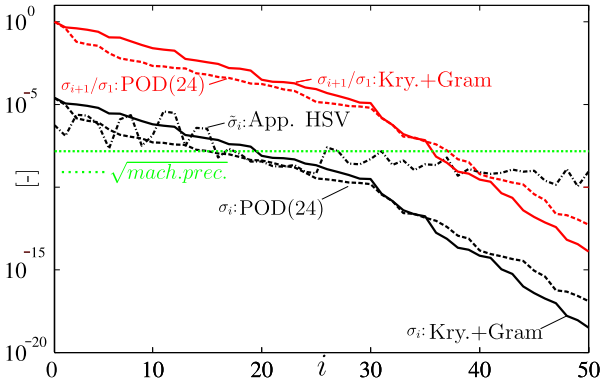


Fig. 11 Error indicator used for the reduction process

is much more accurate than the model of size 40 obtained by an improved modal reduction technique, labeled "imp. modal." The improved modal reduction method chooses those 40 modes with the biggest approximated HSV (25) and is the optimum modal reduction with 40 modes. The other SVD based reduction method, i.e., numerical approximation of the frequency weighted Gramian matrices with POD is also plotted and labeled "POD." The size of the reduced order model $n = 36$ was automatically determined by the reduction process and no user input is necessary. Both SVD based reduction methods leads to the same results.

In [17, 32], an error expression based on neglected eigenvalues of the Gramian matrices is introduced. In Fig. 11, the approximated HSV $\tilde{\sigma}_i$ (26), the eigenvalues of the Gramian matrix $\boldsymbol{P}_p^{\omega}$ and the error indicator σ_{i+1}/σ_1 for the SVD based reduction methods are plotted over the dimension i of the reduced order model. Both the POD based and the two-step approach calculate a similar error indicator. The rapid decay of the eigenvalues σ_i of the Gramian matrices can be seen. This means that the error indicator σ_{i+1}/σ_1 used for the automated reduction procedure is admissible. It can be seen that the approximated HSV $\tilde{\sigma}_i$ decay slower than the eigenvalues of the Gramian matrices.

5.1 Comparisons in the time domain

The time domain is of great importance in flexible multibody dynamics. For this purpose, dynamic simulations are considered and the different approaches are compared with respect to accuracy. The body is clamped to the surrounding at node 2 and the body is actuated at

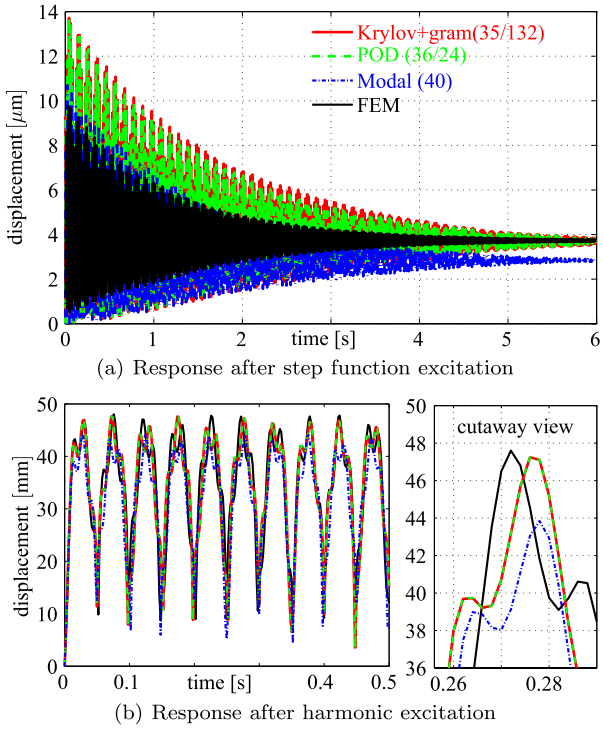


Fig. 12 Results of the dynamic simulation

(a) Response after step function excitation

(b) Response after harmonic excitation

node 1 with a step function force $\boldsymbol{F}_{\text{step}}$ and a harmonic force $\boldsymbol{F}_{\text{harm}}$

$$\boldsymbol{F}_{\text{step}} = -100 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} N, \qquad \boldsymbol{F}_{\text{harm}} = -100 \sin(2 \cdot 10\pi t) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} N.$$

