

A numerical study of different projection-based model reduction techniques applied to computational homogenisation

Dominic Soldner¹ · Benjamin Brands¹ · Reza Zabihyan¹ · Paul Steinmann¹ · Julia Mergheim¹

Received: 15 July 2016 / Accepted: 29 May 2017 / Published online: 8 June 2017
© The Author(s) 2017. This article is an open access publication

Abstract Computing the macroscopic material response of a continuum body commonly involves the formulation of a phenomenological constitutive model. However, the response is mainly influenced by the heterogeneous microstructure. Computational homogenisation can be used to determine the constitutive behaviour on the macro-scale by solving a boundary value problem at the micro-scale for every so-called macroscopic material point within a nested solution scheme. Hence, this procedure requires the repeated solution of similar microscopic boundary value problems. To reduce the computational cost, model order reduction techniques can be applied. An important aspect thereby is the robustness of the obtained reduced model. Within this study reduced-order modelling (ROM) for the geometrically nonlinear case using hyperelastic materials is applied for the boundary value problem on the micro-scale. This involves the Proper Orthogonal Decomposition (POD) for the primary unknown and hyper-reduction methods for the arising nonlinearity. Therein three methods for hyper-reduction, differing in how the nonlinearity is approximated and the subsequent projection, are compared in terms of accuracy and robustness. Introducing interpolation or Gappy-POD based approximations may not preserve the symmetry of the system tangent, rendering the widely used Galerkin projection sub-optimal. Hence, a different projection related to a Gauss-Newton scheme (Gauss-Newton with Approximated Tensors- GNAT) is favoured to obtain an optimal projection and a robust reduced model.

Keywords Computational homogenisation · Hyper-reduction · Reduced-order modelling · Hyperelasticity

1 Introduction

Phenomenological constitutive models are frequently used to compute the material response of a continuum body. However, the main influence of the macroscopic response is driven by the heterogeneous microstructure, whereas the direct modelling of the underlying microstructure is usually not feasible. A variety of analytical methods exists to account for the microstructure, e.g. Eshelby [12] or Mori-Tanaka-Method [27]. Since these methods are often limited, for instance to the linear regime or by the shape of the inhomogeneities that can be modelled, a different approach is necessary for the general nonlinear case and arbitrary shapes of the inhomogeneities. One possibility is given by computational homogenisation, which requires a nested solution scheme involving the computation of a boundary value problem (BVP) at the microscopic level, using a representative volume element (RVE) for every so-called material point.

Using the Finite Element Method to compute approximate solutions to the governing equations on both scales is often referred to as the FE^2 -method and does not rely on macroscopic constitutive models, but on the solution of underlying BVPs and a consistent transition between the two scales [16]. This approach is usually accompanied by high computational costs, due to the repeated solution of numerous BVPs at the micro-scale in a possibly high dimensional space. One approach to lower the computational cost is given by methods that rely on the fast Fourier transformation (FFT) [28,37]. These methods replace the assembly of the system tangent and the subsequent solution of the linear system on e.g. the

✉ Dominic Soldner
dominic.soldner@fau.de

¹ Chair of Applied Mechanics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Egerlandstrasse 5, 91058 Erlangen, Germany

micro-scale by an algorithm that utilizes FFT. This lowers the complexity of the solution process but does not necessarily reduce storage requirements [14].

Within this work, projection-based ROM is considered, which is characterised by taking advantage of the observation that the solutions of the aforementioned boundary value problems often lie in a lower dimensional subspace and different computational tasks in the offline and online stages. During the offline stage all necessary computations aiming to build a basis for the reduced-order model are carried out. This basis is used in the online stage to compute approximate solutions using a lower dimensional description. Several methods for model reduction exist, e.g. Proper Generalized Decomposition [8, 10], Reduced-Basis methods [3, 30, 31], or approaches using the Proper Orthogonal Decomposition (POD). The latter has been widely used, e.g. in homogenisation, fluid mechanics and many other fields [6, 17, 18, 23, 26, 41]. For a deeper discussion and a broader overview the reader is referred to [2, 33] and references therein.

In the context of computational homogenisation the Reduced Model Multiscale Method (R3M) has been presented in [41], applying POD-based model order reduction to the BVP on the micro-scale for the case of hyperelasticity, directly projecting the governing equations onto the truncated POD basis. Due to the missing approximation of the arising nonlinear terms the computational savings were limited. In general, the approach of solely applying e.g. Galerkin projection for nonlinear problems may even lead to higher computational cost compared to the Finite Element model.

A further contribution in the context of geometrically linear homogenisation is given by [18], in which the stress field itself is approximated by a POD basis. Furthermore the authors showed in which circumstances an ill-posed system is obtained in the context of computational homogenisation and possibilities to avoid this problem. A similarly approach, i.e. not approximating the displacement field, is used in the nonuniform transformation field analysis (NTFA) [24], based on [11], and its extensions [14, 15, 25]. These methods apply a decomposition of the internal variables using reduced bases and derive suitable evolution equations for the reduced variables.

A popular approach in projection-based model reduction is the application of a Galerkin projection, as for e.g. the R3M, to solve the arising overdetermined system for the reduced coordinates. In the presence of nonlinearities these problems are often handled using an interpolation technique [1, 7] or Gappy-POD reconstruction [5, 13], both also often referred to as hyper-reduction techniques. Convergence difficulties have been reported for certain ROM configurations applying Galerkin projection coupled with hyper-reduction, e.g. [5, 9, 22, 34]. In [9] the condition

number of the reduced system tangent is suspected to lead to said convergence problems. The authors therefore propose a gappy data reconstruction with Galerkin projection to improve the condition number. While improving the robustness the number of diverging ROMs has only been reduced. In [5] an alternative to the Galerkin projection is proposed, which is related to the Gauss-Newton algorithm (GNAT). This projection lowers the constraints of the arising system tangent matrix and renders a robust model reduction technique.

The aim of this contribution is to apply projection-based ROM techniques in the context of computational homogenisation of hyperelastic materials including hyper-reduction techniques, which requires a robust reduced model for the present multi-query context. The focus thereby lies on the problem on the micro-scale, i.e. the quantities computed on the micro-scale which would be used on the macro-scale in a fully coupled problem formulation. The main objective is to compare three different model reduction approaches in terms of accuracy, robustness and optimality.

The remainder of this contribution is organised as follows: Sect. 2 discusses the fundamentals of first-order computational homogenisation and the governing equations. Section 3 describes the ROM techniques used within this study, including considerations regarding optimality of different projection techniques. Section 4 presents numerical examples, followed by some concluding remarks in Sect. 5.

2 Computational homogenisation in a nutshell

In solid mechanics, phenomenological constitutive models are frequently used to describe the material response of a body under a given load. However, the response is mainly driven by the heterogeneous microstructure. One possible approach to account for the heterogeneous material would be to directly model the substructure, e.g. within a FE model, which usually leads to computationally expensive models. A different approach is given by computational homogenisation. In this context the material at the macroscopic level is modelled without a prescribed constitutive model. Instead, the constitutive response is computed for every so-called material point at the micro-scale, taking into account the heterogeneities through prescribed constitutive behaviour of the constituents. To distinguish between the two scales, the superscripts $(\bullet)^M$ and $(\bullet)^m$ are used to denote quantities on the macro- and micro-scale respectively. Figure 1 illustrates the general concept of this method.

On both scales the balance of linear momentum represents the equation of interest, which reads on the micro-scale

$$\text{Div} \mathbf{P}^m = \mathbf{0} \text{ in } \mathcal{B}_0^m, \quad (2.1)$$

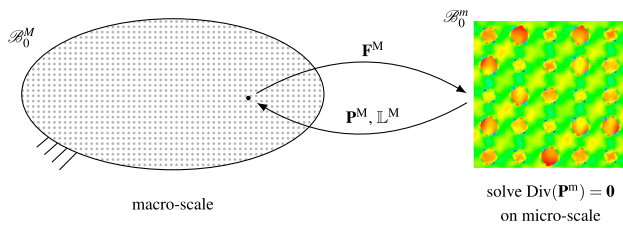


Fig. 1 General concept of first-order computational homogenisation

where \mathbf{P}^m denotes the Piola stress on the micro-scale and \mathcal{B}_0^m the domain of the RVE. The microscopic stresses are computed using a hyperelastic constitutive model, i. e.

$$\mathbf{P}^m = \frac{\partial \Psi^m(\mathbf{F}^m)}{\partial \mathbf{F}^m}, \tag{2.2}$$

with \mathbf{F}^m representing the micro-scale deformation gradient and Ψ^m the strain energy density. The material response on the macro-scale relies on quantities computed on the micro-scale. In order to compute these quantities, certain requirements have to be satisfied.

