An optimal preconditioned FFT-accelerated finite element solver for homogenization

Martin Ladecký [a,*], Richard J. Leute [b], Ali Falsafi [c], Ivana Pultarová [a], Lars Pastewka [b], Till Junge [c], Jan Zeman [a]

[a] Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 166 29 Prague 6, Czech Republic
[b] Department of Microsystems Engineering, University of Freiburg, Georges-Köhler-Allee 103, Freiburg 79110, Germany
[c] Department of Mechanical Engineering, École Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland

ARTICLE INFO

ABSTRACT

We generalize and provide a linear algebra-based perspective on a finite element (FE) homogenization scheme, pioneered by Schneider et al. (2017)[1] and Leuschner and Fritzen (2018)[2]. The efficiency of the scheme is based on a preconditioned, well-scaled reformulation allowing for the use of the conjugate gradient or similar iterative solvers. The geometrically-optimal preconditioner—a discretized Green's function of a periodic homogeneous reference problem—has a block-diagonal structure in the Fourier space which permits its efficient inversion using fast Fourier transform (FFT) techniques for generic regular meshes. This implies that the scheme scales as $\mathcal{O}(n \log(n))$, like FFT, rendering it equivalent to spectral solvers in terms of computational efficiency. However, in contrast to classical spectral solvers, the proposed scheme works with FE shape functions with local supports and does not exhibit the Fourier ringing phenomenon. We show that the scheme achieves a number of iterations that are almost independent of spatial discretization. The scheme also scales mildly with phase contrast. We also discuss the equivalence between our displacement-based scheme and the recently proposed strain-based homogenization technique with finite-element projection.

© 2023 Published by Elsevier Inc.

1. Introduction

Complex macroscopic phenomena such as deformations during plastic yielding or simulating damage to materials are governed by the nonlinear behavior of such materials at meso-, micro-, or nanoscales. This intrinsic multiscale aspect of the behavior of materials has created a demand for the development of specialized techniques that bridge scales [3–5]. We focus here on an image-based homogenization technique [6] that combines the characterization of material microstructures using high-resolution images (originating, e.g., from micro-computed tomography [7] or geometry-based models [8]) and the numerical solution of an underlying partial differential equation (PDE) with a coefficient defined on a regular grid and typically involving periodic boundary conditions.

The solution of such PDEs discretized with the conventional finite element (FE) is challenging even in the simplest scalar elliptic case, because it results in a system of equations with millions to billions of unknowns [9, Section 7.6]. In this regard,

---

iterative solvers are clearly preferable to direct solvers in terms of computation, because they have lower memory footprints and are faster than direct solvers, with the conjugate gradient (CG) method [10] being the optimal iterative solver for symmetric positive definite system matrices to the best of our knowledge. However, the convergence behavior of the CG method depends on the spectral properties of the linear system matrix and deteriorates as FE mesh size decreases [9, Section 7.7].

More than two decades ago, Moulinec and Suquet in their seminal works [11,12], proposed a method that resolved these issues. According to the original interpretation of their method, fixed-point iterations involving convolution with the Green's function of an auxiliary homogeneous problem are employed with data and unknowns defined directly on an input grid. This method is suitable for high resolution homogenization problems thanks to the efficient implementation of the convolution step using the fast Fourier transform (FFT) algorithm [13] and a number of iterations that does not depend upon mesh size.

In the community investigating the computational mechanics of materials, these features attracted great interest, as documented in three recent surveys by Schneider [14], Lucarini et al. [15], and Gierden et al. [16]. We outline the developments most relevant to our work directly below and refer interested readers to Schneider [14], Lucarini et al. [15], Gierden et al. [16] for more in-depth discussion of FFT-based methods.

Conjugate gradient solvers As reported independently by Brisard and Dormieux [17] and Zeman et al. [18], the original spectral scheme [11,12] can be further accelerated when replacing the fixed-point algorithm with the CG method. These computational observations were confirmed by Brisard and Dormieux [19], who showed that their computational scheme [17] follows from the Ritz discretization of the Hashin–Shtrikman variational principles. Vondřejc et al. [20] also showed that the Zeman et al.'s [18] computational scheme follows from the Fourier–Galerkin discretization of the underlying PDE. These results directly extended to nonlinear problems linearized by Newton's method, as first reported by Gélébart and Mondon-Cancel [21] and Kabel et al. [22] for Green's function framework and by Zeman et al. [23] and de Geus et al. [24] for the Fourier–Galerkin framework.

Oscillations Because stress or strain fields may exhibit discontinuities at the interphases between different phases of materials, discretizing the problem by Fourier trigonometric polynomials results in spurious numerical oscillations (also referred to as "Fourier ringing artifacts" in Section 2.5 of [14]) that pollute the approximate results. To reduce these oscillations, Kaßbohm et al. [25] smoothed the data for the material under investigation, and Shanthraj et al. [26] filtered out high Fourier frequencies from the solution fields. A different approach was used by Willot et al. [27], who considered a modified Green's function obtained from a finite difference discretization. Schneider et al. [28] extended this approach by proposing a staggered grid finite difference approximation to the underlying PDE, with a follow-up study [1] on FE discretization employing linear hexahedral elements. A related approach building on bi/trilinear FE basis functions instead of the Fourier basis was proposed by Leuschner and Fritzen [2]. Most recently, Leute et al. [29] developed a compatibility projection-based method in the spirit of Refs. [23,24] while considering several finite difference- and finite element-based discretization stencils. Further discussion on mitigating the oscillation phenomena can be found in a dedicated comparative study from Ma et al. [30] and in Sections 2.5 and 2.6 of Schneider [14].

Our work We developed an alternative FFT-accelerated, oscillation-free computational homogenization strategy based purely on FE discretization that scales quasilinearly with mesh size. We explain the procedure using a nonlinear small-strain elasticity micromechanical problem[1] discretized on a regular periodic grid with the FE method in Section 2 and linearize it with Newton's method in Section 3. Note that localized support of the FE basis functions directly resolves the oscillation issue, see, e.g., Leute et al. [29]. Thus, no additional artificial adjustments of the data or the solution are needed.

In Section 4, we overcome the main drawback of the FE discretization—deteriorating conditioning of the linear system as the size of the discretization grid increases—using a suitable preconditioner. Similar to Schneider et al. [1], Leuschner and Fritzen [2], we construct the preconditioner from the stiffness matrix of a reference problem with generally anisotropic, spatially uniform data about a material discretized on the same regular grid as the original problem. Using classical results, see e.g., Axelsson and Karátson [32, Section 5.1.2], we can guarantee that the condition number of the preconditioned linear system becomes almost independent of the mesh size. Moreover, employing local ratios of the data of the material used in the problem and the data of the material used in the reference problem, we can localize all individual eigenvalues [33–35]. This may help in creating better predictions about the convergence of the CG method, see, e.g., Gergelits et al. [34, Section 2]. Therefore, the iterative CG solver is an optimized choice for the solution of problems with highly resolved microstructures. The application of the preconditioner is presented in detail in Section 5, with emphasis on reducing its computational complexity using the FFT algorithm [36]. Our formulation is built for an unknown displacement field, like those in Willot et al. [27], Grimm-Strele and Kabel [37], Lucarini and Segurado [38]. However, in Section 6.1, we show that these results are fully transferable to formulations with unknown strain-fields, see e.g., Zeman et al. [23], Schneider et al. [28], Leute et al. [29]. Section 6.2 is devoted to a comparison of our scheme to related developments described by Schneider et al. [1] and Leuschner and Fritzen [2].

We demonstrate the main features of the proposed methodology by examples collected in Section 7 that cover two-dimensional linear thermal conduction (with the necessary adjustments outlined in Appendix A), three-dimensional linear small-strain elasticity, and two-dimensional nonlinear finite-strain elasto-plasticity, and we conclude our work in Section 8.

---

[1] Note that this scheme only considers an imposed macroscopic strain. We expect that for other loading scenarios, the methodology proposed in Lucarini and Segurado [31] can be used in its strain-based version.
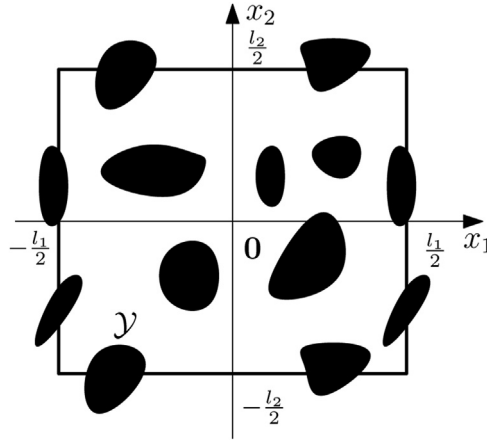
Fig. 1. A rectangular two-dimensional cell $\mathcal{Y} = \left[-\frac{l_1}{2}, \frac{l_1}{2}\right] \times \left[-\frac{l_2}{2}, \frac{l_2}{2}\right]$ with outlined periodic microstructure.

