A physics-based model reduction approach for node-to-segment contact problems in linear elasticity

Diana Manvelyan1 | Bernd Simeon2 | Utz Wever3

1Siemens AG Corporate Technology, Munich, Germany
2Department of Mathematics, TU Kaiserslautern, Germany
3Independent, Munich, Germany

Abstract

The article presents a novel model order reduction method for mechanical problems in linear elasticity with nonlinear contact conditions. Recently, we have proposed an efficient reduction scheme for the node-to-node formulation [Manvelyan et al., Comput Mech 68, 1283–1295 (2021)] that leads to linear complementarity problems (LCPs). Here, we enhance the underlying contact problem to a node-to-segment formulation, which leads to quadratic inequalities as constraints. The adjoint system corresponds to a nonlinear complementarity problem that describes the Lagrange multiplier. The latter is solved by a Newton-type iteration based on a LCP solver in each time step. Since the maximal set of potential contact nodes is predefined, an additional substructuring by Craig–Bampton can be performed. This contact treatment turns out to be necessary and allows exclusion of the Lagrange multipliers and the nodal displacements at contact from reduction. The numerical solutions of the reduced contact problem achieve high accuracy and the dynamic contact carries over the behavior of the full order model. Moreover, if the contact area is small compared to the overall structure, the reduction scheme performs very efficiently. The performance of the resulting reduction method is assessed on two 2D computational examples from linear elasticity.

KEYWORDS

adjoint method, digital twin technology, dynamic contact, model order reduction, Newton-like iteration, nonlinear complementarity problem

1 | INTRODUCTION

During the last couple of years, “digital twins” appear to be one of the key concepts toward digitalization. They provide computer-aided assistance for real-world models allowing to monitor their state at specified positions or make prediction for the future states.1 Model order reduction realizes digital twins by enabling real-time execution. It reduces the degrees of freedom while keeping the essential physical properties of the model. Moreover, it enables the reusability of such models for other scenarios. An extensive literature is devoted to model order reduction applied for different kinds of partial differential equations, see, for example, References 2-7. Model reduction methods are typically categorized between data-driven methods8-10 or physics-based methods such as balanced truncation,11-13 Krylov subspace methods,14-19 and modal reduction.4
Linear-elastic contact problems represent variational problems under unilateral constraints\textsuperscript{20,21} and are often nonlinear due to the unknown moving contact interface, see, for example, References 20 and 22-24 for introduction to contact mechanics. Those problems are commonly solved by the penalty method\textsuperscript{25} or the augmented Lagrange method\textsuperscript{26} after the discretization. Despite their advantages, these methods turn out to be inconvenient for designing reduction schemes. Especially, the nonlinear constraints complicate the procedure of reduction, since in that case reproducing the contact shape behavior in the reduced system gets more challenging. First results on reduction approaches in this matter can be found in Reference 27, where a Signorini contact problem is reduced by a POD-based approach. Furthermore, a data-based reduction method is presented in Reference 28, which reduces not only the state variables but also the additional constraints and thus is applicable for elastic problems with comparatively large contact shape. Other reduction methods for contact problems include Reference 29, which performs the corresponding reduction in the frequency domain and a reduction approach for nonlinear parametric contact problems in flexible multibody dynamics.\textsuperscript{20}

In this work, we present a novel physics-based reduction method for linear-elastic mechanical problems with nonlinear contact conditions. The algorithm does not require simulation data and is highly efficient for contact problems with a small contact zone compared to the total structure. The contact problems of interest are discretized by the finite element method, such that the contact shape has a piece-wise linear description (segments). In case of a linear contact condition, a reduction scheme for node-to-node contact problems was introduced in Reference 31. Such problems allow, however, only normal contact, that is, stick and release at the contact interface in the normal direction without any tangential movement. In the current work, we extend the novel reduction method for the class of node-to-segment contact problems.\textsuperscript{20} This discretization technique of the contact condition is widespread in the field of contact mechanics since it can handle more flexibility in terms of a sliding movement at the contact zone, see References 32-34. Solvers for such contact problems have also been integrated into nonlinear finite element commercial software.\textsuperscript{35,36} For early implementations of node-to-segment techniques, we refer to References 37 and 38.

The essential components of our new algorithm are described as follows. We derive a quadratic formulation for the node-to-segment constraint, which is based on an angular function and describes the nonpenetration condition between a contact node and a contact segment (general formulations of the distance function lead to more advanced nonlinearity such as the square-root function). Due to this formulation, the Kuhn–Tucker condition is linear in the state variable and therefore the primal system can be resolved with respect to the displacement vector. The latter leads to a decoupling of the primal system from the adjoint system. Using this fact, we reduce only the primal system and keep the adjoint system.

The projection matrix can be computed in the offline phase based on the system matrices and the position vector of the external load (without trajectories of the displacements). We note, that the maximal set of the potential contact nodes has to be predefined and does not affect the reduction matrix, because the reduction matrix is computed only for the area outside of the potential contact zone. This step can be achieved by a partitioning of the overall displacement degrees of freedom in the flavor of the Craig–Bampton method,\textsuperscript{39} which separates the interior nodes, that is, the slave nodes, from the contact nodes, that is, the master nodes. Due to the nonlinear constraints, the system stiffness matrix depends on the Lagrange multipliers. However, the partitioning of the nodes provides linear system block-matrices for the slave nodes allowing an accurate reduction. After the reduced slave nodes are computed, the nonlinear block-matrices of the master nodes are utilized in static condensation,\textsuperscript{40} which recovers the coupling between the master and slave nodes.

Once the reduced primal system is established, we turn to the adjoint space of the reduced displacements. In the adjoint space the only unknowns that we seek are the Lagrange multipliers, which are described by a time-dependent nonlinear complementarity problem (NCP). The latter is approximated by a sequence of linear complementarity problems (LCPs), which for each time step can be solved by a Newton-like iteration method. A quadratic convergence rate for the iterations is guaranteed due the general theoretical framework.\textsuperscript{41-43} In combination with the assumption, that the number of the constraints is small, the convergence rate enables the proposed reduction algorithm to solve contact problems in real time. To put it in a nutshell, our proposed reduction scheme combines Krylov subspace methods, the Craig–Bampton partitioning technique and Newton-like methods in combination with adjoint methods.

The overall methodology is general in nature, but within the scope of the manuscript the reduction approach has been worked out in detail and tested in the planar case only. Our contribution is structured as follows: In Section 2, the variational formulation of dynamic contact problems with quadratic constraints is derived. The common node-to-segment technique is described in Section 3 and adapted to our case. Furthermore, Sections 4 and 5 discuss solution methods for the resulted static and dynamic contact problem. The novel reduction method is outlined in Section 6. The performance of the approach is demonstrated by means of one verification example and two numerical applications in Section 7. Finally, a conclusion is drawn in Section 8.
2 | GENERAL PROBLEM SETTING IN CONTACT MECHANICS

In this article, we assume a frictionless, adhesive-free normal contact in combination with small deformation theory and a linear-elastic material. Moreover, following holds:

(i) The Lagrange multiplier method is used to enforce the nonpenetration condition.
(ii) The finite element method is applied for the spatial discretization.
(iii) The node-to-segment contact technique is used for the discretization of the contact zone.

We do not consider approaches like the augmented Lagrangian method or the Nitsche method. The same holds for additional friction effects. In contrast to Reference 31, where linear constraints were derived with node-to-node discretization technique, in the current work, we use the node-to-segment methodology leading to quadratic inequalities as constraints. This technique will be explained in greater detail in Section 3.

2.1 Strong formulation of the dynamic contact problem

We introduce some notational preliminaries. The initial undeformed solid \( \Omega \subset \mathbb{R}^d, d = 2 \) or \( d = 3 \) is a disjoint union of bounded connected subdomains \( \Omega_i, \ i \in I \), that is, \( \Omega = \bigcup_{i \in I} \Omega_i \) with \( \Omega_i \cap \Omega_j = \emptyset \) for \( i \neq j \). On \( \Omega \) we consider a frictionless, adhesive-free normal multi-body contact problem and a linear-elastic material describing the displacement field \( \mathbf{u}(x, t) \in \mathbb{R}^d \) and the contact pressure \( p(x, t) \in \mathbb{R} \) where \( x \in \Omega \) is the spatial variable and \( t \in [t_0, T] \) stands for the temporal variable. The restriction of the displacement field on the corresponding domain we denote as \( \mathbf{u}_i = \mathbf{u}_{|\Omega_i} \). The boundary of each elastic body is decomposed into \( \partial \Omega_i = \Gamma_D^i \cup \Gamma_N^i \cup \Gamma_C^i \), where the latter represents the contact interface. The contact problem in strong form is then given by a dynamic boundary value problem (BVP): For all \( i \in I \) and for the corresponding initial data \( \mathbf{u}_i(\cdot, t_0) = \mathbf{u}_i^0, \mathbf{u}_i(\cdot, t_0) = \dot{\mathbf{u}}_i^0 \) find \( \mathbf{u}_i : \Omega_i \rightarrow \mathbb{R}^d \) such that

\[
\begin{align*}
\rho_i \mathbf{u}_{i,t} - \text{div} \sigma(\mathbf{u}_i) &= F_i \quad \text{in} \ \Omega_i \times [t_0, T], \\
\mathbf{u}_i &= \mathbf{0} \quad \text{on} \ \Gamma_D^i \times [t_0, T], \\
\sigma(\mathbf{u}_i) : \mathbf{n}_i &= \tau_i \quad \text{on} \ \Gamma_N^i \times [t_0, T],
\end{align*}
\]

where the constant \( \rho_i \) denotes the mass density of the body \( \Omega_i \), \( F_i(x, t) \in \mathbb{R}^d \) the volume force and \( \tau_i(x, t) \in \mathbb{R}^d \) the surface traction. The stress tensor \( \sigma(\mathbf{u}_i) \in \mathbb{R}^{d \times d} \) and the linearized strain tensor \( \mathbf{e}(\mathbf{u}_i) \in \mathbb{R}^{d \times d} \) are given by

