Received: 19 December 2023 | Revised: 21 May 2024 | Accepted: 24 June 2024

DOI: 10.1002/nme.7567

RESEARCH ARTICLE

WILEY

Variational formulation and monolithic solution of computational homogenization methods

Christian Hesch | **Felix Schmidt** | **Stefan Schuß**

Chair of Computational Mechanics, University of Siegen, Germany

Correspondence
Christian Hesch, Chair of Computational Mechanics, University of Siegen, Germany
Email: christian.hesch@uni-siegen.de

Funding information
Deutsche Forschungsgemeinschaft (DFG), Grant/Award Numbers: HE5943/24-1, HE5943/26-1

Abstract

In this contribution, we derive a consistent variational formulation for computational homogenization methods and show that traditional FE^2 and IGA^2 approaches are special discretization and solution techniques of this most general framework. This allows us to enhance dramatically the numerical analysis as well as the solution of the arising algebraic system. In particular, we expand the dimension of the continuous system, discretize the higher dimensional problem consistently and apply afterwards a discrete null-space matrix to remove the additional dimensions. A benchmark problem, for which we can obtain an analytical solution, demonstrates the superiority of the chosen approach aiming to reduce the immense computational costs of traditional FE^2 and IGA^2 formulations to a fraction of the original requirements. Finally, we demonstrate a further reduction of the computational costs for the solution of general nonlinear problems.

KEYWORDS

computational homogenization, FE^2, IGA^2, linearization, multiscale, null-space, representative volume element, variational

1 | INTRODUCTION

Complex materials with dedicated morphological features like fibers or additive manufactured microstructures are nowadays used throughout the industry. To obtain a suitable model, continuum mechanical or thermomechanical formulations are often used. Using infinitesimal defined strain measures, we can always resolve finite inhomogenities, however, the computational costs of such a large scale approximative solution may economically not affordable or in certain cases technically not possible. Additionally, a material may appear perfectly homogeneous on the considered scale of approximations, whereas the microstructural information has been lost and often we are only interested in the macroscale solution. Therefore, two different approaches are common to include this information. First, the microstructure can be taken into account in a phenomenological way. The resulting constitutive relations can also incorporate size effects, leading to generalized theories for materials, see References 1–3 for a general overview of gradient extended continua. In contrast, microstructural information about morphology and material properties can be accounted for in a more explicit manner by means of homogenization methods, assuming that the scales of the finite approximation and the morphological features on the microscale are clearly separated.

In the context of first-order homogenization schemes, we refer to References 4–8 for fundamental analytical approaches and to References 9–12 for seminal contributions to two-scale finite-element (FE) simulations. For computational multiscale techniques using the well-known FE^2-methods we refer to the review article in Reference 13 and to Reference 14 for a general overview of computational multiscale techniques. The application to solids at finite strains including physically coupled problems in thermo-elasticity are presented in References 15 and 16 and for electro-elasticity in Reference 17. The general extension on computational homogenization of strain gradient materials using Isogeometric Analysis, labeled IGA^2-method, is presented in our preliminary work in Reference 18. We remark, that there exists a wide range of different homogenization methods like the "variational multiscale method", see References 19 and 20, and derivatives thereof like the "variational consistent homogenization" method as proposed in Reference 21. In this work, we refer to the terminus "variational" in the sense of a Galerkin type method, applied on a higher dimensional space for the microscale fluctuations.

Traditional methods like FE^2 and IGA^2 consider a representative volume element (RVE) at every Gauss point. Since the primary idea was to utilize the RVE and the corresponding set of partial differential equations representing the static equilibrium of some kind of a microstructure in exchange of a constitutive law. However, we already know from concepts like the Hu-Washizu functional (see Zienkiewicz et al.[22,23]) that the relation between the strain energy function and the stress tensor can be considered as an independent equation in a weak form to be solved using again a finite element framework. To this end, the stress tensor is not evaluated at every Gauss point but rather interpolated using the applied shape functions for the solution of the weak form, see Bonet et al.[24] for details on the implementation of Hu-Washizu and Hellinger-Reissner type formulations.

In this work, we propose a dimensional expansion of the macroscopic system for the calculations within the RVE. The arising microscopical fluctuations are interpolated in the expanded, higher-dimensional framework using continuous shape functions on the microscale and discontinuous shape functions on the macroscale to allow again for a static condensation procedure, written here in the form of a null-space reduction scheme. With this monolithic framework, we can reconstruct traditional methods like FE^2 and IGA^2 by using Delta Dirac on the Gauss points and the application of a staggered scheme between the dimensions. However, based on the variational formulation we can now make use of sophisticated methods for the solution of the arising large scale algebraic system of equations. As already shown in Lange et al.,[25] monolithic solution techniques using static condensation procedures are by far superior to the original staggered scheme. Both, suitable choices of the macroscale shape functions and the application of sophisticated solution techniques lead to a dramatic reduction of the computational cost preserving the accuracy of the solution. Additionally, this most general approach allows for the application of highly efficient methods like recursive trust-region multigrid formulations, see, among others, Gross and Krause.[26]

The manuscript is organized as follows. First, the governing equations are presented in Section 2 including the variational form of the multiscale problem. The discretization and the arising algebraic system of the multiscale boundary value problem are shown in Section 3, followed by the different solution procedures in Section 4. Representative examples are given in Section 5 and conclusions are drawn in Section 6.

2 | GOVERNING EQUATIONS

We start with a short summary of the governing equations for the macro- and afterwards for the microscale. At the end of this section, we will introduce a common variational formulation for both scales. With regard to notation, Einstein's summation convention is used in the following for clearer presentation.

2.1 | Macro-continuum

Consider a Lipschitz bounded continuum body in its reference configuration $\mathcal{B}_0 \subset \mathbb{R}^n$ with $n \in \{1, 2, 3\}$, undergoing a motion characterized by a deformation mapping $\boldsymbol{\varphi} : \mathcal{B}_0 \to \mathbb{R}^n, \boldsymbol{X} \mapsto \boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X})$, which maps material points $\boldsymbol{X} \in \mathcal{B}_0$ of the reference configuration $\mathcal{B}_0$ onto points $\boldsymbol{x} \in \mathcal{B}_t$ of the current configuration $\mathcal{B}_t = \boldsymbol{\varphi}(\mathcal{B}_0)$. The material deformation gradient is defined by $\boldsymbol{F} : \mathcal{B}_0 \to \mathbb{R}^{n\times n}, \boldsymbol{F} = \nabla_{\boldsymbol{X}}(\boldsymbol{x})$, which maps infinitesimal vectors $\mathrm{d}\boldsymbol{X}$ at the material point $\boldsymbol{X}$ to the infinitesimal vector $\mathrm{d}\boldsymbol{x}$ at $\boldsymbol{x}$ in the deformed configuration. An infinitesimal area element, oriented at the material point $\boldsymbol{X}$ with the outward, reference normal vector $\boldsymbol{N}$ can be defined by two linearly independent vectors via $\mathrm{d}\boldsymbol{X}^{(1)} \times \mathrm{d}\boldsymbol{X}^{(2)} = \boldsymbol{N}\mathrm{d}A$. The well-known Nanson's formula reads $\boldsymbol{n}\mathrm{d}a = \mathrm{cof}(\boldsymbol{F})\boldsymbol{N}\mathrm{d}A$, where $\mathrm{cof}(\boldsymbol{F}) = \boldsymbol{H}$ is the co-factor, given by

$$\boldsymbol{H} := \frac{1}{2}\left(\boldsymbol{F} \overset{\times}{\times} \boldsymbol{F}\right). \tag{1}$$

Here, we apply the cross product of two second-order tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ with $[\boldsymbol{A} \overset{\times}{\times} \boldsymbol{B}]_{iJ} = \epsilon_{ijk}\epsilon_{JMN}[\boldsymbol{A}]_{jM}[\boldsymbol{B}]_{kN}$, using the third-order Levi-Civita permutation symbol ϵ_{ijk}. The infinitesimal volume element $\mathrm{d}V$ in the material configuration is related to the deformed counterpart via the Jacobian $J := \det[\boldsymbol{F}] > 0$, where the determinant is defined by

$$\det(\boldsymbol{F}) := \frac{1}{6}\boldsymbol{F} : \left(\boldsymbol{F} \overset{\times}{\times} \boldsymbol{F}\right). \tag{2}$$

A graphical representation of the different kinematical values is given in Figure 1, see Reference 27 for further details.

Introducing the space of virtual or admissible test functions for the deformation

$$\mathcal{V}^{\delta\boldsymbol{\varphi}} = \{\delta\boldsymbol{\varphi} \in H^1(\mathcal{B}_0) \mid \delta\boldsymbol{\varphi} = \boldsymbol{0} \text{ on } \Gamma^{\varphi}\}, \tag{3}$$

where $\Gamma^{\varphi} \subset \partial\mathcal{B}_0$ refers to the Dirichlet boundary, the internal virtual work reads

$$\delta\Pi^{\mathrm{int},\varphi} = \int_{\mathcal{B}_0} \boldsymbol{P} : \nabla_{\boldsymbol{X}}(\delta\boldsymbol{\varphi})\, \mathrm{d}V, \tag{4}$$

where $\boldsymbol{P}$ denotes the first Piola–Kirchhoff stress tensor. Applying integration by parts and equating the resulting terms with external contributions yields

$$\begin{aligned} \nabla_{\boldsymbol{X}} \cdot \boldsymbol{P}^{\mathrm{T}} + \boldsymbol{B} &= 0, \\ \boldsymbol{P}\boldsymbol{N} - \boldsymbol{T} &= 0, \\ \boldsymbol{\varphi} - \overline{\boldsymbol{\varphi}} &= 0, \end{aligned} \tag{5}$$

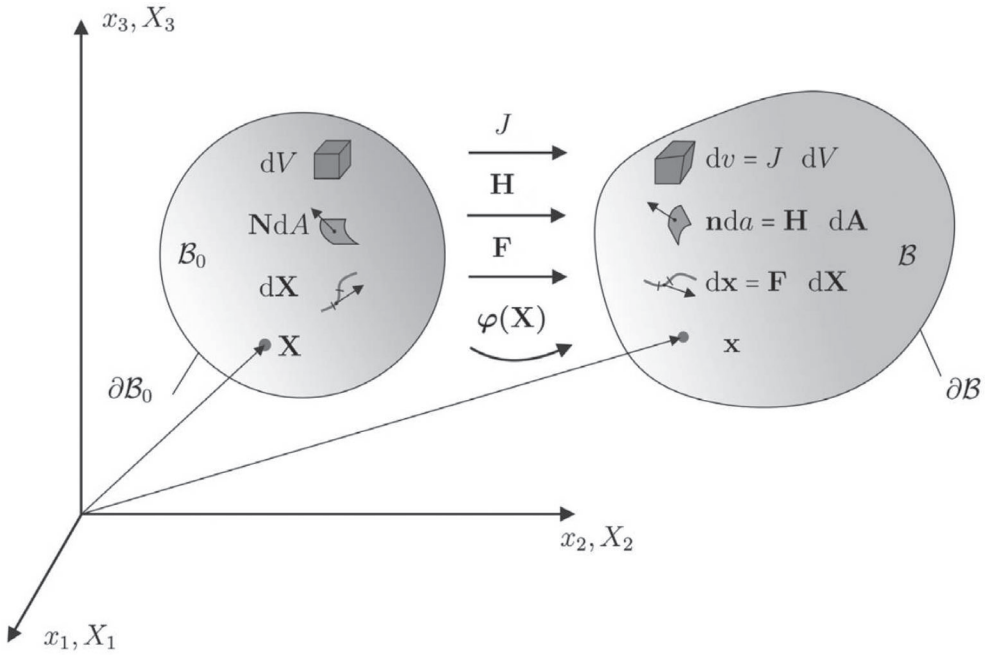

FIGURE 1 Graphical representation kinematical values $\boldsymbol{\varphi}$, $\boldsymbol{F}$, $\boldsymbol{H}$ and J, taken from Reference 27.

where $\boldsymbol{B}$ is a body load, $\boldsymbol{T}$ the Neumann surface load on $\Gamma^\sigma \subset \partial\mathcal{B}_0$ and $\overline{\boldsymbol{\varphi}}$ the predefined deformation on the Dirichlet boundary Γ^φ. Note that $\Gamma^\varphi \cap \Gamma^\sigma = \emptyset$ and $\overline{\Gamma}^\varphi \cup \overline{\Gamma}^\sigma = \partial\mathcal{B}_0$.