Notation We denote d-dimensional vectors and matrices with boldface letters: $\boldsymbol{a} = (a_\alpha)_{\alpha=1}^d \in \mathbb{R}^d$ or $\boldsymbol{A} = (A_{\alpha\beta})_{\alpha,\beta=1}^d \in \mathbb{R}^{d \times d}$. Matrix-matrix and matrix-vector multiplications are denoted as $\boldsymbol{C} = \boldsymbol{BA}$ and $\boldsymbol{c} = \boldsymbol{Ba}$. Vectors and matrices arising from the discretization are denoted as $\mathbf{a}$ and $\mathbf{A}$, to highlight their special structure. The (I)th component of $\mathbf{a}$ is denoted as $\mathbf{a}[I]$ and (I,J)th component of $\mathbf{A}$ is denoted as $\mathbf{A}[I,J]$. The Euclidean norm of vector $\mathbf{a}$, $\sqrt{\mathbf{a}^\mathsf{T}\mathbf{a}}$, is denoted as $\|\mathbf{a}\|$, and the $\mathbf{A}$-norm of vector $\mathbf{a}$, $\sqrt{\mathbf{a}^\mathsf{T}\mathbf{Aa}}$, is denoted as $\|\mathbf{a}\|_\mathbf{A}$. We consider a general d-dimensional setting throughout the paper. However, for the sake of readability, we use $d = 2$ in the expanded form of matrices, such as in Eq. (1).

2. Nonlinear small-strain elasticity

We consider a d-dimensional rectangular periodic cell $\mathcal{Y} = \prod_{\alpha=1}^d \left[-\frac{l_\alpha}{2}, \frac{l_\alpha}{2}\right]$, of volume $|\mathcal{Y}| = \prod_{\alpha=1}^d l_\alpha$, to be a representative volume element, i.e., a typical material microstructure; see Fig. 1 for an illustration. The symmetries of small-strain elasticity allow us to employ the Mandel notation and reduce the dimension of the second-order strain tensor $\nabla_s \boldsymbol{u} = \frac{1}{2}(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^\mathsf{T}) : \mathcal{Y} \to \mathbb{R}_{\mathrm{sym}}^{d \times d}$ to a vector $\boldsymbol{\partial u} : \mathcal{Y} \to \mathbb{R}^{d_\mathrm{m}}$, where $\boldsymbol{\partial}$ is the symmetrized gradient operator such that, for $d = 2$,

$$\boldsymbol{\partial u} = \begin{pmatrix} (\nabla_s \boldsymbol{u})_{11} \\ (\nabla_s \boldsymbol{u})_{22} \\ \sqrt{2}(\nabla_s \boldsymbol{u})_{12} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x_1} & 0 \\ 0 & \frac{\partial}{\partial x_2} \\ \frac{\sqrt{2}}{2}\frac{\partial}{\partial x_2} & \frac{\sqrt{2}}{2}\frac{\partial}{\partial x_1} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{1}$$

Similarly, a fourth-order tensor $\mathbb{C} : \mathcal{Y} \to \mathbb{R}_{\mathrm{sym}}^{d \times d \times d \times d}$ is represented by a matrix $\mathbf{C} : \mathcal{Y} \to \mathbb{R}^{d_\mathrm{m} \times d_\mathrm{m}}$,

$$\mathbf{C} = \begin{pmatrix} \mathbb{C}_{1111} & \mathbb{C}_{1122} & \sqrt{2}\mathbb{C}_{1112} \\ \mathbb{C}_{2211} & \mathbb{C}_{2222} & \sqrt{2}\mathbb{C}_{2212} \\ \sqrt{2}\mathbb{C}_{1211} & \sqrt{2}\mathbb{C}_{1222} & 2\mathbb{C}_{1212} \end{pmatrix},$$

where the number of components of the symmetrized gradient in the Mandel notation is $d_\mathrm{m} = \dfrac{(d+1)d}{2}$, and indices $\alpha_\mathrm{m}, \beta_\mathrm{m}, \gamma_\mathrm{m} \in \{1, \ldots, d_\mathrm{m}\}$.

In the small-strain micromechanical problem, we split the overall strain $\boldsymbol{\varepsilon} : \mathcal{Y} \to \mathbb{R}^{d_\mathrm{m}}$ into an average strain $\boldsymbol{e} = \frac{1}{|\mathcal{Y}|} \int_\mathcal{Y} \boldsymbol{\varepsilon}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \in \mathbb{R}^{d_\mathrm{m}}$ and a periodically fluctuating field $\boldsymbol{\partial \tilde{u}} : \mathcal{Y} \to \mathbb{R}^{d_\mathrm{m}}$,

$$\boldsymbol{\varepsilon}(\boldsymbol{x}) = \boldsymbol{e} + \boldsymbol{\partial \tilde{u}}(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x} \in \mathcal{Y}.$$

Here, $\boldsymbol{\partial \tilde{u}}$ denotes the symmetrized gradient in the Mandel notation, and the fluctuating displacement field $\tilde{\boldsymbol{u}}$ belongs to the space of admissible functions $\mathcal{V} = \left\{ \tilde{\boldsymbol{v}} : \mathcal{Y} \to \mathbb{R}^d, \ \tilde{\boldsymbol{v}} \text{ is } \mathcal{Y}\text{-periodic}, \int_\mathcal{Y} \tilde{\boldsymbol{v}} \, \mathrm{d}\boldsymbol{x} = \mathbf{0} \right\}$. The governing equations for $\boldsymbol{\partial \tilde{u}}$ are the mechanical equilibrium conditions

$$-\boldsymbol{\partial}^\mathsf{T} \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{e} + \boldsymbol{\partial \tilde{u}}(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x})) = \mathbf{0} \quad \text{for all } \boldsymbol{x} \in \mathcal{Y},$$

in which $\boldsymbol{\sigma} : \mathcal{Y} \times \mathbb{R}^{d_\mathrm{m}} \times \mathbb{R}^g \to \mathbb{R}^{d_\mathrm{m}}$ is the stress field and $\boldsymbol{g} : \mathcal{Y} \to \mathbb{R}^g$ designates the vector of internal parameters. The equilibrium equations are converted to the weak form

$$\int_\mathcal{Y} \boldsymbol{\partial \tilde{v}}(\boldsymbol{x})^\mathsf{T} \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{e} + \boldsymbol{\partial \tilde{u}}(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x})) \, \mathrm{d}\boldsymbol{x} = 0 \qquad \text{for all } \tilde{\boldsymbol{v}} \in \mathcal{V}, \tag{2}$$
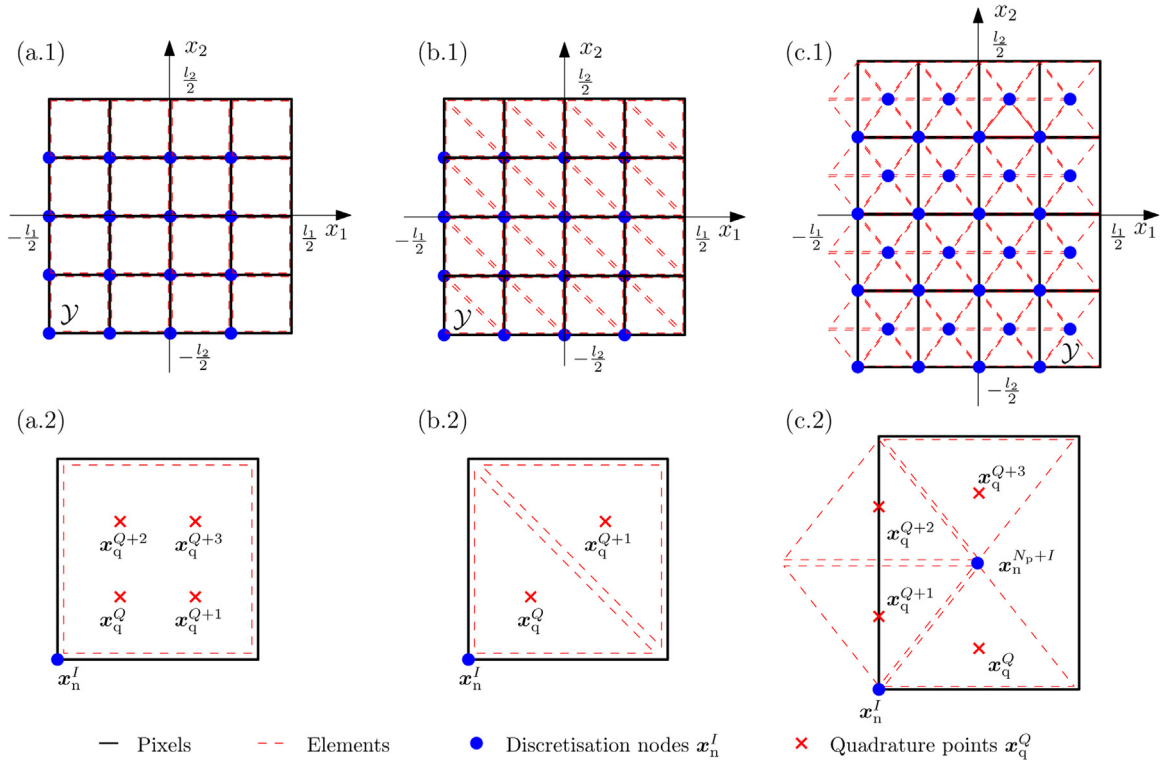
Fig. 2. Example of regular periodic FE grids with associated discretization stencils for a two-dimensional cell $\mathcal{Y}$. All grids consist of 16 pixels ($N_\mathrm{p} = 16$). Row **(1)** shows: (a.1) a grid with 16 discretization nodes ($N_\mathrm{I} = 16$) and quadrature points ($N_\mathrm{Q} = 64$), (b.1) a grid with 16 discretization nodes ($N_\mathrm{I} = 16$) and 32 quadrature points ($N_\mathrm{Q} = 32$), (c.1) a grid with 32 discretization nodes ($N_\mathrm{I} = 32$) and 64 quadrature points ($N_\mathrm{Q} = 64$). Row **(2)** shows: (a.2) a one-node stencil ($N_\mathrm{n} = 1$) with one bilinear rectangular element and four quadrature points with quadrature weights $w^Q = \frac{1}{4}V_\mathrm{p}$, (b.2) a one-node stencil ($N_\mathrm{n} = 1$) with two linear triangular elements and two quadrature points with quadrature weights $w^Q = \frac{1}{2}V_\mathrm{p}$, (c.2) a two-node stencil ($N_\mathrm{n} = 2$) with four linear triangular elements and four quadrature points with quadrature weights $w^Q = \frac{1}{4}V_\mathrm{p}$. Here, V_p denotes pixel volume, such that $V_\mathrm{p}N_\mathrm{p} = |\mathcal{Y}|$.