2.1 Hill-Mandel condition

A main ingredient of scale transition is the equality of the virtual work on the macro- and micro-scale, known as Hill-Mandel condition [19–21],

$$\frac{1}{V_0} \int_{\mathcal{B}_0^m} \mathbf{P}^m : \delta \mathbf{F}^m \, dV = \mathbf{P}^M : \delta \mathbf{F}^M, \tag{2.3}$$

where V_0 denotes the volume of the RVE in the reference configuration. Assuming a linear ansatz for the deformation on the micro-scale,

$$\mathbf{x}^m = \mathbf{F}^M \cdot \mathbf{X}^m + \tilde{\mathbf{u}}, \tag{2.4}$$

with $\tilde{\mathbf{u}}$ denoting the fluctuation field and $\mathbf{X}^m, \mathbf{x}^m$ the coordinates in the reference and spatial configuration respectively, the deformation gradient is given by

$$\mathbf{F}^m = \nabla_X \mathbf{x}^m = \mathbf{F}^M + \nabla_X \tilde{\mathbf{u}}. \tag{2.5}$$

Considering the variation of the deformation gradient

$$\delta \mathbf{F}^m = \delta \mathbf{F}^M + \nabla_X \delta \tilde{\mathbf{u}} \tag{2.6}$$

and inserting Eq. (2.6) into Eq. (2.3) leads to

$$\begin{aligned} 0 &= \frac{1}{V_0} \int_{\mathcal{B}_0^m} \mathbf{P}^m : [\delta \mathbf{F}^M + \nabla_X \delta \tilde{\mathbf{u}}] \, dV - \mathbf{P}^M : \delta \mathbf{F}^M \tag{2.7} \\ &= \left[\frac{1}{V_0} \int_{\mathcal{B}_0^m} \mathbf{P}^m \, dV - \mathbf{P}^M \right] : \delta \mathbf{F}^M \\ &\quad + \left[\frac{1}{V_0} \int_{\mathcal{B}_0^m} \mathbf{P}^m : \nabla_X \delta \tilde{\mathbf{u}} \, dV \right]. \end{aligned}$$

It follows that the volume average of the microscopic Piola stress has to equal its macro-scale counterpart, i. e.

$$\frac{1}{V_0} \int_{\mathcal{B}_0^m} \mathbf{P}^m \, dV = \mathbf{P}^M, \tag{2.8}$$

for the first term to vanish. Furthermore, the second term, containing the fluctuations, has to vanish in order to comply with the Hill-Mandel condition. Reformulating this term as a boundary integral renders admissible boundary conditions. Within this work the fluctuations are set to vanish on the Dirichlet boundary, which yields linear displacement boundary conditions:

$$\mathbf{u}^m = [\mathbf{F}^M - \mathbf{1}] \cdot \mathbf{X}^m \text{ on } \partial \mathcal{B}_0^m \tag{2.9}$$

It should be noted that alternatively periodic boundary conditions could be applied.

2.2 Computation of the tangent

To provide information about the behaviour at a material point due to an increased load, the macroscopic tangent modulus has to be computed. In the present work the approach in [39] is used. Therein, a series of numerical test cases is performed in order to compute the fourth order Lagrangian elasticity tensor

$$\mathbb{L}^M = \frac{\partial \mathbf{P}^M}{\partial \mathbf{F}^M}. \tag{2.10}$$

This leads, depending on the dimension d of the problem, to a total of d^2 linear perturbation test cases at the converged equilibrium state with the perturbations of the macroscopic deformation gradient

$$\delta_{iJ} \mathbf{F}^M = h \mathbf{e}_i \otimes \mathbf{E}_J \text{ with } 0 < h \ll 1, \tag{2.11}$$

with \mathbf{e}_i and \mathbf{E}_J denoting the basis vectors in the spatial and the reference configuration respectively. Hence, the computations

$$[\mathbb{L}^M]_{kLiJ} = \frac{[\delta_{iJ} \mathbf{P}^M]_{kL}}{h} \tag{2.12}$$

are necessary to compute the constitutive constants.

2.3 Weak form and spatial discretisation

Within the scope of this contribution, an approximate solution to the micro-scale problem is computed using the FEM, which requires the weak formulation

$$\int_{\mathcal{B}_0^m} \nabla_X \delta \mathbf{u}^m : \mathbf{P}^m \, dV = 0, \quad \forall \delta \mathbf{u}^m \tag{2.13}$$

with $\mathbf{u}^m = [\mathbf{F}^M - \mathbf{1}] \cdot \mathbf{X}^m$ on $\partial \mathcal{B}_0^m$.

The discretisation of the above equations is carried out using the Bubnov-Galerkin Finite Element Method in a standard manner, i. e.

$$\mathbf{u}^{m,el} = \sum_{i=1}^{n_{en}} N_i \mathbf{u}_i, \quad \delta \mathbf{u}^{m,el} = \sum_{i=1}^{n_{en}} N_i \delta \mathbf{u}_i, \tag{2.14}$$

where the displacement field and the test function are approximated as a sum of shape functions (Lagrange polynomials) and nodal values, with n_{en} denoting the number of element nodes. Using the definitions given in Eq. (2.14), the discretisation of Eq. (2.13) takes the form

$$\int_{\mathcal{B}_0^m} \nabla_X \delta \mathbf{u}^m : \mathbf{P}^m \, dV \approx \sum_{e=1}^{n_{el}} \int_{\mathcal{B}_0^{m,e}} \sum_{i=1}^{n_{en}} [\delta \mathbf{u}_i^m \otimes \nabla_X N_i] : \mathbf{P}^m \, dV^e = 0, \tag{2.15}$$

with the operator $\sum_{e=1}^{n_{el}}$ representing the assembly of the element contributions. Since Eq. (2.15) has to hold for arbitrary $\delta \mathbf{u}_i^m$, it follows that

$$\sum_{e=1}^{n_{el}} \int_{\mathcal{B}_0^{m,e}} \sum_{i=1}^{n_{en}} \nabla_X N_i \cdot \mathbf{P}^m \, dV^e = \mathbf{0}, \tag{2.16}$$

which may be written as a vector-valued equation

$$\mathbf{f}^m(\mathbf{u}^m) = \mathbf{0}. \tag{2.17}$$

Hence, \mathbf{f}^m represents a nonlinear function of the unknown nodal displacement values. In order to find an approximate

solution an iterative Newton-Raphson scheme is used. This requires the linearisation of the nonlinear function

$$\Delta \mathbf{f}^m = \mathbf{K}^m \cdot \Delta \mathbf{u}^m, \tag{2.18}$$

which yields the tangent stiffness matrix

$$\mathbf{K}_{JJ}^m = \sum_{e=1}^{n_{el}} \left[\int_{\mathcal{B}_0^{m,e}} \left[\nabla_X N_i \cdot \frac{\partial \mathbf{P}^m}{\partial \mathbf{F}^m} \cdot \nabla_X N_j \right] dV^e \right]. \tag{2.19}$$

Starting at an initial guess of the solution $\mathbf{u}^{m,k}$, an iterative update of the displacement is computed via

$$\Delta \mathbf{u}^{m,k} = - [\mathbf{K}^{m,k}]^{-1} \cdot \mathbf{f}^{m,k}, \tag{2.20}$$

leading to

$$\mathbf{u}^{m,k+1} = \mathbf{u}^{m,k} + \Delta \mathbf{u}^{m,k}. \tag{2.21}$$

This iterative procedure is repeated until a prescribed convergence criterion is satisfied. Since the focus lies on the micro-scale problem, the superscripts $(\bullet)^m$ will be omitted in the following in order to improve readability and only be used if it is required in the context. Furthermore, Eq. (2.17) will be solved for the unknown fluctuation field $\tilde{\mathbf{u}}$ instead for the micro displacement field \mathbf{u} .

3 Reduced-order modelling

As presented in the previous section, one possibility to compute approximate solutions to the microscopic problem defined by Eq. (2.1) involves the use of the Finite Element Method. Depending on the discretisation of the considered domain this may lead to large-scale systems. Especially in a multi-query context, as is the case in computational homogenisation, this leads to high computational costs. The solutions of such systems often lie in an affine subspace of lower dimension and therefore techniques to reduce the dimensionality of such problems are desired. Within the scope of this study a POD-based ROM approach is applied, including various hyper-reduction techniques.