Without loss of generality of the chosen approach, we restrict ourselves here to elastic materials, assuming that ductile and general inelastic models can be applied as well. To describe elastic materials in the large strain regime, the concept of polyconvexity as a mathematically well-accepted requirement must be satisfied by the strain energy density function, see Reference 24 for details. In particular, we require the strain energy density function Ψ to be a convex multi-valued and inhomogeneous function given by

$$\Psi := \Psi(\boldsymbol{F}) = \overline{\Psi}(\boldsymbol{F}, \boldsymbol{H}, J, \boldsymbol{X}), \tag{6}$$

where $\overline{\Psi}$ is convex with respect to the $2n^2 + 1$ values of $\boldsymbol{F}$, $\boldsymbol{H}$ and J. For isotropic materials such as a compressible Moonley-Rivlin material we can assume the constitutive relation to take the form

$$\overline{\Psi}_{MR} := \alpha(\boldsymbol{X})\boldsymbol{F} : \boldsymbol{F} + \beta(\boldsymbol{X})\boldsymbol{H} : \boldsymbol{H} + f(J, \boldsymbol{X}), \tag{7}$$

where α and β are positive material parameters locally defined at $\boldsymbol{X}$ and f is a convex function of J and $\boldsymbol{X}$. The first Piola–Kirchhoff stress tensor follows immediately from

$$\boldsymbol{P} := \frac{\partial\Psi}{\partial\boldsymbol{F}} = \frac{\partial\overline{\Psi}}{\partial\boldsymbol{F}} + \frac{\partial\overline{\Psi}}{\partial\boldsymbol{H}} \overset{\times}{\underset{\times}{}} \boldsymbol{F} + \frac{\partial\overline{\Psi}}{\partial J}\boldsymbol{H}. \tag{8}$$

With regard to (5), we obtain for the virtual work of the internal and external contributions

$$G^{\boldsymbol{\varphi}} := \int_{\mathcal{B}_0} \boldsymbol{P} : \nabla_{\boldsymbol{X}}(\delta\boldsymbol{\varphi})\, \mathrm{d}V - \int_{\mathcal{B}_0} \delta\boldsymbol{\varphi} \cdot \boldsymbol{B}\, \mathrm{d}V - \int_{\Gamma^\sigma} \delta\boldsymbol{\varphi} \cdot \boldsymbol{T}\, \mathrm{d}A, \tag{9}$$

and require $G^{\boldsymbol{\varphi}} = 0$ for all $\delta\boldsymbol{\varphi} \in \mathcal{V}^{\delta\boldsymbol{\varphi}}$.

2.2 | Micro-continuum

To account for the influence of the underlying microstructure on the material behavior, we introduce a representative elementary volume (RVE) $\Omega_0 \subseteq \mathbb{R}^n$, link it to each point $\boldsymbol{X} \in \mathcal{B}_0$ via a first-order Taylor approximation*

$$\tilde{\boldsymbol{\varphi}}(\boldsymbol{X}, \tilde{\boldsymbol{X}}) = \boldsymbol{F}(\boldsymbol{X})\tilde{\boldsymbol{X}} + \tilde{\boldsymbol{w}}(\boldsymbol{X}, \tilde{\boldsymbol{X}}), \quad \boldsymbol{X} \in \mathcal{B}_0,\ \tilde{\boldsymbol{X}} \in \Omega_0, \tag{10}$$

and interpret $\tilde{\boldsymbol{\varphi}}(\boldsymbol{X}, \cdot) : \Omega_0 \to \mathbb{R}^n$ as a deformation map on Ω_0. The associated quantities like the corresponding deformation gradient are derived analogous to the previous section. To distinguish between macro- and micro-entities, all quantities of the micro-continuum will be marked with a superimposed tilde.

Per definition, the micro deformation consists of a homogeneous part $\boldsymbol{F}(\boldsymbol{X})\tilde{\boldsymbol{X}}$ and a nonhomogeneous field $\tilde{\boldsymbol{w}} : \mathcal{B}_0 \times \Omega_0 \to \mathbb{R}^n$ referred to as microscopic fluctuations. We point out, that this is the most important definition, as the fluctuations $\tilde{\boldsymbol{w}}$ are defined in the higher dimensional space $\mathcal{B}_0 \times \Omega_0$. This has dramatic consequences for the approximation, as we will show subsequently for the spatial discretization and in the example section.

According to Eq. (10), we obtain

$$\tilde{\boldsymbol{F}} = \nabla_{\tilde{\boldsymbol{X}}}(\tilde{\boldsymbol{\varphi}}) = \boldsymbol{F}(\boldsymbol{X}) + \tilde{\boldsymbol{F}}', \quad \tilde{\boldsymbol{F}}' := \nabla_{\tilde{\boldsymbol{X}}}(\tilde{\boldsymbol{w}}), \tag{11}$$

so that for homogeneous materials with zero fluctuations, $\boldsymbol{F}(\boldsymbol{X})$ is recovered for the micro deformation gradient.

The material of Ω_0 is described by a polyconvex strain energy density function $\tilde{\Psi}$ and we assume that the macroscopic and microscopic gradients, $\boldsymbol{F}$ and $\tilde{\boldsymbol{F}}$, as well as the elastic potentials, Ψ and $\tilde{\Psi}$, are related via the volume averages

$$\boldsymbol{F} = \frac{1}{|\Omega_0|}\int_{\Omega_0} \tilde{\boldsymbol{F}}\, \mathrm{d}\tilde{V} \quad \text{and} \quad \Psi = \frac{1}{|\Omega_0|}\int_{\Omega_0} \tilde{\Psi}\, \mathrm{d}\tilde{V}, \tag{12}$$

where $|\Omega_0|$ is the volume of Ω_0. In addition, insertion of (11) into $(12)_1$ yields the constraint

$$\boldsymbol{F} = \boldsymbol{F} + \frac{1}{|\Omega_0|}\int_{\Omega_0} \nabla_{\tilde{\boldsymbol{X}}}(\tilde{\boldsymbol{w}})\, \mathrm{d}\tilde{V} \;\Leftrightarrow\; \int_{\Omega_0} \nabla_{\tilde{\boldsymbol{X}}}(\tilde{\boldsymbol{w}})\, \mathrm{d}\tilde{V} = \boldsymbol{0} \tag{13}$$

for the superimposed deformation field $\tilde{\boldsymbol{w}}$. Applying Gauss divergence theorem, the constraint (13) can be alternatively expressed as

$$\int_{\partial\Omega_0} \tilde{\boldsymbol{w}} \otimes \tilde{\boldsymbol{N}}\, \mathrm{d}\, \tilde{A} = \boldsymbol{0}, \tag{14}$$

showing that (13) is satisfied by the alternative conditions

$$(i)\ \tilde{\boldsymbol{w}} = \boldsymbol{0} \quad \text{in } \Omega_0, \qquad (ii)\ \tilde{\boldsymbol{w}} = \boldsymbol{0} \quad \text{on } \partial\Omega_0, \qquad (iii)\ \tilde{\boldsymbol{w}}^{+} = \boldsymbol{w}^{-} \quad \text{on } \partial\Omega_0, \tag{15}$$

cf. Miehe .[10] The trivial condition (i) enforces a homogeneous deformation of the entire domain and represents the basic assumption of a Voight-type homogenization (see Taylor[28]). Condition (ii) demands homogeneous deformations on the boundary and (iii) addresses periodic boundary conditions. Assuming a static equilibrium state of the micro-continuum governed by the field equation

$$\nabla_{\tilde{\boldsymbol{X}}} \cdot \tilde{\boldsymbol{P}}^{\mathrm{T}} = \boldsymbol{0}, \quad \text{in} \quad \Omega_0, \tag{16}$$

with Equation $(12)_2$ we obtain a relationship between macro- and micro-stress as follows

$$\boldsymbol{P} = \frac{\partial \Psi}{\partial \boldsymbol{F}} = \frac{1}{|\Omega_0|}\int_{\Omega_0} \frac{\partial \tilde{\Psi}}{\partial \boldsymbol{F}}\, \mathrm{d}\tilde{V} = \frac{1}{|\Omega_0|}\int_{\Omega_0} \frac{\partial \tilde{\Psi}}{\partial \tilde{\boldsymbol{F}}}\frac{\partial \tilde{\boldsymbol{F}}}{\partial \boldsymbol{F}}\, \mathrm{d}\tilde{V} = \frac{1}{|\Omega_0|}\int_{\Omega_0} \tilde{\boldsymbol{P}}\, \mathrm{d}\tilde{V}, \tag{17}$$

where we have utilized in the last equation that according to (11), $\partial\tilde{\boldsymbol{F}}/\partial\boldsymbol{F}$ is the identity.

Remark 1. Each of the three conditions (15), together with (16), ensures that the macro-homogeneity condition

$$\frac{1}{|\Omega_0|}\int_{\Omega_0} \tilde{\boldsymbol{P}} : \delta\tilde{\boldsymbol{F}}\, \mathrm{d}\tilde{V} = \boldsymbol{P} : \delta\boldsymbol{F}, \tag{18}$$

also known as Hill–Mandel criterion, is fulfilled. The virtual work applied in a specific point $\boldsymbol{X} \in \mathcal{B}_0$ to the system thus corresponds to the volumetric average of the virtual work in Ω_0. In a different interpretation, the Hill–Mandel criterion along with the kinematic constraint in (13) justifies the use of the volume averages of the strain energy function $(12)_2$ and subsequently the volume averages of the micro stresses in (17). Hence, using the virtual work instead of the strain energy function in $(12)_2$ as basic assumption is more general, as it allows for nonelastic materials as well, see Miehe et al.[10] for the application on crystal plasticity.