where $\tilde{\boldsymbol{v}}$ is the test displacement field. The weak form (2) serves as the starting point for the FE method.

3. Finite element discretization

For the discretization of the weak form (2), we use a uniform mesh and conforming FE basis functions. In our setting, the discretization mesh does not necessarily follow the regular pixel/voxel structure, but can correspond to a space-filling pattern of finite elements; see the first row in Fig. 2. The discretization mesh is generated by a periodic repetition of a discretization stencil in cell $\mathcal{Y}$; see the second row in Fig. 2. Such flexibility in discretization is useful, e.g., for strain-softening damage or for plasticity material models that exhibit sensitivity to mesh-grid anisotropy.

Strain and stress fields are evaluated at quadrature points $\boldsymbol{x}_\mathrm{q}^Q$, $Q \in \{1, 2, \ldots, N_\mathrm{Q}\}$, cf. Fig. 2, and the displacement fields are sampled at discretization nodes $\boldsymbol{x}_\mathrm{n}^I$, $I \in \{1, 2, \ldots, N_\mathrm{I}\}$. The number of discretization nodes $N_\mathrm{I} = N_\mathrm{p}N_\mathrm{n}$ is given by the number of pixel/voxel-associated discretization stencils N_p and the number of nodes per stencil N_n, as explained in Fig. 2. The number of degrees of freedom per stencil is thus dN_n and the total number of degrees of freedom per domain is dN_I.

Following standard FE theory, $\tilde{\boldsymbol{v}}$ and $\tilde{\boldsymbol{u}}$ are approximated by continuous element-wise polynomials $\mathcal{P}_k$ of the degree k; their symmetrized gradients $\partial\tilde{\boldsymbol{v}}$ and $\partial\tilde{\boldsymbol{u}}$ then become element-wise polynomials of the degree up to k. Furthermore, the integral (2) can be approximated with a suitable quadrature rule,

$$
\int_\mathcal{Y} \partial\tilde{\boldsymbol{v}}(\boldsymbol{x})^\top \, \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{e} + \partial\tilde{\boldsymbol{u}}(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x})) \, \mathrm{d}\boldsymbol{x}
$$
$$
\approx \sum_{Q=1}^{N_\mathrm{Q}} \partial\tilde{\boldsymbol{v}}(\boldsymbol{x}_\mathrm{q}^Q)^\top \, \boldsymbol{\sigma}(\boldsymbol{x}_\mathrm{q}^Q, \boldsymbol{e} + \partial\tilde{\boldsymbol{u}}(\boldsymbol{x}_\mathrm{q}^Q), \boldsymbol{g}(\boldsymbol{x}_\mathrm{q}^Q)) \, w^Q,
\tag{3}
$$

where the positions of the quadrature points $\boldsymbol{x}_\mathrm{q}^Q$ and the quadrature weights w^Q depend on the choice of the quadrature rule[2]

---

[2] Note, that under-integrated quadrature rules can be used to reduce the computational memory footprint. However, the quality of the solution field can then deteriorate; see Section 7 for explicit examples.

Every component $\tilde{u}_\alpha$ of the unknown vector $\tilde{\boldsymbol{u}}$ is approximated by a linear combination

$$\tilde{u}_\alpha(\boldsymbol{x}) \approx \tilde{u}_\alpha^N(\boldsymbol{x}) = \sum_{I=1}^{N_I} \tilde{u}_\alpha^N(\boldsymbol{x}_n^I) \phi^I(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x} \in \mathcal{Y},$$

where the coefficients $\tilde{u}_\alpha^N(\boldsymbol{x}_n^I)$ are the nodal values of $\tilde{u}_\alpha^N$ at discretization nodes $\boldsymbol{x}_n^I$ and ϕ^I are FE basis functions. A partial derivative of this approximation

$$\frac{\partial \tilde{u}_\alpha^N(\boldsymbol{x})}{\partial x_\beta} = \sum_{I=1}^{N_I} \tilde{u}_\alpha^N(\boldsymbol{x}_n^I) \frac{\partial \phi^I(\boldsymbol{x})}{\partial x_\beta} \quad \text{for all } \boldsymbol{x} \in \mathcal{Y},$$

evaluated in the quadrature points is given by

$$\frac{\partial \tilde{u}_\alpha^N(\boldsymbol{x}_q^Q)}{\partial x_\beta} = \sum_{I=1}^{N_I} \tilde{u}_\alpha^N(\boldsymbol{x}_n^I) \frac{\partial \phi^I(\boldsymbol{x}_q^Q)}{\partial x_\beta} \quad \text{for } Q = 1, \ldots, N_Q.$$

Therefore, if we store the nodal values of displacement $\tilde{\boldsymbol{u}}(\boldsymbol{x}_n^I)$ into a vector $\tilde{\mathbf{u}} \in \mathbb{R}^{dN_I}$, the gradient vector $\partial\tilde{\mathbf{u}} \in \mathbb{R}^{d_m N_Q}$ at all quadrature points is given with

$$\partial\tilde{\mathbf{u}} = \mathbf{D}\tilde{\mathbf{u}} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \\ \frac{\sqrt{2}}{2}\mathbf{D}_2 & \frac{\sqrt{2}}{2}\mathbf{D}_1 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{u}}_1 \\ \tilde{\mathbf{u}}_2 \end{bmatrix}, \tag{4}$$

where the matrix $\mathbf{D} \in \mathbb{R}^{d_m N_Q \times dN_I}$ consists of sub-matrices of the partial derivatives

$$\mathbf{D}_\beta[Q, I] = \frac{\partial \phi^I(\boldsymbol{x}_q^Q)}{\partial x_\beta} \quad \text{for } Q = 1, \ldots, N_Q \text{ and } I = 1, \ldots, N_I, \tag{5}$$

and $\tilde{\mathbf{u}}_\alpha$ stores values of the displacement in the direction α. Due to the local supports of the basis functions ϕ^I, these sub-matrices exhibit significant sparsity, e.g., for the element-wise linear approximation shown in the middle of Fig. 2, each row of $\mathbf{D}_\beta$ contains only two nonzero entries. Since both the interpolating and quadrature points are periodically distributed in $\mathcal{Y}$, the matrix $\mathbf{D}_\beta$ has a block circulant structure.

Now, the discretized weak form (2) using quadrature (3) can be rewritten in the matrix notation as

$$\tilde{\mathbf{v}}^\mathsf{T} \mathbf{D}^\mathsf{T} \mathbf{W} \boldsymbol{\sigma}(\mathbf{e} + \mathbf{D}\tilde{\mathbf{u}}, \mathbf{g}) = 0 \quad \text{for all } \tilde{\mathbf{v}} \in \mathbb{R}^{dN_I}, \tag{6}$$

where $\tilde{\mathbf{v}}$ stores the nodal values of test displacements, $\mathbf{e} \in \mathbb{R}^{d_m N_Q}$ stands for the discretized average strain, $\boldsymbol{\sigma} : \mathbb{R}^{d_m N_Q} \times \mathbb{R}^{g N_Q} \to \mathbb{R}^{d_m N_Q}$ is a nonlinear map transforming, locally at quadrature points, a vector of discrete strains and internal parameters $\mathbf{g} \in \mathbb{R}^{g N_Q}$ to discrete stresses, and the diagonal matrix $\mathbf{W} \in \mathbb{R}^{d_m N_Q \times d_m N_Q}$

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_m & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{W}_m \end{bmatrix} \tag{7}$$

consists of d_m identical diagonal matrices $\mathbf{W}_m \in \mathbb{R}^{N_Q \times N_Q}$ storing quadrature weights, $\mathbf{W}_m[Q, Q] = w^Q$.