As a reference frame for the dynamic simulation a Buckens-frame is chosen such that the origin of the reference frame coincides with the center of gravity; compare Fig. 4 black. The reduced order models are simulated with the multibody dynamics simulation tool Simpack. The resulting system of ODEs is solved numerically with the standard solver of Simpack. In Figs. 12(a) and 12(b), the displacement of the actuated lower plate is shown for models of different size obtained with different reduction methods.

Furthermore, we compare these results with the translation of the nonlinear finite element model. The accuracy of the reduced order model is good. It shows approximately the same response as the nonlinear FEM model with 35,000 degrees of freedom. The reduced order model of size 35 obtained with the two step frequency-weighted Gramian matrix approach has the same accuracy as a model of size 36 obtained by a POD based model reduction. As the error in the frequency response already indicated in Fig. 10, the results from the modal model of order 40 are worse compared to the others. Besides this, the static response after the step response does not fit. The simulation of the flexible multibody system in Simpack is about 90 times faster than the simulation in Ansys.

6 Conclusion

During the last years, an extensive set of new model order reduction techniques have been developed, implemented, and tested at our institute. Each is characterized by certain advantages and disadvantages. The convergence for modal reduction is usually very slow and an extension of the projection space with constraint or other modes is a crucial task that requires much experience and insight into the specific problem. Moreover, the choice of modes based on the approximated HSV did not lead to a big improvement. By using reduction techniques based on frequency weighted Gramian matrices, the error of the reduced system is known a-priori. With these methods, excellent reduction results are obtained. Furthermore, because the error is known a-priori the necessary order of the reduced model can be chosen which simplifies and supports the reduction process for the user. However, the calculation of Gramian matrices is expensive. For large scale models, the dominant eigenspace of the Gramian matrices is approximated by two different approaches which both showed good results. For the two-step approach, a new second-order adaptive global Arnoldi algorithm was used as first reduction step. The SOAGA algorithm further assists the user in the reduction process by providing an error estimator for Krylov based reduction. By giving the user an easy-to-use tool for testing several model order reduction techniques and giving assistance how to choose an appropriate reduction space, the simulation process of EMBS can be improved. This is demonstrated for a demanding technical system in the frequency domain as well as in the time domain. The use of a flexible multibody system speeds up the calculation by a factor of 90 compared to the simulation in Ansys.

However, part of the methods are still in the development stage and some time is needed for improvements and testing. The Matlab implementation MatMorembs speeds up the testing and debugging process for new algorithms. Furthermore, given the steeper learning curve, MatMorembs is easier accessible for students and beginners to become familiar with the concepts. However, the C++ implementation Morembs is based on advanced numeric libraries leading to short computation times and can be used for the reduction of real industrial problems with more than 100,000 degrees of freedom.

Different interface definitions were introduced. Appropriate interface definitions are necessary for good simulation results. However, the model reduction code Morembs can handle all the different interface definitions in a mathematically correct way leaving the user the choice. The identification of the most appropriate interface definition always depends on the actual system and requires verification by measurements.

Acknowledgements We want to thank the reviewers for their thorough review and highly appreciate the comments and suggestions, which significantly contributed to improving the quality of the publication.