2.3 | Variational formulation

Taking into account relations $(12)_2$ and (11) between the elastic potentials, Ψ and $\tilde{\Psi}$, and the gradients, $\boldsymbol{F}$, $\tilde{\boldsymbol{F}}$, and $\tilde{\boldsymbol{F}}'$, the macro-deformation and micro-fluctuation are determined as minimizer of the system's total energy, that is, $\boldsymbol{\varphi}$ and $\tilde{\boldsymbol{w}}$ have to fulfil the condition

$$(\boldsymbol{\varphi}, \tilde{\boldsymbol{w}}) = \underset{(\boldsymbol{u},\boldsymbol{v})\in\mathcal{V}^{\boldsymbol{\varphi}}\times\mathcal{V}^{\tilde{\boldsymbol{w}}}}{\arg\inf} \left\{\Pi^{\text{int}}(\boldsymbol{u},\boldsymbol{v}) - \Pi^{\text{ext}}(\boldsymbol{u})\right\}, \tag{19}$$

where $\mathcal{V}^{\boldsymbol{\varphi}}$ and $\mathcal{V}^{\tilde{\boldsymbol{w}}}$ are suitable spaces of admissible solutions. The internal and external energy contributions are given by

$$\Pi^{\text{int}}(\boldsymbol{u},\boldsymbol{v}) = \int_{\mathcal{B}_0} \frac{1}{|\Omega_0|} \int_{\Omega_0} \tilde{\Psi}[\tilde{\boldsymbol{F}}(\boldsymbol{F}[\boldsymbol{u}], \tilde{\boldsymbol{F}}'[\boldsymbol{v}])] \, \mathrm{d}\tilde{V} \, \mathrm{d}V, \quad \Pi^{\text{ext}}(\boldsymbol{u}) = \int_{\mathcal{B}_0} \boldsymbol{B}\cdot\boldsymbol{u} \, \mathrm{d}V + \int_{\Gamma^\sigma} \boldsymbol{T}\cdot\boldsymbol{u} \, \mathrm{d}A. \tag{20}$$

It should be noted that according to the assumptions in the previous section, no external contributions from the micro-continuum need to be considered. Furthermore, it is assumed here that both the body load and the Neumann surface load are independent of the deformation for the ease of presentation. Introducing the spaces of virtual deformations and fluctuations

$$\begin{aligned} \mathcal{V}^{\delta\boldsymbol{\varphi}} &= \{\delta\boldsymbol{\varphi} \in H^1(\mathcal{B}_0) \mid \delta\boldsymbol{\varphi} = \boldsymbol{0} \text{ on } \Gamma^\varphi\}, \\ \mathcal{V}^{\delta\tilde{\boldsymbol{w}}} &= \{\delta\tilde{\boldsymbol{w}} \in L_2(\mathcal{B}_0; H^1(\Omega_0)) \mid \delta\tilde{\boldsymbol{w}} = \boldsymbol{0} \text{ on } \delta\Omega_0\}, \end{aligned} \tag{21}$$

standard variational calculus finally yields the variational problem:

Find $(\boldsymbol{\varphi}, \tilde{\boldsymbol{w}}) \in \mathcal{V}^{\boldsymbol{\varphi}} \times \mathcal{V}^{\tilde{\boldsymbol{w}}}$, so that for all $(\delta\boldsymbol{\varphi}, \delta\tilde{\boldsymbol{w}}) \in \mathcal{V}^{\delta\boldsymbol{\varphi}} \times \mathcal{V}^{\delta\tilde{\boldsymbol{w}}}$

$$\begin{aligned} \int_{\mathcal{B}_0} \frac{1}{|\Omega_0|} \int_{\Omega_0} \tilde{\boldsymbol{P}} : \nabla_{\boldsymbol{X}}(\delta\boldsymbol{\varphi}) \, \mathrm{d}\tilde{V} \, \mathrm{d}V + \int_{\mathcal{B}_0} \frac{1}{|\Omega_0|} \int_{\Omega_0} \tilde{\boldsymbol{P}} : \nabla_{\tilde{\boldsymbol{X}}}(\delta\tilde{\boldsymbol{w}}) \, \mathrm{d}\tilde{V} \, \mathrm{d}V = \\ \int_{\mathcal{B}_0} \boldsymbol{B}\cdot\delta\boldsymbol{\varphi} \, \mathrm{d}V + \int_{\Gamma^\sigma} \boldsymbol{T}\cdot\delta\boldsymbol{\varphi} \, \mathrm{d}A, \quad \forall\, (\delta\boldsymbol{\varphi}, \delta\tilde{\boldsymbol{w}}) \in \mathcal{V}^{\delta\boldsymbol{\varphi}} \times \mathcal{V}^{\delta\tilde{\boldsymbol{w}}}, \end{aligned} \tag{22}$$

or equivalently

$$\int_{\mathcal{B}_0} \frac{1}{|\Omega_0|} \int_{\Omega_0} \tilde{\boldsymbol{P}} \, \mathrm{d}\tilde{V} : \nabla_{\boldsymbol{X}}(\delta\boldsymbol{\varphi}) \, \mathrm{d}V = \int_{\mathcal{B}_0} \boldsymbol{B}\cdot\delta\boldsymbol{\varphi} \, \mathrm{d}V + \int_{\Gamma^\sigma} \boldsymbol{T}\cdot\delta\boldsymbol{\varphi} \, \mathrm{d}A, \quad \forall\, \delta\boldsymbol{\varphi} \in \mathcal{V}^{\delta\boldsymbol{\varphi}}, \tag{23}$$

$$\int_{\mathcal{B}_0} \frac{1}{|\Omega_0|} \int_{\Omega_0} \tilde{\boldsymbol{P}} : \nabla_{\tilde{\boldsymbol{X}}}(\delta\tilde{\boldsymbol{w}}) \, \mathrm{d}\tilde{V} \, \mathrm{d}V = 0 \quad \forall\, \delta\tilde{\boldsymbol{w}} \in \mathcal{V}^{\delta\tilde{\boldsymbol{w}}}. \tag{24}$$

Note, that the last equation (24) automatically fulfils the required field equation in (16) with regard to the boundary conditions (15), that is, we obtain a static equilibrium of the micro-continuum.

3 | DISCRETE MULTISCALE BOUNDARY VALUE PROBLEM

To achieve a numerical solution for the problem, we apply a finite element framework to solve for $\boldsymbol{\varphi}$ and $\tilde{\boldsymbol{w}}$. In particular, we consider a standard displacement-based finite element approach, where we first introduce finite dimensional approximations of $\boldsymbol{\varphi}$ and $\delta\boldsymbol{\varphi}$, so that

$$\boldsymbol{\varphi}^{\text{h}}(\boldsymbol{X}) = \sum_{A\in\omega} N^A(\boldsymbol{X})\boldsymbol{q}_A \quad \text{and} \quad \delta\boldsymbol{\varphi}^{\text{h}}(\boldsymbol{X}) = \sum_{A\in\omega} N^A(\boldsymbol{X})\delta\boldsymbol{q}_A, \quad \text{with} \quad \boldsymbol{X} \in \mathcal{B}_0^{\text{h}}, \tag{25}$$

where $\mathcal{B}_0^{\text{h}}$ represents a (possibly approximate) parameterization of $\mathcal{B}_0$ according to the isoparametric concept. Here, $\omega = \{1, \ \dots \ , \ n_{\text{node}}\}$, such that $\boldsymbol{q}_A \in \mathbb{R}^n$ denotes the position vector of node A and $N^A : \mathcal{B}_0^{\text{h}} \to \mathbb{R}$ are global shape functions. Inserting the discrete approximation in (9) gives

$$\delta\boldsymbol{q}_A \cdot \int_{\mathcal{B}_0^{\text{h}}} \boldsymbol{P}^{\text{h}}(\boldsymbol{X}, \tilde{\boldsymbol{X}}) \nabla_{\boldsymbol{X}}(N^A(\boldsymbol{X})) \mathrm{d}V = \delta\boldsymbol{q}_A \cdot \boldsymbol{F}^{\text{ext},A}, \tag{26}$$

which has to hold for every $\delta \boldsymbol{q}_A, A = 1, \dots, n_{\text{node}}$. Here, $\boldsymbol{F}^{\text{ext},A}$ is the nodal force vector of the external contributions, given by

$$\boldsymbol{F}^{\text{ext},A} = \int_{\mathcal{B}_0^{\text{h}}} N^A \boldsymbol{B} \mathrm{d}V + \int_{\Gamma^{\sigma,\text{h}}} N^A \, \boldsymbol{T} \mathrm{d}A. \tag{27}$$

Applying a suitable quadrature formula yields for every $\delta \boldsymbol{q}_A$

$$\sum_{k=1}^{n_{\text{ip}}} \eta_k \left[\boldsymbol{P}^{\text{h}}(\boldsymbol{X}_{\text{ip}}^k, \tilde{\boldsymbol{X}}) \nabla_{\boldsymbol{X}} N^A(\boldsymbol{X}_{\text{ip}}^k) \right] = \boldsymbol{F}_{\text{quad}}^{\text{ext},A}, \tag{28}$$

where $\boldsymbol{F}_{\text{quad}}^{\text{ext},A}$ represents the quadrature of the external contributions and η_k the integration weights.

To obtain $\boldsymbol{P}^{\text{h}}(\boldsymbol{X}_{\text{ip}}^k, \tilde{\boldsymbol{X}})$, a suitable discretization of the microscopic fluctuations $\tilde{\boldsymbol{w}}$, defined on $\mathcal{B}_0^{\text{h}} \times \Omega_0$, is required. Therefore, we introduce the tensor product of macro- and microscale shape functions as follows

$$\begin{aligned} \tilde{\boldsymbol{w}}^{\text{h}}(\boldsymbol{X}, \tilde{\boldsymbol{X}}) &= \sum_{b \in \tilde{\omega}_1} \sum_{c \in \tilde{\omega}_2} R^b(\boldsymbol{X}) \tilde{R}^c(\tilde{\boldsymbol{X}}) \tilde{\boldsymbol{w}}_{b,c} = \sum_{B \in \tilde{\omega}} \overline{R}^B(\boldsymbol{X}, \tilde{\boldsymbol{X}}) \tilde{\boldsymbol{w}}_B \quad \text{and} \\ \delta \tilde{\boldsymbol{w}}^{\text{h}}(\boldsymbol{X}, \tilde{\boldsymbol{X}}) &= \sum_{b \in \tilde{\omega}_1} \sum_{c \in \tilde{\omega}_2} R^b(\boldsymbol{X}) \tilde{R}^c(\tilde{\boldsymbol{X}}) \delta \tilde{\boldsymbol{w}}_{b,c} = \sum_{B \in \tilde{\omega}} \overline{R}^B(\boldsymbol{X}, \tilde{\boldsymbol{X}}) \delta \tilde{\boldsymbol{w}}_B, \quad \boldsymbol{X} \in \mathcal{B}_0^{\text{h}}, \ \tilde{\boldsymbol{X}} \in \Omega_0^{\text{h}}, \end{aligned} \tag{29}$$

where the sets $\tilde{\omega}_i = \{1, \dots, n_{\text{micro}}^i\}, i = 1, 2$, summarize the tensor indices describing the position in the tensor product structure and the natural scheme $B(b, c) = (c-1) n_{\text{micro}}^1 + b$ is used for the global numbering $\tilde{\omega} = \{1, \dots, n_{\text{micro}}\}$, with $n_{\text{micro}} = n_{\text{micro}}^1 n_{\text{micro}}^2$. An example of such a construction in the case of one-dimensional macro- and microscales is shown in Figure 2.