Because vector $\tilde{\mathbf{v}}$ is arbitrary, the discretized weak form (6) is equivalent to a system of discrete nonlinear equilibrium conditions

$$\mathbf{D}^\mathsf{T} \mathbf{W} \boldsymbol{\sigma}(\mathbf{e} + \mathbf{D}\tilde{\mathbf{u}}, \mathbf{g}) = \mathbf{0}. \tag{8}$$

3.1. Linearization

We employ Newton's method to solve the nonlinear system (8) iteratively. For this purpose, the $(i+1)$th approximation of the nodal displacement $\tilde{\mathbf{u}}_{(i+1)} \in \mathbb{R}^{N_I}$ is given by the previous approximation $\tilde{\mathbf{u}}_{(i)} \in \mathbb{R}^{N_I}$ adjusted by a finite displacement increment $\delta\tilde{\mathbf{u}}_{(i+1)} \in \mathbb{R}^{N_I}$,

$$\tilde{\mathbf{u}}_{(i+1)} = \tilde{\mathbf{u}}_{(i)} + \delta\tilde{\mathbf{u}}_{(i+1)},$$

with an initial approximation $\tilde{\mathbf{u}}_{(0)} \in \mathbb{R}^{N_I}$. The displacement increment $\delta\tilde{\mathbf{u}}_{(i+1)}$ follows from the solution of the linear system

$$\underbrace{\mathbf{D}^\mathsf{T} \mathbf{W} \mathbf{C}_{(i)} \mathbf{D}}_{\mathbf{K}_{(i)}} \delta\tilde{\mathbf{u}}_{(i+1)} = \underbrace{-\mathbf{D}^\mathsf{T} \mathbf{W} \boldsymbol{\sigma}(\mathbf{e} + \mathbf{D}\tilde{\mathbf{u}}_{(i)}, \boldsymbol{g}_{(i)})}_{\mathbf{b}_{(i)}}, \tag{9}$$

where the algorithmic tangent matrix $\mathbf{C}_{(i)} = \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}}(\mathbf{e} + \mathbf{D}\tilde{\mathbf{u}}_{(i)}, \mathbf{g}_{(i)}) \in \mathbb{R}^{d_m N_Q \times d_m N_Q}$,

$$\mathbf{C}_{(i)} = \begin{bmatrix} \mathbf{C}_{(i)11} & \mathbf{C}_{(i)12} & \mathbf{C}_{(i)13} \\ \mathbf{C}_{(i)21} & \mathbf{C}_{(i)22} & \mathbf{C}_{(i)23} \\ \mathbf{C}_{(i)31} & \mathbf{C}_{(i)32} & \mathbf{C}_{(i)33} \end{bmatrix},$$

is obtained from the constitutive tangent $\mathbf{C}_{(i)}(\boldsymbol{x}) = \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{x}, \boldsymbol{e} + \partial \tilde{\boldsymbol{u}}_{(i)}(\boldsymbol{x}), \mathbf{g}_{(i)}(\boldsymbol{x}))$, evaluated at the quadrature points. Therefore, the sub-matrices $\mathbf{C}_{(i)\alpha_m\beta_m} \in \mathbb{R}^{N_Q \times N_Q}$ are diagonal with entries $\mathbf{C}_{(i)\alpha_m\beta_m}[Q, Q] = C_{(i)\alpha_m\beta_m}(\boldsymbol{x}_q^Q)$. Traditionally, $\mathbf{K}_{(i)} \in \mathbb{R}^{dN_I \times dN_I}$ denotes the matrix of the linear system (9), and $\mathbf{b}_{(i)} \in \mathbb{R}^{dN_I}$ stands for the right-hand side of (9).

4. Preconditioning

Recall that we focus on micromechanical problems with a finely described microstructure that involves many degrees of freedom dN_I. We aim to use an optimal matrix-free iterative method to find the solution of the linear system (9). The system matrix $\mathbf{K}_{(i)}$ is symmetric and positive definite for the symmetric algorithmic tangent $\mathbf{C}_{(i)}$, which renders the CG method as the method of choice, when combined with an appropriate preconditioner. For non-symmetric algorithmic tangents, an alternative iterative linear solver must be used, e.g., the generalized minimal residual method (GMRES). This section discusses the construction of a Green's function preconditioner and its properties in the FE framework.

4.1. Reference material-based preconditioner

The idea of preconditioning, see, e.g., Golub and Van Loan [13, Section 10.3] and Saad [39, Chapters 9 and 10], is based on assumptions that the matrix of the preconditioned linear system

$$\mathbf{M}_{(i)}^{-1}\mathbf{K}_{(i)}\delta\tilde{\mathbf{u}}_{(i+1)} = \mathbf{M}_{(i)}^{-1}\mathbf{b}_{(i)}, \tag{10}$$

has more favorable spectral properties than the original system $\mathbf{K}_{(i)}\delta\tilde{\mathbf{u}}_{(i+1)} = \mathbf{b}_{(i)}$. At the same time, the preconditioning matrix $\mathbf{M}_{(i)} \in \mathbb{R}^{dN_I \times dN_I}$ should be relatively easy to invert, such that the faster convergence of the iterative method compensates for the computational overhead of the preconditioning. Please note that system matrix $\mathbf{M}_{(i)}^{-1}\mathbf{K}_{(i)}$ is no longer symmetric. However, for symmetric $\mathbf{M}_{(i)}$ and $\mathbf{K}_{(i)}$, system (10) is equivalent with the system preconditioned in the symmetric form $\mathbf{M}_{(i)}^{-1/2}\mathbf{K}_{(i)}\mathbf{M}_{(i)}^{-1/2}\delta\mathbf{z}_{(i+1)} = \mathbf{M}_{(i)}^{-1/2}\mathbf{b}_{(i)}$, where $\delta\mathbf{z}_{(i+1)} = \mathbf{M}_{(i)}^{1/2}\delta\tilde{\mathbf{u}}_{(i+1)}$. The latter form is in fact solved using the PCG method; see Saad [39, Section 9.2.1] for more details. Nonetheless, we prefer a notation with left preconditioning (10) for brevity.

Our approach is based on a preconditioner constructed in the same manner as the original matrix of the linear system (9),

$$\mathbf{M}_{(i)} = \mathbf{K}_{(i)}^{\mathrm{ref}} = \mathbf{D}^{\mathsf{T}}\mathbf{W}\mathbf{C}_{(i)}^{\mathrm{ref}}\mathbf{D} \in \mathbb{R}^{dN_I \times dN_I}, \tag{11}$$

where the reference algorithmic tangent matrix $\mathbf{C}_{(i)}^{\mathrm{ref}} \in \mathbb{R}^{d_m N_Q \times d_m N_Q}$ corresponds to spatially uniform material data $\mathbf{C}_{(i)}^{\mathrm{ref}} \in \mathbb{R}^{d_m \times d_m}$. Finally, substituting (11) into (10) leads to the preconditioned linear system

$$\left(\mathbf{K}_{(i)}^{\mathrm{ref}}\right)^{-1}\mathbf{K}_{(i)}\delta\tilde{\mathbf{u}}_{(i+1)} = \left(\mathbf{K}_{(i)}^{\mathrm{ref}}\right)^{-1}\mathbf{b}_{(i)}, \tag{12}$$

referred to as the "reference material-based preconditioned problem" below. Notice that the spectrum of $\mathbf{K}_{(i)}^{\mathrm{ref}}$ contains null eigenvalue(s), associated with the infinitesimal rigid body modes; thus, instead of the inverse of $\mathbf{K}_{(i)}^{\mathrm{ref}}$, we consider its (Moore–Penrose) pseudo-inverse,[3] but we still denote it with $\left(\mathbf{K}_{(i)}^{\mathrm{ref}}\right)^{-1}$ for notation simplicity.

In the next section, we explain why we advocate this choice for the preconditioner. First, we derive a computationally efficient pseudo-inverse of $\mathbf{K}_{(i)}^{\mathrm{ref}}$ and second, we explain how preconditioning impacts the spectral properties of the matrix of the system (12).