3.1 Proper Orthogonal Decomposition

Consider discrete values $\mathbf{a}(\bullet)_j$, e.g. displacement fluctuation values of the BVP of the micro-scale problem, solved using the Finite Element Method. These so-called snapshots are arranged into a matrix \mathbf{A}

$$\mathbf{A}(\tilde{\mathbf{u}}) = [\mathbf{a}(\tilde{\mathbf{u}})_1, \dots, \mathbf{a}(\tilde{\mathbf{u}})_{n_s}] \in \mathbb{R}^{n \times n_s}, \tag{3.1}$$

with rank $d \leq \min(n, n_s)$, n denoting the number of degrees of freedom of the FEM model and n_s the number of snapshots. These solutions span a certain space denoted by \mathcal{V} . The snapshot POD [36] is then used to filter out the dominant characteristics, allowing the computation of an orthonormal basis, that best suites a rank $l \ll d$ approximation of the snapshots in a least-squares sense. This task can be formulated as a constraint optimization problem. It can be shown that the solution of this optimization problem is given by the first l left singular vectors $\mathbf{U}(:, 1:l)$ of $\mathbf{A}(\bar{\mathbf{u}}) = \mathbf{U} \cdot \boldsymbol{\Sigma} \cdot \mathbf{V}^T$ called the POD basis of rank l [2,40]. The basis vectors optimally represent the snapshots in a least-squares sense for the given rank l approximation. For the choice of a suitable l , it is useful to consider a truncation criterion ε in order to select the first l POD modes. For a system of rank d the criterion may be defined in terms of the singular values σ_i as

$$\frac{\sum_{i=1}^l \sigma_i^2}{\sum_{i=1}^d \sigma_i^2} \geq 1 - \varepsilon, \tag{3.2}$$

which gives information about the ability of the truncated basis to reproduce the snapshots. In the following, POD bases will be abbreviated by $\mathbf{U}_{(\bullet)}^r$, e.g. $\mathbf{U}_{(\bar{\mathbf{u}})}^r \in \mathbb{R}^{n \times l}$ for an l -dimensional POD basis of the displacement fluctuation field $\bar{\mathbf{u}}$. In case of large snapshot sets, a nested POD as given in [4,32] may be used, which in essence partitions the snapshots into smaller sets, computes a lower rank approximation of each set and eventually computes a POD of the low rank approximations of the snapshot sets.

3.2 Projection approaches

Introducing the dimensionality reduction for the primary unknown via the POD renders an overdetermined system of equations and therefore suitable projection techniques are required. The widely used Galerkin projection and an alternative Petrov-Galerkin projection are briefly reviewed in this section. For a more detailed discussion the reader is referred to [5,35].

3.2.1 Galerkin projection

Employing the Galerkin projection the fluctuation field $\bar{\mathbf{u}}$ and the test function are approximated using the POD basis vectors, i. e.

$$\bar{\mathbf{u}} = \underbrace{\mathbf{U}_{(\bar{\mathbf{u}})}^r}_{n \times l} \cdot \underbrace{\hat{\mathbf{u}}}_{l \times 1} \quad \text{and} \quad \delta \mathbf{u} = \underbrace{\mathbf{U}_{(\bar{\mathbf{u}})}^r}_{n \times l} \cdot \underbrace{\delta \hat{\mathbf{u}}}_{l \times 1}, \tag{3.3}$$

with the generalised coordinates $\hat{\mathbf{u}}$ for the reduced model. Inserting the definitions from Eq. (3.3) into Eq. (2.17) renders the reduced nonlinear term

$$\hat{\mathbf{f}} = \underbrace{\mathbf{U}_{(\bar{\mathbf{u}})}^r}_{l \times n}^T \cdot \underbrace{\mathbf{f}(\mathbf{U}_{(\bar{\mathbf{u}})}^r \cdot \hat{\mathbf{u}})}_{n \times 1} \tag{3.4}$$

and the corresponding reduced tangent stiffness matrix

$$\hat{\mathbf{K}} = \underbrace{\mathbf{U}_{(\bar{\mathbf{u}})}^r}_{l \times n}^T \cdot \underbrace{\mathbf{K}(\mathbf{U}_{(\bar{\mathbf{u}})}^r \cdot \hat{\mathbf{u}})}_{n \times n} \cdot \underbrace{\mathbf{U}_{(\bar{\mathbf{u}})}^r}_{n \times l} \tag{3.5}$$

Recalling the dimension of the POD basis matrix $\mathbf{U}_{(\bar{\mathbf{u}})}^r \in \mathbb{R}^{n \times l}$, with n equal to the number of degrees of freedom and l the number of selected modes according to a chosen error criterion, the obvious benefit of this procedure is that now a system of equations for only l unknowns has to be solved. This decreases the computational cost especially for $l \ll n$. Using the iterative Newton-Raphson scheme

$$\Delta \hat{\mathbf{u}} = - \left[\hat{\mathbf{K}}^k \right]^{-1} \cdot \hat{\mathbf{f}} \tag{3.6}$$

leads to the updated approximate solution

$$\hat{\mathbf{u}}^{k+1} = \hat{\mathbf{u}}^k + \Delta \hat{\mathbf{u}}. \tag{3.7}$$

As shown in [5,35] this projection, which produces Eq. (3.6), is optimal in the sense that it minimizes the error between the solutions of the reduced and the full order model in the \mathbf{K} -norm:

$$\Delta \hat{\mathbf{u}} = \arg \min_{\mathbf{w} \in \mathbb{R}^l} \left\| \mathbf{U}_{(\bar{\mathbf{u}})}^r \cdot \mathbf{w} - \left[-\mathbf{K}^{-1} \cdot \mathbf{f} \right] \right\|_{\mathbf{K}} \tag{3.8}$$

Thereby the tangent stiffness matrix has to be symmetric positive definite. It will be shown in Sect. 3.3 that neither of the hyper-reduction techniques discussed in this work guarantees the symmetry of the unreduced tangent stiffness matrix given in Eq. (3.20). Hence, the Galerkin projection combined with the hyper-reduction approaches as discussed within the work is not optimal in the sense of Eq. (3.8).

3.2.2 Petrov-Galerkin projection

As shown in Eq. (3.4) and Eq. (3.5) the Galerkin projection multiplies $\mathbf{U}_{(\bar{\mathbf{u}})}^r$ from the left. An alternative approach is given by selecting $\left[\mathbf{K} \cdot \mathbf{U}_{(\bar{\mathbf{u}})}^r \right]^T$ to be multiplied from the left, which results in

$$\hat{\mathbf{f}} = \underbrace{\mathbf{U}_{(\bar{\mathbf{u}})}^r}_{l \times n}^T \mathbf{K}^T \cdot \underbrace{\mathbf{f}}_{n \times 1} \tag{3.9}$$

and

$$\hat{\mathbf{K}} = \underbrace{\mathbf{U}_{(\hat{\mathbf{u}})}^r \mathbf{T} \mathbf{K}^T}_{l \times n} \cdot \underbrace{\mathbf{K}}_{n \times n} \left(\underbrace{\mathbf{U}_{(\hat{\mathbf{u}})}^r \cdot \hat{\mathbf{u}}}_{n \times 1} \right) \cdot \underbrace{\mathbf{U}_{(\hat{\mathbf{u}})}^r}_{n \times 1}. \quad (3.10)$$

This approach renders

$$\Delta \hat{\mathbf{u}} = \arg \min_{\mathbf{w} \in \mathbb{R}^l} \left\| \mathbf{U}_{(\hat{\mathbf{u}})}^r \cdot \mathbf{w} - \left[-\mathbf{K}^{-1} \cdot \mathbf{f} \right] \right\|_{\mathbf{K}^T \mathbf{K}} \quad (3.11)$$

and corresponds to the least-square problem

$$\Delta \hat{\mathbf{u}} = \arg \min_{\mathbf{w} \in \mathbb{R}^l} \left\| \mathbf{K} \cdot \mathbf{U}_{(\hat{\mathbf{u}})}^r \cdot \mathbf{w} + \mathbf{f} \right\|_2, \quad (3.12)$$

requiring the tangent stiffness matrix solely to be regular [5, 35].

While reducing the number of unknowns, the nonlinear terms still have to be evaluated at the full scale and projected onto the subspace at every iteration step, which clearly limits the computational savings. Hence, further reduction techniques have to be applied in order to significantly reduce the computational cost.

3.3 Hyper-reduction

As previously highlighted, the direct projection approach still depends on the full scale dimension n , due to the evaluation of the nonlinear terms. There exists a variety of approximation techniques for nonlinearities such as Empirical Interpolation Method (EIM) [1], its extension Discrete Empirical Interpolation Method (DEIM) [7] or the Gappy-POD [5, 13], amongst others. It should be noted that, as shown in [18], employing hyper-reduction may lead to ill-posed systems, since the internal force vector, which is approximated using hyper-reduction, is zero at a converged state. Within our studies the bases for the subsequent hyper-reduction techniques are computed using snapshots of the internal force vector during the iterative solution process (the vector is non-zero). Hence, using a non-truncated basis for the approximated nonlinearity, one obtains the same internal force vector as that of the full order model, which justifies this approach. Within the present study the DEIM and the Gappy-POD in combination with the discussed projection approaches are compared and will therefore be shortly discussed.