\[
\begin{align*}
\sigma(\mathbf{u}_i) &= \frac{E_i}{1 + \nu_i} \mathbf{e}(\mathbf{u}_i) + \frac{\nu_i E_i}{(1 + \nu_i)(1 - 2\nu_i)} \text{trace}(\mathbf{e}(\mathbf{u}_i)) \mathbf{I}, \\
\mathbf{e}(\mathbf{u}_i) &= \frac{1}{2}(\nabla \mathbf{u}_i + \nabla \mathbf{u}_i^T).
\end{align*}
\]

with Young’s modulus \( E_i \geq 0 \), Poisson’s ratio \(-1 < \nu_i < 0.5\), and the identity matrix \( \mathbf{I} \in \mathbb{R}^{d \times d} \). Furthermore, the BVP (1) is subject to the contact conditions

\[
\begin{align*}
g_i \geq 0, \quad p_i \geq 0, \quad g_i \cdot \mathbf{p}_i &= 0 \quad \text{on} \ \Gamma_C^i \times [t_0, T]
\end{align*}
\]

with \( g_i = g_i(\mathbf{u}_i) \), where \( g(\cdot, \mathbf{u}(\cdot)) : \partial \Omega_i \rightarrow \mathbb{R} \) is the gap function described as

\[
g_i|\partial \Omega_i(x_i, \mathbf{u}_i) = \min_{j \in I \setminus \{i\}} (\xi_j(x_i, \mathbf{u}_i) - (x_i + \mathbf{u}_i))^T \mathbf{n}_i(x_i, \mathbf{u}_i), \quad x_i \in \partial \Omega_i
\]

and \( \mathbf{p}_i \) is the contact pressure on \( \partial \Omega_i \), respectively (Figure 1). Note, that \( \mathbf{n}_i \) is the outer normal vector on \( \partial \Omega_i \) and \( \xi_j(x_i, \mathbf{u}_i) \) is the projection point of \( x_i \) to the body \( \Omega_j \) in the outer normal direction, fulfilling

\[
\xi_j(x_i, \mathbf{u}_i) = \arg\min_{\xi \in \partial \Omega_j(t)} \| (x_i + \mathbf{u}_i(x_i, t)) - \xi \|, \quad x_i \in \partial \Omega_i.
\]
The gap function (5) corresponds to the distance between \( x_i \in \partial \Omega_i \) and \( \xi_j \in \partial \Omega_j \). But depending on the underlying geometry, (5) and therefore (6) can lead to complex expressions such as nonlinearities including the root-function. Moreover, in case of the nonuniqueness of (6), the gap function (5) cannot turn into a nonsmooth function. In Reference 31, we considered a nonpenetration condition allowing only normal contact such that no movement in the tangential direction was possible, see also Reference 20. In particular, the gap function \( g \) was linear. In this work, however, using the assumption (7) we want to generalize the constraint to a quadratic condition allowing also sliding movement such that a point can move along the contact area. For simplicity, we assume that there exist bounded functionals \( H : \partial \Omega \to \mathbb{R}^{d \times d} \) and \( \eta : \partial \Omega \to \mathbb{R}^d \) such that for all \( i \in I \) it holds

\[
\mathbf{n}_i = H(x_i)\mathbf{u}_i + \eta(x_i), \quad x_i \in \partial \Omega_i, \tag{7}
\]

that is, \( \mathbf{n}_i \) points out in the normal outward direction, does not necessarily have a unit length and depends on the displacement vector \( \mathbf{u}_i \) affine-linearly. Note, that inserting the orthogonal nonunit vector (7) into the gap function (5), still lead to a valid nonpenetration condition (4) in case of a sliding contact.

### 2.2 Weak formulation of the dynamic contact problem

In order to state the weak formulation of the contact problem, we introduce the function spaces

\[
\mathcal{V} := \{ v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D \}, \tag{8}
\]

\[
\mathcal{L} := \left\{ \mu \in H^{1/2}(\Gamma_C)' : \int_{\Gamma_C} \mu w \, ds \geq 0 \ \forall w \in H^{1/2}(\Gamma_C), w \geq 0 \right\}, \tag{9}
\]

where \( \mathcal{V} \) is the set of admissible displacements and is decomposed body-wise such that for all \( v \in \mathcal{V} \) it holds \( v_{|\Omega_i} \in H^1(\Omega_i)^d \) and \( v_{|\Omega_i} = 0 \) on \( \Gamma_{D_i} \). The set \( \mathcal{L} \) is the convex cone of the Lagrange multipliers defined on \( \Gamma_C \), which is the union of all \( \Gamma_C^i \) and \( H^{1/2}(\Gamma_C)' \) is the dual space of \( H^{1/2}(\Gamma_C) \). Furthermore, the following abstract notation is introduced:

\[
\langle \rho \mathbf{u}_r, \mathbf{v} \rangle := \int_\Omega \mathbf{v}^T \rho \mathbf{u}_r \, dx, \quad a(\mathbf{u}, \mathbf{v}) := \int_\Omega \sigma(\mathbf{u}) : \mathbf{e}(\mathbf{v}) \, dx, \quad \langle \mathbf{l}', \mathbf{v} \rangle := \int_\Omega \mathbf{v}^T \mathbf{F} \, dx + \int_{\Gamma_N} \mathbf{v}^T \mathbf{r} \, ds. \tag{10}
\]

The nonpenetration condition in weak form with test function \( \mu \in \mathcal{L} \) is in general given by

\[
0 \leq \int_{\Gamma_C} g \mu \, ds = \int_{\Gamma_C} g(\mu - p) \, ds. \tag{11}
\]

In particular for a two-body problem, that is, \( I = \{1, 2\} \), by inserting the gap function (5) combined with the assumption (7) in the first integral of (11), the right-hand side of (11) reads:
\[ \int_{\Gamma_c} g \mu ds = \int_{\Gamma_c} (u^T H u + u^T (\eta + H^T (x - \xi)) + (x - \xi)^T \eta) \mu ds. \]  

(12)

Note, that we provide the weak formulation solely for the two-body problem. The extension to the general multi-body case is straightforward. Using (12), we define the trilinear form on \( V \times V \times \mathcal{L} \)

\[ d(v, v, p) := \int_{\Gamma_c} v^T H v \cdot p \, ds, \]  

(13)

the bilinear form on \( V \times \mathcal{L} \)

\[ c(v, p) := \int_{\Gamma_c} v^T (\eta + H^T (x - \xi)) \cdot p \, ds \]  

(14)

and the linear form on \( \mathcal{L} \)

\[ \langle b, p \rangle := \int_{\Gamma_c} (x - \xi)^T \eta \cdot p \, ds. \]  

(15)

After these preparations, the weak form of the dynamic contact problem is stated as follows: For each \( t \in [t_0, T] \) find the displacement field \( u(\cdot, t) \in V \) and the contact pressure \( p(\cdot, t) \in \mathcal{L} \) such that

\[ \langle \rho u_t, v \rangle + a(u, v) - d(u, v, p) = \langle \varepsilon, v \rangle + c(v, p) \quad \text{for all } v \in V, \]  

\[ d(u, u, \mu - p) + c(u, \mu - p) + \langle b, \mu - p \rangle \geq 0 \quad \text{for all } \mu \in \mathcal{L}. \]  

(16)

The system (16) shows, that the contact condition includes a quadratic term with respect to the displacement variable. Note, that the system (16) originates from employment of the Hamiltonian principle for the Lagrangian \( J \) defined as the sum of the kinetic and potential energies as well as the contact contribution, that is,

\[ J(u, p) = \int_{t_0}^T \left( \frac{1}{2} \langle \rho u_t, u \rangle - \frac{1}{2} a(u, u) + \langle \varepsilon, u \rangle - \int_{\Gamma_c} gpds \right) \, dt. \]  

(17)

The technical details on variational calculus are omitted here.

3 | DISCRETIZED PROBLEM SETTING

The system (16) is discretized with respect to the spatial variable by applying the standard Galerkin projection \( u(x, t) \equiv \sum_{i=1}^{N} \phi_i(x) q_i(t) = \Phi(x) q(t) \) with basis functions \( \phi_i \) and nodal variables \( q = (q_1, \ldots, q_N)^T \in \mathbb{R}^N \) of the displacement field. The discretization of the contact condition requires special attention. In general, the integrals over \( \Gamma_c \) in the weak dynamic equation (16) can be approximated by a discrete sum, that is,

\[ \int_{\Gamma_c} v^T H u \cdot p \, ds + \int_{\Gamma_c} v^T (\eta + H^T (x - \xi)) \cdot p \, ds \approx \sum_{\varepsilon = 1}^{m} A_{\varepsilon} \left( v(x_{\varepsilon})^T H u(x_{\varepsilon}) + v(x_{\varepsilon})^T (\eta(x_{\varepsilon}) - H(x_{\varepsilon} + \xi(x_{\varepsilon}))) \right) \cdot p(x_{\varepsilon}, t). \]  

(18)

where \( x_{\varepsilon} \in \Gamma_c \) for \( \varepsilon = 1, \ldots, m \) are the contact nodes and \( A_{\varepsilon} \) is the area around node \( x_{\varepsilon} \). Furthermore, the Lagrange multiplier can be defined, for example, as the discrete pressure variable, that is, \( \lambda(t) := (p(x_1, t), \ldots, p(x_m, t))^T \in \mathbb{R}^m \). However, this relation might involve additional constant factors, depending on the discretization technique used for the contact condition.