4.2. Fourier pseudo-inversion of

Regular FE discretization of the problem with periodic boundary conditions leads to the same stencil for every pixel. Thus, for the uniform $\mathbf{C}_{(i)}^{\mathrm{ref}}$ in the whole $\mathcal{Y}$ (at every quadrature point $\boldsymbol{x}_q^Q$), the resulting preconditioning matrix $\mathbf{K}_{(i)}^{\mathrm{ref}} \in \mathbb{R}^{dN_n N_p \times dN_n N_p}$,

$$\mathbf{K}_{(i)}^{\mathrm{ref}} = \begin{bmatrix} \mathbf{K}_{(i)11}^{\mathrm{ref}} & \mathbf{K}_{(i)12}^{\mathrm{ref}} \\ \mathbf{K}_{(i)21}^{\mathrm{ref}} & \mathbf{K}_{(i)22}^{\mathrm{ref}} \end{bmatrix} \in \mathbb{R}^{2N_p \times 2N_p}, \quad (\text{for } dN_n = 2) \tag{13}$$

---

[3] For details about the Moore–Penrose pseudo-inverse, refer to Golub and Van Loan [13].
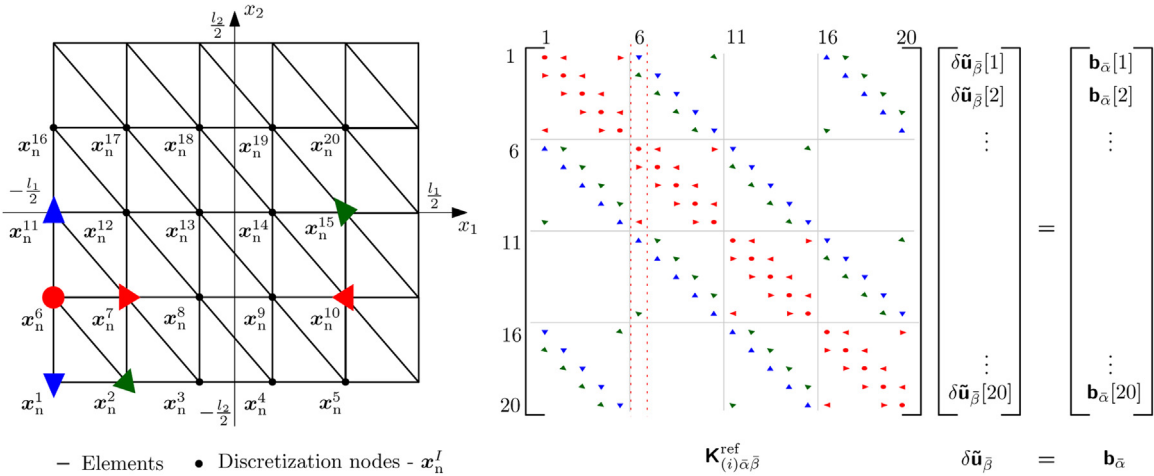
Fig. 3. The block-circulant structure of block $\mathsf{K}^{\mathrm{ref}}_{(i)\bar\alpha\bar\beta}$ from the preconditioner $\mathsf{K}^{\mathrm{ref}}_{(i)}$ for spatially uniform material data $\mathbf{C}^{\mathrm{ref}}_{(i)}$ and the periodic boundary condition. The two-dimensional ($d = 2$) discretization grid consisting of 20 pixels ($N_\mathrm{p} = 20$) with a one-node stencil ($N_\mathrm{n} = 1$) and 20 discretization nodes ($N_\mathrm{l} = 20$) is shown left. Contributions of unit nodal displacement $\delta\tilde{\mathbf{u}}_{\bar\beta}[l] = 1$ to nodal components of the right-hand side vector, graphically shown in node $\boldsymbol{x}^6_\mathrm{n}$, are given as follows: (●) self contribution, contributions (▶) to the right node, (◀) to the left node, (▼) to the upper left node, (▲) to the upper node, (▼) to the bottom node, and (◀) to the bottom right node.

consists of $(dN_\mathrm{n})^2$ block-circulant blocks $\mathsf{K}^{\mathrm{ref}}_{(i)\bar\alpha\bar\beta} \in \mathbb{R}^{N_\mathrm{p} \times N_\mathrm{p}}$, where $\bar\alpha, \bar\beta \in \{1, \ldots, dN_\mathrm{n}\}$. All row vectors of a block-circulant block $\mathsf{K}^{\mathrm{ref}}_{(i)\bar\alpha\bar\beta}$ contain the same information, and each row is block-periodically shifted with respect to the preceding one. This directly reflects the periodically repeated discretization pattern; recall Fig. 2, and that the action of $\mathsf{K}^{\mathrm{ref}}_{(i)\bar\alpha\bar\beta}$ is a discrete convolution of the displacement $\delta\tilde{\mathbf{u}}_{\bar\beta}$ with the discretization kernel, as schematically shown in Fig. 3.

Note that in the one-dimensional ($d = 1$) case with one node per interval ($N_\mathrm{n} = 1$), $\mathsf{K}^{\mathrm{ref}}_{(i)}$ has only one circulant block, $\mathsf{K}^{\mathrm{ref}}_{(i)} = \mathsf{K}^{\mathrm{ref}}_{(i)11}$. The block structure of $\mathsf{K}^{\mathrm{ref}}_{(i)}$ appears whenever more than one type of degree of freedom is involved, i.e., $d > 1$, or $N_\mathrm{n} > 1$.

To make the inversion of $\mathsf{K}^{\mathrm{ref}}_{(i)}$ efficient, let us define the discrete d-dimensional Fourier transform matrix $\mathbf{F} \in \mathbb{R}^{N_\mathrm{p} \times N_\mathrm{p}}$ such that $\mathbf{F}^\mathsf{H} = \mathbf{F}^{-1}$, where $\mathbf{F}^\mathsf{H}$ is the conjugate transpose of $\mathbf{F}$. Then the Fourier counterpart

$$\widehat{\mathsf{K}}^{\mathrm{ref}}_{(i)\bar\alpha\bar\beta} = \mathbf{F}\mathsf{K}^{\mathrm{ref}}_{(i)\bar\alpha\bar\beta}\mathbf{F}^\mathsf{H}$$

to any block-circulant $\mathsf{K}^{\mathrm{ref}}_{(i)\bar\alpha\bar\beta}$ is diagonal, and has the same spectrum (eigenvalues) as $\mathsf{K}^{\mathrm{ref}}_{(i)\bar\alpha\bar\beta}$, see, e.g., Frazier [40]. Therefore, $\widehat{\mathsf{K}}^{\mathrm{ref}}_{(i)}$ is block-diagonal and cheaply (pseudo) invertible

$$\left(\mathsf{K}^{\mathrm{ref}}_{(i)}\right)^{-1} = \mathbf{F}^\mathsf{H}_d\left(\widehat{\mathsf{K}}^{\mathrm{ref}}_{(i)}\right)^{-1}\mathbf{F}_d = \begin{bmatrix} \mathbf{F}^\mathsf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}^\mathsf{H} \end{bmatrix}\begin{bmatrix} \widehat{\mathsf{K}}^{\mathrm{ref}}_{(i)11} & \widehat{\mathsf{K}}^{\mathrm{ref}}_{(i)12} \\ \widehat{\mathsf{K}}^{\mathrm{ref}}_{(i)21} & \widehat{\mathsf{K}}^{\mathrm{ref}}_{(i)22} \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{bmatrix}, \tag{14}$$

where $\mathbf{F}_d = I_{dN_\mathrm{n}} \otimes \mathbf{F}$ and $I_{dN_\mathrm{n}} \in \mathbb{R}^{dN_\mathrm{n} \times dN_\mathrm{n}}$ is the identity matrix. The expanded form in (14) applies for $dN_\mathrm{n} = 2$.

Finally, inserting (14) as the preconditioner in (12) leads to

$$\underbrace{\mathbf{F}^\mathsf{H}_d\left(\widehat{\mathsf{K}}^{\mathrm{ref}}_{(i)}\right)^{-1}\mathbf{F}_d}_{\left(\mathsf{K}^{\mathrm{ref}}_{(i)}\right)^{-1}}\mathbf{K}_{(i)}\delta\tilde{\mathbf{u}}_{(i+1)} = \underbrace{\mathbf{F}^\mathsf{H}_d\left(\widehat{\mathsf{K}}^{\mathrm{ref}}_{(i)}\right)^{-1}\mathbf{F}_d}_{\left(\mathsf{K}^{\mathrm{ref}}_{(i)}\right)^{-1}}\mathbf{b}_{(i)}, \tag{15}$$

which reads in the expanded form as

$$\underbrace{\mathbf{F}^H_d\left(\mathbf{F}_d\mathbf{D}^T\mathbf{W}\mathbf{C}^{\mathrm{ref}}_{(\mathbf{i})}\mathbf{D}\mathbf{F}^H_d\right)^{-1}\mathbf{F}_d}_{\left(\mathsf{K}^{\mathrm{ref}}_{(i)}\right)^{-1}}\underbrace{\mathbf{D}^T\mathbf{W}\mathbf{C}_{(i)}\mathbf{D}}_{\mathbf{K}_{(i)}}\delta\tilde{\mathbf{u}}_{(i+1)}$$

$$= -\underbrace{\mathbf{F}^H_d\left(\mathbf{F}_d\mathbf{D}^T\mathbf{W}\mathbf{C}^{\mathrm{ref}}_{(\mathbf{i})}\mathbf{D}\mathbf{F}^H_d\right)^{-1}\mathbf{F}_d}_{\left(\mathsf{K}^{\mathrm{ref}}_{(i)}\right)^{-1}}\underbrace{\mathbf{D}^T\mathbf{W}\sigma\left(\mathbf{e} + \mathbf{D}\tilde{\mathbf{u}}_{(i)}, \boldsymbol{g}_{(i)}\right)}_{-\mathbf{b}_{(i)}}. \tag{16}$$