3.3.1 Discrete empirical interpolation method

In essence, this method approximates a nonlinear function as

$$\mathbf{f}(\tilde{\mathbf{u}}(\mu)) \approx \mathbf{U}_{(\mathbf{f})}^r \cdot \mathbf{c}(\tilde{\mathbf{u}}(\mu)), \quad (3.13)$$

where the parameter μ is introduced to denote the dependence of the fluctuation field on the macroscopic deformation

gradient \mathbf{F}^M , used to compute the macroscopic displacement field. The parameter stems from a suitable parameter space $\mu \in \mathcal{D} \subset \mathbb{R}^d$, e.g. $\mu = [F_{11}^M, F_{12}^M, F_{21}^M, F_{22}^M] \in \mathcal{D} \subset \mathbb{R}^4$, for the two dimensional case. The direct projection approach requires the collection of snapshots $\mathbf{a}_{(\tilde{\mathbf{u}})_i}$ of the Finite Element approximated fluctuation field arranged into $\mathbf{A}_{(\tilde{\mathbf{u}})}$ in order to compute the POD basis. Using DEIM, snapshots of the corresponding nonlinear function $\mathbf{a}_{(\mathbf{f})_i}$, see Eq. (2.17), are collected during the iterative solution procedure in the offline phase and assembled into $\mathbf{A}_{(\mathbf{f})}$,

$$\mathbf{A}_{(\mathbf{f})} = \left[\mathbf{a}_{(\mathbf{f})_1}, \dots, \mathbf{a}_{(\mathbf{f})_{n_s}} \right], \quad (3.14)$$

where n_s equals the number of considered snapshots. Performing the POD of $\mathbf{A}_{(\mathbf{f})}$ renders the matrix $\mathbf{U}_{(\mathbf{f})}^r \in \mathbb{R}^{n \times k}$, representing a k -dimensional orthonormal basis, i.e. k modes are considered, for the space spanned by the snapshots of the nonlinear term. The coefficients of \mathbf{c} in Eq. (3.13) are computed using k rows of $\mathbf{f}(\tilde{\mathbf{u}}(\mu))$

$$\mathcal{P}^T \cdot \mathbf{f} = \left[\mathcal{P}^T \cdot \mathbf{U}_{(\mathbf{f})}^r \right] \cdot \mathbf{c} \quad (3.15)$$

where \mathcal{P} denotes an extraction operator. This may be considered as a matrix composed of k vectors

$$\mathcal{P} = [\mathbf{i}_{\rho_1}, \dots, \mathbf{i}_{\rho_k}] \in \mathbb{R}^{n \times k}, \quad (3.16)$$

where $\mathbf{i}_{\rho_i} = [0, \dots, 0, 1, 0, \dots, 0]^T$ denotes a vector in which the position of the only nonzero entry corresponds to the index ρ_i [7]. Since the matrix $\left[\mathcal{P}^T \cdot \mathbf{U}_{(\mathbf{f})}^r \right]$ is always regular [7] the coefficients of \mathbf{c} can be uniquely determined. This leads together with Eq. (3.15) to the DEIM approximation

$$\mathbf{f} \approx \mathbf{U}_{(\mathbf{f})}^r \cdot \mathbf{c} = \underbrace{\mathbf{U}_{(\mathbf{f})}^r \cdot \left[\mathcal{P}^T \cdot \mathbf{U}_{(\mathbf{f})}^r \right]^{-1}}_{n \times k} \cdot \underbrace{\mathcal{P}^T \cdot \mathbf{f}}_{k \times 1}. \quad (3.17)$$

The nonlinear term \mathbf{f} now only needs to be evaluated at k entries specified by \mathcal{P} . The corresponding DEIM indices ρ are determined using algorithm 1, proposed in [7], which computes the indices ρ based on the basis $\mathbf{U}_{(\mathbf{f})}^r$. The reduced nonlinear term reads after Galerkin projection

$$\hat{\mathbf{f}} = \underbrace{\mathbf{U}_{(\hat{\mathbf{u}})}^r \mathbf{T} \cdot \mathbf{U}_{(\mathbf{f})}^r}_{l \times k} \cdot \left[\mathcal{P}^T \cdot \mathbf{U}_{(\mathbf{f})}^r \right]^{-1} \cdot \underbrace{\mathcal{P}^T \cdot \mathbf{f}}_{k \times 1}, \quad (3.18)$$

where the first term, of dimension $l \times k$, represents a constant quantity and is thus computed during the offline phase. Online only k components, corresponding to the k DEIM indices, need to be computed. The tangent is obtained as the derivative of Eq. (3.18),

$$\hat{\mathbf{K}} = \underbrace{\mathbf{U}_{(\hat{\mathbf{u}})}^r \mathbf{T} \cdot \mathbf{U}_{(\mathbf{f})}^r}_{l \times k} \cdot \left[\mathcal{P}^T \cdot \mathbf{U}_{(\mathbf{f})}^r \right]^{-1} \cdot \underbrace{\mathcal{P}^T \cdot \mathbf{K} \cdot \mathbf{U}_{(\hat{\mathbf{u}})}^r}_{k \times l} \tag{3.19}$$

The part of (3.19) which represents the tangent approximation,

$$\tilde{\mathbf{K}} = \underbrace{\mathbf{U}_{(\mathbf{f})}^r}_{n \times k} \cdot \left[\mathcal{P}^T \cdot \mathbf{U}_{(\mathbf{f})}^r \right]^{-1} \cdot \underbrace{\mathcal{P}^T \cdot \mathbf{K}}_{k \times n} \tag{3.20}$$

may not be symmetric as pointed out by [5,34]. Hence, applying a Galerkin projection is not optimal in the sense of Eq. (3.8). The same holds for the Gappy-POD in combination with a Galerkin projection. Consider therefore the following short example with the quantities

$$\mathbf{K} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{U}_{(\mathbf{f})}^r = \begin{bmatrix} 0 & 0 \\ -0.7071 & 0 \\ 0.7071 & 0 \\ 0 & 1 \end{bmatrix}$$

and the sampling matrix

$$\mathcal{P} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Using these matrices to evaluate Eq. (3.20) produces

$$\tilde{\mathbf{K}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3.21}$$

which is not symmetric. This small example shows that it can not be guaranteed that the tangent approximation in Eq. (3.20) preserves symmetry.

3.3.2 Gappy-POD

Contrary to the interpolation in Eq. (3.17) the Gappy-POD [5,9,13] uses regression to approximate the nonlinear function. The approximation of the nonlinear term results in

$$\mathbf{f} \approx \mathbf{U}_{(\mathbf{f})}^r \cdot \mathbf{c} = \underbrace{\mathbf{U}_{(\mathbf{f})}^r}_{n \times k_s} \cdot \left[\mathcal{P}^T \cdot \mathbf{U}_{(\mathbf{f})}^r \right]^\dagger \cdot \underbrace{\mathcal{P}^T \cdot \mathbf{f}}_{k_s \times 1} \tag{3.22}$$

Algorithm 1 Compute the DEIM indices ρ_i

Input: POD basis $\mathbf{U}_{(\mathbf{f})}^r$

Output: $\boldsymbol{\rho} = [\rho_1, \dots, \rho_k]$ containing the k DEIM indices

```

 $\mathbf{u}_{(\mathbf{f})j} = \mathbf{U}_{(\mathbf{f})}^r(:, j)$ 
set  $\rho_1$  such that  $|\mathbf{u}_{(\mathbf{f})1}(\rho_1)| = \|\mathbf{u}_{(\mathbf{f})1}\|_\infty$ 
 $\mathbf{U}(:, 1) = \mathbf{u}_{(\mathbf{f})1}, \mathcal{P}(:, 1) = \mathbf{i}_{\rho_1}$ 

```

```

for  $j=2; j \leq k; j++$  do
   $\mathbf{c} = (\mathcal{P}^T \cdot \mathbf{U})^{-1} \cdot \mathcal{P}^T \cdot \mathbf{u}_{(\mathbf{f})j}$ 
   $\mathbf{r} = \mathbf{u}_{(\mathbf{f})j} - (\mathbf{U} \cdot \mathbf{c})$ 
  set  $\rho_j$  such that  $|\mathbf{r}(\rho_j)| = \|\mathbf{r}\|_\infty$ 
  Update  $\mathbf{U}, \mathcal{P}$ :
   $\mathbf{U}(:, j) = \mathbf{u}_{(\mathbf{f})j}, \mathcal{P}(:, j) = \mathbf{i}_{\rho_j}$ 
end for

```

Here, k_s indicates the number of sampling points with $k_s \geq k$, i.e. more sampling points than modes (keeping in mind that $\mathbf{U}_{(\mathbf{f})}^r \in \mathbb{R}^{n \times k}$) and \dagger denotes the pseudo-inverse. The tangent is computed analogously to Eq. (3.19). Similar to [5,9], algorithm 2 represents the point selection algorithm used in this work.