References

1. Schwertassek, R., Wallrapp, O.: Dynamik flexibler Mehrkörpersysteme. Vieweg, Braunschweig (1999) (in German)
2. Shabana, A.A.: Dynamics of Multibody Systems. Cambridge University Press, Cambridge (1998)
3. Wu, S.-C., Haug, E.J., Kim, S.-S.: A variational approach to dynamics of flexible multibody systems. Mech. Struct. Mach., **17**, 3–32 (1989)
4. Cardona, A., Géradin, M.: A superelement formulation for mechanism analysis. Comput. Methods Appl. Mech. Eng. **100**, 1–29 (1992)
5. Schwertassek, R., Wallrapp, O., Shabana, A.: Flexible multibody simulation and choice of shape functions. Nonlinear Dyn. **20**, 361–380 (1999)
6. Wallrapp, O., Wiedemann, S.: Comparison of results in flexible multibody dynamics using various approaches. Nonlinear Dyn. **34**, 189–206 (2003)

7. Craig, R., Bampton, M.: Coupling of substructures for dynamic analyses. AIAA J. **6**, 1313–1319 (1968)
8. Wallrapp, O.: Standardization of flexible body modeling in multibody system codes, Part I: Definition of standard input data. Mech. Struct. Mach. **22**, 283–304 (1994)
9. Lehner, M., Eberhard, P.: A two-step approach for model reduction in flexible multibody dynamics. Multibody Syst. Dyn. **17**, 157–176 (2007)
10. Fehr, J., Eberhard, P., Lehner, M.: Improving the reduction process in flexible multibody dynamics by the use of 2nd order position Gramian matrices. In: Proceedings 6th EUROMECH Nonlinear Dynamics Conference, St. Petersburg, Russia (2008)
11. Lu, J., Ast, A., Eberhard, P.: Modeling and active vibration control of flexible structures using multibody system theory. In: Proceedings of ICMEM, Wuxi, China (2007)
12. Nikravesh, P.E., Lin, Y.-S.: Use of principal axes as the floating reference frame for a moving deformable body. Multibody Syst. Dyn. **13**, 211–231 (2005)
13. Heckmann, A.: On the choice of boundary conditions for mode shapes in flexible multibody systems. Multibody Syst. Dyn. **23**, 141–163 (2010)
14. Release 11.0 Documentation for Ansys. http://www.ansys.com (2009)
15. Dresig, H., Holzweißig, F.: Maschinendynamik, vol. 5. Springer, Berlin (2004) (in German)
16. Bavely, C.A., Stewart, G.W.: An algorithm for computing reducing subspaces by block diagonalization. SIAM J. Numer. Anal. **16**, 359–367 (1979)
17. Lehner, M.: Modellreduktion in elastischen Mehrkörpersystemen. In: Dissertation, Schriften aus dem Institut für Technische und Numerische Mechanik der Universität Stuttgart, vol. 10. Shaker, Aachen (2007) (in German)
18. MSC.Software Corporation: Theoretical Background of ADAMS/Flex, Part of the ADAMS/Flex Training Course, 2003rd ed. MSC Software, Santa Ana (2003)
19. Antoulas, T.C.: Approximation of Large-Scale Dynamical Systems. SIAM, Philadelphia (2005)
20. Salimbahrami, B., Lohmann, B.: Order reduction of large scale second-order systems using Krylov subspace methods. Linear Algebra Appl. **415**, 385–405 (2006)
21. Stykel, T.: Balanced truncation model reduction of second-order systems. In: Troch, I., Breitenecker, F. (eds.) 5th Vienna Symposium on Mathematical Modelling, Vienna (2006)
22. Bai, Z.: Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. Appl. Numer. Math. **43**, 9–44 (2002)
23. Beattie, C.A., Gugercin, S.: Krylov-based model reduction of second-order systems with proportional damping. In: 44th IEEE Conference on Decision and Control, 2005 European Control Conference, pp. 2278–2283 (2005)
24. Han, J.S., Rudnyi, E.B., Korvink, J.: Efficient optimization of transient dynamic problems in MEMS devices using model order reduction. J. Micromech. Microeng. **15**, 822–832 (2005)
25. Gawronski, W.K.: Advanced Structural Dynamics and Active Control of Structures. Springer, New York (2004)
26. Salimbahrami, S.B.: Structure preserving order reduction of large scale second order models. Ph.D. Thesis, Technische Universität München (2005)
27. Bai, Z., Slone, R., Smith, W., Ye, Q.: Error bound for reduced system model by Pade approximation via the Lanczos process. IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **18**, 133–141 (1999)
28. Gugercin, S., Antoulas, A.C., Beattie, C.A.: $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. SIAM J. Matrix Anal. Appl. **30**, 609–638 (2008)
29. Eid, R., Salimbahrami, B., Lohmann, B.: Equivalence of Laguerre-based model order reduction and moment matching. IEEE Trans. Autom. Control **52**, 1104–1108 (2007)
30. Eid, R.: Time Domain Model Reduction by Moment Matching. Dr. Hut, München (2009)
31. Chu, C.-C., Lai, M.-H., Feng, W.-S.: MIMO Interconnects order reductions by using the multiple point adaptive-order rational global Arnoldi algorithm. IEICE Trans. Electron. **E89-C**, 792–802 (2006)
32. Fehr, J., Eberhard, P.: Error-controlled model reduction in flexible multibody dynamics. J. Comput. Nonlinear Dyn. **5**, 031005 (2010)
33. Heyouni, M., Jbilou, K.: Matrix Krylov subspace methods for large scale model reduction problems. Appl. Math. Comput. **181**, 1215–1228 (2006)
34. Grimme, E.J.: Krylov projection methods for model reduction. Ph.D. Thesis, University of Illinois at Urbana-Champaign (1997)
35. Konkel, Y., Farle, O., Dyczij-Edlinger, R.: Ein Fehlerschätzer für die Krylov-Unterraum basierte Ordnungsreduktion zeitharmonischer Anregungsprobleme. In: Lohmann, B., Kugi, A. (eds.) Tagungsband GMA-Fachausschuss 1.30 "Modellbildung, Identifikation und Simulation in der Automatisierungstechnik", VDI/VDE-GMA, Technische Universität Wien, Institut für Automatisierungs- und Regelungstechnik, Salzburg (2008) (in German)
36. Laub, A.J., Laub, A.J., Heath, M.T., Paige, C.C., Ward, R.C.: Computation of system balancing transformations. In: 25th IEEE Conference on Decision and Control, Athens, Greece, pp. 548–553 (1986)

