



HAL
open science

Hybrid hyper-reduced modeling for contact mechanics problems

Jules Fauque, Isabelle Ramière, David Ryckelynck

► **To cite this version:**

Jules Fauque, Isabelle Ramière, David Ryckelynck. Hybrid hyper-reduced modeling for contact mechanics problems. *International Journal for Numerical Methods in Engineering*, 2018, 115 (1), pp.117-139. 10.1002/nme.5798 . hal-01853252

HAL Id: hal-01853252

<https://minesparis-psl.hal.science/hal-01853252>

Submitted on 11 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hybrid hyper-reduced modeling for contact mechanics problems

J. Fauque^{1,2} | I. Ramière¹  | D. Ryckelynck²

¹CEA, DEN, DEC, SESC, LSC,
Saint-Paul-lez-Durance, France

²MINES ParisTech, PSL Research
University, Centre des Matériaux, CNRS,
Évry, France

Correspondence

I. Ramière, CEA, DEN, DEC, SESC, LSC,
Bâtiment 151, 13108
Saint-Paul-lez-Durance, France.
Email: isabelle.ramiere@cea.fr

Funding information

CEA, EDF and Framatome

Summary

The model reduction of mechanical problems involving contact remains an important issue in computational solid mechanics. In this article, we propose an extension of the hyper-reduction method based on a reduced integration domain to frictionless contact problems written by a mixed formulation. As the potential contact zone is naturally reduced through the reduced mesh involved in hyper-reduced equations, the dual reduced basis is chosen as the restriction of the dual full-order model basis. We then obtain a hybrid hyper-reduced model combining empirical modes for primal variables with finite element approximation for dual variables. If necessary, the inf-sup condition of this hybrid saddle-point problem can be enforced by extending the hybrid approximation to the primal variables. This leads to a hybrid hyper-reduced/full-order model strategy. This way, a better approximation on the potential contact zone is further obtained. A posttreatment dedicated to the reconstruction of the contact forces on the whole domain is introduced. In order to optimize the offline construction of the primal reduced basis, an efficient error indicator is coupled to a greedy sampling algorithm. The proposed hybrid hyper-reduction strategy is successfully applied to a 1-dimensional static obstacle problem with a 2-dimensional parameter space and to a 3-dimensional contact problem between two linearly elastic bodies. The numerical results show the efficiency of the reduction technique, especially the good approximation of the contact forces compared with other methods.

KEYWORDS

contact mechanics, hybrid-order model, hyper-reduction, mixed formulation, model order reduction, reduced integration domain (reduced mesh)

1 | INTRODUCTION

The numerical solution of partial differential equations (PDEs), especially variational inequalities, can become really expensive when accurate predictions are required. The size of the full-order model (FOM) obtained by standard discretization methods (finite difference, finite element (FE), and finite volume) is proportional to the number of discrete unknowns. To tackle this limitation, some techniques based on reduced bases (RBs) have been developed since the 1970s and generalized to various systems. The principle of this kind of technique is to restrict the solution space to a smaller subspace defined from the RB.¹ The reduced-order model (ROM) is then obtained by projecting the initial PDE on the RB.

For nonlinear PDE, one of the most popular methods to a posteriori construct the RB is the snapshot proper orthogonal decomposition (POD).^{2,3}

We focus here on the model reduction of mechanical problems involving contact written by a mixed formulation (saddle-point problem) and numerically solved using Lagrange multipliers that physically represent the contact forces.^{4,5} Building a dual RB (for the Lagrange multipliers) with POD is not convenient because of their positivity constraint. Two recent articles^{6,7} focus on the construction of a positive dual RB. The first article⁶ consists in extending the so-called RB method⁸ to variational inequalities solved due to a mixed formulation. Similar to the primal RB (for the displacements), the dual RB is directly composed of well-chosen Lagrange multiplier snapshots. This dual RB represents precisely the contact forces but can rapidly become of large dimension. In the work of Balajewicz et al,⁷ the projection-based (PB) method proposes two different approaches to building the primal RB and the dual RB. In this method, an efficient primal RB is obtained due to POD. The dual RB is built applying the nonnegative matrix factorization (NNMF) algorithm⁹ to the Lagrange multiplier snapshots. This latter algorithm guarantees positive basis vectors and a limited RB dimension, but the obtained dual RB is far less accurate than the primal RB.

In this paper, we propose to extend another type of model order reduction method called hyper-reduction (HR)¹⁰⁻¹² to contact problems. Hyper-reduction methods have been proposed to limit to few entries the computation of implicit nonlinear balanced equations and, hence, to speed up the computation time. Various approaches have been developed for that purpose, as follows:

- Interpolation methods that estimate the nonlinear terms of interest by interpolating their values at a few spatial locations. For the empirical interpolation method (EIM)¹⁰ and the discrete EIM (DEIM),¹³ the set of interpolation points is generated by the method itself using the snapshots and a POD basis, respectively. In some other model reduction techniques as missing point estimation¹⁴ or Gauss-Newton with approximated tensors,¹⁵ the construction is done using the gappy POD method.¹⁶ The EIM and, respectively, the DEIM have been recently extended to variational inequalities treated by penalty methods in the works of Bader et al¹⁷ and Balajewicz and Toivanen.¹⁸
- Cubature methods^{12,19} that estimate the spacial integrals involved in the weak formulation, by using only few unassembled elemental contributions. The elements of interest and their weighting coefficients to the global contribution are determined using an optimization process.
- Boundary value problems restricted to a reduced integration domain (RID).²⁰ The RID usually involve elements connected to interpolation points computed by the DEIM algorithm,¹³ by considering several POD RBs including the primal one.

We adopt the last approach as it had been successfully applied to elastoplastic problems^{20,21} that usually appear under extreme contact conditions, although we restrict our attention in this paper to elastic problems. For simplicity, the HR-RID approach is termed HR in the sequel. As for the classical HR method, the primal solution is obtained on the whole domain since the related RB is defined on the whole domain.

The proposed extension of the HR method to contact problems consists in conserving a few vectors of the FOM dual basis as the number of contacts is naturally reduced by the RID. Hence, only the contacts in the RID are treated but with a local high fidelity. A strategy based on the solution of a nonnegative least squares (NNLS) problem is introduced to reconstruct contact forces on the whole domain by postprocessing the HR predictions. As for the previous cited ROM for mixed contact problems, the so-called LBB (Ladyzenskaia-Babuška-Brezzi)^{22,23} or inf-sup condition has to be respected for the obtained ROM mixed problem to have a unique solution. The condition imposes the projected contact matrix restrained to the active contacts to have full row rank. As the active contacts are a priori unknown, we propose to extend the condition to the potential contact. For the proposed hybrid HR method for contact problems, the verification of this condition depends on the primal RB and the RID. If required, the respect of this condition can be enforced by coupling the hyper-reduced approximation to a restricted FOM approximation due to a domain decomposition technique of the RID as proposed in the work of Baiges et al.²⁴ Originally, hybrid reduced-order/full-order models have been proposed in the literature to circumvent the lack of accuracy of RB approximations; see, for example, the application to elliptic problems in elasticity,²⁵ to nonlinear structural problems involving plasticity or damage^{26,27} or to Navier-Stokes equations.²⁴ The method introduced here can be viewed as a hybrid HR/FOM approach to deal with contact problems solved with a mixed formulation.

In order to choose pertinent snapshots, a greedy sampling algorithm can be applied as initially proposed in the works of Prud'Homme et al.^{28,29} This strategy ensures the robustness of the ROM with respect to variations of the model parameters.

A simple but efficient error indicator adapted to the HR is proposed. This indicator consists in the difference between the FOM solved on the RID submitted to interpolated Dirichlet boundary conditions and the HR solution.

This article is structured as follows. The next section is devoted to the introduction of some notations. In Section 3, the contact problem written by a mixed formulation and its solution with the FE method is briefly recalled. Section 4 emphasizes the main points of the HR method, whereas Section 5 is devoted to the extension of the HR method to contact problems. In Section 6, the proposed hybrid HR approach is verified on a 1-dimensional (1D) (obstacle-type inequality) test case derived from the literature.^{6,7,17} A comparison between the hybrid HR method and the PB method introduced in the work of Balajewicz et al⁷ is done. Finally, in Section 7, a two-elastic-solid unilateral contact 3-dimensional (3D) problem is treated with the proposed approach.

2 | NOTATION

Tensors and hypermatrices are underlined as many times as their order. For example, vectors are once underlined $\underline{\cdot}$, whereas matrices are twice underlined $\underline{\underline{\cdot}}$, etc. Moreover, $\underline{0}$ (respectively, $\underline{\underline{0}}$) identifies a vector (respectively, a matrix) of zeros. Subscripts identify the components of a tensor. For example, $\underline{\underline{A}}_{ij}$ denotes the component located at the i th row and the j th column of the matrix $\underline{\underline{A}}$.

$\underline{\underline{I}}_N$ denotes the identity matrix of size $N \times N$.

We use the Python notation to restrict vectors/matrices to some rows or columns. For example, the restriction of a matrix $\underline{\underline{A}}$ to some rows with the set of indices \mathcal{L} is written as $\underline{\underline{A}}[\mathcal{L}, :]$.

We denote by \odot the element-wise multiplication. For example, $\underline{\underline{A}} \odot \underline{\underline{B}}_{ij} = (\underline{\underline{A}})_{ij}(\underline{\underline{B}})_{ij}$ or $(\underline{a} \odot \underline{B})_{ij} = (\underline{a})_i(\underline{B})_{ij}$.

The symbol “:” designates the double inner product that consists in a double contraction over the last two indices of the first tensor and the first two indices of the second tensor. For example, $\underline{\underline{A}} : \underline{\underline{B}} = \sum_{ij} (\underline{\underline{A}})_{ij}(\underline{\underline{B}})_{ij}$ is a scalar, whereas $\underline{\underline{C}} : \underline{\underline{B}} = \sum_{kl} (\underline{\underline{C}})_{ijkl}(\underline{\underline{B}})_{kl}$ gives a matrix.

The positive and the negative part of $x \in \mathbb{R}$ are respectively defined as $[x]_+ = \max(x, 0)$ and $[x]_- = \min(x, 0)$ and those of a vector $\underline{x} \in \mathbb{R}^N$ as $[\underline{x}]_+ = ([(\underline{x})_i]_+)_{i=1}^N$ and $[\underline{x}]_- = ([(\underline{x})_i]_-)_{i=1}^N$.

The Euclidian norm or 2-norm of a vector is denoted by $\|\cdot\|_2$

$$\|\underline{u}\|_2 = (\underline{u}^T \underline{u})^{\frac{1}{2}} = \left(\sum_i (\underline{u}_i)^2 \right)^{\frac{1}{2}}, \quad (1)$$

where the superscript T designates the transpose operation.

The L^2 inner product defined on Ω is denoted by $\langle \cdot, \cdot \rangle_{L^2(\Omega)}$, such that

$$\langle \underline{u}, \underline{v} \rangle_{L^2(\Omega)} = \int_{\Omega} \underline{u}^T \underline{v} \, d\Omega. \quad (2)$$

The related L^2 norm defined over the domain Ω is denoted by $\|\cdot\|_{L^2(\Omega)}$ and defined as

$$\|\underline{u}\|_{L^2(\Omega)} = (\langle \underline{u}, \underline{u} \rangle_{L^2(\Omega)})^{\frac{1}{2}} = \left(\int_{\Omega} \underline{u}^T \underline{u} \, d\Omega \right)^{\frac{1}{2}},$$

whereas the H^1 norm writes

$$\|\underline{u}\|_{H^1(\Omega)} = \left(\|\underline{u}\|_{L^2(\Omega)}^2 + \sum_{k=1}^D \|\partial_{x_k} \underline{u}\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}$$

with D as the space dimension.

Let E denote a Hilbert space and F its dual space. Their inner products are respectively denoted by $\langle \cdot, \cdot \rangle_E$ and $\langle \cdot, \cdot \rangle_F$. Introducing a basis of E , $\{\underline{\beta}_i\}_{i=1}^{H_E}$, we can write $\forall \underline{u}, \underline{u}' \in E$, $\langle \underline{u}, \underline{u}' \rangle_E = \underline{\xi}_u^T \underline{\underline{M}}^E \underline{\xi}_{u'}$ with $\underline{\xi}_u$ and $\underline{\xi}_{u'}$ as the coefficients associated to \underline{u} and \underline{u}' in the E basis and $\underline{\underline{M}}^E = (\langle \underline{\beta}_i, \underline{\beta}_j \rangle_E)_{i,j=1}^{H_E}$. We then have $\langle \underline{v}, \underline{v}' \rangle_F = \underline{\xi}_v^T (\underline{\underline{M}}^E)^{-1} \underline{\xi}_{v'}$, $\forall \underline{v}, \underline{v}' \in F$ with $\underline{\xi}_v$ and $\underline{\xi}_{v'}$ as the coefficients associated to \underline{v} and \underline{v}' in the F basis.

3 | CONTACT PROBLEM – FINITE ELEMENT MODEL

We are interested in the solution of elastic frictionless unilateral contact problems inherently nonlinear, involving variational inequalities and constrained minimizations.

For the sake of clarity and without any loss of generality, we consider here a model problem involving two deformable elastic bodies defined by the two open-bounded domains $\Omega_1 \in \mathbb{R}^D$ and $\Omega_2 \in \mathbb{R}^D$ with D as the dimension of the problem. The whole domain is defined by $\Omega = \Omega_1 \cup \Omega_2$ (see Figure 1). We denote by $\partial\Omega = \partial\Omega_1 \cup \partial\Omega_2$ the boundary of Ω . This boundary can be divided into three parts Γ^D , Γ^N , and Γ^C such that $\partial\Omega = \Gamma^D \cup \Gamma^N \cup \Gamma^C$ with $\Gamma^D \cap \Gamma^N = \emptyset$ and $\Gamma^D \cap \Gamma^C = \emptyset$.