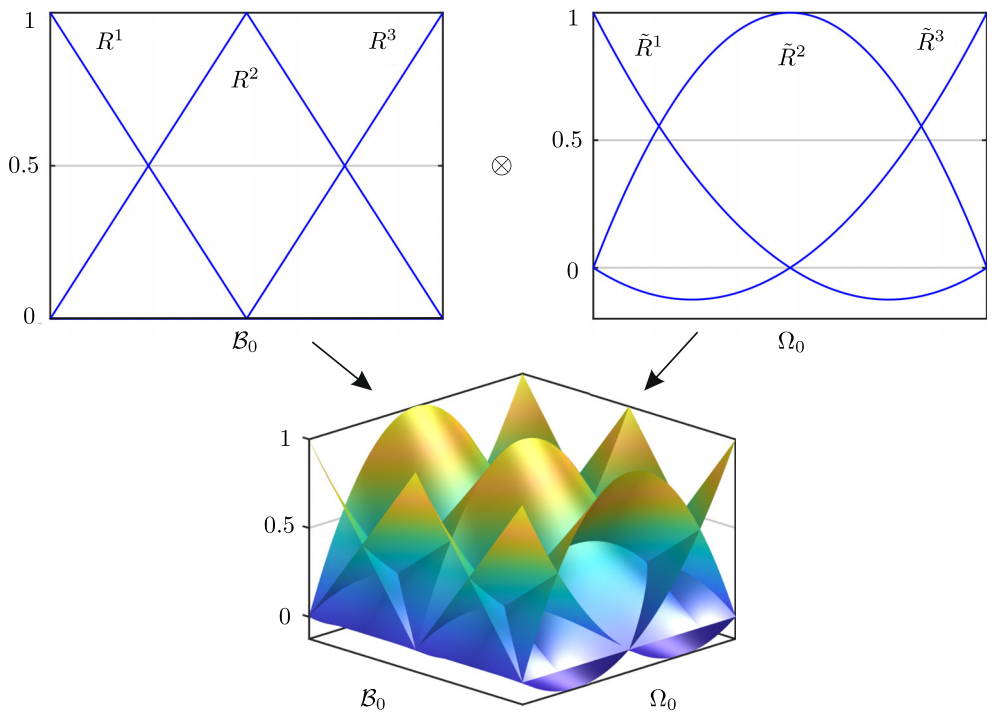

FIGURE 2 Example of a tensor product basis on a two-dimensional product space $\mathcal{B}_0 \times \Omega_0$. The upper images show the basis functions on the macro and micro level, respectively, and the lower image shows the resulting basis on $\mathcal{B}_0 \times \Omega_0$. A polynomial degree of 1 was selected at the macro level and a polynomial degree of 2 at the micro level.

As usual in multiscale techniques, we apply for the macroscale basis $\{R^b\}_{b=1}^{n^1_{\text{micro}}}$ a Delta Dirac, *that can be formally constructed as the limit of a sequence of smooth functions with compact support converging to a distribution and satisfying the so-called shifting property, that is,*[†]

$$\int_{\mathcal{G}} f(\boldsymbol{X})\delta(\boldsymbol{X}-\boldsymbol{X}_i)\mathrm{d}V = f(\boldsymbol{X}_i), \tag{30}$$

for every function f continuous at $\boldsymbol{X}_i \in \mathcal{G}$. In particular, computational multiscale methods like FE^2 evaluate the microscopic fluctuations at the integration points $\boldsymbol{X}^b_{\text{ip}}$, which is achieved by formally setting $R^b(\boldsymbol{X}) = \delta(\boldsymbol{X}-\boldsymbol{X}^b_{\text{ip}})$, $b = 1, \ldots n_{\text{ip}}$ (i.e., $n^1_{\text{micro}} = n_{\text{ip}}$), such that we obtain virtual fluctuations of the form

$$\delta\tilde{\boldsymbol{w}}^{\text{h}}(\boldsymbol{X},\tilde{\boldsymbol{X}}) = \sum_{b=1}^{n_{\text{ip}}}\sum_{c\in\tilde{\omega}_2}\delta(\boldsymbol{X}-\boldsymbol{X}^b_{\text{ip}})\tilde{R}^c(\tilde{\boldsymbol{X}})\delta\tilde{\boldsymbol{w}}_{b,c}. \tag{31}$$

Hence, (24) reads now

$$\sum_{b=1}^{n_{\text{ip}}}\delta\tilde{\boldsymbol{w}}_{b,c}\cdot\int_{\Omega^{\text{h}}_0}\tilde{\boldsymbol{P}}^{\text{h}}(\boldsymbol{X}^b_{\text{ip}},\tilde{\boldsymbol{X}})\nabla_{\tilde{\boldsymbol{X}}}\tilde{R}^c(\tilde{\boldsymbol{X}})\,\mathrm{d}\tilde{V} = 0, \quad \forall\delta\tilde{\boldsymbol{w}}_{b,c}\in\mathbb{R}^n, \tag{32}$$

where we multiplied the equation by the factor $|\Omega^{\text{h}}_0|$. Note, that the last statement explicitly reproduces a static equilibrium with regard to the boundary conditions presented in (15) on the microscale at each macroscale integration point.

According to the choice of the virtual fluctuations, we use Kronecker delta functions, the discrete analogs of the Delta-Dirac pulses, to interpolate the solutions, that is, we introduce interpolations of the form

$$\tilde{\boldsymbol{w}}^{\text{h}}(\boldsymbol{X},\tilde{\boldsymbol{X}}) = \sum_{b=1}^{n_{\text{ip}}}\sum_{c\in\tilde{\omega}_2}\delta'(\boldsymbol{X}-\boldsymbol{X}^b_{\text{ip}})\tilde{R}^c(\tilde{\boldsymbol{X}})\tilde{\boldsymbol{w}}_{b,c}, \quad \delta'(\boldsymbol{X}) := \begin{cases}1, & \text{if } \boldsymbol{X}=\boldsymbol{0},\\ 0, & \text{otherwise,}\end{cases} \tag{33}$$

see Figure 3 for illustration.

Insertion in (28) and (32) along with a suitable quadrature on the microscale yields

$$\sum_{b=1}^{n_{\text{ip}}}\eta_k\left[\frac{1}{|\Omega^{\text{h}}_0|}\sum_{l=1}^{n^{micro}_{\text{ip}}}\left[\tilde{\eta}_l\,\tilde{\boldsymbol{P}}^{\text{h}}\left(\boldsymbol{F}^{\text{h}}\left(\boldsymbol{X}^b_{\text{ip}}\right),\tilde{\boldsymbol{F}}'^{,\text{h}}\left(\boldsymbol{X}^b_{\text{ip}},\tilde{\boldsymbol{X}}^l_{\text{ip}}\right)\right)\right]\nabla_{\boldsymbol{X}}N^A\left(\boldsymbol{X}^b_{\text{ip}}\right)\right] = \boldsymbol{F}^{\text{ext},A}_{\text{quad}}, \quad A\in\omega, \tag{34}$$

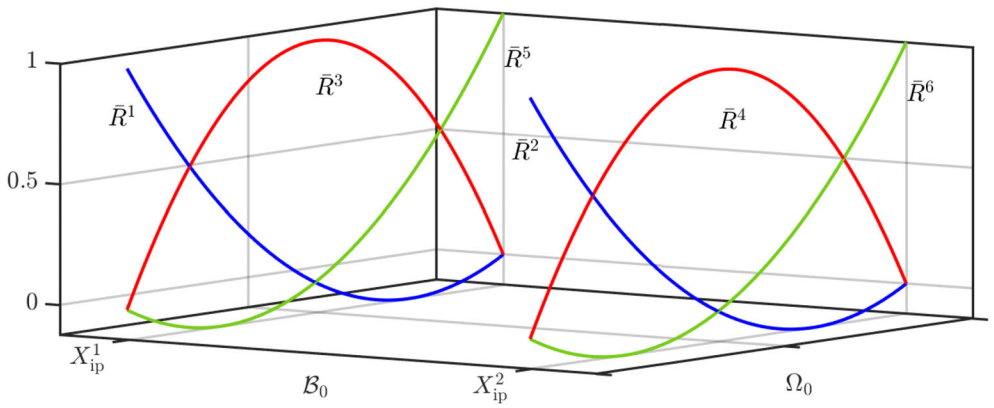

FIGURE 3 Example of a tensor product basis on a two-dimensional product space $\mathcal{B}_0 \times \Omega_0$ using Kronecker delta functions with two integration points on the macroscale and piecewise polynomials of degree 2 on the microscale. The six global functions are numbered according to the global numbering scheme introduced in Equation (29).

$$\sum_{l=1}^{n_{\text{ip}}^{micro}} \tilde{\eta}_l \left[\tilde{\boldsymbol{P}}^{\text{h}} \left(\boldsymbol{F}^{\text{h}} \left(\boldsymbol{X}_{\text{ip}}^{b} \right), \tilde{\boldsymbol{F}}'^{,\text{h}} \left(\boldsymbol{X}_{\text{ip}}^{b}, \tilde{\boldsymbol{X}}_{\text{ip}}^{l} \right) \right) \nabla_{\tilde{\boldsymbol{X}}} \tilde{R}^{c} \left(\tilde{\boldsymbol{X}}_{\text{ip}}^{l} \right) \right] = \boldsymbol{0}, \quad (b,c) \in \tilde{\omega}_1 \times \tilde{\omega}_2, \tag{35}$$

where $\tilde{\eta}_l$ are the quadrature weights at the n_{ip}^{micro} integration point to evaluate the integral on Ω_0^{h}.

Remark 2. Equations (34) and (35) represent the classical FE^2 method, evaluated in a staggered solution process to be discussed in the next section.