In the following subsections, we will apply the node-to-segment discretization technique, which will lead to quadratic inequalities as constraints.
3.1 Node-to-segment contact condition

In this section, contact conditions based on the node-to-segment technique are studied in more detail. It is essential for our algorithm that the nonlinearity of the contact condition is at most quadratic. In that case the Kuhn–Tucker condition is linear, and the primal system can be resolved with respect to the displacement vector, which leads to a decoupling from the adjoint system. Moreover, an arbitrary distance function can lead in general to higher or weaker nonlinearities such as the square-root function, thus we designed a quadratic contact condition based on an angular function, which defines a nonpenetration condition between a node and a segment. In the following we assume a 2D setting, where linear finite elements are used for the discretization. Let

\[ \mathbf{p} = (p_1, p_2)^T \in \mathbb{R}^2, \quad \mathbf{\tilde{p}} = (\tilde{p}_1, \tilde{p}_2)^T \in \mathbb{R}^2, \quad \mathbf{p} \neq \mathbf{\tilde{p}} \]  

be two neighboring contact nodes. They define a contact segment, which can be written as

\[ s(\mathbf{p}, \mathbf{\tilde{p}}) = \{ s(t) = \mathbf{p} + t(\mathbf{\tilde{p}} - \mathbf{p}) \mid t \in [0, 1] \}. \]  

Furthermore, we assume that after the discretization a node-segment correspondence is established at the contact zone. Let \( r \in \mathbb{R}^2 \) be the corresponding node of the segment \( s \). The distance of the segment \( s \) to the node \( r \) is given by (see Figure 2A)

\[ g_s(r) = \min_{t \in [0, 1]} \| s(t) - r \|_2. \]  

Similarly as for the gap function in (5), the distance function in (21) can lead in general to a quite complex expression such as higher or weaker nonlinearities. Therefore we define an angular function representing an alternative to (21), that leads to a quadratic condition in terms of the displacements. For a node \( r \), let \( s(\mathbf{p}, \mathbf{\tilde{p}}) \) be the nearest segment to \( r \) and let \( \mathbf{n} \) denote an orthogonal vector to the line containing \( s \). Then the angular function is given by (see Figure 2B)

\[ \mathcal{M}(\mathbf{p}, \mathbf{\tilde{p}}, r) = \langle \mathbf{n}, r - \mathbf{p} \rangle = \mathbf{n}^T(r - \mathbf{p}). \]  

Note, that for two nonzero vectors \( \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2 \setminus \{0\} \) the equality \( \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \|\mathbf{v}_1\|_2 \cdot \|\mathbf{v}_2\|_2 \cdot \cos(\angle(\mathbf{v}, \mathbf{w})) \) always holds, where \( \angle(\mathbf{v}, \mathbf{w}) \in [0, \pi] \) denotes the angle enclosed by \( \mathbf{v} \) and \( \mathbf{w} \).

For (22), three different scenarios can be distinguished:

- \( \mathcal{M}(\mathbf{p}, \mathbf{\tilde{p}}, r) > 0 \), which means that \( r \) lies in the half of the line with positive normal.
- \( \mathcal{M}(\mathbf{p}, \mathbf{\tilde{p}}, r) = 0 \), which means that \( r \) lies on the line containing the segment \( s \).
- \( \mathcal{M}(\mathbf{p}, \mathbf{\tilde{p}}, r) < 0 \), which means that \( r \) lies in the half of the line with negative normal.

\[ \text{FIGURE 2 Geometrical illustration. (A) } g_s(r) > 0 \text{ and (B) } \mathcal{M}(\mathbf{p}, \mathbf{\tilde{p}}, r) > 0 \]
Since we are only interested in the sign of the angular function (22) for the description of the nonpenetration condition, the unity of the orthogonal vector \( \mathbf{n} \) is not necessary. By using the orthogonality of the vector \( \mathbf{n} \) to the segment \( \mathbf{s} \), we derive \( \mathbf{n} \) as follows:

\[
\mathbf{n} (\mathbf{p}, \mathbf{\hat{p}}) = \mathbf{N} (\mathbf{\hat{p}} - \mathbf{p}).
\]

where

\[
\mathbf{N} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.
\]

In order to derive a formulation for (22) in terms of the displacements and also to keep the notation concise, we write

\[
i = \mathbf{u}' + \mathbf{u}'_0, \quad i \in \{ \mathbf{p}, \mathbf{\hat{p}}, \mathbf{r} \}
\]

with the vectors \( \mathbf{u}'_0 \in \mathbb{R}^2 \) standing for the fixed initial position of the node \( i \) and \( \mathbf{u}' \in \mathbb{R}^2 \) for the corresponding displacement. Based on this notation and (23), the function (22) can be reformulated as

\[
\mathcal{M} (\mathbf{p}, \mathbf{\hat{p}}, \mathbf{r}) = (\mathbf{\hat{p}} - \mathbf{p})^T \mathbf{N}^T (\mathbf{r} - \mathbf{p})
\]

\[
= (\mathbf{u}' - \mathbf{u}'_0)^T \mathbf{N}^T (\mathbf{r}' - \mathbf{u}'_0) + \left( \mathbf{u}' - \mathbf{u}'_0 \right)^T \mathbf{N}^T (\mathbf{u}' - \mathbf{u}'_0)
\]

\[
+ \left( \mathbf{u}' - \mathbf{u}'_0 \right)^T \mathbf{N} (\mathbf{u}' - \mathbf{u}'_0) + \left( \mathbf{u}' - \mathbf{u}'_0 \right)^T \mathbf{N}^T (\mathbf{u}' - \mathbf{u}'_0).
\]

The function \( \mathcal{M} \) defined in (26) is at most of quadratic order in terms of the displacements \( \mathbf{u}' \), \( \mathbf{u}'_0 \), \( \mathbf{u}' \).

### 3.2 The fully assembled semi-discretized contact problem

For each grid node the corresponding \( x \)- and \( y \)-displacements are subsequently stored in the global displacement vector \( \mathbf{q} \in \mathbb{R}^N \). (The degrees of freedom of the global system is twice as much as the grid size.) The initial displacement vector is denoted as \( \mathbf{q}_0 \). For a contact problem with \( k \) bodies that fits into the framework described so far, the mass matrix \( \mathbf{M} \) and the stiffness matrix \( \mathbf{K} \) consist of \( k \) diagonal block matrices that stem from the discretization of each individual body. The contact is discretized by the node-to-segment technique that allows the sliding of a node over segments at the contact zone. Based on the angular function (26) we derive the global constraint matrices. We assume that the discretized contact zone is represented by the sets \( \mathcal{N} = \{ p_1, p_2, \ldots, p_k \} \) and \( S = \{ s_1, s_2, \ldots, s_k \} \). For each node \( r \in \mathcal{N} \), there is a related segment \( s_r (\mathbf{p}, \mathbf{\hat{p}}) \in S \), defined as in (20). For related nodes \( r, \mathbf{p}, \mathbf{\hat{p}} \) the corresponding indices of the displacement vector elements are denoted by \( i_1, i_2, i_3, i_4, i_5, i_6 \). For \( k \)th node-segment pair the block-matrix of the constraint matrix \( \mathbf{D}_k \in \mathbb{R}^{N \times N} \) is given by

\[
(D_k)^{i_1, \ldots, i_6} = \begin{pmatrix} 0 & 0 & 0 \\ -\mathbf{N} & \mathbf{N} & 0 \\ \mathbf{N} & -\mathbf{N} \end{pmatrix} \in \mathbb{R}^{6 \times 6}.
\]

Note, that per definition \( \mathbf{q}^T \mathbf{D}_k \mathbf{q} \) corresponds to the quadratic term in (26). Analogically, the constraint vectors \( \mathbf{c}_k \in \mathbb{R}^N \) and the constraint scalars \( b_k \in \mathbb{R} \) are defined:

\[
\mathbf{c}_k^{i_1, \ldots, i_6} = \left( (\mathbf{q}_0^{i_1,i_2} - \mathbf{q}_0^{i_1,i_3})^T \mathbf{N} | (2\mathbf{q}_0^{i_1,i_4} - \mathbf{q}_0^{i_1,i_5} - \mathbf{q}_0^{i_1,i_6})^T \mathbf{N} | (\mathbf{q}_0^{i_1,i_5} - \mathbf{q}_0^{i_1,i_4})^T \mathbf{N} \right) \in \mathbb{R}^6,
\]

\[
b_k = (\mathbf{q}_0^{i_1,i_4} - \mathbf{q}_0^{i_1,i_5})^T \mathbf{N}(\mathbf{q}_0^{i_1,i_5} - \mathbf{q}_0^{i_1,i_4}) \in \mathbb{R}.
\]
For further details of various contact models and discretization techniques we refer to References 20 and 23. For m contact nodes the contact condition reads

\[ q^T D_k q + c_k^T q + b_k \geq 0, \quad k = 1, \ldots, m. \]  

(30)

The latter represents a quadratic contact condition derived by the node-to-segment discretization technique. Assembling \( D_k, c_k, \) and \( b_k \) for all \( k \), we obtain the global constraint tensor \( D \in \mathbb{R}^{mN \times N} \), the global constraint matrix \( C \in \mathbb{R}^{mN} \), and the clearance vector \( b \in \mathbb{R}^m \). Then the generic model for the system of semi-discretized equations reads

\[
M \ddot{q} + Kq - (D + D^T) : \lambda - f - C^T \lambda = 0, \\
(D : q)q + Cq + b \geq 0, \\
\lambda \geq 0, \\
\lambda^T (D : q)q + Cq + b = 0,
\]

(31a-e)

where \( f(t) \in \mathbb{R}^N \) denotes the external loads of the system. In (31b) and (31c) for the quadratic term, we make use of the notation \( [D : q]q = q^T D_k q \) for \( k = 1, \ldots, m \). Note, that the constraint tensors \( D \) and the constraint matrix \( C \) exhibit a sparse structure with each row standing for a node-to-segment contact condition. Analogously, the inequality constraints are to be interpreted component-wise and stand for the discretized nonpenetration conditions of all variables in the contact interface where the vector of Lagrange multipliers \( \lambda(t) \in \mathbb{R}^m \) enforces the constraint in the dynamic equation (31a). The multiplier \( \lambda \), standing for the discretized contact pressure apart from some geometrical scaling factor, is always non-negative and its inner product with the constraints satisfies the complementarity condition (31c). Note also that

\[
(D + D^T) : \lambda = \sum_{k=1}^m (D_k + D_k^T) q \lambda_k = \left( \sum_{k=1}^m \lambda_k (D_k + D_k^T) \right) q,
\]

(32)

which means that this term can be viewed as a nonlinear contribution to the stiffness matrix that depends on the Lagrange multipliers.

## 4 SOLVING STATIC CONTACT PROBLEMS

A common approach to solve mechanical contact problems is the augmented Lagrangian method.\textsuperscript{20,26,44} The method is very flexible to different descriptions of the contact region. However, here we apply a dual approach, allowing us to turn to the adjoint problem resulting into a NCP. In the context of model order reduction for constrained problems, dual approaches have a great benefit, which will be highlighted in the following sections. First, we address the dual formulation and the solving procedure of a NCP in case of static contact problems.