4.3. Spectrum of the preconditioned problem

To support the claim that the system matrix of the linear system (15) is well conditioned, we rely on the results published recently in Pultarová and Ladecký [33], Gergelits et al. [34], Ladecký et al. [35] that provide simple algorithms for obtaining guaranteed two-sided bounds for all individual eigenvalues of the preconditioned operator by using element-by-element estimates. Note that bounds on extreme eigenvalues obtained by such an element-by-element algorithm were introduced first in Axelsson and Karátson [32], Eijkhout and Vassilevski [41] and have found use, e.g., in algebraic multilevel methods [42]. Recently, motivated by Nielsen et al. [43], Gergelits et al. [34] published a new method yielding the bounds to all individual eigenvalues. This allows not only the condition number of the preconditioned system to be estimated, but also enables its spectrum to be characterized, which can provide more specific insights into the convergence of the CG method; see, e.g., Gergelits et al. [34, Section 2] for more details. In [33,35] an alternative algorithm was presented that can be applied to a variety of problems and discretization methods.

Let us recall Ladecký et al.'s approach [35]. Thanks to the local supports of FE basis functions ϕ^I all eigenvalues of the preconditioned linear system matrix can be estimated (15). For each ϕ^I, we calculate

$$\lambda_I^L = \min_{\mathbf{x}_q^Q \in \text{supp } \phi^I} \lambda_{\min}\left(\left(\mathbf{C}_{(i)}^{\text{ref}}(\mathbf{x}_q^Q)\right)^{-1}\mathbf{C}_{(i)}(\mathbf{x}_q^Q)\right), \quad I = 1, \ldots, N_I,$$

$$\lambda_I^U = \max_{\mathbf{x}_q^Q \in \text{supp } \phi^I} \lambda_{\max}\left(\left(\mathbf{C}_{(i)}^{\text{ref}}(\mathbf{x}_q^Q)\right)^{-1}\mathbf{C}_{(i)}(\mathbf{x}_q^Q)\right), \quad I = 1, \ldots, N_I,$$

where supp ϕ^I denotes the support of ϕ^I, and $\lambda_{\min}, \lambda_{\max}$ are the minimal and maximal generalized eigenvalues, respectively. For element-wise constant materials $\mathbf{C}_{(i)}$ and $\mathbf{C}_{(i)}^{\text{ref}}$, any quadrature point $\mathbf{x}_q^Q$ can be used to evaluate $\lambda_{\min}$ and $\lambda_{\max}$ on an element. Therefore, only one pair $\lambda_{\min}, \lambda_{\max}$ has to be calculated for each element. Considering all λ_I^L and λ_I^U d-times and sorting these two sets into non-decreasing sequences gives the desired lower and upper eigenvalue bounds.

The resulting eigenvalue bounds are therefore independent of the characteristic element diameter h, which suggests that the condition number[4] $\kappa((\mathbf{K}_{(i)}^{\text{ref}})^{-1}\mathbf{K}_{(i)})$ of the preconditioned linear system (15) will be independent of the size of the problem. In contrast to $\kappa(\mathbf{K}_{(i)}) = \mathcal{O}(h^{-2})$ for the unpreconditioned problem, e.g., Johnson [9, Section 7.7]. The ratio between the maximum and minimum eigenvalues of the preconditioned problem (15) will increase with an increasing ratio between extreme eigenvalues of $\mathbf{C}_{(i)}$ (so-called material contrast) and decrease as the data for the reference material $\mathbf{C}_{(i)}^{\text{ref}}$ approach the data for the material $\mathbf{C}_{(i)}$ of the problem. Therefore, we can state that our preconditioner is optimal, or more precisely, as *geometrically optimal*, which emphasizes that by keeping the discretization and changing only the data of the preconditioner, a matrix can be generated in which all eigenvalues are the same, i.e., the condition number is 1. However, in such cases, the inversion of the preconditioner becomes more expensive. The effects of phase contrast and the choice of $\mathbf{C}_{(i)}^{\text{ref}}$ on the CG performance are further illustrated with examples presented in Section 7.2.

5. Implementation

The pseudo-algorithm of the incremental Newton-PCG solver for FE discretization on a regular grid is outlined in Algorithm 1. In the first part, we detail a matrix-free implementation. The second part deals with the assembly of the preconditioner via matrix-free operators, and the third part focuses on the efficient pseudo-inversion of the preconditioner.[5]

5.1. Matrix-free implementation

As mentioned in the previous sections, the explicit matrix structure is useful in terms of explanation, but computations can be performed more efficiently in a matrix-free manner.

The Gradient. Computational efficiency of a matrix-free implementation relies on the fast evaluation of the gradient vector $\boldsymbol{\partial}\tilde{\mathbf{u}} = \mathbf{D}\tilde{\mathbf{u}}$. For regular periodic discretizations, multiplication $\mathbf{D}\tilde{\mathbf{u}}$ can be implemented as a convolution of $\tilde{\mathbf{u}}$ with a short kernel, namely the gradient stencil. To emphasize this, we replace matrix notations $\mathbf{D}$ and $\mathbf{D}^T$ with the (matrix-free) operator notation $\mathcal{D}: \mathbb{R}^{dN_I} \to \mathbb{R}^{d_m N_Q}$ and $\mathcal{D}^T: \mathbb{R}^{d_m N_Q} \to \mathbb{R}^{dN_I}$, such that

$$\mathcal{D}\delta\tilde{\mathbf{u}}_{(i+1)} = \mathbf{D}\delta\tilde{\mathbf{u}}_{(i+1)}, \quad \text{and} \quad \mathcal{D}^T\mathbf{W}\mathbf{C}_{(i)}\mathcal{D}\delta\tilde{\mathbf{u}}_{(i+1)} = \mathbf{D}^T\mathbf{W}\mathbf{C}_{(i)}\mathbf{D}\delta\tilde{\mathbf{u}}_{(i+1)}.$$

These operations are equal from the viewpoint of linear algebra, but algorithmically, $\mathcal{D}$ and $\mathcal{D}^T$ are of linear $\mathcal{O}(N_I)$ cost.

The fast Fourier transform. In the same manner, multiplication with the discrete Fourier transform matrix can be replaced with the forward and the inverse fast Fourier transform algorithm

$$\mathcal{F}\delta\tilde{\mathbf{u}}_{\alpha,(i+1)} = \mathbf{F}\delta\tilde{\mathbf{u}}_{\alpha,(i+1)} \quad \text{and} \quad \mathcal{F}^{-1}\delta\tilde{\mathbf{u}}_{\alpha,(i+1)} = \mathbf{F}^H\delta\tilde{\mathbf{u}}_{\alpha,(i+1)},$$

---

[4] Please note that by the condition number $\kappa((\mathbf{K}_{(i)}^{\text{ref}})^{-1}\mathbf{K}_{(i)})$ we mean the ratio of the largest and the smallest eigenvalues of $(\mathbf{K}_{(i)}^{\text{ref}})^{-1}\mathbf{K}_{(i)}$.

[5] Please note that the matrix-free implementation is one of the options for implementing this FE scheme, and it does not always have to be more efficient than assembling local FE matrices for every element in advance and then multiplying the displacement field directly in the CG. This holds especially for FE with multiple quadrature points, e.g., trilinear elements with 8 quadrature points.