- On Γ^D , displacement conditions are imposed (Dirichlet boundary conditions).
- On Γ^N , surface forces are applied (Neumann boundary conditions).
- Γ^C is the possible and a priori unknown contact zone (contact boundary conditions). We have $\Gamma^C = \Gamma_1^C \cup \Gamma_2^C$ with $\Gamma_i^C, i \in \{1, 2\}$ as the potential contact boundary on $\Omega_i, i \in \{1, 2\}$.

Let \underline{n}_i denote the external unit normal to Ω_i on Γ^C . We take \underline{u}_i (respectively, $\underline{\sigma}_i$) to denote the displacement field (respectively, stress field) in Ω_i . The points on Γ^C are paired by minimal distance in order to evaluate the gap denoted by u_N on each point of Γ^C . Let us denote by $\{\underline{x}, \underline{x}'\}$ a pair of coupled points between Γ_1^C and Γ_2^C . Then, the gap u_N writes

$$u_N = \underline{u}_1(\underline{x})^T \underline{n}_1(\underline{x}) + \underline{u}_2(\underline{x}')^T \underline{n}_2(\underline{x}'), \quad \forall \text{ paired couple } \{\underline{x}, \underline{x}'\} \in (\Gamma_1^C, \Gamma_2^C). \quad (3)$$

Under small perturbation hypothesis, $\underline{n}_1(\underline{x}) = -\underline{n}_2(\underline{x}'), \forall \{\underline{x}, \underline{x}'\} \in (\Gamma_1^C, \Gamma_2^C)$. The unknown contact forces \underline{F}^i on Γ_i^C are decomposed into normal and tangential parts, ie,

$$\underline{F}^i(\underline{x}) = \underline{\sigma}_i(\underline{x}) \underline{n}_i = F_N^i(\underline{x}) \underline{n}_i(\underline{x}) + \underline{F}_T^i(\underline{x}), \quad \forall \underline{x} \in \Gamma_i^C. \quad (4)$$

In this paper, we restrict our attention to frictionless contact. Signorini's law of unilateral contact³⁰ is given on Γ^C by

$$\underline{F}_T^i = \underline{0} \quad (\text{frictionless}), \quad (5)$$

$$u_N \leq d \quad (\text{nonpenetration}), \quad (6)$$

$$F_N = F_N^1 = F_N^2 \leq 0 \quad (\text{nonadhesion}), \quad (7)$$

$$(u_N - d)F_N = 0 \quad (\text{complementary slackness}), \quad (8)$$

where d generically denotes the initial gap on each node of Γ^C .

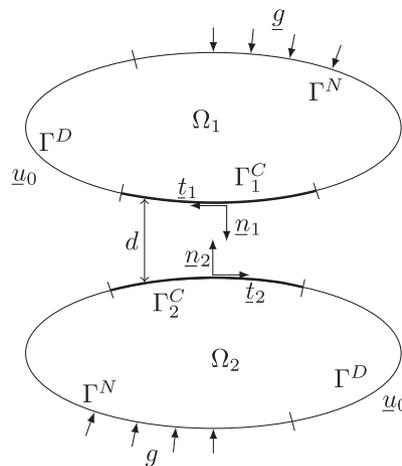


FIGURE 1 Geometry of a generic 2-body contact problem

The strong formulation of the elastostatic frictionless unilateral contact problem can then be expressed as

$$\left\{ \begin{array}{ll} -\operatorname{div} \underline{\underline{\sigma}} = \underline{\underline{f}} & \text{in } \Omega, \\ \underline{\underline{\sigma}} = \underline{\underline{C}} : \underline{\underline{\varepsilon}} & \text{in } \Omega, \\ \underline{\underline{\varepsilon}} = \frac{1}{2} (\operatorname{grad} \underline{\underline{u}} + (\operatorname{grad} \underline{\underline{u}})^T) & \text{in } \Omega, \\ \underline{\underline{u}} = \underline{\underline{u}}_0 & \text{on } \Gamma^D, \\ \underline{\underline{\sigma}} \underline{\underline{n}} = \underline{\underline{g}} & \text{on } \Gamma^N, \\ u_N \leq d, F_N \leq 0, (u_N - d)F_N = 0 & \text{on } \Gamma^C, \\ \underline{\underline{\sigma}} \underline{\underline{n}} = F_N \underline{\underline{n}} & \text{on } \Gamma^C. \end{array} \right. \quad (9)$$

The primal variational formulation of (9) can be written as a variational inequality³¹

$$\left\{ \begin{array}{l} \text{Find } \underline{\underline{u}} \in \mathcal{V}_C \text{ such that} \\ a(\underline{\underline{u}}, \underline{\underline{v}} - \underline{\underline{u}}) \geq l(\underline{\underline{v}} - \underline{\underline{u}}), \quad \forall \underline{\underline{v}} \in \mathcal{V}_C \end{array} \right. \quad (10)$$

with $\mathcal{V}_C = \{\underline{\underline{v}} \in \mathcal{V} \mid v_N \leq d \text{ on } \Gamma^C\}$, $\mathcal{V} = \{\underline{\underline{v}} = \underline{\underline{r}}(\underline{\underline{u}}_0) + \underline{\underline{u}} \mid \underline{\underline{u}} \in \mathcal{V}_0\}$, $\underline{\underline{r}}(\underline{\underline{u}}_0)$ as the lifting of $\underline{\underline{u}}_0$ over Ω , $\mathcal{V}_0 = \{\underline{\underline{v}} \in (H^1(\Omega))^D \mid \underline{\underline{v}} = \underline{\underline{0}} \text{ on } \Gamma^D\}$, and

$$a(\underline{\underline{u}}, \underline{\underline{v}}) = \int_{\Omega} \underline{\underline{\varepsilon}}(\underline{\underline{u}}) : \underline{\underline{C}} : \underline{\underline{\varepsilon}}(\underline{\underline{v}}) d\Omega, \quad (11)$$

$$l(\underline{\underline{v}}) = \int_{\Omega} \underline{\underline{f}}^T \underline{\underline{v}} d\Omega + \int_{\Gamma^N} \underline{\underline{g}}^T \underline{\underline{v}} d\Gamma. \quad (12)$$

Due to the variational inequality, the unilateral contact problem is nonlinear. In the work of Duvaut and Lions,³¹ it is demonstrated that the solution of inequality (10) can be obtained by minimizing the potential energy, ie,

$$\min_{\underline{\underline{v}} \in \mathcal{V}_C} J(\underline{\underline{v}}) = \frac{1}{2} a(\underline{\underline{v}}, \underline{\underline{v}}) - l(\underline{\underline{v}}). \quad (13)$$

The minimization of J is then equivalent to a saddle-point problem when introducing Lagrange multipliers.^{32,33} The saddle-point problem can be written as

$$\left\{ \begin{array}{ll} \text{Find } (\underline{\underline{u}}, \lambda) \in \mathcal{V} \times \mathcal{W} \text{ such that} \\ a(\underline{\underline{u}}, \underline{\underline{v}}) + b(\underline{\underline{v}}, \lambda) = l(\underline{\underline{v}}), & \forall \underline{\underline{v}} \in \mathcal{V}_0 \\ b(\underline{\underline{u}}, \kappa - \lambda) \leq \langle d, \kappa - \lambda \rangle_{H^{\frac{1}{2}}(\Gamma^C), H^{-\frac{1}{2}}(\Gamma^C)}, & \forall \kappa \in \mathcal{W} \end{array} \right. \quad (14)$$

with $\mathcal{W} = \{\kappa \in H^{-\frac{1}{2}}(\Gamma^C) \mid \kappa \geq 0\}$, where $H^{-\frac{1}{2}}$ denotes the dual of $H^{\frac{1}{2}}$, $b(\underline{\underline{v}}, \kappa) = \langle v_N, \kappa \rangle_{H^{\frac{1}{2}}(\Gamma^C), H^{-\frac{1}{2}}(\Gamma^C)}$, and $\langle \cdot, \cdot \rangle_{H^{\frac{1}{2}}, H^{-\frac{1}{2}}}$ is the duality product. We can notice that the solution $(\underline{\underline{u}}, \lambda)$ recovers the complementary slackness condition. Also, the resulting Lagrange multiplier is physically representative of the contact forces, $\lambda = -F_N$.

We now apply the FEM. The shape functions of the FE basis are denoted by $(\varphi_j)_{j=1}^m$ with m as the number of discretization nodes. For the displacement discretization, we introduce $\underline{\underline{\varphi}}_i = \varphi_j \underline{\underline{e}}_k$ with $i = (j-1)D + k, j \in \{1, \dots, m\}$, $k \in \{1, \dots, D\}$ and $\underline{\underline{e}}_k$ as the canonical vectors of a Cartesian coordinate system. The decomposition of the displacement on the $(\underline{\underline{\varphi}}_i)_{i=1}^N, N = mD$ basis is written as

$$\underline{\underline{u}}_h = \underline{\underline{r}}(\underline{\underline{u}}_0) + \sum_{i=1}^N \underline{\underline{\varphi}}_i u_i, \quad (15)$$

where $\underline{\underline{u}}_h$ is the approximate FE solution and $\underline{\underline{U}} = (u_i)_{i=1}^N$ is the vector of the related degrees of freedom (DOFs). The nonhomogeneous Dirichlet condition is treated using usual techniques existing in the FE solvers (Lagrange multipliers, penalty, elimination, lifting, etc). For the sake of clarity, it will not appear in the discrete approximations written

in what follows. The FOM, with N as the total number of primal DOFs and N_λ as the total number of potentially active contacts, can be written as

$$\begin{cases} \text{Find } (\underline{U}, \underline{\Lambda}) \in \mathbb{R}^N \times (\mathbb{R}^+)^{N_\lambda} \text{ such that} \\ \underline{K}\underline{U} + \underline{B}^T \underline{\Lambda} = \underline{F} \\ \underline{\Lambda} \odot (\underline{B}\underline{U} - \underline{D}) = \underline{0} \\ \underline{B}\underline{U} \leq \underline{D}, \end{cases} \quad (16)$$

where $\underline{K} \in \mathbb{R}^{N \times N}$, $\underline{B} \in \mathbb{R}^{N_\lambda \times N}$, $\underline{F} \in \mathbb{R}^N$, $\underline{D} \in \mathbb{R}^{N_\lambda}$, \underline{U} , and $\underline{\Lambda}$ are, respectively, the discretization of a , b , l , d , u , and λ in (14). \underline{K} is commonly called the stiffness matrix, and \underline{B} is the potential contact matrix where each line is associated to a potential contact node.

4 | HYPER-REDUCTION METHOD

A common approach for model order reduction is projection-based reduction, which aims to reduce the number of DOFs by projecting the balance equation on a smaller subspace. Here, HR is applied in the framework of a posteriori model reduction methods. In this category of methods, it is common to follow a decomposition of computational tasks in two phases, namely, offline and online. The offline phase is the part where parameter-independent quantities like RB or RID are precomputed. It is time consuming because it requires solutions of the FOM for different parameters but permits to the online phase, where the ROM is used, to be faster. This kind of method becomes useful for real-time simulation or when a large number of simulations needs to be done like for a parametric study or the solution of an optimization problem.

The HR²¹ method uses projection on RB to reduce the number of DOFs, but it also uses the fact that solving the equations on a reduced mesh is sufficient to find the reduced DOFs. This particularity improves the computational time savings, especially for nonlinearities that cannot be precomputed offline. Indeed, the RID, built during the offline part, reduces the cost of the projections to get the reduced nonconstant stiffness matrix from the Jacobian matrix.

4.1 | Reduced basis construction

For nonlinear problems, the most famous method for building an RB is the POD method. It gives the optimal low-rank approximation of the minimization problem, which consists in compressing the simulation data generated during the offline phase. This method will hence be used here to reduce the primal basis. The 2-norm POD basis can be easily computed by collecting snapshots and applying the singular value decomposition (SVD) algorithm or computing the eigendecomposition of the correlation matrix. Let us consider a snapshot matrix \underline{S} of dimension $N \times N_s$. Computing the SVD on \underline{S} gives

$$\underline{S} = \underline{V}_s \underline{\Sigma}_s \underline{W}_s^T, \quad (17)$$

where $\underline{V}_s \in \mathbb{R}^{N \times N}$ and $\underline{W}_s \in \mathbb{R}^{N_s \times N_s}$ are orthonormal matrices and $\underline{\Sigma}_s \in \mathbb{R}^{N \times N_s}$ is a diagonal matrix containing singular values σ_j ordered by decreasing size. The 2-norm POD basis of rank l_s is then denoted \underline{V} and obtained, according to the Eckart-Young theorem,³⁴ by taking the l_s first left-singular vectors, $\underline{V} = \underline{V}_s[:, \{1, \dots, l_s\}]$. To choose l_s , we introduce the information rate $\mathcal{E}(l_s)$ of the l_s first singular vectors defined as follows:

$$\mathcal{E}(l_s) = \frac{\sum_{j=1}^{l_s} \sigma_j^2}{\sum_{j=1}^{\min(N, N_s)} \sigma_j^2}. \quad (18)$$

l_s is then obtained by solving $\text{argmin}_{l_s \in \mathbb{N}} (\mathcal{E}(l_s) \geq 1 - \epsilon_{\text{tol}})$ with ϵ_{tol} as a given threshold.