### 4.1 NCP for the static contact problem

In the following, we consider the quadratic contact condition

\[
\begin{pmatrix}
q^T D_1 q \\
\vdots \\
q^T D_m q
\end{pmatrix} + Cq + b \geq 0.
\]

(33)

Then, the variational formulation of the mechanical contact problem starts by considering the energy expression

\[
\min_{q \in \mathbb{R}^N} \frac{1}{2} q^T Kq - q^T f - \sum_{k=1}^m \lambda_k \left( q^T D_k q + c_k^T q + b_k \right).
\]

(34)
The corresponding KKT-conditions read

\[ Kq - f - \left( \sum_{k=1}^{m} \lambda_k (D_k + D_k^T) \right) q - C^T \lambda = 0, \]

\[ q^T D_k q + c_k^T q + b_k \geq 0, \quad k = 1, \ldots, m, \]

\[ \lambda \geq 0, \]

\[ \sum_{k=1}^{m} \lambda_k (q^T D_k q + c_k^T q + b_k) = 0. \]  (35)

Solving for the displacements \( q \) in the KKT-conditions leads to the representation

\[ q = \left( K - \sum_{k=1}^{m} \lambda_k (D_k + D_k^T) \right)^{-1} (f + C^T \lambda). \]  (36)

Next, we introduce the matrix

\[ S(\lambda) = K - \sum_{k=1}^{m} \lambda_k (D_k + D_k^T). \]  (37)

The stiffness matrix \( K \) is assumed to be regular which implies that the structure is fixed (stable) and can be solely deformed. The additional matrix terms \( \lambda_k (D_k + D_k^T) \) of the sum in (37) can be interpreted as an additional contact force applied to the structure and therefore the fixation of the structure still remains (no rigid motion). As a consequence the matrix (37) is invertible. This physical argument is confirmed by our numerical experiments.

\[ q(\lambda) = S(\lambda)^{-1} (f + C^T \lambda) \]  (38)

and insert (36) into the second part of the KKT-conditions. This yields

\[ q(\lambda)^T D_k q(\lambda) + c_k^T q(\lambda) + b_k \geq 0, \quad k = 1, \ldots, m, \]

\[ \lambda \geq 0, \]

\[ \sum_{k=1}^{m} \lambda_k (q(\lambda)^T D_k q(\lambda) + c_k^T q(\lambda) + b_k) = 0. \]  (39)

which is an NCP for \( \lambda \). After computing the Lagrange multiplier \( \lambda \), the displacements are given by Equation (36).

### 4.2 Solving a sequence of LCPs for the static contact problem

A well-known and successful technique of solving NCPs is to consider their linear approximation, representing an LCP system, for which an LCP solver can be applied repeatedly. This method has been introduced in Reference 43 where the problem statement is given in a generalized framework. More about iterative techniques for solving variational inequalities can be found in References 41 and 42. A general form of an LCP is defined as follows:

\[ B + A \lambda \geq 0, \]

\[ \lambda \geq 0, \]

\[ \lambda^T (B + A \lambda) = 0. \]  (40)

LCP problems may be solved by applying either Fischer–Burmeister type algorithms45 or more efficiently Lemke’s algorithm (see Reference 46, 4.4.5). A detailed discussion of Lemke’s algorithm as well as a much more comprehensive study of LCPs and can be found in Reference 46. Specifically, we use a Python implementation of the algorithm.47
The core idea, which below is also used for our reduction technique, is to exchange the original NCP problem

\[ \lambda \geq 0, \quad F(\lambda) \geq 0, \quad \lambda^T F(\lambda) = 0 \]  

(41)

by the linear approximation

\[ \lambda \geq 0, \quad F(\lambda_0) + \nabla F(\lambda_0)(\lambda - \lambda_0) \geq 0, \quad \lambda^T (F(\lambda_0) + \nabla F(\lambda_0)(\lambda - \lambda_0)) = 0, \]  

(42)

where

\[ F : \mathbb{R}^m \to \mathbb{R}^m, \quad F(\lambda) = (F_1(\lambda), \ldots, F_m(\lambda))^T \]  

(43)

with

\[ F_k(\lambda) = q(\lambda)^T D_k q(\lambda) + c_k^T q(\lambda) + b_k, \quad k = 1, \ldots, m \]  

(44)

represents the quadratic constraint function. The computation of the gradient \( \nabla F \) is discussed in Section 5.4. The system (42) is solved by a LCP solver (e.g., the Lemke method) until a carefully chosen stopping criterion depending on the initial starting point \( \lambda_0 \) is satisfied. Then \( \lambda_0 \) is updated by \( \lambda \) and the LCP problem is solved again. Finally, this procedure yields an approximation for the solution vector \( \lambda \) of the nonlinear problem (41).

5 | SOLVING DYNAMIC CONTACT PROBLEMS

After the static contact problems have been addressed, we turn our attention to dynamic contact problems. Similar to the static case, the key idea for solving the time-dependent contact problems is to replace the nonlinear dual system by its linear approximation.

5.1 | Notational convention

For subsequent use, we introduce a new notation for our solution variables. We denote \( \dot{\lambda}_i = \lambda(t_i) \) and \( q_i = q(t_i, \lambda_i) \), where the subscript \( i \) stands for the time step \( t_i \). In case of the linearization the Lagrange multiplier \( \lambda_i : = \dot{\lambda}_i^l \) carries two indices, where the superscript \( l \) denotes the corresponding Newton-like iteration of the LCP. If it is clear from the context that the time instance \( t_i \) is kept fixed, we make use of the variable \( z_i = \dot{\lambda}_i^l \).

5.2 | NCP for the dynamic contact problem

The dynamical model combined with the Kuhn–Tucker conditions (35) leads to the system

\[ Mq + Kq - f - \left( \sum_{k=1}^m \lambda_k (D_k + D_k^T) \right)q - C^T \lambda = 0, \]  

(45a)

\[ q^T D_k q + c_k^T q + b_k \geq 0, \quad k = 1, \ldots, m, \]  

(45b)

\[ \lambda \geq 0, \]  

(45c)

\[ \sum_{k=1}^m \lambda_k (q^T D_k q + c_k^T q + b_k) = 0. \]  

(45d)

All four conditions in (45) must be fulfilled in each time step. The balance of momentum (45a) is discretized in time by the implicit Euler scheme, a straightforward and unconditionally stable method, also known as the first order backward differentiation formula. To this end, we replace the second order derivative by the finite difference
where $h$ is the time step size. Note that for the velocity it holds $\ddot{q}(t+h) \approx (q(t+h) - q(t))/h$, a fact that is hidden behind (46). Furthermore, due to the presence of the second order time derivative, the implicit Euler leads to a two-step method. Since (46) possesses first order of accuracy only, other integration schemes can be applied instead, for example, the generalized-$\alpha$ method, which is very common in structural mechanics, see Reference 40. In this work, however, we mainly focus on designing a reduction scheme and therefore, we continue working with the implicit Euler method. Inserting (46) into Equation (45) and keeping in mind the abbreviation (37), we obtain

$$
M(q(t+h) - 2q(t) + q(t-h)) + h^2 S(\dot{\lambda})q(t+h) = h^2 f(t+h) + h^2 C^T \dot{\lambda}.
$$

(47)

Assuming that the previous two time steps are known, (47) is solved for $q(t+h)$, which leads to

$$
q(t+h) = (M + h^2 S(\dot{\lambda}))^{-1} (h^2 f(t+h) + h^2 C^T \dot{\lambda} + 2Mq(t) - Mq(t-h)).
$$

(48)

Note, that the invertibility of the matrix $M + h^2 S(h)$ can be always guaranteed, since the mass matrix $M$ is invertible and therefore the whole matrix is invertible for a sufficient small time step $h$. We remark that the two-step time discretization requires initial values $q(t_0)$ and $q(t_0 + h)$ to start. If $q_0$ and $q_1$ as initial displacement and velocity are given, one can compute $q(t_0 + h) = q_0 + hq_1$, by an explicit Euler step and then continue with the two-step formula (47). The initialization of the Lagrange multiplier $\lambda$ is shortly addressed in Section 6. If the inequality constraints were replaced by the equality constraints $q^T D_k q + c_k^T q + b_k = 0$, $k = 1, \ldots, m$, the index of the resulting differential-algebraic equation would equal 3, which means that special care must be taken for the time integration. Inserting (48) into the constraints (45b)–(45d) we obtain an NCP problem similar to (39). After computing the Lagrange multiplier $\lambda$, the displacements are given by (48). Note that the NCP has to be solved in each time step.

5.3 Solving a sequence of LCPs for the dynamic contact problem

After having introduced the transient model, we address the corresponding solution procedure in more detail. The core idea of the algorithm is the generalization of the iterative solver for the static problems to an iterative solver for the dynamic problems by substituting the displacement vector in (38) by

$$
q_{i+1}(\lambda_{i+1}) = (M + h^2 S(\dot{\lambda}_{i+1}))^{-1} (h^2 f_{i+1} + h^2 C^T \dot{\lambda}_{i+1} + 2Mq_i - Mq_{i-1}).
$$

(49)

The quadratic constraint function $F : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is defined as

$$
F_k(\lambda_{i+1}) = q_{i+1}(\lambda_{i+1})^T D_k q_{i+1}(\lambda_{i+1}) + c_k^T q_{i+1}(\lambda_{i+1}) + b_k, \quad k = 1, \ldots, m.
$$

(50)

The function (50) defines for each time step the adjoint problem of (45), which represents an NCP problem, that is,

$$
\lambda_{i+1} \geq 0, \quad F(\lambda_{i+1}) \geq 0, \quad \lambda_{i+1}^T F(\lambda_{i+1}) = 0.
$$

(51)

Instead of directly solving (51), we prefer to solve the linear approximation of (51) within Newton-like iterations for each time step. For this purpose, we approximate the nonlinear constraint function (50) by its linearization. As soon as the linearized constraint system comes into play, we make use of the abbreviation $z_l$, see also Section 5.1. This leads to the linearized system

$$
z_l \geq 0, \quad F(z_{l-1}) + \nabla F(z_{l-1})(z_l - z_{l-1}) \geq 0,
$$

$$
z_l^T (F(z_{l-1}) + \nabla F(z_{l-1})(z_l - z_{l-1})) = 0.
$$

(52)

Note that at the time step $t_{i+1}$ it holds for $z_0 = \lambda_i$, which is assumed to be known. For each $l > 0$, the linearization is performed at $z_{l-1}$. Furthermore, we write $A^{l-1} = \nabla F(z_{l-1})$ and $B^{l-1} = F(z_{l-1}) - \nabla F(z_{l-1})z_{l-1}$ such that the system (52) can be transformed into the LCP problem
\[
B^{l-1} + A^{l-1}z_l \geq 0,
\]
\[
z_l \geq 0.
\]
\[
z_l^T(B^{l-1} + A^{l-1}z_l) = 0.
\]

The system (53) is repeatedly solved within a Newton-type iteration, until the error between \(z_l\) and \(z_{l-1}\) is small enough and the system converges, yielding the final \(z_{l^*}\) for a certain \(l^* > 0\).