Algorithm 2 Compute the indices ρ_i for Gappy-POD

Input: POD basis $\mathbf{U}_{(\mathbf{f})}^r \in \mathbb{R}^{n \times k_s}$

Output: $\boldsymbol{\rho} = [\rho_1, \dots, \rho_{k_s}]$ containing the k_s sampling indices

```

set  $\rho_1$  such that  $|\mathbf{U}_{(\mathbf{f})}^r(\rho_1, 1)| = \|\mathbf{U}_{(\mathbf{f})}^r(:, 1)\|_\infty$ 

```

```

 $\mathcal{P}(:, 1) = \mathbf{i}_{\rho_1}, \mathbf{Q} = \mathbf{U}_{(\mathbf{f})}^r(:, 1)$ 

```

```

for  $j=2; j \leq k_s; j++$  do
   $\mathbf{V} = \mathcal{P}^T \cdot \mathbf{Q}$ 
   $\mathbf{W} = \mathbf{V}^T \mathbf{V}$ 
   $\mathbf{z} = \mathbf{U}_{(\mathbf{f})}^r(:, j)$  // closest truncated basis vectors if  $j > k$ 
   $\mathbf{c} = \mathbf{W}^{-1} \cdot \mathbf{V}^T \cdot \mathcal{P}^T \cdot \mathbf{z}$ 
   $\mathbf{r} = \mathbf{z} - (\mathbf{Q} \cdot \mathbf{c})$ 
  set  $\rho_j$  such that  $|\mathbf{r}(\rho_j)| = \|\mathbf{r}\|_\infty$ 
  Update  $\mathcal{P}, \mathbf{Q}$ :
   $b := \min(j, k)$  // max dimension of  $\mathbf{Q} = k$ 
   $\mathcal{P}(:, j) = \mathbf{i}_{\rho_j}, \mathbf{Q} = \mathbf{U}_{(\mathbf{f})}^r(:, 1:b)$ 
end for

```

3.3.3 Gauss-Newton with approximated tensors (GNAT)

Instead of combining Galerkin projection and Gappy-POD, the GNAT solves the least-square problem in Eq. (3.12) using a Gappy-POD approximation of the nonlinear terms, which reads

$$\Delta \hat{\mathbf{u}} = \arg \min_{\mathbf{w} \in \mathbb{R}^l} \left\| \mathbf{Y} \cdot \mathcal{P}^T \cdot \mathbf{K} \cdot \mathbf{U}_{(\hat{\mathbf{u}})}^r \cdot \mathbf{w} + \mathbf{X} \cdot \mathcal{P}^T \cdot \mathbf{f} \right\|_2 \tag{3.23}$$

with the matrices

$$\mathbf{X} = \mathbf{U}_{(\mathbf{K})}^{\mathbf{r}\mathbf{T}} \cdot \mathbf{U}_{(\mathbf{f})}^{\mathbf{r}} \cdot \left[\mathcal{D}^{\mathbf{T}} \cdot \mathbf{U}_{(\mathbf{f})}^{\mathbf{r}} \right]^{\dagger} \in \mathbb{R}^{k \times k_s}$$

$$\mathbf{Y} = \left[\mathcal{D}^{\mathbf{T}} \cdot \mathbf{U}_{(\mathbf{K})}^{\mathbf{r}} \right]^{\dagger} \in \mathbb{R}^{k \times k_s},$$

using $\mathbf{U}_{(\mathbf{K})}^{\mathbf{r}\mathbf{T}} \cdot \mathbf{U}_{(\mathbf{K})}^{\mathbf{r}} = \mathbf{I} \in \mathbb{R}^{k \times k}$, while being independent of the dimension of the FEM model n . Here, the quantity $\mathbf{U}_{(\mathbf{K})}^{\mathbf{r}}$ is introduced to account for the possibility of different snapshot selection strategies for the gappy approximation of the residual and the system tangent as presented in [5,6]. Within the scope of the present work snapshots of the residual obtained from the Finite Element model (including the iterative states during the Newton-Raphson solution procedure) are used. These serve as the input to build the reduced basis for both the residual and the tangent, i.e. $\mathbf{U}_{(\mathbf{K})}^{\mathbf{r}} = \mathbf{U}_{(\mathbf{f})}^{\mathbf{r}}$.

4 Numerical examples

For the subsequent numerical examples a representative volume element of a Neo-Hookean hyperelastic material is used as depicted in Fig. 2. The matrix material and the inclusions differ in terms of the shear modulus, i.e. $\mu_m = 3.4 \times 10^7$ and $\mu_i = 2.0 \times 10^8$, while the Poisson's ratio is set to be $\nu = 0.23$. The subscripts $\{\bullet\}_m$ and $\{\bullet\}_i$ denote the matrix and the inclusion respectively. The Finite Element discretisation renders of a total of 2.312 unknowns and the computations are carried out using a plane strain configuration. As mentioned in Sect. 3 a snapshot POD is used to construct a reduced basis for the unknown fluctuation field. Hence, a training set $\mathcal{D}_{\text{train}}$ is necessary for which the full order model has to be computed. For the subsequent examples the training set was specified to be

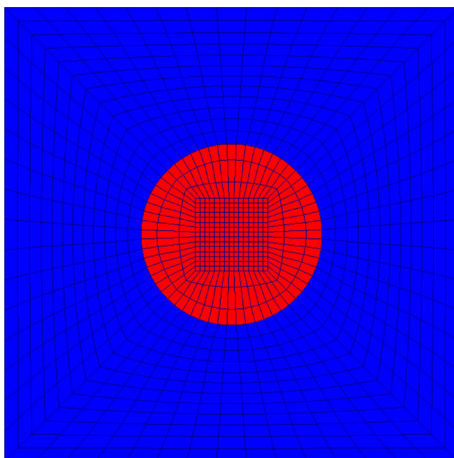


Fig. 2 Finite Element model of fiber reinforced material

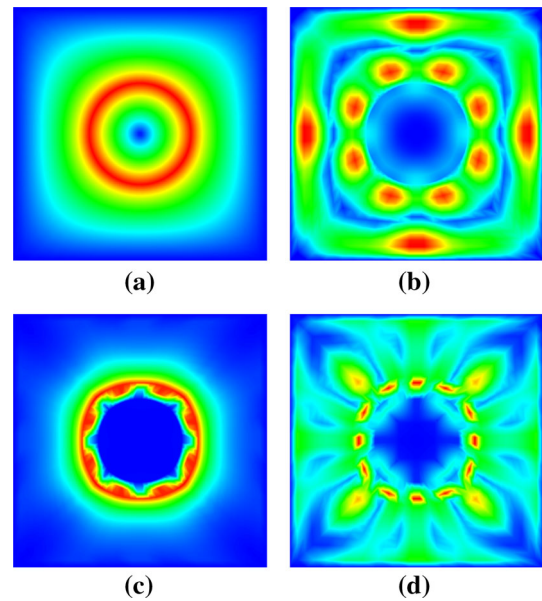


Fig. 3 Examples for POD modes: **a** and **b** show the 1st and 19th mode of $\mathbf{U}_{(\mathbf{u})}^{\mathbf{r}}$ respectively; **c** and **d** show the 5th and 22th mode of $\mathbf{U}_{(\mathbf{f})}^{\mathbf{r}}$ respectively

$$\mathcal{D}_{\text{train}} = \mathbf{1} + \Delta \mathbf{F}^{\mathbf{M}}$$

with $\left[\Delta \mathbf{F}^{\mathbf{M}} \right]_{ij} \in \{-0.2, -0.12, -0.04, 0.04, 0.12, 0.2\}$

(4.1)

Based on this training set a few POD basis modes for the fluctuation field and the nonlinear function are depicted in Fig. 3. One may observe the influence of taking snapshots of \mathbf{f} during the iterative Newton-Raphson procedure where the internal force vector is non-zero.

4.1 Robustness considerations

In this section the aforementioned different model reduction approaches, i.e. DEIM, Gappy-POD and GNAT are tested for various dimensionalities of $\mathbf{U}_{(\mathbf{u})}^{\mathbf{r}}$ and $\mathbf{U}_{(\mathbf{f})}^{\mathbf{r}}$ using the test set

$$\mathcal{D}_{\text{test}} = \mathbf{1} + \Delta \mathbf{F}^{\mathbf{M}}$$

with $\left[\Delta \mathbf{F}^{\mathbf{M}} \right]_{ij} \in \{-0.1986, 0.1886, -0.1321, 0.1461, -0.0521, 0.0921\}$,

(4.2)

which is different to the set $\mathcal{D}_{\text{train}}$. Therefore each configuration is tested against 1296 testcases. In case that any of said test cases does not converge using the reduced model the configuration is highlighted with an “x” in the subsequent result plots. Otherwise the color indicates the relative error of the fluctuation field in the L_2 norm