This way, we obtain an RB composed of a few vectors able to approximate the space spanned by the snapshots. The FE basis enables the recovery of the continuous empirical modes denoted by $(\underline{\psi}_k)_{k=1}^{l_s}$:

$$\underline{\psi}_k = \sum_{i=1}^N \underline{\varphi}_i v_{ik} \quad \text{with} \quad v_{ik} = (\underline{V})_{ik}. \quad (19)$$

4.2 | Reduced integration domain construction

There are different ways to build an RID. One of them is to use a manufacturing construction knowing important zones needed to be included. A more generic method, used in this article, is based on the DEIM procedure.¹³ The procedure takes as input an RB, denoted $\underline{\underline{V}} \in \mathbb{R}^{N \times l_s}$, and returns a set of interpolation indices \mathcal{L} selected so that each of them corresponds to the entry with the largest magnitude of the error between each vector of the input basis and its approximation doing an empirical interpolation of the vectors already treated. The final number of indices is equal to the number of vectors in the RB: $\text{card}(\mathcal{L}) = l_s$. A POD basis is a suitable choice for this algorithm as the vectors are ordered by importance. An important result is that the restriction of the RB $\underline{\underline{V}}$ to the interpolation indices list \mathcal{L} , $\underline{\underline{V}}[\mathcal{L}, :] \in \mathbb{R}^{l_s \times l_s}$, is a full rank matrix and then any extension of the interpolation indices list \mathcal{L}_{ext} of size $l_{\text{ext}} \geq l_s$ would retain $\underline{\underline{V}}[\mathcal{L}_{\text{ext}}, :] \in \mathbb{R}^{l_{\text{ext}} \times l_s}$ an RB. The RID is defined as

$$\Omega_A = \cup_{i \in \mathcal{L}} \text{supp}(\underline{\varphi}_i), \quad (20)$$

where $\text{supp}(\underline{\varphi}_i)$ denotes the FE basis $\underline{\varphi}_i$ support. Usually, as proposed in the work of Ryckelynck et al,²¹ the DEIM algorithm is applied to several RBs, including the RB related to dual variable, to obtain an RID apprehending more information. If required to improve the accuracy of the HR predictions, the RID can also be supplemented by one or several layers of connected elements, as well as a region of interest.

We denote by Ω_B the counterpart of Ω_A such that $\Omega = \Omega_A \cup \Omega_B$ and $\Gamma^I = \Omega_A \cap \Omega_B$, where Γ^I is the interface between the two subdomains. We introduce \mathcal{I} as

$$\mathcal{I} = \left\{ i \in \{1, \dots, N\} \mid \int_{\Gamma^I} \underline{\varphi}_i^T \underline{\varphi}_i \, d\Gamma \neq 0 \right\}. \quad (21)$$

We introduce \mathcal{A} , which denotes the set of all DOFs that are not connected to Ω_B , as follows:

$$\mathcal{A} = \left\{ i \in \{1, \dots, N\} \mid \int_{\Omega_B} \underline{\varphi}_i^T \underline{\varphi}_i \, d\Omega = 0 \right\} \quad (22)$$

and $\underline{\psi}_k^A$, which is the setting to $\underline{0}$ of $\underline{\psi}_k$ over Ω_B , as

$$\underline{\psi}_k^A(\underline{x}) = \sum_{i \in \mathcal{A}} \underline{\varphi}_i(\underline{x}) v_{ik}, \quad \forall \underline{x} \in \Omega. \quad (23)$$

The counterpart of the restricted mode $\underline{\psi}_k^A$ is denoted by $\underline{\psi}_k^B$ such that

$$\underline{\psi}_k^B(\underline{x}) = \underline{\psi}_k(\underline{x}) - \underline{\psi}_k^A(\underline{x}), \quad \forall \underline{x} \in \Omega, \quad k = 1, \dots, l_s. \quad (24)$$

4.3 | Usual hyper-reduced balance equations without contact conditions

Let us consider problem (9) without contact ($\Gamma^C = \emptyset$). The primal variational formulation reads:

$$\begin{cases} \text{Find } \underline{u} \in \mathcal{V} \text{ such that} \\ a(\underline{u}, \underline{v}) = l(\underline{v}), \quad \forall \underline{v} \in \mathcal{V}_0 \end{cases} \quad (25)$$

with a and l defined as before in (11) and (12).

In the work of Ryckelynck et al,²¹ the setting of the hyper-reduced balance equations starts by defining a surrogate space for \mathcal{V} in order to solve a reduced problem in Ω_A . This subspace reads $\mathcal{V}_{\text{HR}}^{\Omega_A} = \{ \underline{v} = \underline{r}_{\Omega_A}(\underline{u}_0) + \tilde{r}_{\Omega_A}(\underline{u}_I) + \tilde{\underline{u}} \mid \tilde{\underline{u}} \in \text{span}\{\underline{\psi}_k^A\}_{k=1}^{l_u} \}$ with \underline{u}_I as an additional boundary condition on Γ^I , $\tilde{r}_{\Omega_A}(\underline{u}_I)$ as its lifting over Ω_A , $\underline{r}_{\Omega_A}(\underline{u}_0)$ as the restriction of $\underline{r}(\underline{u}_0)$ to Ω_A , and $\{\underline{\psi}_k^A\}_{k=1}^{l_u}$ as the primal restricted POD modes. The proposed hyper-reduced balance equation is defined as

$$\begin{cases} \text{Find } \underline{u} \in \mathcal{V}_{\text{HR}}^{\Omega_A} \text{ such that} \\ a_{\Omega_A}(\underline{u}, \underline{v}) = l_{\Omega_A}(\underline{v}), \quad \forall \underline{v} \in \text{span}\left\{ \underline{\psi}_k^A \right\}_{k=1}^{l_u}, \end{cases} \quad (26)$$

where a_{Ω_A} and l_{Ω_A} denote the forms a and l where the integration is reduced to Ω_A since \underline{v} vanishes on Ω_B . Then, the approximate displacement reads:

$$\underline{u}(\underline{x}) = \underline{r}_{\Omega_A}(\underline{u}_0)(\underline{x}) + \tilde{\underline{r}}_{\Omega_A}(\underline{u}_I)(\underline{x}) + \sum_{k=1}^{l_u} \underline{\psi}_k^A(\underline{x}) \gamma_k, \quad \forall \underline{x} \in \Omega_A. \quad (27)$$

In the standard HR method,²¹ the lifting of \underline{u}_I is set up over Ω by

$$\tilde{\underline{r}}(\underline{u}_I)(\underline{x}) = \sum_{k=1}^{l_u} \underline{\psi}_k^B(\underline{x}) \gamma_k^B, \quad \forall \underline{x} \in \Omega, \quad (28)$$

where γ_k^B are additional parameters chosen as $\gamma_k^B = \gamma_k$, $k = 1, \dots, l_u$ in order to recover the usual RB approximation. In particular, $\tilde{\underline{r}}_{\Omega_A}(\underline{u}_I)$ is the restriction to Ω_A of $\tilde{\underline{r}}(\underline{u}_I)$, and \underline{u}_I is the trace on Γ^I of $\tilde{\underline{r}}(\underline{u}_I)(\underline{x})$:

$$\underline{u}_I(\underline{x}) = \sum_{k=1}^{l_u} \underline{\psi}_k^B(\underline{x}) \gamma_k^B \quad \underline{x} \in \Gamma^I. \quad (29)$$

Then, the approximate displacement is defined over Ω by

$$\underline{u}(\underline{x}) = \underline{r}(\underline{u}_0)(\underline{x}) + \sum_{k=1}^{l_u} \underline{\psi}_k(\underline{x}) \gamma_k, \quad \forall \underline{x} \in \Omega. \quad (30)$$

Then, $u \in \mathcal{V}_{\text{POD}} = \{\underline{v} = \underline{r}(\underline{u}_0) + \tilde{\underline{u}} \mid \tilde{\underline{u}} \in \text{span}\{\underline{\psi}_k\}_{k=1}^{l_u}\}$.

After discretization, we obtain problem (31) written with a Petrov-Galerkin formulation.

$$\left\{ \begin{array}{l} \text{Find } \underline{\gamma} \in \mathbb{R}^{l_u} \text{ such that} \\ \underline{V}[\mathcal{A}, :]^T \underline{K}[\mathcal{A}, \mathcal{A} \cup \mathcal{I}] \underline{V}[\mathcal{A} \cup \mathcal{I}, :] \underline{\gamma} = \underline{V}[\mathcal{A}, :]^T \underline{F}[\mathcal{A}] \end{array} \right. \quad (31)$$

Remark. The restriction of $\underline{V} \underline{\gamma}$ on $\mathcal{A} \cup \mathcal{I}$ is due to the structure of \underline{K} . The support of the FE basis functions implies that only DOFs in $\mathcal{A} \cup \mathcal{I}$ are connected to DOFs in \mathcal{A} through \underline{K} .

5 | EXTENSION OF THE HR METHOD TO CONTACT PROBLEMS

5.1 | Additional boundary for the HR model

For contact problems, with the standard RID construction described Section 4.2, the gap may not be computable on all points contained in $\Gamma^C \cap \partial\Omega_A$ (see Figure 2). We then need to extend Γ^I to Γ_I^C , which denotes the part of $\Gamma^C \cap \partial\Omega_A$ where the gap cannot be computed because of missing opposite points or, in other words, when only one point of the pair of coupled points is in $\Gamma^C \cap \partial\Omega_A$. Indeed, no Lagrange multiplier can be computed on a boundary where the gap cannot be estimated. The extended interface permits to enforce additional pseudo-Dirichlet boundary conditions (see Equation (18)) on Γ_I^C instead of writing balance conditions. It reads:

$$\Gamma^I = (\Omega_A \cap \Omega_B) \cup \Gamma_I^C. \quad (32)$$

We must then adapt the definition of \mathcal{A} . The convenient restriction reads:

$$\mathcal{A} = \left\{ i \in \{1, \dots, N\} \mid \int_{\Omega_B} \underline{\varphi}_i^T \underline{\varphi}_i \, d\Omega + \int_{\Gamma_I^C} \underline{\varphi}_i^T \underline{\varphi}_i \, d\Gamma = 0 \right\}, \quad (33)$$

where \mathcal{A} is the set of all displacement DOFs that are neither connected to Ω_B nor located on Γ_I^C . Then, the reduced Hertz-Signorini-Moreau equations are set only over $\Gamma_A^C = (\Gamma^C \cap \partial\Omega_A) \setminus \Gamma_I^C$ (see Figure 2).

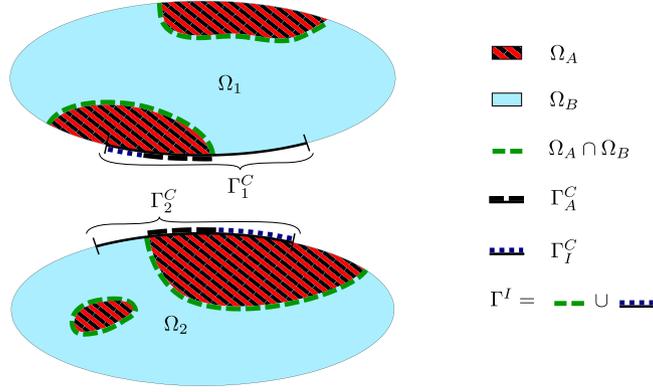


FIGURE 2 Reduced integration domain representation for a generic 2-body contact problem [Colour figure can be viewed at wileyonlinelibrary.com]

5.2 | Hybrid hyper-reduced balance equations for contact

We propose to extend the HR method to the mixed contact problem by preserving a local FE approximation of the Lagrange multipliers. The number of Lagrange multipliers is then bounded by construction due to the restriction of the potential contact zone to the RID. The hybrid HR saddle-point problem reads:

$$\left\{ \begin{array}{l} \text{Find } (\underline{u}, \lambda) \in \mathcal{V}_{\text{HR}}^{\Omega_A} \times \mathcal{W}_{\text{HR}}^{\Omega_A} \text{ such that} \\ a_{\Omega_A}(\underline{u}, \underline{v}) + b_{\Gamma_A^C}(\underline{v}, \lambda) = l_{\Omega_A}(\underline{v}), \quad \forall \underline{v} \in \text{span} \left\{ \underline{\psi}_k^A \right\}_{k=1}^{l_u} \\ b_{\Gamma_A^C}(\underline{u}, \kappa - \lambda) \leq \langle d, \kappa - \lambda \rangle_{H^{\frac{1}{2}}(\Gamma_A^C), H^{-\frac{1}{2}}(\Gamma_A^C)}, \quad \forall \kappa \in \mathcal{W}_{\text{HR}} \end{array} \right. \quad (34)$$

with $\mathcal{W}_{\text{HR}}^{\Omega_A} = \{\kappa \in H^{-\frac{1}{2}}(\Gamma_A^C) \mid \kappa \geq 0\}$ and $b_{\Gamma_A^C}(\underline{v}, \kappa) = \langle v_N, \kappa \rangle_{H^{\frac{1}{2}}(\Gamma_A^C), H^{-\frac{1}{2}}(\Gamma_A^C)}$. This way, the solution \underline{u} verifies $\langle u_N - d, \lambda \rangle_{H^{\frac{1}{2}}(\Gamma_A^C), H^{-\frac{1}{2}}(\Gamma_A^C)} = 0$.

We introduce \mathcal{A}_c as the set of DOFs of \mathcal{A} located on the potential contact zone, \mathcal{A}_λ as the set of dual DOFs connected to \mathcal{A}_c by the matrix $\underline{\underline{B}}$, and \mathcal{A}'_c as the set of primal DOFs connected to \mathcal{A}_λ by $\underline{\underline{B}}$, ie,

$$\mathcal{A}_c = \left\{ j \in \mathcal{A} \mid \exists i \in \{1, \dots, N_\lambda\} \text{ s.t. } \underline{\underline{B}}_{ij} \neq 0 \right\} \quad (35)$$

$$\mathcal{A}_\lambda = \left\{ i \in \{1, \dots, N_\lambda\} \mid \exists j \in \mathcal{A}_c \text{ s.t. } \underline{\underline{B}}_{ij} \neq 0 \right\} \quad (36)$$

$$\mathcal{A}'_c = \left\{ j \in \{1, \dots, N\} \mid \exists i \in \mathcal{A}_\lambda \text{ s.t. } \underline{\underline{B}}_{ij} \neq 0 \right\}. \quad (37)$$

Then, we have $\mathcal{A}_c \subseteq \mathcal{A}'_c$ but not necessarily $\mathcal{A}'_c \subseteq \mathcal{A}$. If the contact problem is an obstacle problem (see Section 6.1) or the contact is numerically treated with a node-to-node algorithm requiring matching meshes, then $\mathcal{A}'_c = \mathcal{A}_c$. We denote by N_λ^C the cardinal of \mathcal{A}_λ .