Algorithm 1 Pseudo-algorithm of the displacement-based Newton-PCG solver

1: **Initialize:**
2: $\tilde{\mathbf{u}}_{(0)}$, $\mathbf{e}$ ▷ initial displacement, macroscopic strain
3: η^{NW}, η^{CG} ▷ Newton- and CG-tolerance
4: $it_{\max}^{\mathrm{NW}}$, $it_{\max}^{\mathrm{CG}}$ ▷ max. iterations Newton and CG
5:
6: **for** $i = 0, 1, 2, \ldots, it_{\max}^{\mathrm{NW}}$ **do**   ▷ Newton iteration
7:      $\mathbf{g}_{(i)} = \ldots$ ▷ update internal parameters
8:      $\mathbf{b}_{(i)} = -\mathcal{D}_{\mathbf{W}}^{\mathsf{T}} \boldsymbol{\sigma}(\mathbf{e} + \mathcal{D}\tilde{\mathbf{u}}_{(i)}, \mathbf{g}_{(i)})$ ▷ right-hand side
9:      $\mathbf{C}_{(i)} = \dfrac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}}(\mathbf{e} + \mathcal{D}\tilde{\mathbf{u}}_{(i)})$ ▷ material tangent
10:      Assembly $(\widehat{\mathbf{K}}_{(i)}^{\mathrm{ref}})^{-1}$ ▷ Preconditioner assembly - Algorithm 2
11:      Solve for $\delta\tilde{\mathbf{u}}_{(i+1)}$ with **PCG**:
12:          $\mathbf{K}_{(i)} \delta\tilde{\mathbf{u}}_{(i+1)} = \mathbf{b}_{(i)}$ with preconditioner $(\mathbf{K}_{(i)}^{\mathrm{ref}})^{-1}$ in $it_{\max}^{\mathrm{CG}}$ steps
13:               or until the termination criterion of Eq. (20) is reached.
14:      $\tilde{\mathbf{u}}_{(i+1)} = \tilde{\mathbf{u}}_{(i)} + \delta\tilde{\mathbf{u}}_{(i+1)}$ ▷ iterative update
15:      **if** $\left\| \delta\tilde{\mathbf{u}}_{(i+1)} \right\| \leq \eta^{\mathrm{NW}} \left\| \tilde{\mathbf{u}}_{(i+1)} \right\|$ **then**
16:          Proceed to line 19   ▷ Newton's method converged
17:      **end if**
18: **end for**
19: **return** $\tilde{\mathbf{u}}_{(i+1)}$

---

of $\mathcal{O}(N_{\mathrm{I}} \log N_{\mathrm{I}})$ complexity.

Quadrature weights. Quadrature weights do not change through the process, so we fuse them with the transposition of the gradient operator

$$\mathcal{D}_{\mathbf{W}}^{\mathsf{T}} = \mathbf{D}^{\mathsf{T}} \mathbf{W},$$

where $\mathcal{D}_{\mathbf{W}}^{\mathsf{T}} : \mathbb{R}^{d_{\mathrm{m}} N_{\mathrm{Q}}} \to \mathbb{R}^{d N_{\mathrm{I}}}$ can be interpreted as a weighted discrete divergence operator.

5.2. Assembly of $\widehat{\mathbf{K}}_{(i)}^{\mathrm{ref}}$

It may be useful to reassemble the preconditioner with updated $\mathbf{C}_{(i)}^{\mathrm{ref}}$, whenever $\mathbf{C}_{(i)}$ significantly changes with respect to the previous Newton step, with $\mathbf{C}_{(i-1)}$. However, the use of matrix-free operators $\mathcal{D}, \mathcal{D}_{\mathbf{W}}^{\mathsf{T}}, \mathcal{F}$ and $\mathcal{F}^{-1}$ prohibits the direct assembly of $\widehat{\mathbf{K}}_{(i)}^{\mathrm{ref}}$ through matrices, like in (13). Thus, we suggest an efficient algorithm for the assembly of $\widehat{\mathbf{K}}_{(i)}^{\mathrm{ref}}$ in analogy to Leuschner and Fritzen [2], that is outlined in Algorithm 2.

First, take a look at (block-periodic) $\bar{\alpha}\bar{\beta}$-block $\mathbf{K}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}} \in \mathbb{R}^{N_{\mathrm{p}} \times N_{\mathrm{p}}}$ of $\mathbf{K}_{(i)}^{\mathrm{ref}} \in \mathbb{R}^{d N_{\mathrm{n}} N_{\mathrm{p}} \times d N_{\mathrm{n}} N_{\mathrm{p}}}$. Thanks to the convolution theorem, the whole diagonal $\mathrm{diag}(\widehat{\mathbf{K}}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}) \in \mathbb{R}^{N_{\mathrm{p}}}$ can be obtained by the FFT of any, say the first, row or, because of the symmetry, column of $\mathbf{K}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}$,

$$\mathrm{diag}(\widehat{\mathbf{K}}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}) = \mathcal{F}(\mathbf{K}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}[1, :])^{\mathsf{T}} = \mathcal{F}(\mathbf{K}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}[:, 1])$$

where a colon indicates a complete column or row, see, e.g., Frazier [40]. Before the FFTs, we have to compute one column $\mathbf{K}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}[1, :]$ for each of $(d N_{\mathrm{n}})^2$ blocks $\mathbf{K}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}$ of $\mathbf{K}_{(i)}^{\mathrm{ref}}$. Consider a unit impulse vector $\mathbf{i}^p \in \mathbb{R}^{d N_{\mathrm{n}} N_{\mathrm{p}}}$ that has only one non-zero element equal to 1 on the pth position. When we apply $\mathbf{K}_{(i)}^{\mathrm{ref}}$ to vector $\mathbf{i}^1$, we obtain the first columns of $d N_{\mathrm{n}}$ blocks $\mathbf{K}_{(i)\bar{\alpha}1}^{\mathrm{ref}}$. From the structure of $\mathbf{K}_{(i)}^{\mathrm{ref}}$ visible in (13) it is obvious that we need $d N_{\mathrm{n}}$ vectors $\mathbf{i}^p$ to obtain all $(d N_{\mathrm{n}})^2$ columns $\mathbf{K}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}[1, :]$, where $p = (\bar{\beta} - 1) N_{\mathrm{p}} + 1$ and $\bar{\beta} \in \{1, \ldots, d N_{\mathrm{n}}\}$. The whole procedure is schematically shown in Fig. 4.

5.3. Pseudo-inverse of $\widehat{\mathbf{K}}_{(i)}^{\mathrm{ref}}$

Once we have all diagonal blocks $\widehat{\mathbf{K}}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}$, we may proceed to the computation of the pseudo-inverse of $\widehat{\mathbf{K}}_{(i)}^{\mathrm{ref}}$. By a proper row and column reordering, it can be seen that the pseudo-inverse of the block diagonal matrix $\widehat{\mathbf{K}}_{(i)}^{\mathrm{ref}}$ is equivalent to the pseudo-inverse of N_{p} (number of pixels/stencils) submatrices

$$\begin{bmatrix} \widehat{\mathbf{K}}_{(i)11}^{\mathrm{ref}}[J, J] & \cdots & \widehat{\mathbf{K}}_{(i)1\bar{\beta}}^{\mathrm{ref}}[J, J] \\ \vdots & \ddots & \vdots \\ \widehat{\mathbf{K}}_{(i)\bar{\alpha}1}^{\mathrm{ref}}[J, J] & \cdots & \widehat{\mathbf{K}}_{(i)\bar{\alpha}\bar{\beta}}^{\mathrm{ref}}[J, J] \end{bmatrix}^{-1} \in \mathbb{R}^{d N_{\mathrm{n}} \times d N_{\mathrm{n}}}, \quad \text{where} \quad J \in \{1, \ldots, N_{\mathrm{p}}\}. \tag{17}$$

Algorithm 2 Pseudo-algorithm of reference material-based preconditioner assembly.

1: **Initialize:**
2: $\mathbf{C}_{(i)}^{\text{ref}}$ $\triangleright$ spatially uniform reference material
3:
4: **for** $\bar{\beta} = 1, \ldots, dN_{\text{n}}$ **do** $\triangleright$ loop over d vectors
5:   $p = (\bar{\beta} - 1)N_{\text{p}} + 1$ $\triangleright$ column index
6:   $\boldsymbol{c}_{\bar{\beta}} = \mathcal{D}_{\mathbf{W}}^{\mathsf{T}} \mathbf{C}_{(i)}^{\text{ref}} \mathcal{D} \mathbf{i}^p$ $\triangleright$ p-th column of $\mathbf{K}_{(i)}^{\text{ref}}$
7:   **for** $\bar{\alpha} = 1, \ldots, dN_{\text{n}}$ **do**
8:    $\text{diag}(\widehat{\mathbf{K}}_{(i)\bar{\alpha}\bar{\beta}}^{\text{ref}}) = \mathcal{F}(\boldsymbol{c}_{\bar{\beta}}[(\bar{\alpha} - 1)N_{\text{p}} + 1 : \bar{\alpha}N_{\text{p}}])$ $\triangleright$ assign to $\widehat{\mathbf{K}}_{(i)\bar{\alpha}\bar{\beta}}^{\text{ref}}$ diagonals
9:   **end for**
10: **end for**
11:                  $\triangleright$ pseudo-inverse of singular submatrix of $\widehat{\mathbf{K}}_{(i)}^{\text{ref}}$

12: $\left(\widehat{\mathbf{K}}_{(i)\bar{\alpha}\bar{\beta}}^{\text{ref}}\right)^{-1}[1,1] = \begin{bmatrix} \widehat{\mathbf{K}}_{(i)\,11}^{\text{ref}}[1,1] & \cdots & \widehat{\mathbf{K}}_{(i)\,1\bar{\delta}}^{\text{ref}}[1,1] \\ \vdots & \ddots & \vdots \\ \widehat{\mathbf{K}}_{(i)\,\bar{\gamma}1}^{\text{ref}}[1,1] & \cdots & \widehat{\mathbf{K}}_{(i)\,\bar{\gamma}\bar{\delta}}^{\text{ref}}[1,1] \end{bmatrix}_{\bar{\alpha}\bar{\beta}}^{\dagger}$ $\triangleright$ 1-th block