Other choices to approximate $\tilde{\boldsymbol{w}}$ and $\delta\tilde{\boldsymbol{w}}$ within $\mathcal{B}_0$ using suitable interpolation functions are certainly possible. We refer to Reference 24 among others, where discontinuous interpolation functions are introduced in the context of Hu-Washizu and Hellinger-Reissner formulations. As the microscopic fluctuations are applied to calculate the stresses, we obtain similar continuity requirements for both. The next step after the Dirac approach described above may be the use of piecewise constant approximation functions on the macroscale. In particular, the construction of suitable bases no longer has to be based on the quadrature formula used and the underlying system of equations generally has the form

$$\int_{\mathcal{B}_0^{\text{h}}} \frac{1}{|\Omega_0^{\text{h}}|} \int_{\Omega_0^{\text{h}}} \tilde{\boldsymbol{P}}(\boldsymbol{F}^{\text{h}}(\boldsymbol{X}), \tilde{\boldsymbol{F}}'^{,\text{h}}(\boldsymbol{X}, \tilde{\boldsymbol{X}}))\, \mathrm{d}\tilde{V}\, \nabla_{\boldsymbol{X}} N^{A}(\boldsymbol{X})\, \mathrm{d}V = \boldsymbol{F}^{\text{ext},A}, \quad A \in \omega, \tag{36}$$

$$\int_{\mathcal{B}_0^{\text{h}}} \frac{1}{|\Omega_0^{\text{h}}|} \int_{\Omega_0^{\text{h}}} \tilde{\boldsymbol{P}}(\boldsymbol{F}^{\text{h}}(\boldsymbol{X}), \tilde{\boldsymbol{F}}'^{,\text{h}}(\boldsymbol{X}, \tilde{\boldsymbol{X}})) \nabla_{\tilde{\boldsymbol{X}}} \overline{R}^{B}(\boldsymbol{X}, \tilde{\boldsymbol{X}})\, \mathrm{d}\tilde{V}\, \mathrm{d}V = \boldsymbol{0}, \quad B \in \tilde{\omega}. \tag{37}$$

Using linear or even higher-order shape functions for $\{R^b\}_{b=1}^{n_{\text{micro}}^1}$ can be done in a straight-forward manner; however, we have to ensure that the shape functions are discontinuous at the element boundaries to allow for a block diagonal submatrix of the Hessian to be used within the following null-space reduction scheme. Moreover, we have to ensure that we do not obtain stability issues, as a higher-order interpolation of the micro-fluctuations (and thus, of the stresses to be calculated) may conflict with a lower-order interpolation of the macroscale deformation map. Both equations (36) and (37) can be rewritten as $\delta\boldsymbol{q}_A \cdot \boldsymbol{R}^A_{\text{macro}} = 0$ and $\delta\tilde{\boldsymbol{w}}_B \cdot \boldsymbol{R}^B_{\text{micro}} = 0$, where $\boldsymbol{R}^A_{\text{macro}}$ and $\boldsymbol{R}^B_{\text{micro}}$ are the corresponding residual vector, to be used within a Newton–Raphson iteration as shown in the next section.

4 | SOLUTION OF THE MULTISCALE PROBLEM

The nonlinear multiscale framework at hand is solved by introducing a Newton–Raphson iteration, noting that further enhancement like line-search or trust-region methods can be applied as well, see Reference 30 for details. In particular, we solve

$$[\delta\boldsymbol{q}_A,\ \delta\tilde{\boldsymbol{w}}_B] \cdot \begin{bmatrix} \boldsymbol{K}^{AC} & \boldsymbol{D}^{AD} \\ \boldsymbol{E}^{BC} & \boldsymbol{L}^{BD} \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{q}_C \\ \Delta\tilde{\boldsymbol{w}}_D \end{bmatrix} = -[\delta\boldsymbol{q}_A,\ \delta\tilde{\boldsymbol{w}}_B] \cdot \begin{bmatrix} \boldsymbol{R}^A_{\text{macro}} \\ \boldsymbol{R}^B_{\text{micro}} \end{bmatrix}. \tag{38}$$

Afterwards, both values are updated via $\boldsymbol{q}_A \leftarrow \boldsymbol{q}_A + \Delta\boldsymbol{q}_A$ and $\tilde{\boldsymbol{w}}_A \leftarrow \tilde{\boldsymbol{w}}_A + \Delta\tilde{\boldsymbol{w}}_A$ and the iteration restarts until $\|[\boldsymbol{R}^{\text{T}}_{\text{macro}},\ \boldsymbol{R}^{\text{T}}_{\text{micro}}]^{\text{T}}\| < \epsilon_{NR}$ with a predefined stop criterion ϵ_{NR}. The tangent matrix is composed of four terms, given by

$$\boldsymbol{K}^{AC} = \int_{\mathcal{B}_0^{\text{h}}} \nabla_{\boldsymbol{X}}(N^A) \cdot \frac{1}{|\Omega_0^{\text{h}}|} \int_{\Omega_0^{\text{h}}} \frac{\partial^2 \tilde{\Psi}}{\partial \tilde{\boldsymbol{F}} \partial \tilde{\boldsymbol{F}}} \mathrm{d}\tilde{V}\, \nabla_{\boldsymbol{X}}(N^C) \mathrm{d}V \tag{39}$$

and

$$\boldsymbol{D}^{AD} = \int_{\mathcal{B}_0^{\text{h}}} \nabla_{\boldsymbol{X}}(N^A) \cdot \frac{1}{|\Omega_0^{\text{h}}|} \int_{\Omega_0^{\text{h}}} \frac{\partial^2 \tilde{\Psi}}{\partial \tilde{\boldsymbol{F}} \partial \tilde{\boldsymbol{F}}} \nabla_{\tilde{\boldsymbol{X}}}(\overline{R}^D) \mathrm{d}\tilde{V} \mathrm{d}V. \tag{40}$$

As the tangent matrix is the Hessian resulting from the condition in (19), the matrix is symmetric and $\boldsymbol{E} = \boldsymbol{D}^{\mathrm{T}}$. Eventually, the last term is given by

$$\boldsymbol{L}^{BD} = \int_{\mathcal{B}_0^{\mathrm{h}}} \frac{1}{|\Omega_0^{\mathrm{h}}|} \int_{\Omega_0^{\mathrm{h}}} \nabla_{\tilde{\boldsymbol{X}}}(\overline{R}^{B}) \cdot \frac{\partial^2 \tilde{\Psi}}{\partial \tilde{\boldsymbol{F}} \partial \tilde{\boldsymbol{F}}} \nabla_{\tilde{\boldsymbol{X}}}(\overline{R}^{D}) \mathrm{d}\tilde{V} \mathrm{d}V \tag{41}$$

and the corresponding global algebraic system reads

$$\begin{bmatrix} \boldsymbol{K} & \boldsymbol{D} \\ \boldsymbol{E} & \boldsymbol{L} \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{q} \\ \Delta \tilde{\boldsymbol{w}} \end{bmatrix} = - \begin{bmatrix} \boldsymbol{R}_{\mathrm{macro}} \\ \boldsymbol{R}_{\mathrm{micro}} \end{bmatrix}, \tag{42}$$

where $\Delta \boldsymbol{q}$, $\Delta \tilde{\boldsymbol{w}}$ summarize the unknown node data $\Delta \boldsymbol{q}_A$, $\Delta \tilde{\boldsymbol{w}}_B$ and the individual blocks are composed according to Equations (38)-(41). Note, that the matrix in (42) is in general smaller than in (38), as the Dirichlet boundary conditions on both scales are now incorporated by removing the corresponding rows and columns.

4.1 | Classical solution strategy

In classical FE^2 methods as presented in, for example, Reference 17 for nonlinear electromechanical multiscale problems, a staggered scheme is applied. First, (42) is solved using fixed $\boldsymbol{q}^{\ddagger}$, until $\|\boldsymbol{R}_{\mathrm{micro}}\| < \epsilon_{NR}^{\mathrm{micro}}$. Then, the second line of (42) reads

$$\boldsymbol{E} \Delta \boldsymbol{q} + \boldsymbol{L} \Delta \tilde{\boldsymbol{w}} = \boldsymbol{0}. \tag{43}$$

Since the matrix $\boldsymbol{L}$ contains all tangent matrices of all RVEs, which are strictly separated, the matrix is block-wise diagonal. As we have to solve the linear system in the second line of (42) (either using direct or iterative solver), we can write for this step in the solution procedure[§]

$$\Delta \tilde{\boldsymbol{w}} = -\boldsymbol{L}^{-1} \boldsymbol{E} \Delta \boldsymbol{q}, \tag{44}$$

and insert this in the first line of (42), such that

$$\left[\boldsymbol{K} - \boldsymbol{D}\boldsymbol{L}^{-1}\boldsymbol{E}\right] \Delta \boldsymbol{q} = -\boldsymbol{R}_{\mathrm{macro}}, \tag{45}$$

update $\boldsymbol{q}$ with $\Delta \boldsymbol{q}$, update $\tilde{\boldsymbol{w}}$ using (44) and restart the staggered scheme again until $\|\boldsymbol{R}_{\mathrm{macro}}\| < \epsilon_{\mathrm{macro}}$.

4.2 | Generalization of the solution strategy

The solution strategy introduced in the previous section can be rewritten as premultiplication with the rectangular null-space matrix

$$\boldsymbol{P} = \left[\boldsymbol{I}, \quad -\boldsymbol{D}\boldsymbol{L}^{-1}\right]^{\mathrm{T}}, \tag{46}$$

where $\boldsymbol{I}$ is the unity matrix of dimension $n_{\mathrm{node}} \times n_{\mathrm{node}}$. Now, instead of applying a staggered scheme, we multiply the linear system composed of both scales using the null-space matrix

$$\boldsymbol{P}^{\mathrm{T}} \cdot \begin{bmatrix} \boldsymbol{K} & \boldsymbol{D} \\ \boldsymbol{E} & \boldsymbol{L} \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{q} \\ \Delta \tilde{\boldsymbol{w}} \end{bmatrix} = -\boldsymbol{P}^{\mathrm{T}} \cdot \begin{bmatrix} \boldsymbol{R}_{\mathrm{macro}} \\ \boldsymbol{R}_{\mathrm{micro}} \end{bmatrix}, \tag{47}$$

which yields

$$\left[\boldsymbol{K} - \boldsymbol{D}\boldsymbol{L}^{-1}\boldsymbol{E}\right] \Delta \boldsymbol{q} = -\left[\boldsymbol{R}_{\mathrm{macro}} - \boldsymbol{D}\boldsymbol{L}^{-1}\boldsymbol{R}_{\mathrm{micro}}\right]. \tag{48}$$

Finally, we need to calculate $\Delta\tilde{\boldsymbol{w}}$ via

$$\Delta\tilde{\boldsymbol{w}} = -\boldsymbol{L}^{-1}\big[\boldsymbol{R}_{\text{micro}} + \boldsymbol{E}\Delta\boldsymbol{q}\big], \tag{49}$$

and update both values $\boldsymbol{q}$ and $\tilde{\boldsymbol{w}}$ within the Newton–Raphson iteration.

Remark 3. For the nonlinear case, $\|\boldsymbol{R}_{\text{micro}}\|$ is usually unequal zero due to numerical issues, hence (43) is never valid and may affect the quadratic convergence of the Newton–Raphson iteration. Thus, the proposed approach is in general advantageous as we can apply the staggered scheme to (48). In comparison to classical (staggered) schemes, the term $-\boldsymbol{D}\boldsymbol{L}^{-1}\boldsymbol{R}_{\text{micro}}$ on the right-hand side of (48) is introduced, noting again, that $\boldsymbol{R}_{\text{micro}}$ is never exact zero.

5 | NUMERICAL EXAMPLES

In this section, we demonstrate the advantages of the proposed formulation. In particular, we start with a convergence study using an analytical solution of a 1 + 1 dimensional problem, followed by a typical problem emanating from nonlinear elasticity.