After discretization, we obtain problem (38) to be solved in the online phase, as follows:

$$\left\{ \begin{array}{l} \text{Find } (\underline{\gamma}, \underline{\Delta}[\mathcal{A}_\lambda]) \in \mathbb{R}^{l_u} \times (\mathbb{R}^+)^{N_\lambda^C} \text{ such that} \\ \underline{\underline{V}}[\mathcal{A}, :]^T \underline{\underline{K}}[\mathcal{A}, \mathcal{A} \cup \mathcal{I}] \underline{\underline{V}}[\mathcal{A} \cup \mathcal{I}, :] \underline{\gamma} + \underline{\underline{V}}[\mathcal{A}_c, :]^T \underline{\underline{B}}[\mathcal{A}_\lambda, \mathcal{A}_c]^T \underline{\Delta}[\mathcal{A}_\lambda] = \underline{\underline{V}}[\mathcal{A}, :]^T \underline{\underline{F}}[\mathcal{A}] \\ \underline{\Delta}[\mathcal{A}_\lambda] \odot \left(\underline{\underline{B}}[\mathcal{A}_\lambda, \mathcal{A}'_c] \underline{\underline{V}}[\mathcal{A}'_c, :] \underline{\gamma} - \underline{\underline{D}}[\mathcal{A}_\lambda] \right) = \underline{0} \\ \underline{\underline{B}}[\mathcal{A}_\lambda, \mathcal{A}'_c] \underline{\underline{V}}[\mathcal{A}'_c, :] \underline{\gamma} \leq \underline{\underline{D}}[\mathcal{A}_\lambda], \end{array} \right. \quad (38)$$

where $\underline{\underline{B}}[\mathcal{A}_\lambda, :]$, $\underline{\underline{D}}[\mathcal{A}_\lambda]$, and $\underline{\Delta}[\mathcal{A}_\lambda]$ are, respectively, the potential contact matrix, the vector of the initial gap, and the approximation of the Lagrange multipliers restricted to \mathcal{A}_λ . One of the benefits of this formulation is that we are verifying the Signorini contact conditions in the RID (without projection).

Property 1. In the case $\mathcal{A}'_c = \mathcal{A}_c$, if $\underline{\underline{B}}[\mathcal{A}_\lambda, \mathcal{A}_c] \underline{\underline{V}}[\mathcal{A}_c, :]$ has full row rank, then Problem (38) is consistent with the FE formulation (16).

Proof. We suppose $\mathcal{A}'_c = \mathcal{A}_c$. A necessary condition for the saddle-point contact Problem (38) to be well posed is given by the so-called inf-sup or LBB condition, which is equivalent to saying that $\underline{\underline{B}}[\mathcal{A}_{\lambda, \text{active}}, \mathcal{A}_c] \underline{\underline{V}}[\mathcal{A}_c, :]$ must have full row rank, where $\mathcal{A}_{\lambda, \text{active}}$ denotes the DOFs of \mathcal{A}_λ associated to an active contact ($\underline{\underline{B}}[\mathcal{A}_{\lambda, \text{active}}, \mathcal{A}_c] \underline{\underline{V}}[\mathcal{A}_c, :] \underline{\underline{\gamma}} = \underline{\underline{D}}[\mathcal{A}_{\lambda, \text{active}}]$). As the active contacts are a priori unknowns, we extend that condition to a stronger one, which is $\underline{\underline{B}}[\mathcal{A}_\lambda, \mathcal{A}_c] \underline{\underline{V}}[\mathcal{A}_c, :]$ having full row rank. A necessary condition for verifying the discrete extended LBB is $l_u \geq N_\lambda^C$. If the discrete LBB condition is verified and because the term $\underline{\underline{V}}[\mathcal{A}, :]^T \underline{\underline{K}}[\mathcal{A}, \mathcal{A} \cup \mathcal{I}] \underline{\underline{V}}[\mathcal{A} \cup \mathcal{I}, :]$ is well defined by the classical HR (see Section 4.3), then Problem (38) admits a unique solution. Moreover, Problem (38) is founded on the FE formulation (16) by restraining matrices and vectors to the RID, projecting the balanced equation on $\underline{\underline{V}}[\mathcal{A}, :]$ and then replacing $\underline{\underline{U}}[\mathcal{A} \cup \mathcal{I}]$ by $\underline{\underline{V}}[\mathcal{A} \cup \mathcal{I}, :] \underline{\underline{\gamma}}$. \square

If the extended LBB condition ($\underline{\underline{B}}[\mathcal{A}_\lambda, \mathcal{A}_c] \underline{\underline{V}}[\mathcal{A}_c, :]$ has full row rank) is not verified, we propose to adopt a hybrid full-order/reduced-order model strategy.^{24,26,27} This approach follows the usual principles of the HR method by using a mixed primal RB. The mixed primal RB contains the usual POD modes and few FE shape functions. The selected FE shape functions have to be related to the nodes on Γ_A^C in order to increase the rank of $\underline{\underline{B}}[\mathcal{A}_\lambda, \mathcal{A}_c] \underline{\underline{V}}[\mathcal{A}_c, :]$. We denote by $(i_\alpha^c)_{\alpha=1}^{\text{card}(\mathcal{A}_c)}$ the indices in \mathcal{A}_c . Then, the mixed primal RB reads:

$$\underline{\underline{\psi}}_k = \begin{cases} \underline{\underline{\psi}}_k, & \text{if } k \leq l_u \\ \underline{\underline{\varphi}}_{k-l_u}^c, & \text{if } l_u < k \leq \bar{l}_u \end{cases}, \quad k = 1, \dots, \bar{l}_u \quad (39)$$

with $\text{card}(\mathcal{A}) \geq \bar{l}_u \geq \max(l_u, N_\lambda^C)$ defined such that $\underline{\underline{B}}[\mathcal{A}_\lambda, \mathcal{A}_c] \underline{\underline{V}}[\mathcal{A}_c, :]$ has full row rank, the matrix $\underline{\underline{V}}[\mathcal{A}_c, :]$ being defined as the discretization of $\underline{\underline{\psi}}_k$:

$$\underline{\underline{V}} = \left(\underline{\underline{V}} \mid \underline{\underline{I}}_N \left[:, \left(i_{k-l_u}^c \right)_{k=l_u+1}^{\bar{l}_u} \right] \right). \quad (40)$$

Problem (38) is then defined with $\underline{\underline{V}}$ playing the role of $\underline{\underline{V}}$.

When $D > 1$ and the number of nodes on Γ_A^C is small enough, it may be interesting to treat every primal DOF on the potential contact zone with FE shape functions in order to simply ensure the full row rank condition, good stability, and accurate solutions on the contact zone.

5.3 | Reconstruction of the Lagrange multipliers from the hyper-reduced solution

Due to the choice of the lifting of $\underline{\underline{u}}_I$ (28), the primal solution is defined everywhere by $\underline{\underline{U}} = \underline{\underline{V}} \underline{\underline{\gamma}}$. However, a drawback of the proposed HR approach is that the nonpenetration condition (6) may not be verified outside the RID.

The proposed reduced model (38) gives contact forces inside the RID only. The reconstruction of the contact forces over Ω is possible with a postprocessing but brings us back to the problem of finding a reliable dual RB. We propose here a postprocessing that consists in extracting from the dual snapshots N_s linearly independent snapshots such that $N_s \leq N_\lambda^C + N_\lambda$, and we denote it $\underline{\underline{S}}_\lambda$. Our strategy is to find the coefficients $\underline{\underline{\gamma}}_\lambda$, with positivity constraints to ensure the nonadhesion condition, associate to $\underline{\underline{S}}_\lambda$ minimizing the complementary slackness condition as well as the distance with the solution obtained by solving the HR model. The proposed reconstruction is obtained by solving the NNLS problem (41), where $\underline{\underline{U}}$ and $\underline{\underline{\Delta}}[\mathcal{A}_\lambda]$ are the primal and the dual solution obtained by solving the HR model (38).

$$\min_{\underline{\underline{\gamma}}_\lambda} \left\| \begin{pmatrix} \underline{\underline{S}}_\lambda[\mathcal{A}_\lambda; :] \\ [\underline{\underline{B}} \underline{\underline{D}} - \underline{\underline{D}}]_- \odot \underline{\underline{S}}_\lambda \end{pmatrix} \underline{\underline{\gamma}}_\lambda - \begin{pmatrix} \underline{\underline{\Delta}}[\mathcal{A}_\lambda] \\ \underline{\underline{0}} \end{pmatrix} \right\|_2^2 \quad \text{subject to } \underline{\underline{\gamma}}_\lambda \geq 0 \quad (41)$$

The Lagrange multipliers are then defined outside the RID due to the relation $\underline{\underline{\Delta}}[\{1, \dots, N_\lambda\} \setminus \mathcal{A}_\lambda, :] = \underline{\underline{S}}_\lambda[\{1, \dots, N_\lambda\} \setminus \mathcal{A}_\lambda, :] \underline{\underline{\gamma}}_\lambda$.

5.4 | Greedy algorithm coupled with an error indicator

The reliability of an RB strictly depends on the snapshots (see Section 4.1 for an RB made of POD modes). In order to accurately choose the snapshots, two possibilities are at least possible. The first possibility is to take the snapshots in

the parameter space randomly or uniformly distributed. The second possibility is to apply a greedy algorithm,³⁵ which is going to select iteratively the snapshots maximizing an error indicator. The greedy algorithm is usually used as it enables choosing pertinent snapshots and obtaining a minimal robust RB. However, this algorithm may be slow (but in the offline part) and requires an error indicator. This error indicator will be used to have a stopping criterion.

In the work of Balajewicz et al,⁷ the error indicator used is based on the nonsatisfaction of the contact conditions. However, for the hybrid HR model (38) proposed here, the contact conditions are respected by construction on the RID. We hence propose a simple but efficient a posteriori error indicator for HR. We first interpolate at $\underline{\mu} \in \mathcal{S}$ the primal snapshots collection $\mathcal{S}_{\text{snap}} \subset \mathcal{S}$ restricted to the DOFs \mathcal{I} located on Γ^I (see Equation (21)):

$$\tilde{U}(\underline{\mu})[\mathcal{I}] = \sum_{\rho \in \mathcal{S}_{\text{snap}}} \alpha_{\rho}(\underline{\mu}) \underline{U}(\rho)[\mathcal{I}] \quad (42)$$

with $\alpha_{\rho}(\underline{\mu})$ as the interpolation coefficients. The interpolated solution is imposed as a Dirichlet boundary condition on the RID interface Γ^I . We can then solve the FOM problem defined on Ω_A , in order to obtain an approximation of the displacements and contact forces reference solutions inside the RID. The primal and dual approximations are respectively denoted $\tilde{U}[\mathcal{A}]$ and $\tilde{\Lambda}[\check{\mathcal{A}}_{\lambda}]$ with $\check{\mathcal{A}}_{\lambda} \subseteq \mathcal{A}_{\lambda}$ since the problem is directly defined on Ω_A . The reduced FE approximation tends to the FOM approximation as the interpolation error tends to zero. Our error indicator is the relative discrepancy between the HR solution $(\underline{U}_{\text{hr}}, \underline{\Lambda}_{\text{hr}})$ and the reduced FE approximation in the RID, where $(H^1)'$ denotes the dual of H^1 .

$$\eta(\underline{\mu}) = \eta_u(\underline{\mu}) + \eta_{\lambda}(\underline{\mu}) \quad \text{with} \quad \eta_u(\underline{\mu}) = \frac{\|(\underline{U}_{\text{hr}} - \tilde{U})(\underline{\mu})[\mathcal{A}]\|_{H^1}}{\|\tilde{U}(\underline{\mu})[\mathcal{A}]\|_{H^1}}, \quad \eta_{\lambda}(\underline{\mu}) = \frac{\|(\underline{\Lambda}_{\text{hr}} - \tilde{\Lambda})(\underline{\mu})[\check{\mathcal{A}}_{\lambda}]\|_{(H^1)'}}{\|\tilde{\Lambda}(\underline{\mu})[\check{\mathcal{A}}_{\lambda}]\|_{(H^1)'}} \quad (43)$$

In the following applications, we use a bilinear interpolation in the parameter space, but it could be interesting to use an interpolation on the manifold as proposed in the work of Amsallem et al.³⁶

6 | ACADEMIC TEST CASE

6.1 | Model

The application we will focus on comes from the work of Haasdonk et al⁶ and has been used in the work of Balajewicz et al.^{7,17} It is a 1D static obstacle problem. We consider a 1D elastic rope described by problem (9) with $\underline{\sigma} = \nu \text{grad } u$ in $\Omega = [0, 1]$. Homogeneous Dirichlet boundary conditions are applied on $\Gamma^D = \{0, 1\}$. Moreover, $\Gamma^N = \emptyset$ and $\Gamma^C =]0, 1[$. The function f is nonparametric and corresponds to a gravity, $f(x) = -1$. The problem depends on a parameter vector $\underline{\mu}$ living in the domain $\mathcal{S} = [0.05, 0.25] \times [-0.05, 0.5]$. The elasticity function ν depends on the first parameter μ_1 as

$$\nu(x, \mu_1) = \mu_1 \mathbb{1}_{[0,0.5]}(x) + 0.15 \times \mathbb{1}_{]0.5,1]}(x) \quad \text{with} \quad \mathbb{1}_{\Xi}(x) = \begin{cases} 1, & \text{if } x \in \Xi \\ 0, & \text{elsewhere.} \end{cases}$$

The second parameter μ_2 controls the obstacle function involved in the gap, ie,

$$d(x, \mu_2) = 0.2(\sin(\pi x) - \sin(3\pi x)) + 0.5 - \mu_2(x - 0.5), \quad \forall x \in \Omega.$$

We represent in Figure 3 the primal and dual solutions of the FOM obtained using linear FE for different parameters $\underline{\mu}$ in \mathcal{S} . The choice of the parameters μ_1 and μ_2 then influences the active contact zone and the magnitude of the contact forces. The domain Ω is discretized in 100 nodes uniformly distributed.