According to the general theoretical framework,\(^{41-43}\) quadratic convergence is guaranteed when iteratively solving (52). This means that only a few iterations suffice to obtain a solution for (51) at each time step. Within this work, we do not pursue the theoretical aspects any further, but rather show a very good agreement with the theory based on our computational examples, see Section 7.

5.4 Computation of the Jacobian

In order to proceed with (52), we need to compute the Jacobian matrix of the function

\[
F : \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad F(\lambda) = (F_1(\lambda), \ldots, F_m(\lambda))^T
\]

with

\[
F_k(\lambda) = q(\lambda)^T D_k q(\lambda) + c_k^T q(\lambda) + b_k.
\]

The derivative matrix is given by

\[
DF(\lambda) = \begin{bmatrix}
\frac{\partial}{\partial \lambda_1} F_1(\lambda) & \cdots & \frac{\partial}{\partial \lambda_m} F_1(\lambda) \\
\vdots & \ddots & \vdots \\
\frac{\partial}{\partial \lambda_1} F_m(\lambda) & \cdots & \frac{\partial}{\partial \lambda_m} F_m(\lambda)
\end{bmatrix} \in \mathbb{R}^{mxm}.
\]

Due to the chain rule for the partial derivative of \(F_k\) with respect to \(\lambda_j\) it holds:

\[
\frac{\partial}{\partial \lambda_j} F_k(\lambda) = (q(\lambda)^T D_k + D_k^T q(\lambda) + c_k^T) \frac{\partial}{\partial \lambda_j} q(\lambda).
\]

In turn, the partial derivative of \(q\) with respect to \(\lambda_j\) results in

\[
\frac{\partial}{\partial \lambda_j} q(\lambda) = h^2(M + h^2S(\lambda))^{-1} \left( c_j + (D_j + D_j^T)q(\lambda) \right).
\]

Finally, inserting (58) into (57), for \(i, j \in \{1, \ldots, m\}\) we obtain

\[
\frac{\partial}{\partial \lambda_j} F_k(\lambda) = h^2 \left( q(\lambda)^T (D_j + D_j^T) + c_j^T \right) (M + h^2S(\lambda))^{-T} \left( c_k + (D_k + D_k^T)q(\lambda) \right).
\]

For the sake of a more compact representation, we introduce the term

\[
Z_k = Z_k(\lambda) = q(\lambda)^T(D_k + D_k^T) + c_k^T, \quad k = 1, \ldots, m
\]

leading to the matrix

\[
Z = \begin{bmatrix}
Z_1 \\
\vdots \\
Z_m
\end{bmatrix} \in \mathbb{R}^{mxN}.
\]
Using the new abbreviation the partial derivatives can be computed as follows

\[ \frac{\partial}{\partial \lambda_j} F_k(\lambda) = h^2 Z_k(M + h^2 S(\lambda))^{-1} Z_j^T. \] (62)

Finally, the Jacobian matrix can be stated as

\[ DF(\lambda) = h^2 Z(M + h^2 S(\lambda))^{-1} Z^T. \] (63)

### 6 | REDUCTION OF THE CONTACT PROBLEM

In many real world applications, the numerical integration of constrained systems such as (45) is not possible in real time due to the large number \( N \) of degrees of freedom. Model order reduction strategies\(^{50}\) introduce a reduced state \( \hat{q} \in \mathbb{R}^n \) with \( n \ll N \), where \( \hat{q} \) is defined by

\[ q = Q \hat{q}. \quad Q \in \mathbb{R}^{N \times n}. \] (64)

#### 6.1 | Computation of ROM

One way to obtain the reduction matrix \( Q \) for the system (45a) is to use modal reduction. Setting up the eigenvalue problem of (45a)

\[ \omega^2 M v = K v \] (65)

and taking the first \( n \) eigenvectors, the matrix \( Q \) may be defined by

\[ Q = \{ v_1, \ldots, v_n \}. \] (66)

A preferable technique may be the Krylov subspace methods.\(^3,14\) The subspace is defined by

\[ Q = \{ K^{-1} f, K^{-1} MK^{-1} f, \ldots, (K^{-1} M)^{n-1} K^{-1} f \}. \] (67)

The Krylov base may be computed by the Arnoldi algorithm, which delivers an orthonormal base of the subset. Inserting the reduction (64), where \( Q \) is defined by (66) or (67) into the differential equations system (45a) and multiplying by \( Q^T \), one obtains the reduced system

\[ \hat{M} \ddot{\hat{q}} + \hat{S}(\lambda) \dot{\hat{q}} = Q^T (f(t) + C^T \lambda(t)), \] (68)

where

\[ \hat{M} = Q^T M Q \in \mathbb{R}^{n \times n}, \] (69)

\[ \hat{S}(\lambda) = Q^T S(\lambda) Q = Q^T K Q - \sum_{k=1}^{m} \lambda_k Q^T (D_k + D_k^T) Q \in \mathbb{R}^{n \times n}, \] (70)

are the reduced mass and stiffness matrices.

Similar to the full dimensional case, we apply implicit Euler to the reduced system (68), which leads to

\[ \hat{q}(t + h) = (\hat{M} + h^2 \hat{S}(\lambda))^{-1} (h^2 Q^T f(t + h) + h^2 Q^T C^T \lambda + 2\hat{M}\hat{q}(t) - \hat{M}\hat{q}(t - h))). \] (71)

Note that (71) involves only operations with matrices and vectors in smaller dimensions \( n \) or \( m \). Although \( Q \in \mathbb{R}^{N \times n} \), the acting force \( f(t + h) \) has only a few entries and for the Lagrange multiplier it holds \( \lambda \in \mathbb{R}^n \) with \( m \ll N \).
6.2 Contact treatment

After the general reduction method has been outlined, we introduce a more accurate way of choosing the basis vectors for the reduced space of the displacements. The underlying idea for increasing the accuracy of a reduction approach for contact problems is to partition the nodal variables into, so called, master and slave nodes. One of the most prominent partitioning methods is the Guyan reduction, also called static condensation, see Reference 51. The slave nodes often possess much larger local stiffness compared to the master nodes and do not inherit any noticeable change of motion, a fact that is hidden behind this technique.40 The combination of static condensation with a reduction method performed on the slave nodes results into the Craig–Bampton method,39 a well-known technique usually applied to contact problems with linear constraints, see Reference 31. We, however, want to extend this method in a way that will allow us to use the same principle for reducing (45) including the nonlinear term $S(\lambda)$.

More precisely, let us assume that the displacement variables and the system matrices in (45) are already partitioned, that is, $q = (q_M, q_S)^T$, where $q_M$ and $q_S$ denote the master and the slave displacement variables, respectively. Then it holds

$$
M = \begin{pmatrix} M_{MM} & M_{MS} \\ M_{SM} & M_{SS} \end{pmatrix}, \quad K = \begin{pmatrix} K_{MM} & K_{MS} \\ K_{SM} & K_{SS} \end{pmatrix}, \quad f = (f_M, f_S)^T. \tag{72}
$$

Since we want to exclude the constraints from reduction, a similar partitioning is performed for the constraint matrices

$$
D_k = \begin{pmatrix} D_{k,MM} & D_{k,MS} \\ D_{k,SM} & D_{k,SS} \end{pmatrix}, \quad c_k = (c_{k,M}, c_{k,S})^T, \quad k = 1, \ldots, m. \tag{73}
$$

In Section 3, the constraint matrices (73) were derived. We can observe that the constraints apply only to the master nodes, implying

$$
D_{k,MS} = 0, \quad D_{k,SM} = 0, \tag{74}
$$

$$
D_{k,SS} = 0, \quad c_{k,S} = 0 \tag{75}
$$

for all $k = 1, \ldots, m$. Then the nonpenetration condition (30) is transferred to the following inequality:

$$
q_M^T D_{k,MM} q_M + c_{k,M}^T q_M + b_{k,M} \geq 0, \quad k = 1, \ldots, m. \tag{76}
$$

Therefore, it follows that

$$
S_{SS}(\lambda) = K_{SS} - \sum_{k=1}^{m} \lambda_k (D_{k,SS} + D_{k,SS}^T) = K_{SS}. \tag{77}
$$

Equation (77) combined with the vanishing block-matrices corresponding the slave nodes indicate that, as expected, the dynamics of the slave nodal variables are not directly constrained. Even more, if we keep the configuration of the contact nodes unchanged, that is, $q_M = 0$, the dynamic motion of the slave nodes is then given by

$$
M_{SS} q_S + K_{SS} q_S = f_S. \tag{78}
$$

Instead of reducing the full system, we want to reduce only the slave nodal variables. Therefore, we compute the transformation matrix of the slave system (78) using the Arnoldi method, which yields

$$
Q_S := \left\{ K_{SS}^{-1} f_S, K_{SS}^{-1} M_{SS} K_{SS}^{-1} F_S, \ldots, (K_{SS}^{-1} M_{SS})^{n-1} K_{SS}^{-1} f_S \right\}. \tag{79}
$$

Now we can construct a smaller space approximating the displacement variables outside the contact area. Nevertheless we still need a coupling condition between the slave and the master variables. In order to achieve this, we assume that the impact of the acting force on the slave nodes is smaller compared to the impact on the master nodes. Following the
standard argument for the Craig–Bampton,\textsuperscript{52} we obtain a static condensation between the slave and the master nodes, which due to (77) reads

\[ K_{SM} q_M + K_{SS} q_S = 0. \]  

(80)

Using (80), we obtain the complete transformation matrix that includes the coupling term of the master and slave nodes,

\[ Q_{CB} = \begin{pmatrix} I_M & 0 \\ -K_{SS}^{-1} K_{SM} & Q_S \end{pmatrix}. \]  

(81)

The partitioning of the master and slave nodes in the fashion of the Craig–Bampton method provides an alternative way to compute the reduced displacements in (71). Although the contact condition is quadratic, the static condensation remains unaffected and does not depend on the Lagrange multiplier, adopting the linear form (80). On the other hand, since the contact nodes are not reduced, the individual contact impact is included within the reduced model. A similar result for contact problems with linear constraint is obtained in Reference 31.