5.1 | Convergence studies

In the following, some numerical convergence analyses are presented using a one-dimensional benchmark problem on the macroscale and a one-dimensional microscale. The macro- and micro-continuum are given by $\mathcal{B}_0 = (0, 1000)$ and $\Omega_0 = (0, 1)$, respectively, where the lengths here and in the remainder of this example are given in millimeters unless stated otherwise. The material is defined by the strain energy density

$$\tilde{\Psi}(\tilde{F}, \tilde{X}) = \lambda(\tilde{X})(\tilde{F}^2 - 1), \quad \lambda(\tilde{X}) = 20\Big[\cos\Big(\frac{2\pi}{3}\tilde{X} - \frac{\pi}{3}\Big)\Big]^{-1}, \tag{50}$$

where λ is given in the unit Jmm^{-1} and represents a location-dependent material property, cf. Figure 4. It should be noted that $\tilde{\Psi}$ is selected so that an analytical solution of the system can be specified. Due to the simple shape, however, residual stresses occur, so the reference configuration is not stress-free. Assuming the body load and boundary conditions

$$B = 1\ \text{Jmm}^{-1}, \quad \varphi(0) = 0, \quad \varphi(1000) = \frac{37{,}575\sqrt{3}}{2\pi} \tag{51}$$

for the macro-deformation and homogeneous Dirichlet conditions in accordance with Eq. $(15)_{\text{ii}}$ for the micro-fluctuation, the solutions $\varphi^a : \mathcal{B}_0 \to \mathbb{R}$ and $\tilde{w}^a : \mathcal{B}_0 \times \Omega_0 \to \mathbb{R}$ of the coupled boundary value problems (5), (16) are given by

$$\varphi^a(X) = -\frac{3\sqrt{3}}{160\pi}\,X^2 + \frac{3003\sqrt{3}}{80\pi}\,X, \tag{52}$$

$$\tilde{w}^a(X, \tilde{X}) = F^a(X)\left[\frac{1}{\sqrt{3}}\sin\Big(\frac{2\pi}{3}\,\tilde{X} - \frac{\pi}{3}\Big) - \tilde{X} + \frac{1}{2}\right], \tag{53}$$

where $F^a = \mathrm{d}\varphi^a/\mathrm{d}X$, see Figure 4 for illustration.

Taking into account the given data, the two functions are characterized according to Equation(22) by the conditions

$$\int_0^{1000}\int_0^1 \lambda\left(\frac{\mathrm{d}\varphi}{\mathrm{d}X} + \frac{\partial\tilde{w}}{\partial\tilde{X}}\right)\frac{\mathrm{d}\,\delta\varphi}{\mathrm{d}X}\,\mathrm{d}\tilde{X}\mathrm{d}X = \frac{1}{2}\int_0^{1000}\delta\varphi\,\mathrm{d}X, \quad \forall\,\delta\varphi \in \mathcal{V}^{\delta\varphi}, \tag{54}$$

$$\int_0^{1000}\int_0^1 \lambda\left(\frac{\mathrm{d}\varphi}{\mathrm{d}X} + \frac{\partial\tilde{w}}{\partial\tilde{X}}\right)\frac{\partial\delta\tilde{w}}{\partial\tilde{X}}\mathrm{d}\tilde{X}\mathrm{d}X = 0, \quad \forall\,\delta\tilde{w} \in \mathcal{V}^{\delta\tilde{w}}, \tag{55}$$

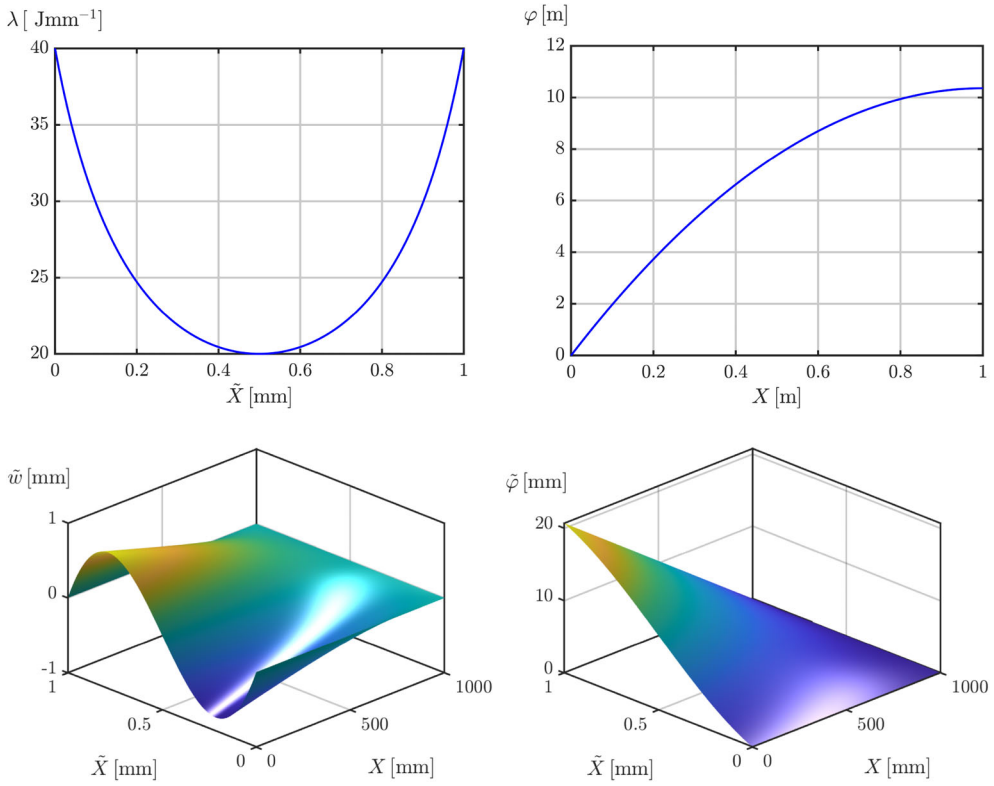

FIGURE 4 Upper left: material parameter λ, upper right: macro-deformation φ, lower left: micro-fluctuation $\tilde{w}$, lower right: micro-deformation $\tilde{\varphi}$.

where the various carried out finite element analyses are based on. In the following, a total of six settings is considered, each of which differs in the choice of approximation functions. In detail, bilinear ($p = [1, 1]$), biquadratic ($p = [2, 2]$) and mixed approaches of the form $p = [0, 1]$, $p = [0, 2]$ are used for the fluctuation field, whereby in the last two cases $\tilde{w}$ is approximated by piecewise constant functions in $\boldsymbol{X}$-direction. In addition, approaches of the form $p = [\delta, 1]$, and $p = [\delta, 2]$ are considered, that is, two classical approaches in which Dirac pulses are used in X-direction of $\tilde{w}$, compare Equations (31) and (33). Regarding the macro deformation, shape-functions of the same order as in $\tilde{X}$-direction of $\tilde{w}$ are used in each case.

Figure 5 shows the results of a convergence study where the relative L_2 errors

$$err_{\varphi} := \frac{\|\varphi^{\mathrm{h}} - \varphi^{a}\|_{L_2(\mathcal{B}_0)}}{\|\varphi^{a}\|_{L_2(\mathcal{B}_0)}}, \qquad err_{\tilde{w}} := \frac{\|\tilde{w}^{\mathrm{h}} - \tilde{w}^{a}\|_{L_2(\mathcal{B}_0\times\Omega_0)}}{\|\tilde{w}^{a}\|_{L_2(\mathcal{B}_0\times\Omega_0)}} \tag{56}$$

corresponding to each setting are plotted as functions of the mesh size h which is defined as the maximal diameter of the elements in the reference configuration. Moreover, we denote by $\|\bullet\|_{L_2(B)}$ the usual L_2 norm on a domain B. For the calculations, resolutions of 2^k elements are used for the macro-domain $\mathcal{B}_0$ and corresponding resolutions of $2^k \times 2^k$ elements are used for the extended region $\mathcal{B}_0 \times \Omega_0$, where $k = 4, \ldots, 9$. As can be seen in the left-hand image of Figure 5, err_{φ} reduces according to an order greater than 2 in the case of linear elements regardless of which approach is used for $\tilde{w}$ in X-direction. The situation is different when using quadratic approaches. While with an approximation of $\tilde{w}$ with $p = [2, 2]$, $[\delta, 2]$ the error decreases with an order of 3.5, an order of 2 can be observed using the piecewise constant/quadratic approach $p = [0, 2]$. Eventually, the behavior of the error $err_{\tilde{w}}$ is shown in Figure 5 on the right. There it can be seen that a convergence order of 1.6 is achieved with the bilinear approach, while an order of 3 is achieved with both the biquadratic and the quadratic Dirac approach $p = [\delta, 2]$. In addition, the mixed approaches $p = [0, 1]$, $[0, 2]$ and the linear Dirac approach $p = [\delta, 1]$ each achieve a convergence order of 1. We point out, that the corresponding error curves overlap in the graphic so that only one curve can be recognized for these three cases.

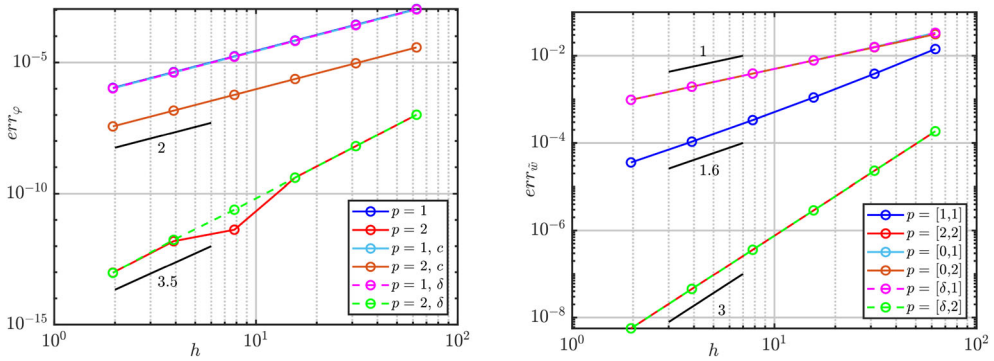

FIGURE 5 Left: err_φ as a function of the mesh-size h. The two cases in which piecewise constant shape-functions where used in X-direction of $\tilde{w}$ are labeled witch "c", and the cases in which Dirac pulses were used are labeled with "δ" accordingly. Right: $err_{\tilde{w}}$ as a function of the mesh-size h. The error curves of $p = [0,1]$, $[0,2]$ and $p = [\delta,1]$ overlap so that only one curve for these three cases is recognizable.