6.2 | Active-set algorithm

All contact saddle-point problems (eg, (16), (38), and (43)) are solved here with an active-set strategy. This strategy of resolution consists in explicitly verifying the complementary slackness condition by imposing contact condition $((\underline{B} \underline{U})_i = (\underline{D})_i)$ where Lagrange multipliers are active. The set of active Lagrange multipliers is updated due to the violation of the Signorini laws. Lagrange multipliers are respectively activated and deactivated where the nonpenetration $((\underline{B} \underline{U})_i \leq (\underline{D})_i)$ and the nonadhesion $((\underline{\Lambda})_i \geq 0)$ conditions are violated.

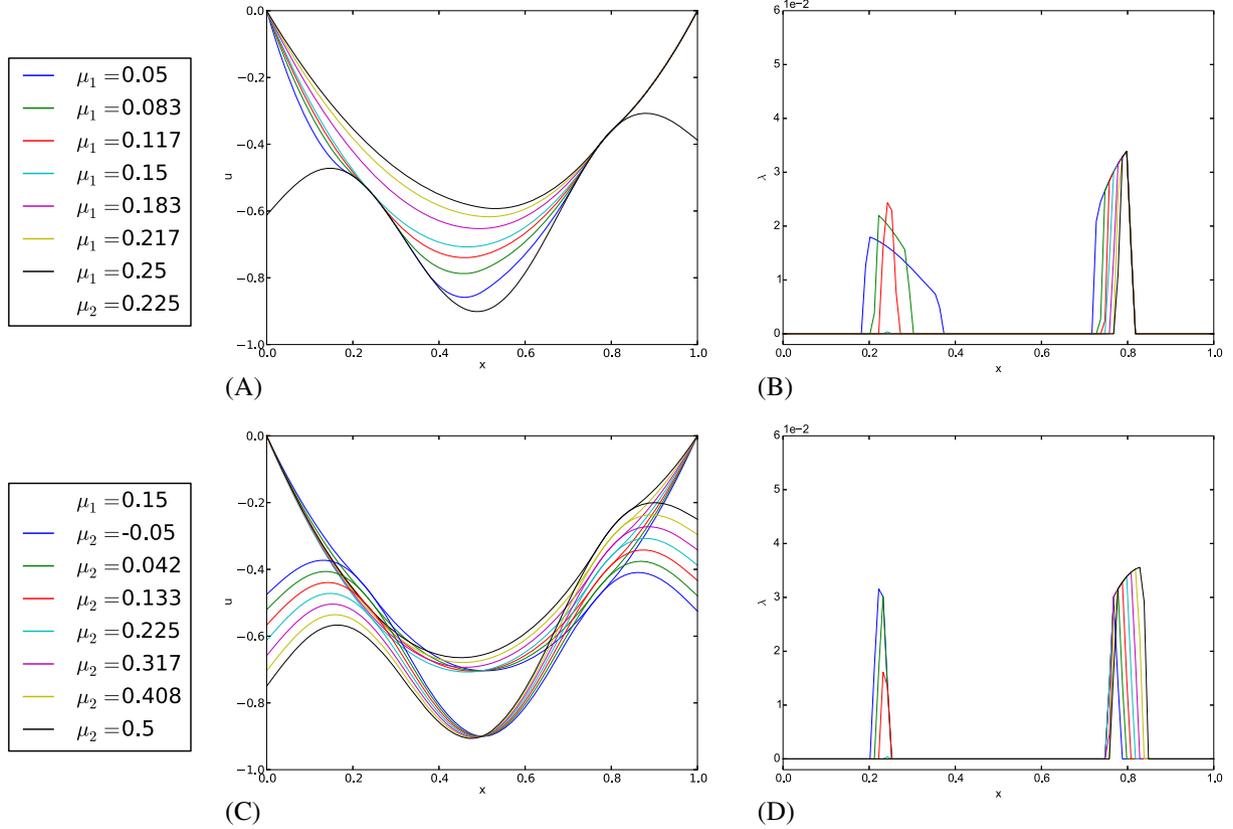


FIGURE 3 Full-order model solutions, where (A) and (B) are respectively the primal and dual solutions with variations of the parameter μ_1 in $[0.05, 0.25]$ and (C) and (D) are respectively the primal and dual solutions with variations of μ_2 in $[-0.05, 0.5]$

6.3 | Projection-Based method

We briefly recall the PB method introduced in the work of Balajewicz et al⁷ as we are going to compare the proposed hybrid HR method for contact to this method. No comparison between the hybrid HR and the RB method introduced in the work of Haasdonk et al⁶ is done here as this latter method does not use data compression to build the RB (like POD) and, hence, can lead to a large RB when using a lot of snapshots, as explained in the work of Balajewicz et al.⁷ The PB method, introduced for contact problems, is based on the POD to construct the primal RB $\underline{\underline{V}}_u$ and on the NNMF algorithm to build a dual RB $\underline{\underline{V}}_\lambda$, which preserves the positivity of Lagrange multipliers after projection. The integers l_u and l_λ denote respectively the number of vectors in $\underline{\underline{V}}_u$ and $\underline{\underline{V}}_\lambda$. The Galerkin projection of the FOM (16) on these two bases gives

$$\left\{ \begin{array}{l} \text{Find } (\underline{\underline{\gamma}}_u, \underline{\underline{\gamma}}_\lambda) \in \mathbb{R}^{l_u} \times (\mathbb{R}^+)^{l_\lambda} \text{ such that} \\ \underline{\underline{V}}_u^T \underline{\underline{K}} \underline{\underline{V}}_u \underline{\underline{\gamma}}_u + \underline{\underline{V}}_u^T \underline{\underline{B}}^T \underline{\underline{V}}_\lambda \underline{\underline{\gamma}}_\lambda = \underline{\underline{V}}_u^T \underline{\underline{F}} \\ \underline{\underline{\gamma}}_\lambda \odot \left(\underline{\underline{V}}_\lambda^T \underline{\underline{B}} \underline{\underline{V}}_u \underline{\underline{\gamma}}_u - \underline{\underline{V}}_\lambda^T \underline{\underline{D}} \right) = \underline{\underline{0}} \\ \underline{\underline{V}}_\lambda^T \underline{\underline{B}} \underline{\underline{V}}_u \underline{\underline{\gamma}}_u \leq \underline{\underline{V}}_\lambda^T \underline{\underline{D}}. \end{array} \right. \quad (44)$$

The displacement field and Lagrange multipliers are finally obtained by $\underline{\underline{U}} = \underline{\underline{V}}_u \underline{\underline{\gamma}}_u$ and $\underline{\underline{\Lambda}} = \underline{\underline{V}}_\lambda \underline{\underline{\gamma}}_\lambda$. The NNMF takes as input the number of vectors in the RB. In order to have the largest number of dual DOFs while respecting the extended necessary condition ($l_u \geq l_\lambda$), we impose $l_\lambda = l_u$ for this method.

Using these RB approximations, only the nonadhesion condition on $\underline{\underline{\Lambda}}$ is ensured by the positive dual RB $\underline{\underline{V}}_\lambda$ built with NNMF. The nonpenetration and complementary slackness conditions are not verified for $\underline{\underline{U}}$ and $\underline{\underline{\Lambda}}$. Indeed, one can prove that $\underline{\underline{\Lambda}}^T (\underline{\underline{B}} \underline{\underline{U}} - \underline{\underline{D}}) = \underline{\underline{0}}$ but not componentwise ($\underline{\underline{\Lambda}} \odot (\underline{\underline{B}} \underline{\underline{U}} - \underline{\underline{D}}) = \underline{\underline{0}}$) since $\underline{\underline{B}} \underline{\underline{U}} - \underline{\underline{B}}$ is not necessarily negative or zero.

Unlike the POD and its unique factorization, there is no unique NNMF factorization. Moreover, NNMF algorithms can converge to different local minima (and even this convergence to local minima is not guaranteed), which is why initialization of the algorithm becomes critical.³⁷

The method used here for the computation of the NNMF dual RB is an alternating NNLS matrix factorization using the projected gradient (bound-constrained optimization) method for each subproblem. It can be found in the Python library introduced in the work of Zitnik and Zupan.³⁸ This method takes as input several parameters as the initialization, the rank approximation, the maximum number of factorization iterations, the number of runs of the algorithm, etc. We opt for an NNDSVD initialization.

6.4 | On the difficulty of finding a reliable dual RB

In order to illustrate the difficulty of finding a reliable dual RB, we consider snapshots taken on a coarse grid of 10×12 points uniformly distributed in S . The difficulty of the dual FOM solutions to be approximated by an RB can be understood looking at Figure 4, where we draw the relative error in the Frobenius norm between the snapshots and its k -rank approximation by an RB. For the POD RBs built with SVD, this is equivalent to $1 - \mathcal{E}(k)$ as defined in Equation (18). Compared to a primal POD RB, the dual POD RB needs much more vectors to reach the same error threshold. The POD giving an optimal k -rank basis, any other algorithm would give a less decreasing curve as we can see for the dual NNMF RB. For this latter algorithm, we study the influence of the maximum number of factorization iterations. Two maxima have been tested: 10^2 and 10^4 iterations as reported in Figure 4. For low k , whatever the maximum number of factorization iterations, the approximation errors done by the dual NNMF RB are really close to the one done by the dual POD RB. Differences appear for large k , where the level of stagnation error varies with the maximum number of factorization iterations. As low k corresponds to the domain of interest for model order reduction and because the CPU time for the construction of the dual NNMF RB grows with the number of iterations, we set the maximum number of factorization iterations to 10^2 in the sequel of the article.

6.5 | Hyper-Reduction method

In order to illustrate the construction of the hybrid HR model for the rope test case, we first consider snapshots taken on the coarse grid of 10×12 points as introduced in Section 6.4. The verification of the hypothesis of Property 1 implies that $\bar{l}_u = N_\lambda^C$ in the contact problem with the obstacle under study ($\underline{\underline{B}} = \underline{\underline{I}}_{\text{card}(\mathcal{A})}$). The chosen full-order DOFs have to ensure that $\underline{\underline{V}}$ has full row rank.

6.5.1 | Primal RB and RID construction

The primal RB is built with the POD. We draw in Figure 5 some of the $l_u = 9$ first POD modes selected for a POD threshold $\epsilon_{\text{tol}} = 10^{-7}$. In practice, the higher the mode number is, the higher gradients there are in the mode.

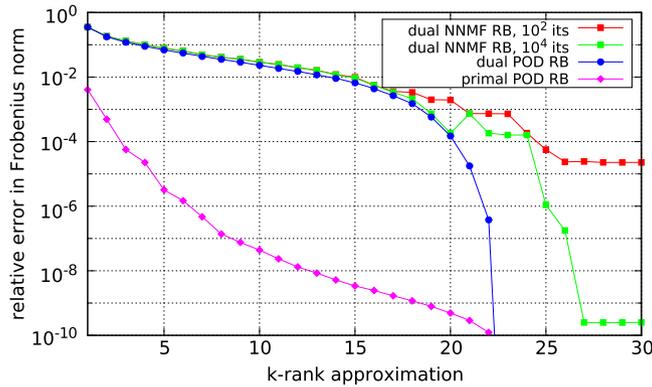


FIGURE 4 Relative error between snapshots and best approximation using the reduced basis (RB). NNMF, nonnegative matrix factorization; POD, proper orthogonal decomposition

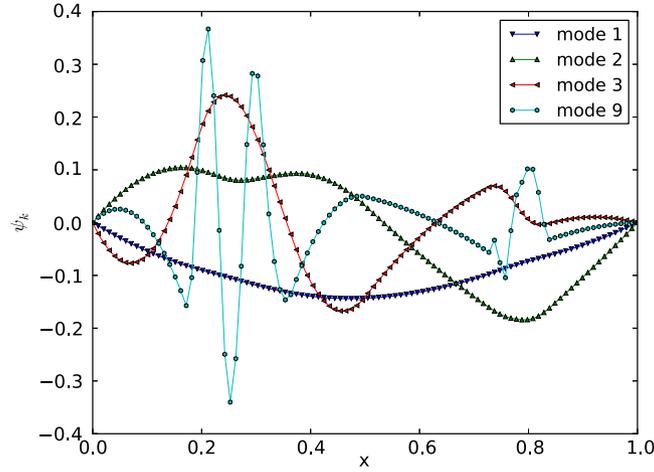


FIGURE 5 Proper orthogonal decomposition modes contained in the primal reduced basis built by using 10×12 snapshots [Colour figure can be viewed at wileyonlinelibrary.com]

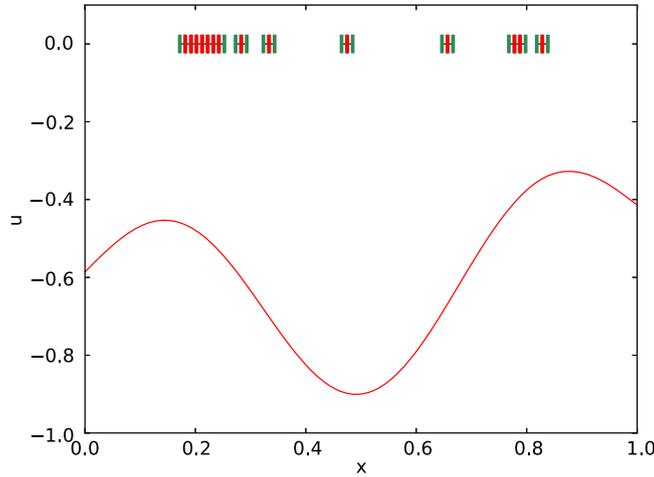


FIGURE 6 Top: reduced integration domain built by using 10×12 snapshots, where red ticks represent set \mathcal{A} , whereas green ticks represent set \mathcal{I} . Bottom: example of the obstacle for $\mu_2 = 0.17$ [Colour figure can be viewed at wileyonlinelibrary.com]

We detail here the procedure used to construct the RID. We first apply the DEIM algorithm on the primal RB. We next apply the DEIM on the POD basis obtained for the Lagrange multipliers to make sure that the most important contact zones are included in the RID. In this example, we take a dual POD threshold $\epsilon_{\text{tol}} = 10^{-1}$ giving $l_\lambda = 1$ vector. The RID is then the union of the two domains obtained with the DEIM. The RID does not need to be extended. Because it is an obstacle problem, the gap can be evaluated in all points of the RID. We represent at the top of Figure 6 the RID obtained and at the bottom an example of the obstacle ($\mu_2 = 0.17$). The reduced mesh is located where the obstacle is at the higher position; hence, the RID seems to include the most part of the contact zone. In this case, the RID contains 28 nodes in Ω_A including 14 in Γ_A^C (the red dots), whereas the initial domain counts 100 nodes. To respect the extended necessary condition, 5 primal DOFs are treated as full-order DOFs.