The reduction matrix \( Q_{CB} \) can be computed apriori in an offline phase, since the maximal set of the potential contact nodes is predefined and therefore does not change within time. During simulation only the nodes within this maximal set undergo a contact-release transition. This does not affect the reduction matrix, because the reduction matrix is computed only for the area outside of the maximal potential contact zone.

### 6.3 The reduction scheme

After these preparations, we present the final reduction scheme. Note that all the reduced matrices can be computed beforehand in an offline procedure. Once the transformation matrix (81) is computed the full matrices can be reduced via

\[ \hat{M} = Q^T_{CB} M Q_{CB} \in \mathbb{R}^{n \times n}, \quad \hat{K} = Q^T_{CB} K Q_{CB} \in \mathbb{R}^{n \times n}, \]  

\[ \hat{S}(\lambda) = Q^T_{CB} S(\lambda) Q_{CB} = \hat{K} - \sum_{k=1}^{m} \lambda_k (\hat{D}_k + \hat{D}_k^T) \in \mathbb{R}^{n \times n}, \]  

(82)

whereas the matrices \( \hat{D}_k \) comprise solely the constraint block-matrices corresponding to the master nodes appended by rank(\( Q_{SS} \))-many zero columns and rows, that is,

\[ \hat{D}_k = Q^T_{CB} D_k Q_{CB} = \begin{pmatrix} D_{k,MM} & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{n \times n} \quad \text{for} \quad k = 1, \ldots, m. \]  

(83)

The same argument applies to the linear term of the constraint (45b),

\[ \hat{c}_k = Q^T_{CB} c_k = (c_{k,MM}, 0)^T \in \mathbb{R}^n, \quad k = 1, \ldots, m. \]  

(84)

The Arnoldi method requires only the position vector of the right-hand side and the rearrangement of the position vector according to the same partitioning results in

\[ \hat{f}_{pos} = Q^T_{CB} f_{pos} Q_{CB} = \left( f_{pos,M}^T, f_{pos,S}^T \right)^T \in \mathbb{R}^n. \]  

(85)

Note that once the matrices \( \hat{K} \) and \( \hat{D}_k \) are computed, the stiffness term \( \hat{S}(\lambda) \) can be recovered be means of (83) for each time step during the online computation.

Based on the notations with respect to the reduced system introduced so far, the reduced displacement vector can be formulated as

\[ \tilde{q}_{i+1}(\lambda_{i+1}) = (\hat{M} + h^2 \hat{S}(\lambda_{i+1}))^{-1} (h^2 \hat{f}_{i+1} + h^2 \hat{C}^T \lambda_{i+1} + 2 \hat{M} \tilde{q}_i - \hat{M} \tilde{q}_{i-1}). \]  

(87)
Finally, the constraint function of the reduced space $\hat{F} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ reads

$$\hat{F}_k(\lambda_{i+1}) = \hat{q}_{i+1}(\lambda_{i+1})^T \hat{D}_k \hat{q}_{i+1}(\lambda_{i+1}) + \hat{c}_k^T \hat{q}_{i+1}(\lambda_{i+1}) + \hat{b}_k, \quad k = 1, \ldots, m. \quad (88)$$

Similarly to the case of the full order model, the NCP formulation for the constraint function $\hat{F}$ of the reduced system can be derived by inserting the reduced system matrices (82)–(85) into (50). The approximating LCP problem can be stated by means of

$$\hat{A}^{l-1} = \nabla \hat{F}(z_{l-1}),$$
$$\hat{B}^{l-1} = \hat{F}(z_{l-1}) - \nabla \hat{F}(z_{l-1}) z_{l-1}, \quad (89)$$

which are computed during the linearization procedure of (88). Similar to the full model case, in (89) it holds $z_1 = \lambda_{i+1}$, implying that the Lagrange multiplier, unlike the displacement variable, remains in the full dimensional form. The latter holds due to the fact that we exclude the contact shape from reduction. The final form of the LCP problem at the $l$th iteration for a fixed time step reads

$$\hat{B}^{l-1} + \hat{A}^{l-1} z_l \geq 0,$$
$$z_l \geq 0,$$
$$z_l^T (\hat{B}^{l-1} + \hat{A}^{l-1} z_l) = 0. \quad (90)$$

The system (90) represents the linearized adjoint problem of the reduced system (68). In the terminology of differential-algebraic equations, $q$ stands for the differential variables while $\lambda$ is the algebraic variable. Based on the numerical analysis performed so far, we can deduce that there is no practical dependency on the starting point $\lambda_0$. However, further investigations of this theoretic aspect might provide extra insight. In case of a constant Jacobian matrix $\nabla \hat{F}(z_0)$ throughout the Newton-type iterations only linear convergence can be achieved. In order to increase the convergence rate, an update of the Jacobian matrix is necessary. In that case quadratic convergence is guaranteed.

---

**Algorithm 1.** Reduction of the quadratic contact problem (45)

1. Input: The system matrices and vectors $M, K, f_{pos}, D_k, c_k, b_k, \ k = 1, \ldots, m, \ TOL > 0$
2. Output: $q_i$ and $\lambda_i$ for $i \in \{1, \ldots, n_T\}$
3. Perform partitioning between master and slave nodes.
4. Compute the reduction matrix $Q_{CR}$ by (81).
5. Compute the reduced system by (82)–(86).
6. Initialization of $\hat{q}_i$ and $\lambda_i$ for $i = 1, 2$.
7. for $i = 2, \ldots, n_T$ do
   8. $z_0 = \hat{\lambda}_{i-1}$
   9. for $l = 1, 2, \ldots$ do
      10. Compute $\hat{A}^{l-1}$ and $\hat{B}^{l-1}$ by (89),
      11. Compute $z_l$ as a solution of (90) by Lemke's method.
      12. if $\|z_l - z_{l-1}\| < TOL$ then
          13. $l^* = l$
          14. break
      end if
   15. end for
   16. $\lambda_i = z_l^*$
   17. Compute the reduced displacement vector $\hat{q}_i$ by (87).
8. end for
9. Compute $q_i = Q_{CR} \hat{q}_i, \ i = 1, \ldots, n_T$. 

We briefly summarize here the properties of Algorithm 1. The reduced model is computed once in an offline phase and it can be utilized for arbitrary right-hand sides. The contact shape is precomputed in the offline phase as well, since it depends on the acting force. The performance of the reduction method depends on the size of the contact zone \( m \). The latter holds due to the fact, that the dual problem of the size \( m \) is solved within several iterations in each time step. However, according to the theory,\(^{42,43}\) quadratic convergence is guaranteed for the Newton-type iterations in Algorithm 1. Therefore, the convergence criterion can be achieved after a few iterations. Based on these observations and facts, we can conclude that Algorithm 1 performs highly efficient for contact problems where the contact zone is small compared to the overall structure.

7 | APPLICATIONS

In this section, three computational examples are demonstrated. The first example represents a Hertz contact problem,\(^{23,53}\) which serves for verification of the solution methods of contact problems presented in Reference 31 and in this work. In the next two examples, our new model order reduction scheme is applied. We consider a plane self-contact problem and a wheel-rail contact problem inspired by applications in the railway industry. The LCP iterations within the reduction scheme are solved by the Lemke’s method taken from the open source library.\(^{47}\) For the stopping criterion of the Newton-type iterations in Algorithm 1, we use a tolerance of \( 10^{-6} \). For all considered contact problems, friction effects are neglected.

Furthermore, for all the applications the underlying geometry is constructed by means of the commercial FEM tool Simcenter12.\(^{35}\) The primal system matrices and the maximal contact grid are extracted from the latter tool. The modeling of the contact condition and the subsequent reduction of the full system are performed within a custom Python framework.

7.1 | Validation of adjoint methods

In the following example, a Hertz contact problem is considered. For such contact problems the theoretical solutions are known and can be used to validate the solution methods. The numerical solutions of the contact pressure and the contact radius are compared with their explicit expressions. Note that during validation no sliding movement is considered, since the theoretical solutions correspond to a static normal contact.

7.1.1 | A cylinder on a rigid foundation

We consider a two-dimensional cross-section of an infinitely long elastic cylinder composed of a homogeneous, isotropic material with the Young’s modulus \( E = 2000 \) and the Poisson’s coefficient \( \sigma = 0.3 \), see Figure 3. Since this example is considered for validation purposes, all parameters are unitless. The cylinder with the radius of \( R = 8 \) is initially located on a rigid foundation with a single point contact, which corresponds to one contact node and is fixed. Moreover, the center of the cylinder is fixed in the horizontal direction in order to prevent rigid motion. A constant load with a magnitude of \( F = 1600 \) is uniformly distributed on the upper boundary segment of the cylinder and is applied in the vertical direction, pressing the cylinder toward the rigid foundation. The contact area can be described in terms of a contact radius \( r \), which in two-dimensional case represents a line and according to the theory is defined by

\[
r = 2\sqrt{\frac{FR(1 - \nu^2)}{E\pi}}. \tag{91}
\]

Using the formula for the contact radius (91) the analytical solution of contact pressure can be computed by

\[
\sigma(x) = \frac{2F}{\pi r^2} \sqrt{r^2 - x^2}. \tag{92}
\]

For the chosen parameters, we obtain a contact radius \( r = 2.723 \) and the maximum contact stress \( \sigma(0) = 374 \).
In order to compare the analytical and numerical solutions, the cylinder is discretized by finite element methods. Static simulations based on node-to-node\textsuperscript{31} and node-to-segment techniques are performed and compared. Since there is no sliding movement but only normal contact between the cylinder and the rigid foundation, both methodologies are equivalent and yield identical solutions for the Lagrange multipliers and the displacements. Hence, we refer to \( r_N \) and \( \sigma_N(x) \) as the numerical counterparts of (91) and (92) computed either by node-to-node or node-to-segment techniques. Note that the Lagrange multipliers are converted to the contact stress \( \sigma_N(x) \) by the tributary area method.\textsuperscript{53}

The deformed cylinder is depicted in Figure 4A, which shows that the penetration at the contact zone is prevented. The total number of DOF is 47,624. The theoretical solution \( \sigma \) is compared with the numerical solution \( \sigma_N \) in Figure 4B. Since the lowest point of the cylinder (corresponds to the center of the contact zone in the deformed state) is fixed, the corresponding node is eliminated during the computation. The deviation of \( \sigma_N \) from \( \sigma \) is below five percent and can be further reduced in case of a finer resolution. Overall, Figure 4B shows, that the numerical solution for both the contact radius and the contact pressure are in a good agreement with the theory.