37. Enns, D.: Model reduction with balanced realizations: an error bound and a frequency weighted generalization. In: Proceedings of the 23rd IEEE Conference on Decision and Control, Las Vegas, 1984, pp. 127–132. IEEE, New York (1984)
38. Chahlaoui, Y., Lemonnier, D., Vandendorpe, A., Dooren, P.V.: Second-order balanced truncation. Linear Algebra Appl. **415**, 378–384 (2006)
39. Reis, T., Stykel, T.: Balanced truncation model reduction of second-order systems. Math. Comput. Model. Dyn. Syst. **14**, 391–406 (2008)
40. Benner, P., Saak, J.: Efficient balancing based MOR for second order systems arising in control of machine tools. In: Troch, I., Breitenecker, F. (eds.) Proceedings of the MathMod 2009 (2009)
41. Kruszinski, L.: Implementierung und Test neuer Methoden zur Approximation des dominanten Eigenraums von Gramschen Matrizen 2ter Ordnung. Thesis DIPL-130, Institute of Engineering and Computational Mechanics, University of Stuttgart (2008) (in German)
42. Gasch, R., Knothe, K.: Strukturdynamik Band 2—Kontinua und ihre Diskretisierung (in German). Springer, Berlin (1989)
43. Gawronski, W.K.: Dynamics and Control of Structures—a Modal Approach. Springer, Berlin (1998)
44. Dietz, S.: Vibration and Fatigue Analysis of Vehicle Systems Using Component Modes. Fortschritt-Berichte VDI, reihe 12(401). VDI, Düsseldorf (1999)
45. Graig, R.R.J.: Structural Dynamics. Wiley, New York (1981)
46. Koutsovasilis, P., Beitelschmidt, M.: Comparison of model reduction techniques for large mechanical systems. Multibody Syst. Dyn. **20**, 111–128 (2008)
47. Dickens, J.M., Nakagawa, J.M., Wittbrodt, M.J.: A critique of mode acceleration and modal truncation augmentation methods for modal response analysis. Comput. Struct. **62**, 985–998 (1997)