6.5.2 | Results obtained using the hybrid HR model

On a fine grid of 100×120 points uniformly distributed in S , we compare the solutions obtained with the reference FOM and with the HR model. We represent in Figure 7 the 2-norm relative errors over the parameter space for both displacements and contact forces. The mean and maximum relative errors for the displacements over Ω respectively read 0.34% and 1.13% (Figure 7A), whereas the mean and maximum relative errors for the contact forces over Γ_A^C respectively read 24% and 148% (Figure 7B). We can conclude that the regular snapshots grid has not caught the maximum relative error for the contact forces.

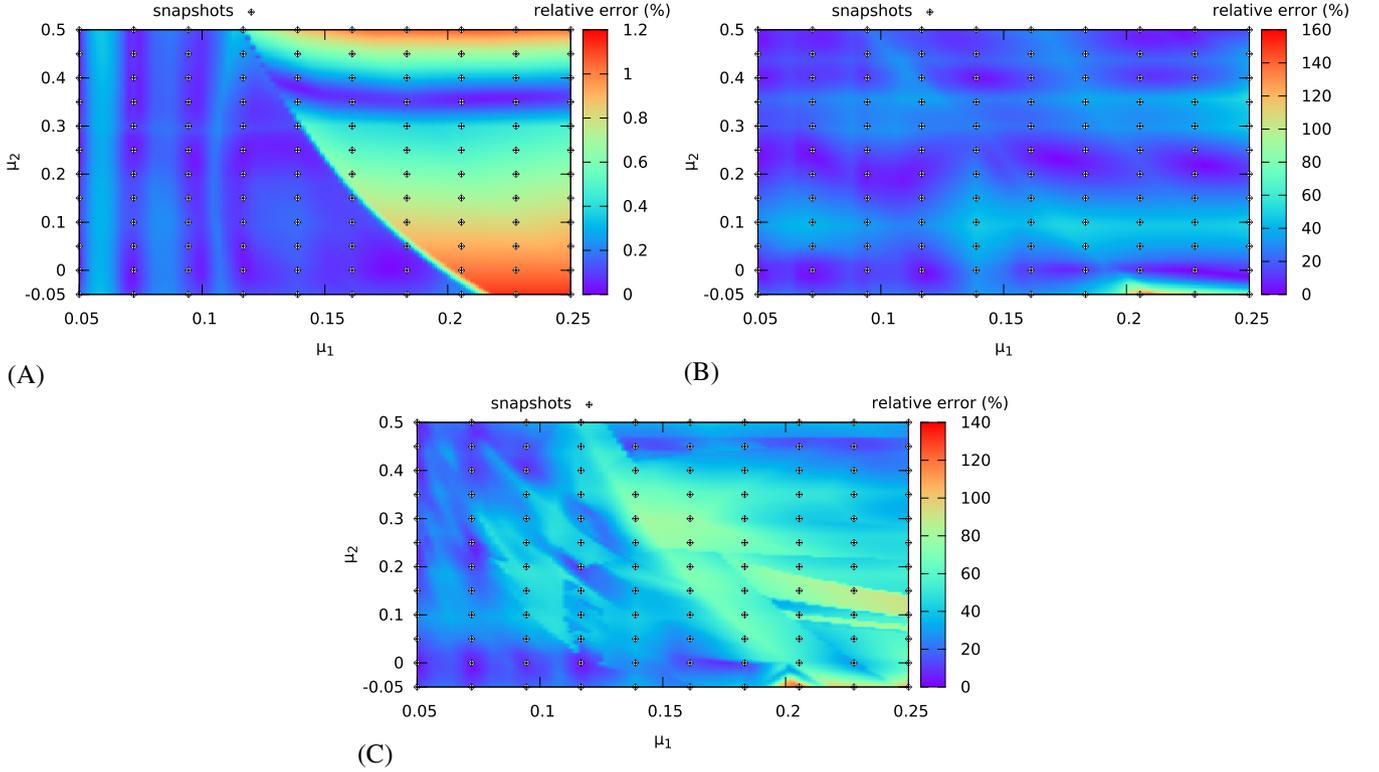


FIGURE 7 Errors of the hybrid hyper-reduction model compared to the full-order model generated by using 10×12 snapshots. The snapshots are represented with the symbol \diamond . A, Primal error on Ω with the 2-norm; B, Dual error on Γ_A^C with the 2-norm; C, Dual error on Ω with the 2-norm (after reconstruction)

Using the reconstruction strategy introduced in Section 5.3, the average and maximum relative errors for the contact forces over Ω respectively read 39% and 127% (Figure 7C). The mean relative error over Ω is higher than before reconstruction but is still satisfactory. The proposed reconstruction seems efficient and reliable.

6.6 | Comparison between the HR and PB methods

6.6.1 | Regular grid

We first compare the hybrid HR model introduced in the previous section with two equivalent models made due to the PB method. The primal RB, built with the POD for the two methods, is made of $l_u = 9$ ($\epsilon_{\text{tol}} = 10^{-7}$) vectors for the HR. Due to the inf-sup condition to be respected, the hybrid HR model may have a hybrid primal basis with 5 more vectors. The two PB models are respectively built with $l_u = 9$ ($\epsilon_{\text{tol}} = 10^{-7}$) and $l_u = 14$ ($\epsilon_{\text{tol}} = 6.0 \times 10^{-9}$) vectors in the primal RB. The dual RB is constructed by NNMF with the same number of vectors as the primal RB while the FOM dual basis is maintained in the HR.

We respectively draw in Figures 8 and 9 the HR and PB solutions to be compared with the FOM solutions for $\underline{\mu} = (0.07, 0.17)^T$ not taken in the sample of snapshots. Details on the calculus are presented in Table 1. As we can see in Figures 8A and 9A, the primal solutions for both methods are truly accurate. Concerning the dual solutions obtained with the PB method and represented in Figure 9B, the Lagrange multipliers are non-null at similar locations, but magnitudes are significantly different compared to the FOM dual solution. In comparison, the dual HR model solution drawn in Figure 8B is better due to the nonprojection on a dual compressed RB. The speed-up values obtained (see Table 1) are rather close to each other if the posttreatment to reconstruct the HR dual solution everywhere using the snapshots basis is not taken into account. These speed-up values are strongly dependent on the number of iterations of the active-set method. We can see in Table 1 that the error on contact forces is of 13% compared to 23% for the PB method with the same number of unknowns. This confirms that the proposed hybrid HR model is really efficient.

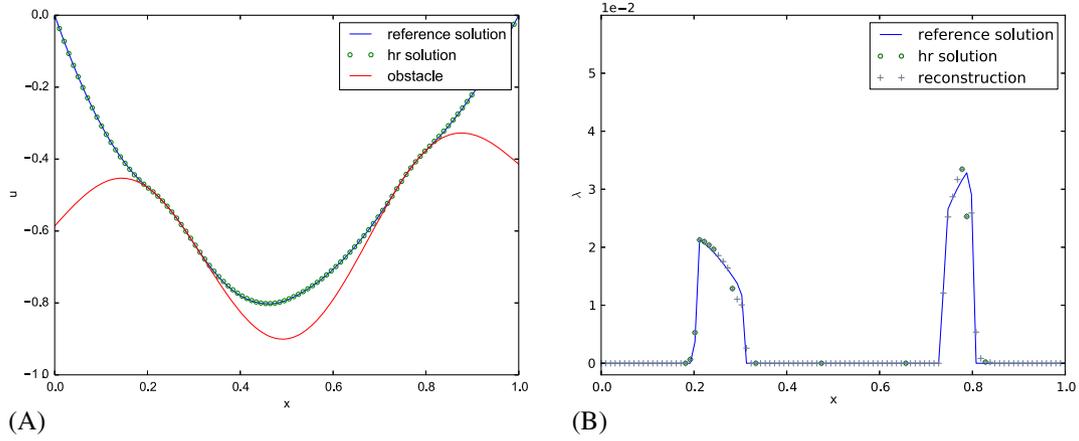


FIGURE 8 Hyper-reduction (HR) solution compared to the full-order model solution for $(\mu_1, \mu_2) = (0.07, 0.17)$. A, Primal solution; B, Dual solution

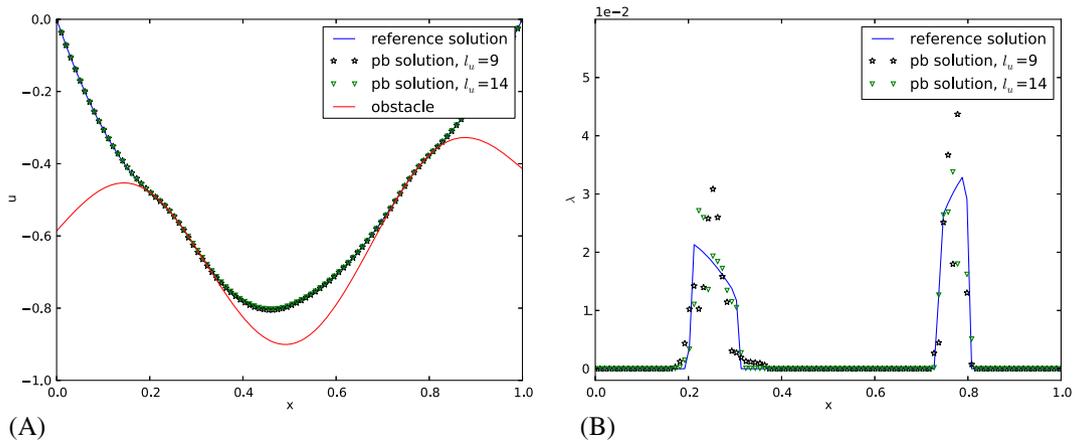


FIGURE 9 Two projection-based (PB) solutions built with $l_u = 9$ and $l_u = 14$ compared to the full-order model solution for $(\mu_1, \mu_2) = (0.07, 0.17)$. A, Primal solution; B, Dual solution

TABLE 1 Comparison of the methods for $(\mu_1, \mu_2) = (0.07, 0.17)$. In the hyper-reduction (HR) column, we specify the results without/with the contact forces reconstruction

	HR ($l_u = 9$)	PB ($l_u = 9$)	PB ($l_u = 14$)
Total number of unknowns	28	18	28
Speed-up	8.5/5.9	8.4	9.1
Active-set iterations	4	6	5
Displacement 2-norm relative error (%)	0.15	0.32	0.04
Contact force 2-norm relative error (%)	13/12	44	23

Abbreviation: PB, projection-based.

6.6.2 | Greedy algorithm

In order to compare both approaches using a greedy algorithm, we will compare the results obtained by choosing the snapshots due either to the error indicator introduced in this paper (see Equation (43)) and dedicated to the HR method or to the error indicator introduced as $I(\gamma, \alpha_1, \alpha_2, \alpha_3)$ in the work of Balajewicz et al⁷ based on the nonverification of the Signorini conditions and dedicated to the PB method. We will use in this comparison only the error indicator based on the nonpenetration $I(\gamma, 1, 0, 0)$ as it is the cheapest and gives the most accurate approximation following the work of Balajewicz et al.⁷ It reads:

$$I(\underline{\mu}) = \left\| \left[\underline{B}(\underline{\mu}) \underline{V} \underline{\gamma}(\underline{\mu}) - \underline{D}(\underline{\mu}) \right]_+ \right\|_2. \quad (45)$$