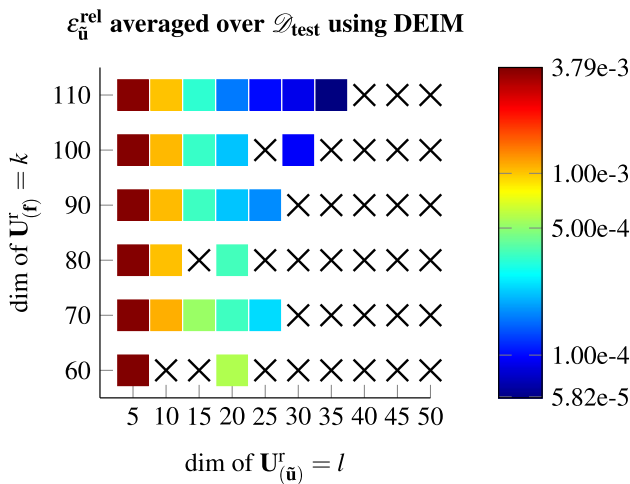


Fig. 4 Using DEIM to compute $\mathcal{D}_{\text{test}}$ for various dimensions of the reduced bases. A “x” denotes the case were at least one test case within $\mathcal{D}_{\text{test}}$ did not converge

$$\epsilon_{\tilde{\mathbf{u}}}^{\text{rel}} = \frac{\|\tilde{\mathbf{u}}_{\text{FEM}} - \tilde{\mathbf{u}}_{\text{ROM}}\|_2}{\|\tilde{\mathbf{u}}_{\text{FEM}}\|_2},$$

averaged over the 1296 testcases. Furthermore the following measures are introduced:

$$\epsilon_{\mathbf{P}^{\text{M}}}^{\text{rel}} = \frac{\|\mathbf{P}^{\text{M}}_{\text{FEM}} - \mathbf{P}^{\text{M}}_{\text{ROM}}\|_{\text{F}}}{\|\mathbf{P}^{\text{M}}_{\text{FEM}}\|_{\text{F}}}, \quad \epsilon_{\mathbb{L}^{\text{M}}}^{\text{rel}} = \frac{\|\mathbb{L}^{\text{M}}_{\text{FEM}} - \mathbb{L}^{\text{M}}_{\text{ROM}}\|_{\text{F}}}{\|\mathbb{L}^{\text{M}}_{\text{FEM}}\|_{\text{F}}}$$

Figure 4 shows the results for the DEIM approach as described in Sect. 3.

One may observe that for various configurations the method rendered scenarios where at least one test case did not converge. In the multi-query context this clearly is an undesirable result as the model should return the quantities of interest for arbitrary input parameters. Considering the errors, they behave as expected and decrease for an increasing number of basis vector $\mathbf{U}^r(\tilde{\mathbf{u}})$. The influence of the dimension of $\mathbf{U}^r(\mathbf{f})$ appears to be rather small after a certain threshold, see also Figs. 8 and 9.

In [9] it has been shown that using a Gappy-POD instead of pure interpolation benefits the condition number of the reduced tangent and should lead to a more robust model. This has also been observed within this work as shown in Fig. 5. While rendering a more robust model with respect to changes of the input parameters, as well as more accurate results, there were still configurations that lead to diverging test cases.

This might be due to a possible loss of symmetry through the application of a hyper-reduction method and the subsequent non-optimal Galerkin projection. To illustrate the asymmetry of $\tilde{\mathbf{K}}$ a comparison of the sparsity patterns for the given example from Fig. 2 is shown in Fig. 6.

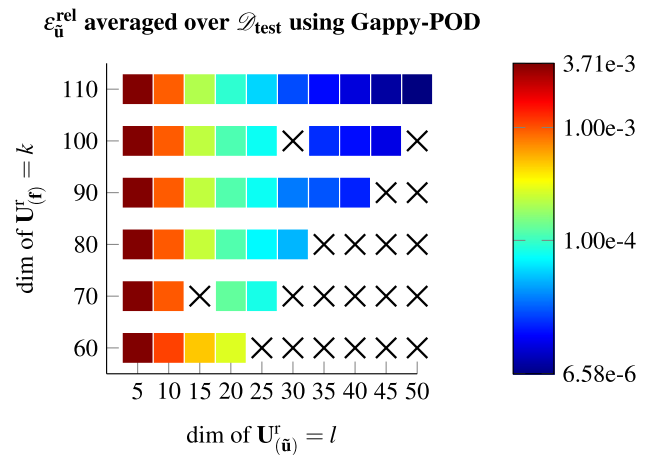


Fig. 5 Using Gappy-POD with Galerkin projection to compute $\mathcal{D}_{\text{test}}$ for various dimensions of the reduced bases with $k_s/k = 2$. A “x” denotes the case were at least one test case within $\mathcal{D}_{\text{test}}$ did not converge

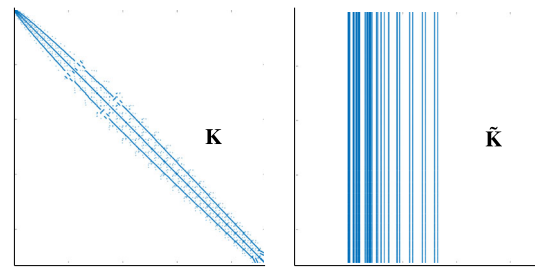


Fig. 6 Sparsity Pattern of \mathbf{K} and $\tilde{\mathbf{K}}$ (with $k = 10$ and rank $(\tilde{\mathbf{K}}) = 10$)

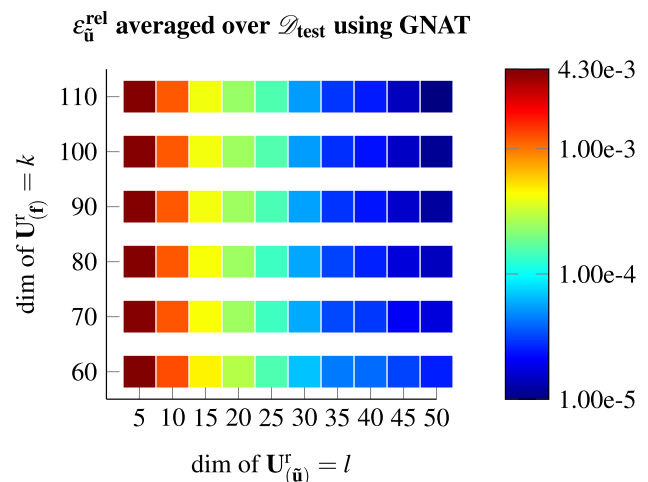


Fig. 7 Using GNAT to compute $\mathcal{D}_{\text{test}}$ for various dimensions of the reduced bases with $k_s/k = 2$. A “x” denotes the case were at least one test case within $\mathcal{D}_{\text{test}}$ did not converge

Employing therefore the GNAT, suited for non-symmetric tangent matrices, lead to the most robust model within the scope of this study as shown in Fig. 7.

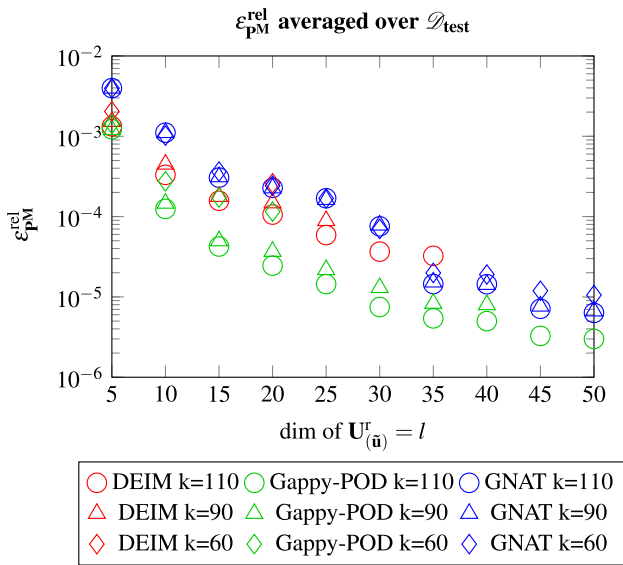


Fig. 8 The relative error of the averaged Piola stress tensor computed by the FEM and the reduced model for an increasing number of l modes of $\mathbf{U}_{(\hat{\mathbf{u}})}^r$. A missing marker denotes a ROM with at least one diverged test case of \mathcal{D}_{test} (see Figs. 4–7)

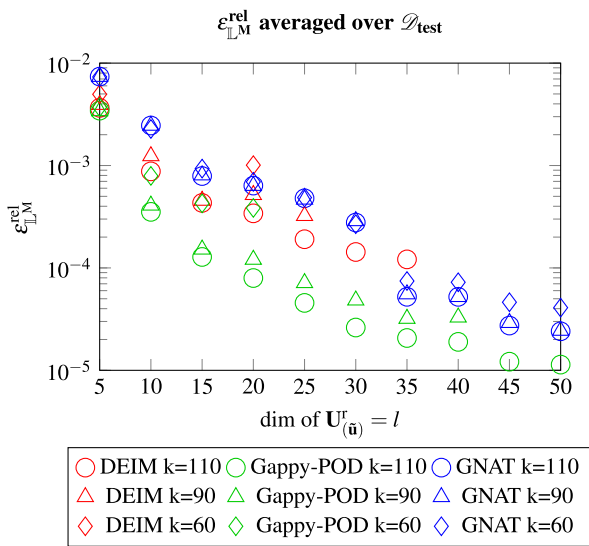


Fig. 9 The relative error of the macroscopic tangent computed by the FEM and the reduced model for an increasing number of l modes of $\mathbf{U}_{(\hat{\mathbf{u}})}^r$. A missing marker denotes a ROM with at least one diverged test case of \mathcal{D}_{test} (see Figs. 4–7)

One can observe that no configuration lead to a diverging test case and the method appears to be very robust. Furthermore, equivalently for DEIM and Gappy-POD, increasing the dimension of $\mathbf{U}_{(\hat{\mathbf{u}})}^r$ decreases the error, while a change of the dimensionality of $\mathbf{U}_{(\hat{\mathbf{f}})}^r$ has only minor effect on the accuracy above a certain threshold. This gets confirmed considering Figs. 8 and 9 which depict the errors for the averaged stresses and the macroscopic tangent.