Finally, the inner plane stress is computed from the gradients of the displacements. The computation of nodal gradients is explained in more detail in Section 7.2.2. The vertical stress is visualized in Figure 5. Again, a good agreement is achieved between the \( y \)-stress and the contact pressure computed both by the explicit expression and by the adjoint methods. In particular, around the contact center a vertical stress of approximately \(-360\) is achieved, which in absolute values represents a reasonable approximation of the contact pressure.
7.2 | Model order reduction for contact problems

The following two applications show the performance of the novel reduction method for sliding contact problems. Note, that since there are no explicit expressions available for the solutions of the corresponding contact problems, we refer to the Lagrange multipliers as the contact pressure apart from scaling (e.g., the tributary area method).

7.2.1 | Simple crack in a square

First, we apply our approach to the two-dimensional unit square $\Omega = [0, 1] \times [0, 1]$ with a self-contact under the assumption of plane stress. For the spatial discretization standard bilinear shape functions are used on quadrilateral elements. The finite element method and the reduction scheme are implemented within a custom-written Python script. The material parameters $\rho = 1$, $E = 1000$, $\nu = 0.3$ are given in dimensionless form. We consider a zero Dirichlet boundary condition on the left edge of the plane square. The square is discretized by 1600 elements, which in total leads to $N = 3386$ DOF. The interior tear defines the contact zone and is given by a fixed number $m = 25$ of discrete grid points. Special care is taken for the data-structure of the nodes at the contact interface. We make use of the node-to-segment technique. In particular, the left edge is composed of a fixed number of segments. Each of such segments is predefined by its start and end nodes. In contrast, the right side of the tear is given by grid points that coincide with the finite element nodes. Similar examples can be found in Reference 31.

This kind of data-structure allows us to formulate a nonpenetration condition during a local sliding movement at the tear of the plane square. For this example, there is a unique assignment between the contact nodes and the contact segments. Note, that at the initial time point the tear is closed, that is, there is no gap between the contact nodes and the contact segments.

The right side of the domain is loaded by a horizontal nodal oscillating force:

$$f(t) = (1.5 \sin(0.1\pi t), 0)^T \in \mathbb{R}^2, \quad t \in [0, 1]. \quad (93)$$

The application of (93) has an impact on the contact interface leading to an opening and closing of the tear over time. Whenever the tear is closed, a sliding motion occurs at the contact zone. The problem setup and the mesh are illustrated in Figure 6. In order to compare the displacements of the full and the reduced model, two sensors are attached to the square. One of those sensors is placed on the contact interface allowing us to track the contact pressure and the gap function.
We eliminate the fixed degrees of freedom a priori before starting with the reduction method. In total, the reduced model has \( n = 53 \) DOF, which results from the combination of 50 contact displacement nodes and 3 additional slave variables.

First, the reduction approach is performed by means of an Arnoldi reduction matrix without any contact treatment. Afterwards, we compute the displacements resulting from the reduction method combined with the Craig–Bampton splitting technique. For the sake of a fair comparison, we choose the same number of basis functions in both cases. When we omit the splitting, the basis consists solely of Krylov basis vectors, whereas in case of splitting, the basis is extended by the contact nodes. Similar to Reference 31, the advantages of additional contact treatment can be clearly observed. Both solutions resulting from the corresponding reduction approach are compared with the full displacement variables in Figures 7 and 8.

Furthermore, the contact pressure and its counterpart, the gap function, deserve attention. We set \( g_{k}^{*} = q^{T}D_{q}q + c_{q}^{T}q + b_{q} \) for the gap function at the sensor node on contact zone and \( \lambda_{k} \) measures the contact pressure apart from the underlying mesh scaling. The index \( \hat{k} \) denotes the sensor node placed at the contact zone. We observe in Figure 9, that, in particular, when computed by the ROM without splitting, the pressure is not recovered, whereas the ROM with a splitting provides the contact pressure, that agrees well with the FOM. Both approaches guarantee positivity of the contact pressure whenever the contact is activated. Furthermore, the switching between an active contact with a nonzero pressure and positive gap with vanishing pressure can be nicely observed, that is, it holds \( g_{k}^{*}\lambda_{k} = 0 \) for all time steps.

Overall, the main reason for this behavior is that by using only Arnoldi for computing the transformation matrix, the reduced space does not comprise the dominant impact of the contact nodes. Moreover, our reduction approach is highly
efficient in case of a small contact zone compared to the total volume and since the reduced space is at least as large as the contact area, we do not lose the computational savings by adding the contact splitting to our reduction approach.

Finally, the convergence behavior of the NCPs is addressed. When the contact is inactive, only one fixed-point iteration of LCP suffices to solve the nonpenetration condition. However, in case of an active contact condition, 3 or 4 iterations are needed for each time-step, see Figure 10. These observations agree well with the theory implying quadratic convergence for solving the NCPs during each time step.

7.2.2 | Wheel-rail contact problem

To conclude, we present a wheel-rail contact problem, a prominent example inspired by real-life applications in railway dynamics. Both the wheel and the rail are composed of alloy steel and discretized within the FEM-framework of NX Simcenter 12.0. Within the scope of this work we solely consider the two-dimensional cross section of the wheel-rail. Furthermore, the rail is fixed on the bottom over time and the rotational degrees of freedom are neglected. Both the wheel and the rail share the same Young modulus, $E = 206,940$ MPa, and the Poisson coefficient $\nu = 0.288$. Two outer loads are applied to the wheel: The first one is the superposition of all vertical forces pointing toward ground (gravity, train mass, etc.). The second force represents the centripetal force defined by a frequency and load magnitude (Figure 11). The force $f_c$ arises naturally during the train ride and the corresponding frequency usually depends on the train velocity. Using linear finite elements and the triangulation as shown in Figure 11, the total number of degrees of freedom is $N = 1074$. The dimension of the reduced space was chosen as $n = 51$. All trajectories are computed at the sensor node placed in the contact zone of the wheel, see Figure 11.
**Figure 10** The number of the fixed-point iterations for solving an NCP problem over time

**Figure 11** Wheel-rail contact
In contrast to the previous example, in this case the underlying dynamical system represents a two-body problem. The mass and stiffness matrices consist of two diagonal blocks where the upper diagonal block refers to the wheel whereas the lower diagonal block refers to the rail body. The coupling between the two bodies is given by the contact condition (45b)–(45d).

In Figure 12, the $x$-displacements of the full and the reduced model, computed at the sensor node, are depicted. The comparison of both trajectories shows a very good agreement between the reduced and the full model. Moreover, the contact pressure and the gap function calculated at the same nodes are shown in Figure 13. Note that between the wheel and the rail a periodic sliding movement occurs. Also, the complementarity between the gap function and the contact pressure retains over time.

Due to the additional contact treatment by Craig–Bampton within our reduction framework, a specific form of contact update is possible. Since the contact nodes and the contact segments are recovered by the reduced model at each time step, the node-segment correspondence can be reestablished, if necessary. Various suitable criterions can be used for such an update. In particular, in case of sliding, nodes can leave the surrounding area of the corresponding segment over time. Therefore, the main idea is based on finding the nearest segment of each contact node.

In Algorithm 2, one iteration of a contact update that can be integrated into our reduction algorithm is introduced. Within this algorithm we make use of the following geometrical observation: For the projection of the node $r \in \mathbb{R}^2$ to the line going through the corresponding segment nodes $p, q \in \mathbb{R}^2$ it holds

$$\Pi(r) = (1 - \alpha)p + \alpha q, \quad \alpha \in \mathbb{R}.$$  \hspace{1cm} (94)

Depending on the value of $\alpha$ we have three different scenarios for the node $r$, see Figure 14B. As soon as $\alpha$ leaves the interval $[0, 1]$, an update step either to the right ($\alpha > 1$) or to the left segment ($\alpha < 0$) is performed. However, a certain
Algorithm 2. Contact update for (45)

1: Input: List of segments $S = (s_1, s_2, \ldots, s_k)$, list of nodes $\mathcal{N} = (p_1, p_2, \ldots, p_{k_N})$.
2: Input: The old selecting function $\Phi : \mathcal{N} \rightarrow S$, TOL $\in (0, 1)$
3: Output: The new selecting function $\tilde{\Phi} : \mathcal{N} \rightarrow S$
4: for $i = 1, \ldots, k_N$ do:
5: Get $s_j = \Phi(p_i)$.
6: Compute the projection $\Pi(p_i)$ w.r.t. the segment $s_j = (p^1_{s_j}, p^2_{s_j})$.
7: Solve $\alpha$ for $\Pi(p_i) = (1 - \alpha)p^1_{s_j} + \alpha p^2_{s_j}$.
8: if $\alpha < -TOL$ and $j > 1$ then
9: Set $\tilde{\Phi}(p_i) = s_{j-1}$
10: end if
11: if $\alpha > 1 + TOL$ and $j < k_S$ then
12: Set $\tilde{\Phi}(p_i) = s_{j+1}$
13: end if
14: end for

tolerance is added for smoothing the discontinuity arising due the update technique, see Algorithm 2. Note that segments are stored within a linear list during the offline procedure. Moreover, there is a linear selecting function that maps the contact nodes to the contact segments. The latter can be updated after each time step and is the output of Algorithm 2.