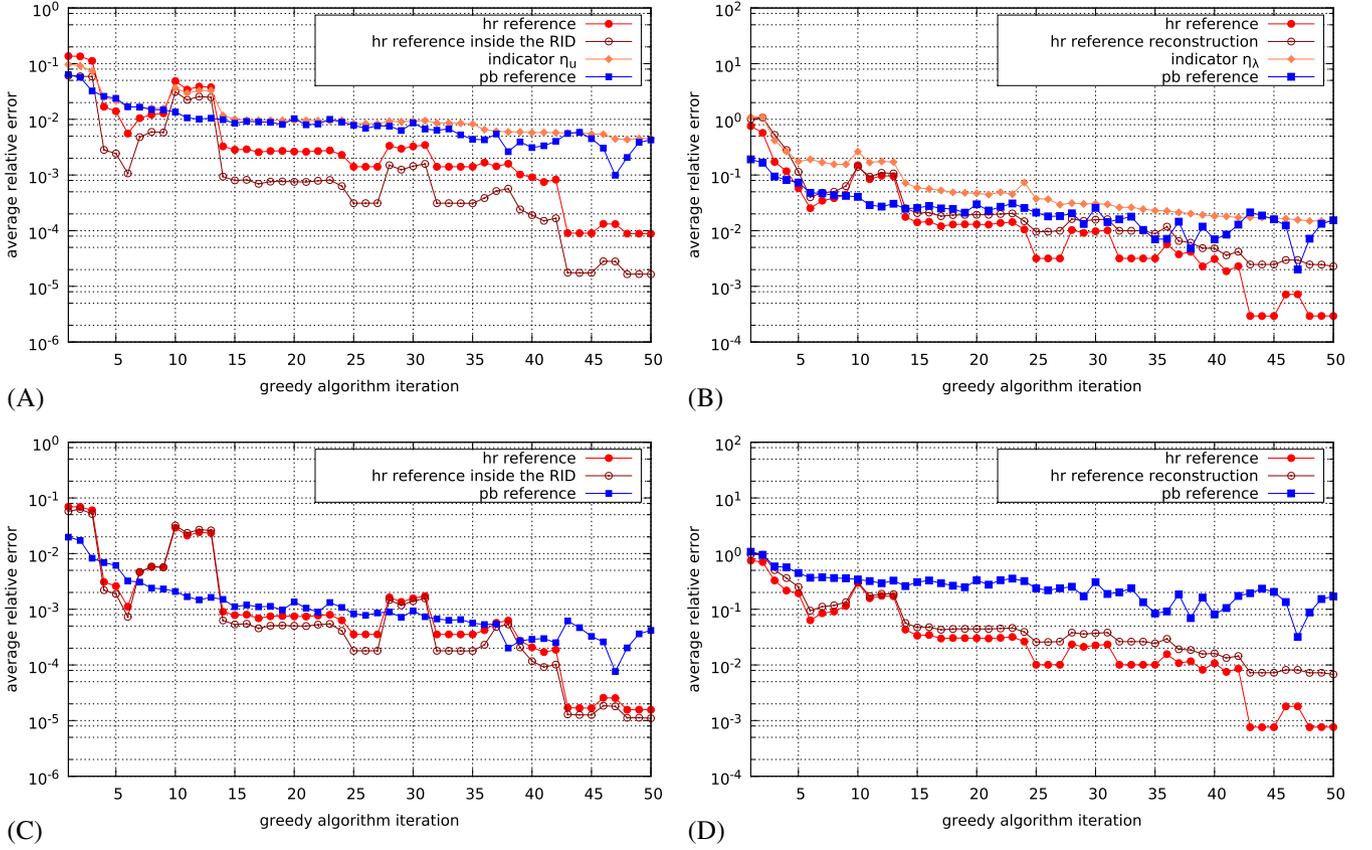


FIGURE 10 Decrease of the average relative errors during the greedy algorithm based on the error indicator η applied to the hyper-reduction (HR) method. A, Primal error with the H^1 norm; B, Dual error with the $(H^1)'$ norm; C, Primal error with the 2-norm; D, Dual error with the 2-norm. PB, projection-based; RID, reduced integration domain [Colour figure can be viewed at wileyonlinelibrary.com]

The primal POD threshold is set at $\varepsilon_{\text{tol}} = 10^{-10}$ for both methods. The dual POD threshold for the HR-RID construction is set to 5.10^{-2} , whereas the number of vectors of the dual RB for the PB method is set to the number of vectors of the primal RB. POD thresholds are taken smaller than for the regular grid in order to have a good RID from the beginning of the algorithm. During the greedy algorithm iteration, the RID can only be enlarged: the RID used in the HR model is then the union of the previous iteration RID and the current iteration RID. The greedy algorithm is stopped at the 50th iteration.

We draw in Figure 10 the evolution of the average relative errors for both methods using the snapshots obtained due to the error indicator η (Equation (43)) applied to the HR method. In this Figure, the reference error means the error between the ROM and the FOM. The HR reference error corresponds to the relative error on Ω for the primal solution and on Γ_A^C for the dual solution, whereas the HR reference reconstruction error corresponds to the relative error on Ω for the dual solution. The error indicators are overestimating the reference errors as the HR model converges faster to the FOM than the FE approximation based on the boundary condition interpolation. This convergence is directly related to the POD threshold, which is rather small here. In the H^1 norm, the primal HR model error is clearly smaller than the primal PB model error as we can see in Figure 10A, whereas it is slightly smaller for the 2-norm (see Figure 10C). This means that the HR model better approximates the gradient of the displacement (strain) than the PB model. In accordance with Table 1, the fact that the HR model 2-norm error is slightly smaller is probably due to a larger number of primal DOFs ($\bar{l}_u \geq l_u$). The dual $(H^1)'$ norm HR model error is equivalent to the dual PB model error (except for the high number of iterations) as we can see in Figure 10B. Because the $(H^1)'$ norm is difficult to apprehend, we also plot the dual relative error obtained with the 2-norm in Figure 10D. With this norm, the dual HR model error is clearly better than the dual PB model error even after reconstruction on Γ^C . We recover here the quantitative first conclusions drawn from Figures 8 and 9.

We draw in Figure 11 the evolution of the average relative error for both methods using snapshots obtained with the error indicator $I(\mu)$ applied to the PB method. The HR reference average relative errors obtained are decreasing faster than in Figure 10 but are more unstable. Once again, the HR model errors are much smaller than the PB model errors,

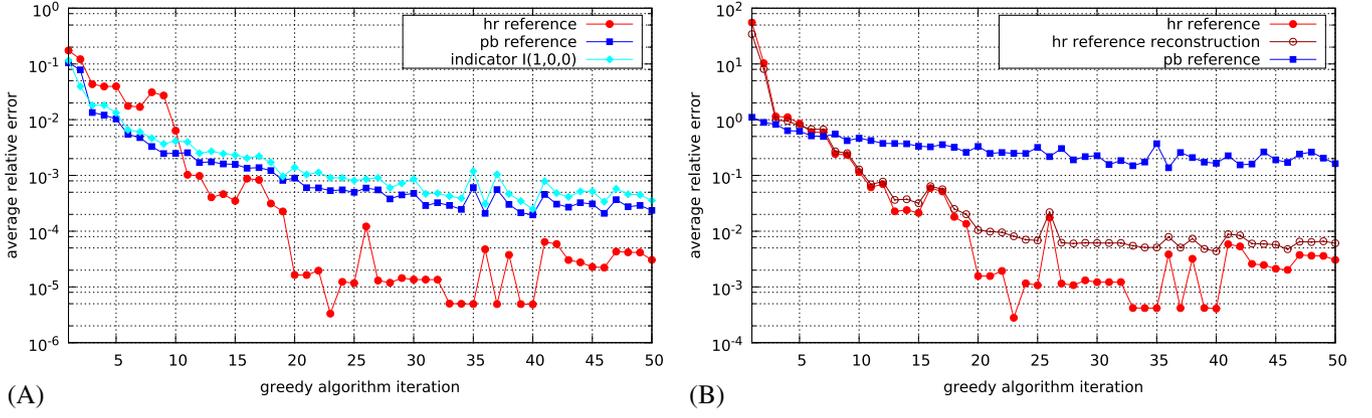


FIGURE 11 Decrease of the average relative errors during the greedy algorithm based on the error indicator $I(\underline{\mu})$ applied to the projection-based (PB) method. A, Primal error with the 2-norm; B, Dual error with the 2-norm. HR, hyper-reduction [Colour figure can be viewed at wileyonlinelibrary.com]

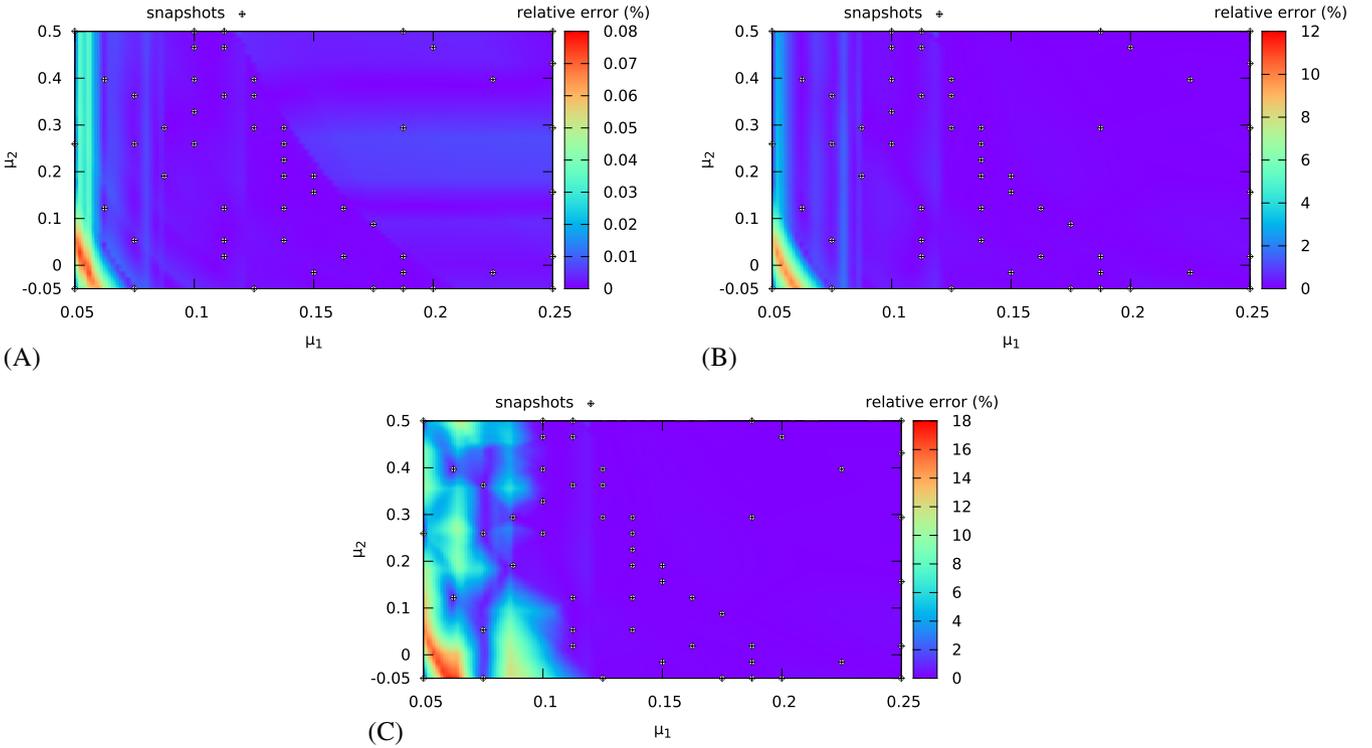


FIGURE 12 Errors of the hybrid hyper-reduction model compared to the full-order model generated by using 10×12 snapshots. The snapshots are represented with the symbol \diamond . A, Primal error on Ω with the 2-norm; B, Dual error on Γ_A^C with the 2-norm; C, Dual error on Ω with the 2-norm (after reconstruction)

especially for contact forces. However, we can see that the error indicator $I(\underline{\mu})$ estimates very well the reference PB error as already shown in the work of Balajewicz et al.⁷

At the end of the greedy algorithm, we report in Figure 12 the 2-norm reference relative errors obtained with the HR model built with its associated error indicator. A fine grid of 100×120 points uniformly distributed in S is used. The mean and maximum relative errors for the displacements over Ω respectively read 0.004% and 0.077% (Figure 12A), whereas the mean and maximum relative errors for the contact forces over Γ_A^C respectively read 0.34% and 10% (Figure 12B). We can see that the error is highly smoother than with a regular grid (see Figure 7). It demonstrates the effectiveness of the proposed error indicator η used in the greedy algorithm. The highest error for both displacements and contact forces is localized where μ_1 is small, eg, where the active contact zone is the most important. Reconstructing the contact forces on

Ω gives average and maximum relative errors that respectively read 1.3% and 18% over Ω (Figure 12C). The error is then higher than before reconstruction and spread to all small μ_1 except on snapshots.

The results obtained with the hybrid HR model give confidence in the proposed method. The hybrid HR model error is generally equivalent to the PB model error or even better depending on the norm, especially to approximate contact forces.

7 | APPLICATION TO A 3D PROBLEM

This 3D example aims to show the supply of the proposed hybrid HR model for 3D contact problems. We will especially focus on the RID construction that predicts to have benefit for nonlinear behaviors. We assume a node-to-node contact.

7.1 | Elastic model

The test case, derived from nuclear fuel simulations, is a 3D extension of the one we can find in the work of Liu et al.³⁹ We consider a cylindrical pellet (13.5-mm height), in a cladding (4.1-mm internal radius, 4.7-mm external radius). To account for symmetries, only one eighth of the mechanical system is simulated, as shown in Figure 13. An axial pressure p_A (5 Mpa) is imposed on the top of the pellet, and 2 radial pressures p_B (50 MPa for $z > 2$ mm) and p_C (250 MPa for $z < 2$ mm) are imposed on the cladding. Three radii of the pellet are considered to investigate the effect of the initial gap on the stresses: radius 4.1 mm (no initial gap), 4.02 mm, and 4.06 mm. The related FE simulations are denoted S1, S2, and S3. The pellet and the cladding have the same elastic properties: 10^5 MPa for Young's modulus and 0.3 for the Poisson coefficient. The loading conditions are applied with 25 time steps according to a linear ramp.

7.2 | Empirical modes and RID

Empirical modes for displacements, stresses, and generalized forces are generated by using the numerical results of simulations S1 and S2 at every time step. Generalized forces denote the Lagrange multipliers and the external forces. Then, the POD method gives 17 displacement modes ($\epsilon_{\text{tol}} = 10^{-5}$). The first 6 displacement modes and the last mode are shown in Figure 14, over the cladding only. As we already saw, the higher the mode number is, the higher there are large gradients in the mode. Similar results are obtained for empirical modes related to stresses and generalized forces. We obtain 20 modes for stresses and 21 for the generalized forces with the POD used to build the RID only.

In Figure 13, we show various contributions to the RID. Ω_u is the domain covered by elements connected to the 17 interpolation indices obtained with the DEIM algorithm applied to the 17 empirical modes related to displacements. These elements are spread on all the domain, either in the pellet or in the cladding. Similar remarks can be drawn for Ω_σ and Ω_F generated by the interpolation indices related to stress modes and generalized force modes, respectively. The elements in Ω_F are on both sides of the contact surface and also on the symmetry plans and surfaces submitted to pressure loading. A region of interest on the top of the pellet is added as Ω_{ROI} in order to better account for inhomogeneous Neumann conditions. The domain $\Omega_A = \Omega_u \cup \Omega_\sigma \cup \Omega_F \cup \Omega_{\text{ROI}}$ is shown on the right of Figure 13.