5.2 | Cook's membrane

Finally, we compare the two solution strategies presented in Section 4 on the basis of a two-dimensional Cook's membrane. The reference configuration of the macro-continuum is defined by the four points $\boldsymbol{P}_1 = (0,0)^{\mathrm{T}}$, $\boldsymbol{P}_2 = (480{,}440)^{\mathrm{T}}$, $\boldsymbol{P}_3 = (480{,}600)^{\mathrm{T}}$, $\boldsymbol{P}_4 = (0{,}440)^{\mathrm{T}}$, where $\boldsymbol{P}_1$ denotes the lower left, $\boldsymbol{P}_2$ the lower right, $\boldsymbol{P}_3$ the upper right and $\boldsymbol{P}_4$ the upper left corner of the membrane. Consequently, the computational macro-domain is given by

$$\mathcal{B}_0 = \left\{ \boldsymbol{X} = (X_1, X_2) \in \mathbb{R}^2 \mid X_1 \in (0{,}480),\ \frac{11}{12} X_1 < X_2 < \frac{1}{3} X_1 + 440 \right\} \tag{57}$$

while we use a square area of the form $\Omega_0 = (-3,3)^2$ to represent the micro-scale. Here and in the following, all length specifications referring to the macro-scale are given in meters and length specifications referring to the micro-scale are given in millimeters. Regarding the boundary conditions, we assume that the left edge $\Gamma^\varphi = \{\boldsymbol{X} \in \mathbb{R}^2 \mid X_1 = 0,\ 0 < X_2 < 440\}$ of the membrane is fixed, whereas the remaining boundary $\Gamma^\sigma = \partial\mathcal{B}_0 \setminus \Gamma^\varphi$ is exposed to the surface load

$$\boldsymbol{T}:\ \Gamma^\sigma \to \mathbb{R}^2, \quad \boldsymbol{T}(\boldsymbol{X}) = \begin{cases} (-5,10)^{\mathrm{T}} & \text{if } \boldsymbol{X} \in \Gamma^\sigma_r, \\ (0,0)^{\mathrm{T}} & \text{otherwise,} \end{cases} \tag{58}$$

acting on the right edge $\Gamma^\sigma_r := \{\boldsymbol{X} \in \mathbb{R}^2 \mid X_1 = 480,\ 440 < X_2 < 600\}$ of $\mathcal{B}_0$, see the left picture in Figure 6 for illustration.

Moreover, we assume a zero body load ($\boldsymbol{B} = \boldsymbol{0}$) and the membrane to consist of two materials described by the strain energy densities

$$\tilde{\Psi}_i(\tilde{\boldsymbol{F}}) = \alpha_i(\tilde{\boldsymbol{F}} : \tilde{\boldsymbol{F}} - 2) + \beta_i(\tilde{\boldsymbol{F}} : \tilde{\boldsymbol{F}} + \tilde{J}^2 - 3) + \frac{\kappa_i}{2}(\tilde{J} - 1)^2 - 2(\alpha_i + 2\beta_i)\log(\tilde{J}), \tag{59}$$

$i = 1,2$, with the material parameters $(\alpha_1, \beta_1, \kappa_1) = (27, 18, 60)$ Jmm^{-2} and $(\alpha_2, \beta_2, \kappa_2) = (13.5, 6.5, 30)$ Jmm^{-2}, respectively. The two materials are distributed within the RVE in such a way that a cross-shaped area of width 2 mm consists of the second material, while the remaining area consists of the first material, see the right picture in Figure 6 for illustration.

To approximate the micro-fluctuation, we apply a classical approach using Dirac pulses for the $\boldsymbol{X}$-direction according to Equations (31), (33) and interpolate both $\tilde{\boldsymbol{w}}$ in $\tilde{X}$-direction and $\boldsymbol{\varphi}$ with bilinear four-node elements. Thereby, the macro continuum is resolved with 20×20, while the RVE is resolved with 30×30 elements, cf. Figure 6. Additionally, we postulate periodic conditions on $\partial\Omega_0$ via a nodal coupling and set $\tilde{\boldsymbol{w}} = \boldsymbol{0}$ on $\mathcal{B}_0 \times \delta\Omega_0^c$, where $\delta\Omega_0^c$ consist of the four corners $\tilde{\boldsymbol{P}}_1 = (-3,-3)^{\mathrm{T}}$, $\tilde{\boldsymbol{P}}_2 = (3,-3)^{\mathrm{T}}$, $\tilde{\boldsymbol{P}}_3 = (3,3)^{\mathrm{T}}$ and $\tilde{\boldsymbol{P}}_4 = (-3,3)^{\mathrm{T}}$ of the RVE, to restrict rigid body movements.[¶] For the load application, we use an incremental scheme with a total of $n_l \in \{2, 12, 22, 32, 42\}$ load steps so that in each case a reduced load of the form $\boldsymbol{T}_k := k\boldsymbol{T}/n_l$ is applied in the kth step. The degrees of freedom of the respecting results are subsequently

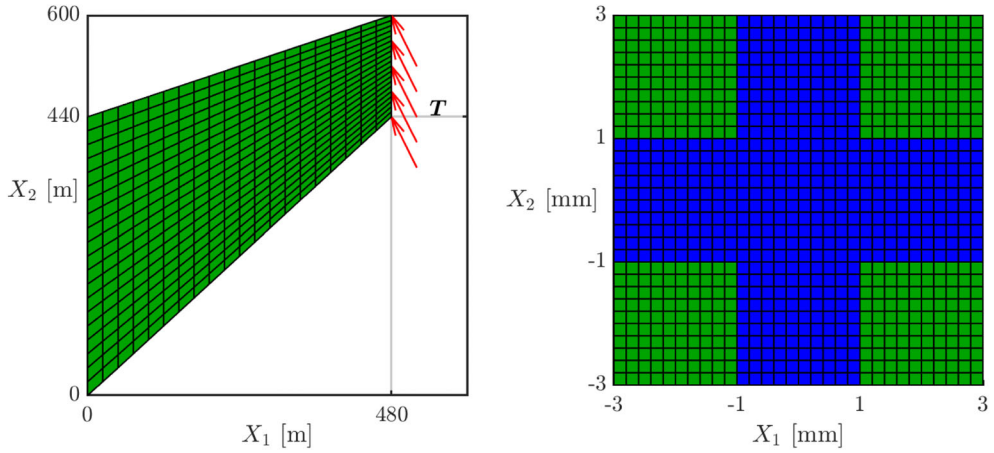

FIGURE 6 Left: reference configuration and computational mesh of the macro-system. The red arrows indicate the impact of the surface load defined in Equation (58). Right: reference configuration and computational mesh of an RVE. The green regions consist of the material defined by $\tilde{\Psi}_1$ and the blue one consist of the material defined by $\tilde{\Psi}_2$.

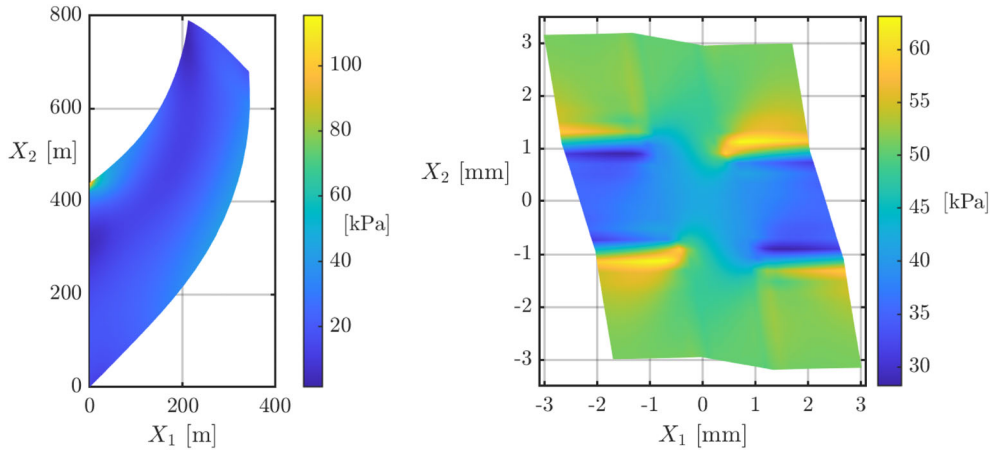

FIGURE 7 Left: von Mises stress distribution at the actual configuration of Cook's membrane. Right: von Mises stress distribution actual configuration of an RVE located in the vicinity of the upper left corner $\boldsymbol{P}_4$ of the membrane.

TABLE 1 Number of global Newton iterations for different numbers n_l of total load steps.

n_l	2	12	22	32	42
Generalized	6	4	4	4	4
Classical	/	/	4	4	4

used as the starting value for the $(k+1)$th load step to ensure the convergence of the Newton iteration. Corresponding results are shown in Figure 7.

In the following, we consider two different settings, wherein the first one we employ the classical solution strategy from Section 4.1. On the micro-level, we utilize the termination criterion $\epsilon = 10^{-13}$, and on the macro-level, we implement the criterion $\epsilon_{\text{macro}} = 10^{-9}$. In the second scenario, we opt for the generalized solution strategy outlined in Section 4.2 and set the termination criterion to $\epsilon_{\text{macro}} = 10^{-9}$. A comparison of the number of Newton iterations required in the two approaches as a function of the number of load steps is shown in Table 1. We remark that the number of iterations remains the same from step to step so that the values represent the number of iterations in each load step. In addition, the cases in which the Newton method was aborted unsuccessfully are indicated by a slash. Thereby, we define the procedure as

TABLE 2 Convergence of the Newton method in different load steps for a load application with a total of 22 load steps. The residual is specified in the unit Newton (N).

k	Load-step 1		
	Generalized	Classical	
		Macro	RVE
0	$1.54 \cdot 10^{1}$	$1.54 \cdot 10^{1}$	$2.07 \cdot 10^{-15}$
1	$3.84 \cdot 10^{1}$	$3.84 \cdot 10^{1}$	$7.00 \cdot 10^{-4}$
			$9.92 \cdot 10^{-7}$
			$1.20 \cdot 10^{-12}$
			$2.34 \cdot 10^{-15}$
2	$1.92 \cdot 10^{-2}$	$1.52 \cdot 10^{-2}$	$2.06 \cdot 10^{-1}$
			$5.24 \cdot 10^{-2}$
			$7.57 \cdot 10^{-3}$
			$2.71 \cdot 10^{-4}$
			$3.97 \cdot 10^{-7}$
			$9.60 \cdot 10^{-13}$
			$2.25 \cdot 10^{-15}$
3	$1.06 \cdot 10^{-7}$	$7.61 \cdot 10^{-7}$	$2.01 \cdot 10^{-1}$
			$4.92 \cdot 10^{-2}$
			$6.79 \cdot 10^{-3}$
			$2.14 \cdot 10^{-4}$
			$2.48 \cdot 10^{-6}$
			$3.70 \cdot 10^{-13}$
			$2.34 \cdot 10^{-15}$
4	$4.61 \cdot 10^{-10}$	$3.73 \cdot 10^{-10}$	$2.01 \cdot 10^{-1}$
			$4.92 \cdot 10^{-2}$
			$6.79 \cdot 10^{-3}$
			$2.14 \cdot 10^{-4}$
			$2.49 \cdot 10^{-7}$
			$3.71 \cdot 10^{-13}$
			$2.40 \cdot 10^{-15}$