Finally, we want to compute the stress at the sensor node both for the full and for the reduced model. From the definition of stress (2), it follows that the latter is a function of the gradient of the displacements. Within an FEM framework the gradient of the displacement function $\nabla u_i$ at one node $p_i$ can be approximated by the averaged sum of the gradients computed on all elements $e_j$ that are adjacent to $p_i$, that is,

$$\nabla u_i = \frac{1}{m} \sum_{j=1}^{m} \nabla u_{i|e_j}(p_i),$$

(95)

where $u_{i|e_j}$ denotes the displacement function restricted on the element $e_j$. In case of linear triangular elements, straightforward calculations show that for an element $e_j$ with the nodes $p^f_{i_1}, p^f_{i_2}, p^f_{i_3}$ it holds

$$\nabla u_{i|e_j} = D_j B_j(p^f_{i_1}, p^f_{i_2}, p^f_{i_3}),$$

(96)

where $D_j = \left( u^j_{i_1}, u^j_{i_2}, u^j_{i_3} \right) \in \mathbb{R}^{2 \times 3}$ contains the nodal displacements of the nodes $p^f_{i_1}, p^f_{i_2}, p^f_{i_3}$ and $B = B(p^f_{i_1}, p^f_{i_2}, p^f_{i_3}) \in \mathbb{R}^{3 \times 2}$ depends solely on the initial coordinates of the same nodes. The entries of the matrix $B$ can be easily calculated and are omitted here.
In summary, in order to compute the stress at a node $p_i$, the coordinates and the displacements of all direct neighboring nodes $p_{k_1}, p_{k_2}, p_{k_{m_i}}$ are required. In particular, in our example for the sensor node at the contact zone, there are $m_i = 3$ neighbor nodes. Finally, there are two possible ways to compute the stress from the reduced displacements. One way is to track the rows $i^x, i^y$ of the reduction matrix $Q_{CB}$ defined in (81) fulfilling

$$u_i = (Q_{CB})_{i^x,i^y} q,$$

where $q \in \mathbb{R}^N$ is the full-dimensional discretized displacement vector and the indices $i^x, i^y$ stand for the $x$- and $y$-coordinates of the node $p_i$. In this case in total $2(m_i + 1)$-many rows of the matrix $Q_{CB}$ has to be tracked. Another way is to make use of the Craig–Bampton technique by adding the three neighboring nodes to the set of the master nodes. The latter method allows a direct computation of the stress within the reduced space without any need for tracking the transformation matrix $Q_{CB}$.

8 | CONCLUSION

In Reference 31, we have presented a novel reduction algorithm for linear structural mechanical equations with linear node-to-node constraints. In the current article, we have extended the reduction algorithm for treating of node-to-segment contact problems, where also sliding movements may be considered. The node-to-segment contact formulation leads to quadratic contact conditions, which in terms of the variational formulation are appended by Lagrange multipliers to the energy functional. Because the Kuhn–Tucker condition remains linear in the state, an explicit decoupled expression for the adjoint equation may be derived. Here, the Lagrange multiplier must fulfill nonlinear adjoint equations, which represent a NCP. The latter has to be solved for each time step. From our computational results, it turns out that the corresponding Newton-like algorithm converges in about three iterations. Therefore, the computational effort of the reduction method during a single time step is still comparatively low. All the benefits of the algorithm for node-to-node problems carry over to the extended algorithm designed for node-to-segment contact problems. The solution method was applied for a Hertz contact problem for validation purposes. Furthermore, the performance of the novel reduction method was demonstrated by two computational examples: The first one illustrates the sliding capability of the reduced model and allows to study the NCP-problem in detail. The second application describes the dynamic contact between rail and wheel in a simplified frictionless setting. Here, we additionally need a variable mapping between nodes and corresponding segments. Due to Craig–Bampton partitioning, the node-segment relation can be updated during the online phase. Overall, the reduction method performs very efficiently for contact problems with a small contact zone compared to the total structure. Moreover, for many application in contact mechanics the required predefined of the maximal set of the potential contact nodes is not restrictive. Further research will address friction in the contact zone and the adaption of the corresponding reduction algorithm.

ACKNOWLEDGMENTS
The authors would like to thank Christoph Heinrich and his analysis group for valuable discussions.

DATA AVAILABILITY STATEMENT
Research data are not shared.

ORCID
Diana Manvelyan https://orcid.org/0000-0003-2218-4915

REFERENCES
1. Benner P, Schilders W, Grivet-Talocia S, Quarteroni A, Rozza G, Miguel Silveira L. Model Order Reduction: Volume 3 Applications. De Gruyter; 2020.
2. Lohmann B, Salimbahrami B. Ordnungsreduktion Mittels Krylov-unterraummethoden (order reduction using Krylov subspace methods). at-Automatisierungstechnik. 2004;52(1):30-38.
3. Salimbahrami S, Lohmann B. Order reduction of large scale second-order systems using Krylov subspace methods. Linear Algebra Appl. 2006;415:385-405.
4. Baur U, Benner P, Feng L. Model order reduction for linear and nonlinear systems: a system-theoretic perspective. Max Planck Institute Magdeburg Preprints, MPIMD/14-07; 2014.
5. Farhat C, Avery P, Chapman T, Cortial J. Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency. Int J Numer Methods Eng. 2014;98(9):625-662.

6. Chaturantabut S, Sorensen DC. Discrete empirical interpolation for nonlinear model reduction. Proceedings of the IEEE Conference on Decision and Control; Vol. 32, 2010:4316-4321.

7. Benner P, Ohlberger M, Cohen A, Willcox K. Model Reduction and Approximation: Theory and Algorithms. SIAM; 2017.

8. Radermacher A, Reese S. Pod-based model reduction with empirical interpolation applied to nonlinear elasticity. Int J Numer Methods Eng. 2016;107(6):477-495.

9. Peherstorfer B, Willcox K. Data-driven operator inference for nonintrusive projection-based model reduction. Comput Methods Appl Mech Eng. 2016;306:196-215.

10. Swischuk R, Mainini L, Peherstorfer B, Willcox K. Projection-based model reduction: formulations for physics-based machine learning. Comput Fluids. 2019;179:704-717.

11. Benner P, Kürschner P, Saak J. Frequency-limited balanced truncation with low-rank approximations. SIAM J Sci Comput. 2016;38(1):A471-A499.

12. Besselink B, van de Wouw N, Scherpen JMA, Nijmeijer H. Model reduction for nonlinear systems by incremental balanced truncation. IEEE Trans Automat Contr. 2014;59(10):2739-2753.

13. Kürschner P. Balanced truncation model order reduction in limited time intervals for large systems. Adv Comput Math. 2018;44(6):1821-1844.

14. Bai Z. Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. Appl Numer Math. 2002;43:9-44.

15. Ilyas Ahmad M, Benner P, Jaimoukha I. Krylov subspace methods for model reduction of quadratic-bilinear systems. IET Control Theory Appl. 2016;10(16):2010-2018.

16. Breiten T, Damm T. Krylov subspace methods for model order reduction of bilinear control systems. Syst Control Lett. 2010;59(8):443-450.

17. Benner P, Feng L. Recycling Krylov subspaces for solving linear systems with successively changing right-hand sides arising in model reduction. In: Benner P, Hinze M, ter Maten E, eds. Model Reduction for Circuit Simulation. Springer; 2011:125-140.

18. Wolf T, Panzer HK, Lohmann B. H2 pseudo-optimality in model order reduction by Krylov subspace methods. Proceedings of the 2013 European Control Conference (ECC); 2013:3427-3432; IEEE.

19. Bentbib AH, Jbilou K, Kaouane Y. A computational global tangential Krylov subspace method for model reduction of large-scale mimo dynamical systems. J Sci Comput. 2018;75(3):1614-1632.

20. Wrighers P. Computational Contact Mechanics. Springer; 2006.

21. Capatina A. Variational Inequalities and Frictional Contact Problems. Vol 31. Springer; 2014.

22. Francavilla A, Zienkiewicz OC. An note on numerical computation of elastic contact problems. Int J Numer Methods Eng. 2015;102(5):1162-1191.

23. Petrov EP. A high-accuracy model reduction for analysis of nonlinear vibrations in structures with contact interfaces. Int J Numer Methods Eng. 2011;133(10):102503.

24. Petrov EP. A high-accuracy model reduction for analysis of nonlinear vibrations in structures with contact interfaces. J Eng Gas Turbines Power. 2011;133(10):102503.

25. Hallquist JO. Nike2d: an implicit, finite-deformation, finite-element code for analyzing the static and dynamic response of two-dimensional solids. Technical report, California University, Livermore. Lawrence Livermore Lab; 1979.
43. Josephy NH. *Newton’s Method for Generalized Equations*. Technical report, Wisconsin University-Madison Mathematics Research Center; 1979.

44. Powell MJD. *A Method for Nonlinear Constraints in Minimization Problems*. Academic Press, R. Fletcher edition; 1969.

45. Fischer A. A special newton-type optimization method. *Optimization*. 1992;24:269-284.

46. Cottle RW, Pang JS, Stone RE. *The Linear Complementarity Problem*. SIAM; 2009.

47. Lamperski A. Python: lemkelcp 0.1. [https://pypi.org/project/lemkelcp/](https://pypi.org/project/lemkelcp/)

48. Brenan KE, Campbell SL, Petzold LR. *The Numerical Solution of Initial Value Problems in Ordinary Differential-Algebraic Equations*. SIAM; 1996.

49. Hairer E, Lubich C, Roche M. *The Numerical Solution of Differential-Algebraic Equations by Runge-Kutta Methods*. Lecture Notes in Mathematics. Vol 1409. Springer; 1989.

50. Baur U, Benner P, Feng L. Model order reduction for linear and nonlinear systems: a system-theoretic perspective. *Arch Comput Methods Eng*. 2014;21:331-358.

51. Guyan RJ. Reduction of stiffness and mass matrices. *AIAA J*. 1965;3(2):380.

52. Craig RR Jr, Kurdila AJ. *Fundamentals of Structural Dynamics*. John Wiley & Sons; 2006.

53. Hughes TJR, Taylor RL, Sackman JL, Curnier A, Kanoknukulchai W. A finite element method for a class of contact-impact problems. *Comput Methods Appl Mech Eng*. 1976;8(3):249-276.

**How to cite this article:** Manvelyan D, Simeon B, Wever U. A physics-based model reduction approach for node-to-segment contact problems in linear elasticity. *Int J Numer Methods Eng*. 2022;123(23):5937-5963. doi: 10.1002/nme.7095