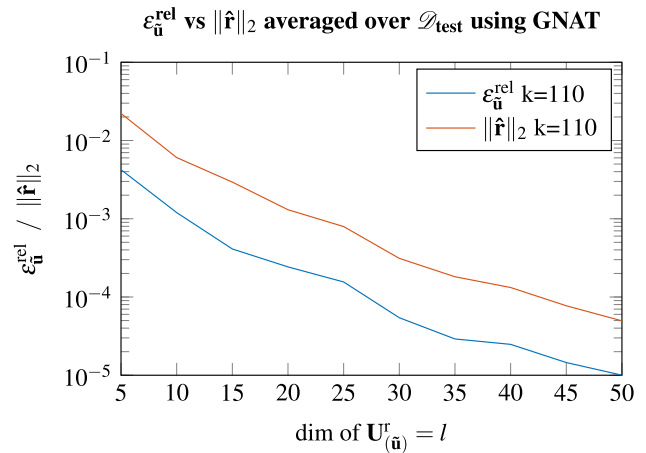


Fig. 10 The reduced residual ($\hat{\mathbf{r}} = -\mathbf{X} \cdot \mathcal{P}^T \cdot \mathbf{f}$) from GNAT versus the relative error of the fluctuation field $\epsilon_{\hat{\mathbf{u}}}^{rel}$ with $k_s/k = 2$; for every l the error is averaged over all test cases within \mathcal{D}_{test}

Employing Gappy-POD with Galerkin projection rendered the most accurate results while the GNAT rendered the most robust model and showed no convergence difficulties. Though the Gappy-POD is more robust compared to the DEIM, there is no guarantee that a different \mathcal{D}_{test} would yield the same behaviour. It should be noted that the reduced model has been used for the perturbation procedure from Sect. 2.2 to obtain the macroscopic tangent components while the stresses were averaged over the full domain using the solution of the reduced model $\mathbf{U}_{(\hat{\mathbf{u}})}^r \cdot \hat{\mathbf{u}}$.

A further benefit of GNAT is that it minimizes the global full order residual. Therefore using the reduced residual $\hat{\mathbf{r}} = -\mathbf{X} \cdot \mathcal{P}^T \cdot \mathbf{f}$ from Eq. (3.23) may serve as an error indicator, considering the relation to the relative error as depicted in Fig. 10. The reduced residual as well as the relative error decrease with an increasing size of the reduced basis and match quite well up to a certain offset. This may be beneficial considering greedy selection algorithms to build the ROM, which rely on the ability to estimate or at least indicate the error of the reduced model approximation.

The relation between the errors $\epsilon_{\hat{\mathbf{u}}}^{rel}$ and the residual is further investigated for a given ROM and all test cases of the parameter set \mathcal{D}_{test} in Fig. 11. Based on the previous results the ROM dimensions are set to $l = 45, k = 90$ and $k_s/k = 2$ using GNAT. It can be seen that the relative error has minor fluctuations but does not show considerable deviations from its mean. The same holds for the norm of the reduced residual from the GNAT model. Hence, the norm of the reduced residual appears to be a computational cheap and reliable error indicator which could be used in future studies in conjunction with an adaptive sampling algorithm, e.g. [32], instead of using \mathcal{D}_{train} to build the reduced model.

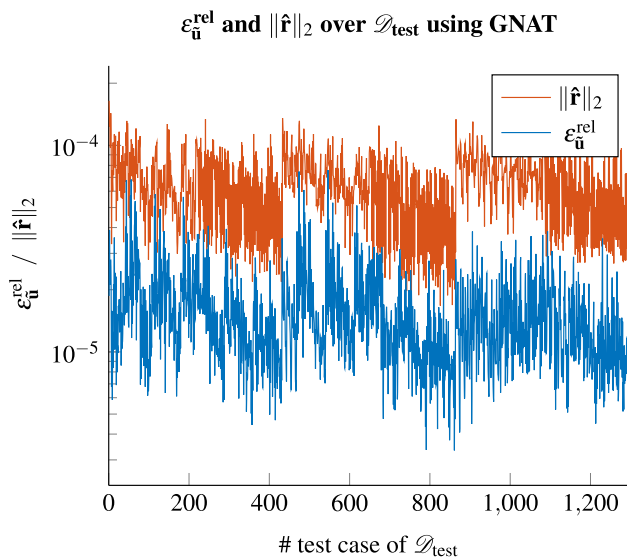


Fig. 11 The norm of the reduced residual compared to the relative error of the fluctuation field for all test cases within \mathcal{D}_{test} using GNAT with $l = 45, k = 90$ and $k_s/k = 2$

4.2 Local fields

Besides acting as a simple input-output model it is possible in computational homogenisation to investigate local fields. Considering a test case with

$$\mathbf{F}^M = \begin{pmatrix} -1.1986 & -0.1986 \\ -0.1986 & -1.1986 \end{pmatrix} \quad (4.3)$$

the Piola stress components computed by the FEM and the reduced model are compared in Fig. 12 for the GNAT with $l = 45, k = 90$ and $k_s/k = 2$. Hence, instead of solving for 2.312 unknowns only 45 unknowns have to be determined reducing the computational cost considerably. It can be seen that the xx -component of \mathbf{P}^m computed by the FEM model in Fig. 12a and the GNAT in Fig. 12b match quite well which is supported considering the relative errors in percent

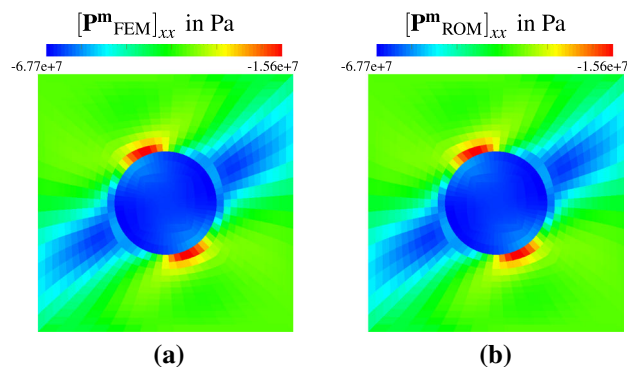


Fig. 12 In **a** and **b** the xx -components of the Piola stress tensor (computed using the full order model and the ROM) are shown

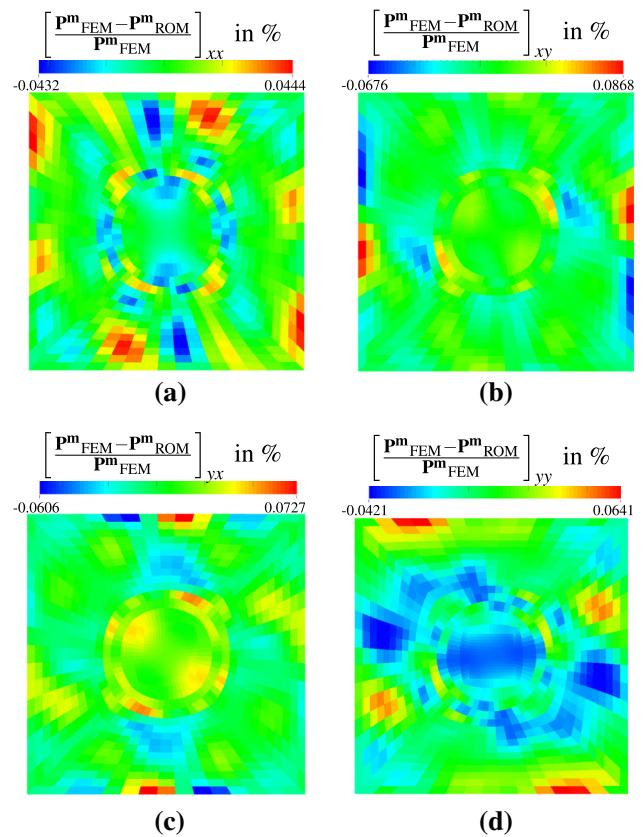


Fig. 13 Figures **a–d** show the relative error of the individual stress components in percent

of the individual components of the stress tensor given in Fig. 13a–d. For the current ROM configuration the errors appear to be acceptable small and correlate with the observations in all the error plots throughout this section.