k	Load-step 22		
	Generalized	Classical	
		Macro	RVE
0	$1.54 \cdot 10^{1}$	$1.54 \cdot 10^{1}$	$2.29 \cdot 10^{-12}$
1	$1.17 \cdot 10^{1}$	$1.17 \cdot 10^{1}$	$5.63 \cdot 10^{-4}$
			$6.11 \cdot 10^{-7}$
			$4.38 \cdot 10^{-13}$
			$2.31 \cdot 10^{-15}$
2	$8.97 \cdot 10^{-3}$	$8.16 \cdot 10^{-3}$	$5.19 \cdot 10^{-6}$
			$3.47 \cdot 10^{-12}$
			$2.43 \cdot 10^{-15}$
3	$7.11 \cdot 10^{-7}$	$5.24 \cdot 10^{-7}$	$2.66 \cdot 10^{-7}$
			$1.02 \cdot 10^{-13}$
			$2.65 \cdot 10^{-15}$
4	$4.18 \cdot 10^{-10}$	$6.00 \cdot 10^{-10}$	$1.75 \cdot 10^{-13}$
			$2.36 \cdot 10^{-15}$

failed if the macro-residual $\|\boldsymbol{R}_{\text{macro}}\|$ exceeds the value 10^{10} during the iteration. As can be seen, the Newton method terminates from $n_l = 22$ using the classical solution method, while the generalized method terminates successfully from $n_l = 2$. It should be noted that the step numbers are minimal in the sense that the scheme does not converge for $n_l = 1$ and $n_l = 21$ using the generalized and the classical solution strategy, respectively. Eventually, Table 2 shows an example of the convergence behavior of the Newton method for the two solution procedures to give an idea of the processes. The value of the macro-residual in the kth iteration step is given under "generalized" and "classical," respectively, while the iteration progression for the classical staggered scheme at the micro level is given in the column "RVE." The quantities shown depict the course of convergence in relation to an exemplary integration point, whereby a point was selected in each case for which the maximum number of iterations was required to achieve the termination criterion.

6 | CONCLUSIONS

The variational formulation of the multiscale system leads to a set of partial differential equations in a higher-dimensional space. Even constant shape functions of the fluctuations in direction of the macroscale match the results of Delta Dirac functions evaluated at every Gauss points, although the number of RVEs is dramatically reduced to one per element. Using higher-order shape functions produces the expected convergence rates, noting that the absolute error for linear shape functions for the fluctuations already surpasses the results of the traditional formulation.

With regard to (19), we can show that the discrete multiscale problem at hand can be considered as a finite-dimensional and nonconvex minimization problem introducing a poly-convex strain energy function on the microscale. Using this as a discrete objective function, sophisticated methods like the recursive trust-region multigrid formulations as proposed in Gross and Krause[26] can be adapted as well, such that state of the art parallelization techniques can be applied instead of the usual "farming" on clusters to solve for the physical equilibrium condition on the microscale within every macro Newton step.

Moreover, it is common to apply a staggered scheme for nonlinear multiscale problems, searching first for a physical equilibrium condition on the microscale within each Newton–Raphson step on the macroscale. However, as long as the macroscale is not in equilibrium, the physical interpretation of the microscale is questionable. In particular, in the simple case of a purely mechanical, first order hyperelastic problem, the forces at the macroscale are unbalanced and the converged solution of the corresponding microscale corresponds to the unbalanced forces. Due to this issue, we observe the limits of the original staggered scheme as presented in Table 1, which clearly demonstrates that the application of larger load steps is impossible in contrast to the monolithic solution. This leads to the curious situation, that a huge amount of additional computational power is required for the microscale solutions, without improving the solution on the macroscale. In the best case, we obtain quadratic convergence on the macroscale, in the worst case, the staggered scheme prevents the solution of the macroscale system for larger external load steps.

In contrast, the proposed null-space reduction scheme allows for a quadratic convergence simultaneously for both, the micro- and the macroscale without using any additional step in a staggered scheme. This yields a massive reduction of the computational effort for each iteration. Additionally, we could demonstrate that the monolithic solution is tremendously more robust. For the Cook's membrane, the Newton scheme does not converge for the staggered scheme until we apply the external load in 22 load steps, whereas the monolithic scheme requires only two. Combined with the massive reduction of RVEs to be evaluated using constant shape functions in direction of the macroscale, the computational power required for the whole simulation is reduced to a fraction compared to the classical formulation.

Moreover, the applied methodology of null-space reduction schemes allows for a rather simple application on all kinds of multiscale problems. The development of a consistent linearization for generalized materials in the context of IGA^2, for example, for second or third-order materials is highly tedious, see the Appendix in our previous paper.[18] The same holds for multiphysical problems. For the proposed null-space reduction scheme, the construction of a consistent null-space matrix can now be done in a straight-forward manner once the global stiffness matrix of the monolithic macro-micro system has been generated.

ACKNOWLEDGMENTS

Support for this research was provided by the Deutsche Forschungsgemeinschaft (DFG), Germany under grant HE5943/24-1 and HE5943/26-1. This support is gratefully acknowledged.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ENDNOTES

*As shown in Reference 18, we need a higher-order Taylor approximation in the case of higher-order problems.

†Taken from Reali and Hughes.[29]

‡Analogous to the introductory considerations, $\boldsymbol{q} \in \mathbb{R}^{n \cdot n_{\text{node}}}$ and $\tilde{\boldsymbol{w}} \in \mathbb{R}^{n \cdot n_{\text{micro}}}$ summarize in the following the node degrees of freedom $\boldsymbol{q}_A$ and $\tilde{\boldsymbol{w}}_B$ respectively.

§In particular, we compute for the direct solver the factorization of the block diagonal matrix $\boldsymbol{L}$ and solve afterwards $\boldsymbol{L}^{-1}\boldsymbol{E}$. For an iterative solver, an incomplete factorization can be applied.

¶Through periodicity, fixating one corner also fixates opposing corners and thus all corners at once.

ORCID

Christian Hesch https://orcid.org/0000-0002-4614-2202

REFERENCES

1. Maugin GA, Metrikine AV. *Mechanics of Generalized Continua*. Springer; 2010.
2. Altenbach H, Maugin GA, Erofeev V. *Mechanics of Generalized Continua*. Vol 7. Springer; 2011.
3. Bertram A, Forest S. *Mechanics of Strain Gradient Materials*. Springer; 2020.
4. Hill R. The elastic behaviour of a crystalline aggregate. *Proce Phys Soc A*. 1952;65(5):349.
5. Hashin Z, Shtrikman S. A variational approach to the theory of the elastic behaviour of multiphase materials. *J Mech Phys Solids*. 1963;11(2):127-140.
6. Hill R. A self-consistent mechanics of composite materials. *J Mech Phys Solids*. 1965;13(4):213-222.
7. Mori T, Tanaka K. Average stress in matrix and average elastic energy of materials with misfitting inclusions. *Acta Metall*. 1973;21(5):571-574.
8. Willis JR. Bounds and self-consistent estimates for the overall properties of anisotropic composites. *J Mech Phys Solids*. 1977;25:185-202.
9. Smit RJM, Brekelmans WAM, Meijer HEH. Prediction of the mechanical behavior of nonlinear heterogeneous systems by multi-level finite element modeling. *Comput Methods Appl Mech Eng*. 1998;155(1-2):181-192.
10. Miehe C, Schröder J, Schotte J. Computational homogenization analysis in finite plasticity simulation of texture development in polycrystalline materials. *Comput Methods Appl Mech Eng*. 1999;171(3-4):387-418.
11. Feyel F, Chaboche J-L. FE2 multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials. *Comput Methods Appl Mech Eng*. 2000;183(3-4):309-330.
12. Kouznetsova VG, Brekelmans WAM, Baaijens FPT. An approach to micro-macro modeling of heterogeneous materials. *Comput Mech*. 2001;27(1):37-48.
13. Schröder J. *A Numerical Two-Scale Homogenization Scheme: The FE^2-Method*. Springer; 2014.
14. Matouš K, Geers MGD, Kouznetsova VG, Gillman A. A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials. *J Comput Phys*. 2017;330:192-220.
15. Temizer İ. On the asymptotic expansion treatment of two-scale finite thermoelasticity. *Int J Eng Sci*. 2012;53:74-84.
16. Temizer İ, Wriggers P. Homogenization in finite thermoelasticity. *J Mech Phys Solids*. 2011;59(2):344-372.
17. Keip M-A, Steinmann P, Schröder J. Two-scale computational homogenization of electro-elasticity at finite strains. *Comput Methods Appl Mech Eng*. 2014;278:62-79.
18. Schmidt F, Krüger M, Keip M-A, Hesch C. Computational homogenization of higher-order continua. *Int J Numer Methods Eng*. 2022;123(11):2499-2529.
19. Hughes TJR, Feijoo GR, Mazzei L, Quincy JB. The variational multiscale method-a paradigm for computational mechanics. *Comput Methods Appl Mech Eng*. 1998;166:3-24.
20. Larson MG, Målqvist A. Adaptive variational multiscale methods based on a posteriori error estimation: Energy norm estimates for elliptic problems. *Comput Methods Appl Mech Eng*. 2007;196(21):2313-2324.
21. Kaessmair S, Runesson K, Steinmann P, Jänicke R, Larsson F. Variationally consistent computational homogenization of chemomechanical problems with stabilized weakly periodic boundary conditions. *Int J Numer Methods Eng*. 2021;122(22):6429-6454.
22. Zienkiewicz OC, Taylor RL. *The Finite Element Method for Solid and Structural Mechanics*. 6th ed. Butterworth Heinemann; 2005.
23. Zienkiewicz OC, Qu S, Taylor RL, Nakazawa S. The patch test for mixed formulations. *Int J Numer Methods Eng*. 1986;23:1873-1883.
24. Bonet J, Gil AJ, Ortigosa R. A computational framework for polyconvex large strain elasticity. *Comput Methods Appl Mech Eng*. 2015;283:1061-1094.
25. Lange N, Hütter G, Kiefer B. An efficient monolithic solution scheme for FE^2 problems. *Comput Methods Appl Mech Eng*. 2021;382:113886.
26. Gross C, Krause R. On the convergence of recursive trust-region methods for multiscale nonlinear optimization and applications to nonlinear mechanics. *SIAM J Numer Anal*. 2009;47(4):3044-3069.

27. Betsch P, Janz A, Hesch C. A mixed variational framework for the design of energy–momentum schemes inspired by the structure of polyconvex stored energy functions. *Comput Methods Appl Mech Eng*. 2018;335:660-696.
28. Taylor GI. Plastic strain in metals. *Plastic Strain Metals*. 1938;62:307-324.
29. Reali A, Hughes TJR. *An Introduction to Isogeometric Collocation Methods*. Springer Vienna; 2015:173-204.
30. Schuß S, Glas S, Hesch C. Non-linear space-time elasticity. *Int J Numer Methods Eng*. 2023;124(9):1965-1990.

How to cite this article: Hesch C, Schmidt F, Schuß S. Variational formulation and monolithic solution of computational homogenization methods. *Int J Numer Methods Eng*. 2024;e7567. doi: 10.1002/nme.7567