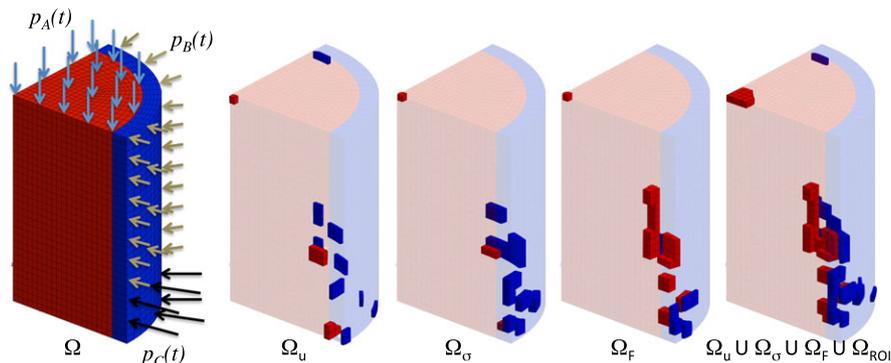


FIGURE 13 Left: the reference geometry and the Neumann boundary condition. Right: various contributions to the reduced integration domain [Colour figure can be viewed at wileyonlinelibrary.com]

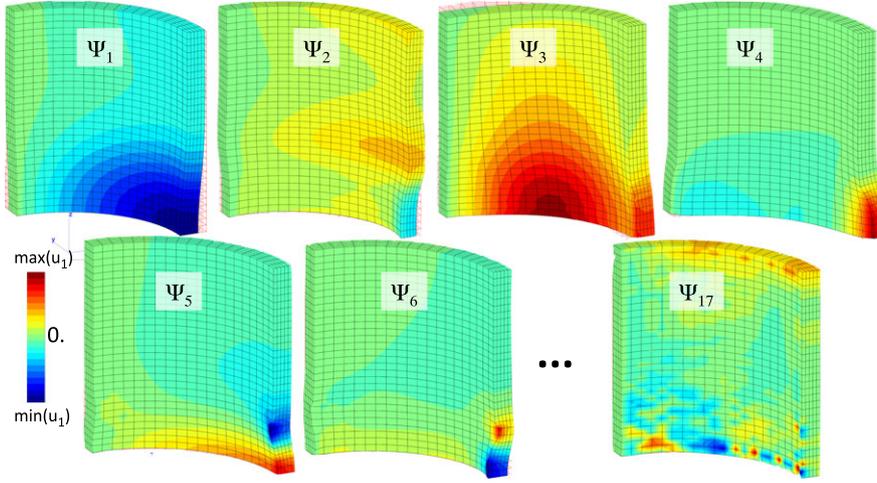


FIGURE 14 Empirical modes for displacement approximation, over the cladding only

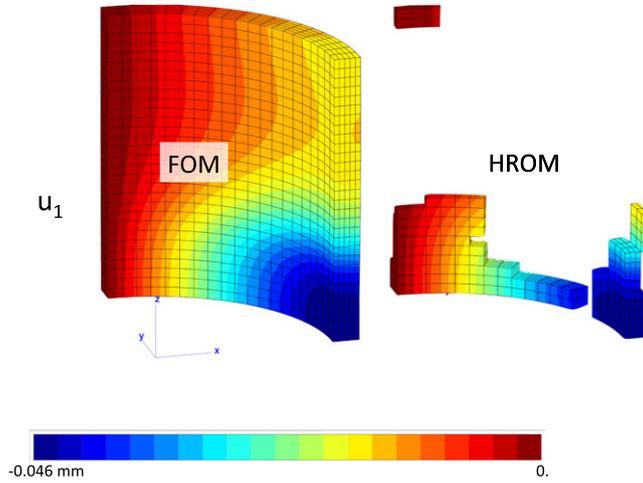


FIGURE 15 Displacement u_1 over the cladding, hybrid reduced-order model (HROM) versus full-order model (FOM)

We can observe that it is not possible to compute the gap for many nodes on the contact surface ($\Gamma^C \cap \partial\Omega_A$) on the red and blue domains. The domain Ω_A is hence extended by elements on Γ^C having a node at a distance less than $d = 0.1$ mm of one node on $\Gamma^C \cap \partial\Omega_A$. If a layer of elements connected to Ω_A is added to this domain, then we obtain the RID shown on the right of Figure 15 for the cladding and on the right of Figure 16 for pellet and cladding. We have clearly increased the number of nodes where the gap can be computed.

7.3 | Hyper-reduced predictions

The hyper-reduced order model obtained by using the simulation results of S1 and S2 is used to predict the numerical results of S3 restrained to the RID. In the hybrid hyper-reduced model, each primal DOF on the potential contact zone is treated with the FE shape function associated. Here, the hyper-reduced simulation is 10 times faster than S3. As for the 1D elastic rope, the hyper-reduced predictions are more accurate for displacements than for stresses. The maximum error over the RID is respectively 1% and 5%. Since the RID contains the region of active contacts, these results are satisfactory. The predictions related to S3 and the hyper-reduced simulation are shown in Figure 15 and Figure 16 for the components u_1 and σ_{11} of displacements and stresses respectively.

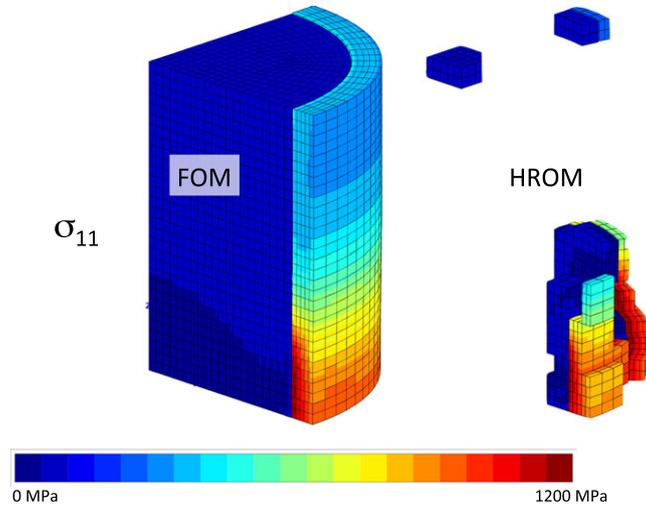


FIGURE 16 Stress σ_{11} , hybrid reduced-order model (HROM) versus full-order model (FOM)

Without the extension of the RID with respect to the d distance, the hyper-reduced prediction are twice less accurate. This is due to the fact that less contact conditions can be correctly treated by the hybrid HR model.

8 | CONCLUSION

We have introduced an extension of the HR method based on an RID able to deal with contact mechanics problems solved due to a mixed formulation. This method is hybrid because the dual basis related to the contact forces is only a restriction of the FOM one. It enables us to obtain a good approximation of the contact forces. If necessary, a hybrid reduced-order/full-order approach is also used for the displacement in order to ensure the inf-sup condition. We propose an efficient error indicator to be coupled with a greedy algorithm in order to choose snapshots to be done. The method has been tested on an academic test case, a reference 1D elastic rope obstacle problem, and compared with the PB method. Results show the efficiency of the proposed strategy in terms of accuracy of the solution (especially contact forces) compared to the PB method. A 3D test case implying two elastic bodies in contact confirms that the RID must account for the contribution of interpolation points related to contact forces. Moreover, it shows the importance for a contact between solids of extending the RID to the elements allowing the computation of the gap. The obtained results enable to appreciate the efficiency of the hybrid HR-contact strategy. A future work will consist in applying the hybrid HR-contact strategy to reduce a node-to-surface contact problem. As the HR method has already shown its capacity to successfully deal with nonlinear mechanical behaviors, the next step will consist in the extension of the HR-contact strategy to viscoplastic behaviors.

ACKNOWLEDGEMENTS

The authors are grateful to the PLEIADES project, financially supported by CEA, EDF, and Framatome, that funded this research work.

ORCID

I. Ramière  <http://orcid.org/0000-0001-7487-0218>

REFERENCES

1. Almroth BO, Stern P, Brogan FA. Automatic choice of global shape functions in structural analysis. *AIAA J.* 1978;16(5):525-528.
2. Lumley JL. The structure of inhomogeneous turbulent flows. *Atmospheric Turbulence and Radio Propagation.* Moscow, Russia: Nauka; 1967:166-178.

3. Sirovich L. Turbulence and the dynamics of coherent structures. Parts I-III. *Q Appl Math*. 1987;45(3):561-590.
4. Alart P, Curnier A. A mixed formulation for frictional contact problems prone to Newton like solution methods. *Comput Meth Appl Mech Eng*. 1991;92(3):353-375.
5. Wriggers P. *Computational Contact Mechanics*. Berlin, Germany: Springer Science & Business Media; 2006.
6. Haasdonk B, Salomon J, Wohlmuth B. A reduced basis method for parametrized variational inequalities. *SIAM J Numer Anal*. 2012;50(5):2656-2676.
7. Balajewicz M, Amsallem D, Farhat C. Projection-based model reduction for contact problems. *Int J Numer Methods Eng*. 2016;106(8):644-663.
8. Quarteroni A, Manzoni A, Negri F. *Reduced Basis Methods for Partial Differential Equations: An Introduction*. Cham, Switzerland: Springer; 2015.
9. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature*. 1999;401(6755):788-791.
10. Barrault M, Maday Y, Nguyen NC, Patera A. An "empirical interpolation" method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Math*. 2004;339(9):667-672.
11. Ryckelynck D. A priori hyperreduction method: an adaptive approach. *J Comput Phys*. 2005;202(1):346-366.
12. An SS, Kim T, James DL. Optimizing cubature for efficient integration of subspace deformations. *ACM Trans Graph*. 2008;27(5):165-175.
13. Chaturantabut S, Sorensen DC. Nonlinear model reduction via discrete empirical interpolation. *SIAM J Sci Comput*. 2010;32(5):2737-2764.
14. Astrid P, Weiland S, Willcox K, Backx T. Missing point estimation in models described by proper orthogonal decomposition. *IEEE Trans Autom Control*. 2008;53(10):2237-2251.
15. Carlberg K, Bou-Mosleh C, Farhat C. Efficient non-linear model reduction via a least-squares Petrov-Galerkin projection and compressive tensor approximations. *Int J Numer Methods Eng*. 2011;86(2):155-181.
16. Everson R, Sirovich L. Karhunen-Loève procedure for gappy data. *J Opt Soc Am A*. 1995;12(8):1657-1664.
17. Bader E, Zhang Z, Veroy K. An empirical interpolation approach to reduced basis approximations for variational inequalities. *Math Comput Modell Dyn Syst*. 2016;22(4):345-361.
18. Balajewicz M, Toivanen J. Reduced order models for pricing European and American options under stochastic volatility and jump-diffusion models. *J Comput Sci*. 2017;20:198-204.
19. 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.
20. Ryckelynck D. Hyper reduction of mechanical models involving internal variables. *Int J Numer Methods Eng*. 2009;77(1):75-89.
21. Ryckelynck D, Lampoh K, Quilicy S. Hyper-reduced predictions for lifetime assessment of elasto-plastic structures. *Meccanica*. 2016;51(2):309-317.
22. Babuška I. The finite element method with Lagrangian multipliers. *Numer Math*. 1973;20(3):179-192.
23. Brezzi F. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *ESAIM: Math Model Numer Anal - Modél Math et Anal Numér*. 1974;8(R2):129-151.
24. Baiges J, Codina R, Idelsohn S. A domain decomposition strategy for reduced order models. Application to the incompressible Navier-Stokes equations. *Comput Meth Appl Mech Eng*. 2013;267:23-42.
25. Ammar A, Chinesta F, Cueto E. Coupling finite elements and proper generalized decompositions. *Int J Multiscale Comput Eng*. 2011;9(1).
26. Kerfriden P, Passieux JC, Bordas SPA. Local/global model order reduction strategy for the simulation of quasi-brittle fracture. *Int J Numer Methods Eng*. 2012;89(2):154-179.
27. Radermacher A, Reese S. Model reduction in elastoplasticity: proper orthogonal decomposition combined with adaptive sub-structuring. *Comput Mech*. 2014;54(3):677-687.
28. Prud'Homme C, Rovas DV, Veroy K, et al. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. *J Fluids Eng*. 2002;124(1):70-80.
29. Prud'Homme C, Rovas DV, Veroy K, Patera AT. A mathematical and computational framework for reliable real-time solution of parametrized partial differential equations. *ESAIM: Math Model Numer Anal*. 2002;36(5):747-771.
30. Signorini A. Questioni di elasticità non linearizzata e semilinearizzata. *Rend Mat e Appl*. 1959;5:95-139.
31. Duvaut G, Lions J-L. *Inequalities in mechanics and physics*. Berlin, Germany: Springer Science & Business Media; 1976.
32. Kikuchi N, Oden JT. *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*. Philadelphia, PA: SIAM; 1988.
33. Chabrand P, Dubois F, Raous M. Various numerical methods for solving unilateral contact problems with friction. *Math Comput Model*. 1998;28(4-8):97-108.
34. Eckart C, Young G. The approximation of one matrix by another of lower rank. *Psychometrika*. 1936;1(3):211-218.
35. Maday Y, Patera A, Turinici G. A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations. *J Sci Comput*. 2002;17(1-4):437-446.
36. Amsallem D, Cortial J, Carlberg K, Farhat C. A method for interpolating on manifolds structural dynamics reduced-order models. *Int J Numer Methods Eng*. 2009;80(9):1241-1258.
37. Nicole LA, Meyer CD, Albright R, Cox J, Duling D. Algorithms, initializations, and convergence for the nonnegative matrix factorization. Paper presented at: Proceedings of the 12Th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2014; Philadelphia, PA.

38. Zitnik M, Zupan B. NIMFA: a python library for nonnegative matrix factorization. *J Mach Learn Res.* 2012;13:849-853.
39. Liu H, Ramière I, Lebon F. On the coupling of local multilevel mesh refinement and ZZ methods for unilateral frictional contact problems in elastostatics. *Comput Meth Appl Mech Eng.* 2017;323:1-26.