5 Conclusion

Within the scope of the present work reduced-order modelling techniques based on the Proper Orthogonal Decomposition and so-called hyper-reduction techniques have been applied in the context of computational homogenisation of hyperelastic materials. The focus has thereby been the accuracy and robustness of the reduced model, comparing different hyper-reduction and projection approaches. It has been shown that introducing an additional approximation of the nonlinear terms via an empirical interpolation or Gappy-POD may not preserve the symmetry of the system tangent. This leads to a non-optimal Galerkin projection. As shown in the numerical examples this can cause convergence problems. This is clearly an undesirable property in the multi-query context as given in computational homogenisation. The Gauss-Newton like approach (GNAT), relying on a Petrov-Galerkin projection suited for non-symmetric tan-

gent matrices, rendered the most robust model and showed no convergence problems while minimizing the global residual.

Future studies should aim towards more effective methods to obtain error estimates for the outputs of interest and averaging procedures. Furthermore, alternative approximations of the system tangent, e.g. [5, 29, 38], could be investigated. Yet some of these methods require additional high dimensional snapshots of the sparse system tangent which becomes computational infeasible rapidly. For instance considering only the non-zero elements of the tangent from the presented examples in Sect. 4 renders an additional snapshot of the size 40.836 for the 2.312 unknowns in every iteration step.

Regarding the snapshot selection adaptive strategies, e.g. [32], could be employed to adaptively select the positions in parameter space for which the full model should be evaluated in order to build the ROM.

Acknowledgements The first and fifth author gratefully acknowledge the financial support by the German Science Foundation (DFG) within the Collaborative Research Center 814: Additive Manufacturing (sub-project C3). The second, third and fourth author gratefully acknowledge the financial support by the ERC Advanced Grant MOCOPLY.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Barrault M, Maday Y, Nguyen NC, Patera AT (2004) An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Math* 339(9):667–672
- Benner P, Gugercin S, Willcox K (2015) A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Rev* 57(4):483–531
- Boyaval S (2008) Reduced-basis approach for homogenization beyond the periodic setting. *Multiscale Model Simul* 7(1):466–494
- Brands B, Mergheim J, Steinmann P (2016) Reduced-order modelling for linear heat conduction with parametrised moving heat sources. *GAMM-Mitt* 39(2):170–188
- Carlberg K, Bou-Mosleh C, Farhat C (2011) Efficient nonlinear model reduction via a least-squares Petrov-Galerkin projection and compressive tensor approximations. *Int J Numer Methods Eng* 86(2):155–181
- Carlberg K, Farhat C, Cortial J, Amsallem D (2013) The GNAT method for nonlinear model reduction: effective implementation and application to computational fluid dynamics and turbulent flows. *J Comput Phys* 242:623–647
- Chaturantabut S, Sorensen DC (2010) Nonlinear model reduction via discrete empirical interpolation. *SIAM J Sci Comput* 32(5):2737–2764
- Chinesta F, Ammar A, Leygue A, Keunings R (2011) An overview of the proper generalized decomposition with applications in computational rheology. *J Non-Newton Fluid Mech* 166(11):578–592
- Cosimo A, Cardona A, Idelsohn S (2014) Improving the k-compressibility of hyper reduced order models with moving sources: applications to welding and phase change problems. *Comput Methods Appl Mech Eng* 274:237–263
- Cremonesi M, Néron D, Guidault PA, Ladevèze P (2013) A PGD-based homogenization technique for the resolution of nonlinear multiscale problems. *Comput Methods Appl Mech Eng* 267:275–292
- Dvorak GJ (1992) Transformation field analysis of inelastic composite materials. In: *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, vol. 437. The Royal Society, pp. 311–327
- Eshelby J (1957) The determination of the elastic field of an ellipsoidal inclusion, and related problems. In: *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, vol. 241. The Royal Society, pp. 376–396
- Everson R, Sirovich L (1995) Karhunen-loeve procedure for gappy data. *JOSA A* 12(8):1657–1664
- Fritzen F, Hodapp M (2016) The finite element square reduced (FE^{2R}) method with GPU acceleration: towards three-dimensional two-scale simulations. *Int J Numer Meth Eng* 107:853–881. doi:10.1002/nme.5188
- Fritzen F, Leuschner M (2013) Reduced basis hybrid computational homogenization based on a mixed incremental formulation. *Comput Methods Appl Mech Eng* 260:143–154
- Geers M, Kouznetsova V, Brekelmans W (2010) Multi-scale computational homogenization: trends and challenges. *J Comput Appl Math* 234(7):2175–2182
- Gunzburger MD, Peterson JS, Shadid JN (2007) Reduced-order modeling of time-dependent PDEs with multiple parameters in the boundary data. *Comput Methods Appl Mech Eng* 196(4):1030–1047
- Hernández J, Oliver J, Huespe AE, Caicedo M, Cante J (2014) High-performance model reduction techniques in computational multiscale homogenization. *Comput Methods Appl Mech Eng* 276:149–189
- Hill R (1963) Elastic properties of reinforced solids: some theoretical principles. *J Mech Phys Solids* 11(5):357–372
- Hill R (1972) On constitutive macro-variables for heterogeneous solids at finite strain. In: *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, vol. 326. The Royal Society, pp. 131–147
- Hill R (1984) On macroscopic effects of heterogeneity in elastoplastic media at finite strain. In: *Mathematical proceedings of the Cambridge philosophical society*, vol. 95. Cambridge University Press, pp. 481–494
- Hinze M, Kunkel M (2012) Discrete empirical interpolation in pod model order reduction of drift-diffusion equations in electrical networks. In: *Scientific computing in electrical engineering SCEE 2010*. Springer, pp. 423–431
- Kerfriden P, Goury O, Rabczuk T, Bordas SPA (2013) A partitioned model order reduction approach to rationalise computational expenses in nonlinear fracture mechanics. *Comput Methods Appl Mech Eng* 256:169–188
- Michel JC, Suquet P (2003) Nonuniform transformation field analysis. *Int J Solids Struct* 40(25):6937–6955
- Michel JC, Suquet P (2016) A model-reduction approach in micromechanics of materials preserving the variational structure of constitutive relations. *J Mech Phys Solids* 90:254–285
- Monteiro E, Yvonnet J, He Q (2008) Computational homogenization for nonlinear conduction in heterogeneous materials using model reduction. *Comput Mater Sci* 42(4):704–712
- Mori T, Tanaka K (1973) Average stress in matrix and average elastic energy of materials with misfitting inclusions. *Acta Metall* 21(5):571–574
- Moulinec H, Suquet P (1998) A numerical method for computing the overall response of nonlinear composites with complex

- microstructure. *Comput Methods Appl Mech Eng* 157(1–2):69–94
29. Negri F, Manzoni A, Amsallem D (2015) Efficient model reduction of parametrized systems by matrix discrete empirical interpolation. *J Comput Phys* 303:431–454
 30. Nguyen NC (2008) A multiscale reduced-basis method for parametrized elliptic partial differential equations with multiple scales. *J Comput Phys* 227(23):9807–9822
 31. Noor AK, Peters JM (1980) Reduced basis technique for nonlinear analysis of structures. *Aiaa J* 18(4):455–462
 32. Paul-Dubois-Taine A, Amsallem D (2015) An adaptive and efficient greedy procedure for the optimal training of parametric reduced-order models. *Int J Numer Methods Eng* 102(5):1262–1292
 33. Quarteroni A, Manzoni A, Negri F (2015) *Reduced basis methods for partial differential equations: an introduction*, vol 92. Springer, Berlin
 34. Radermacher A, Reese S (2015) POD-based model reduction with empirical interpolation applied to nonlinear elasticity. *Int J Numer Meth Eng* 107:477–495. doi:10.1002/nme.5177
 35. Saad Y (2003) *Iterative methods for sparse linear systems*. SIAM, Philadelphia
 36. Sirovich L (1987) Turbulence and the dynamics of coherent structures. part i: coherent structures. *Q Appl Math* 45(3):561–571
 37. Spahn J, Andrä H, Kabel M, Müller R (2014) A multiscale approach for modeling progressive damage of composite materials using fast fourier transforms. *Comput Methods Appl Mech Eng* 268:871–883
 38. Ștefănescu R, Sandu A (2017) Efficient approximation of sparse jacobians for time-implicit reduced order models. *Int J Numer Methods Fluids* 83(2):175–204
 39. Temizer I, Wriggers P (2008) On the computation of the macroscopic tangent for multiscale volumetric homogenization problems. *Comput Methods Appl Mech Eng* 198(3):495–510
 40. Volkwein S (2011) *Model reduction using proper orthogonal decomposition*. Lecture Notes, Institute of mathematics and scientific computing, University of Graz. see <http://www.uni-graz.at/imawww/volkwein/POD.pdf>
 41. Yvonnet J, He QC (2007) The reduced model multiscale method (R3M) for the non-linear homogenization of hyperelastic media at finite strains. *J Comput Phys* 223(1):341–368