13: **for** $J = 2, \ldots, N_{\text{p}}$ **do** $\triangleright$ inverse of remaining submatrices of $\widehat{\mathbf{K}}_{(i)}^{\text{ref}}$

14:   $\left(\widehat{\mathbf{K}}_{(i)\bar{\alpha}\bar{\beta}}^{\text{ref}}\right)^{-1}[J,J] = \begin{bmatrix} \widehat{\mathbf{K}}_{(i)\,11}^{\text{ref}}[J,J] & \cdots & \widehat{\mathbf{K}}_{(i)\,1\bar{\delta}}^{\text{ref}}[J,J] \\ \vdots & \ddots & \vdots \\ \widehat{\mathbf{K}}_{(i)\,\bar{\gamma}1}^{\text{ref}}[J,J] & \cdots & \widehat{\mathbf{K}}_{(i)\,\bar{\gamma}\bar{\delta}}^{\text{ref}}[J,J] \end{bmatrix}_{\bar{\alpha}\bar{\beta}}^{-1}$ $\triangleright$ J-th block

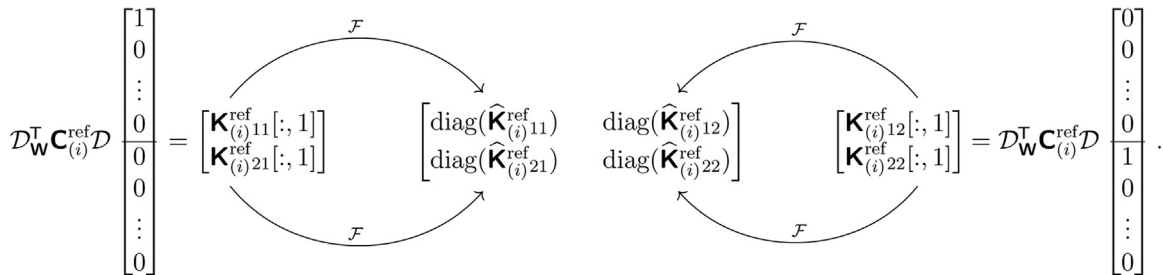15: **end for**

Fig. 4. The schematic procedure for matrix-free assembly of $\widehat{\mathbf{K}}_{(i)}^{\text{ref}}$ for $dN_{\text{n}} = 2$. First columns of blocks $\mathbf{K}_{(i)\,11}^{\text{ref}}$ and $\mathbf{K}_{(i)\,21}^{\text{ref}}$ are obtained as a result of the matrix-free action of $\mathbf{K}_{(i)}^{\text{ref}}$ on the unit impulse vector $\mathbf{i}^1$. Diagonals $\text{diag}(\widehat{\mathbf{K}}_{(i)\,11}^{\text{ref}})$ and $\text{diag}(\widehat{\mathbf{K}}_{(i)\,21}^{\text{ref}})$ are then computed using d-dimensional FFT of $\mathbf{K}_{(i)\,11}^{\text{ref}}[:, 1]$ and $\mathbf{K}_{(i)\,21}^{\text{ref}}[:, 1]$, respectively. By analogy, columns of blocks $\mathbf{K}_{(i)\,12}^{\text{ref}}$ and $\mathbf{K}_{(i)\,22}^{\text{ref}}$ are obtained by the matrix-free action of $\mathbf{K}_{(i)}^{\text{ref}}$ on the unit impulse vector $\mathbf{i}^p$ where $p = (2 - 1)N_{\text{p}} + 1$.

The $(N_{\text{p}} - 1)$ submatrices are of full rank and thus directly invertible. Only one submatrix, corresponding to the zero frequency Fourier mode; this submatrix is singular and has to be treated separately. This block has exactly d null eigenvalues corresponding to d rigid-body modes. We compute the (Moore–Penrose) pseudo-inverse of this block instead of its inversion.[6] The pseudo-inverse can be computed exactly by restriction onto the space orthogonal to the kernel of the singular block. For any specific type of FE and the corresponding discretization stencil, the kernel can be exactly identified.

6. Comparison with related FFT-based schemes

Several FFT-based computational homogenization schemes exist [14,15]. An interested reader may therefore find it useful to compare and place our approach in the context of contemporary literature.

Recall that our approach is derived from the weak form of the mechanical equilibrium condition (2) with an unknown displacement field. Equilibrium (2) is discretized in the standard Galerkin manner with the FE basis functions. The nonlinear nodal equilibrium (8) is linearized by Newton's method, and the system of linear Eq. (9) is solved by the PCG method. For the PCG method, a favorable convergence property is guaranteed by the reference material-based preconditioner (11), whose fast application builds on FFT.

---

[6] Please note that the Moore–Penrose pseudo-inverse is depicted by $\dagger$ in Algorithm 2.

6.1. The connection with strain-based approaches

Unlike DB FE, most spectral methods use strains (gradients) as unknown. SB approaches, like those in Zeman et al. [23], Schneider et al. [28], Leute et al. [29], typically use the projection operator to enforce the compatibility of strain fields. The connection between the DB and SB formulations has been discussed in Leuschner and Fritzen [2, Section 3.3] and of Schneider et al. [1, Section 4.3]. In order for this paper to be self-contained, we explain the link between the DB and SB formulation in our framework. Recall the preconditioned scheme (16),

$$
\underbrace{(\mathbf{D}^{\mathsf{T}}\mathbf{W}\mathbf{C}_{(i)}^{\mathrm{ref}}\mathbf{D})^{-1}}_{(\mathbf{K}_{(i)}^{\mathrm{ref}})^{-1}}\underbrace{\mathbf{D}^{\mathsf{T}}\mathbf{W}\mathbf{C}_{(i)}\mathbf{D}}_{\mathbf{K}_{(i)}}\delta\tilde{\mathbf{u}}_{(i+1)} = -\underbrace{(\mathbf{D}^{\mathsf{T}}\mathbf{W}\mathbf{C}_{(i)}^{\mathrm{ref}}\mathbf{D})^{-1}}_{(\mathbf{K}_{(i)}^{\mathrm{ref}})^{-1}}\underbrace{\mathbf{D}^{\mathsf{T}}\mathbf{W}\boldsymbol{\sigma}(\mathbf{e}+\mathbf{D}\tilde{\mathbf{u}}_{(i)},\mathbf{g}_{(i)})}_{-\mathbf{b}_{(i)}},
\tag{18}
$$

where we omit the FFTs for simplicity. In the case of linear triangles or tetrahedral elements with a single quadrature point per element, all quadrature weights w^Q are equal. Then the multiplication by quadrature weights $\mathbf{W}$ can be left out in (18), leading to

$$
\underbrace{(\mathbf{D}^{\mathsf{T}}\mathbf{C}_{(i)}^{\mathrm{ref}}\mathbf{D})^{-1}}_{(\mathbf{K}_{(i)}^{\mathrm{ref}})^{-1}}\underbrace{\mathbf{D}^{\mathsf{T}}\mathbf{C}_{(i)}\mathbf{D}}_{\mathbf{K}_{(i)}}\delta\tilde{\mathbf{u}}_{(i+1)} = -\underbrace{(\mathbf{D}^{\mathsf{T}}\mathbf{C}_{(i)}^{\mathrm{ref}}\mathbf{D})^{-1}}_{(\mathbf{K}_{(i)}^{\mathrm{ref}})^{-1}}\underbrace{\mathbf{D}^{\mathsf{T}}\boldsymbol{\sigma}(\mathbf{e}+\mathbf{D}\tilde{\mathbf{u}}_{(i)},\mathbf{g}_{(i)})}_{-\mathbf{b}_{(i)}}.
$$

Next, we replace the iterated unknown $\tilde{\mathbf{u}}_{(i)}$ with its gradient $\partial\tilde{\mathbf{u}}_{(i)}$, recognizing that $\partial\tilde{\mathbf{u}}_{(i)} = \mathbf{D}\tilde{\mathbf{u}}_{(i)}$. After multiplying by $\mathbf{D}$ from the left-hand side, we finally obtain

$$
\underbrace{\mathbf{D}(\mathbf{D}^{\mathsf{T}}\mathbf{C}_{(i)}^{\mathrm{ref}}\mathbf{D})^{-1}\mathbf{D}^{\mathsf{T}}}_{\boldsymbol{\Gamma}_{(i)}^{0}}\mathbf{C}_{(i)}\delta\partial\tilde{\boldsymbol{u}}_{(i+1)} = -\underbrace{\mathbf{D}(\mathbf{D}^{\mathsf{T}}\mathbf{C}_{(i)}^{\mathrm{ref}}\mathbf{D})^{-1}\mathbf{D}^{\mathsf{T}}}_{\boldsymbol{\Gamma}_{(i)}^{0}}\boldsymbol{\sigma}(\mathbf{e}+\partial\tilde{\boldsymbol{u}}_{(i)},\mathbf{g}_{(i)}),
\tag{19}
$$

where $\boldsymbol{\Gamma}_{(i)}^{0} : \mathbb{R}^{d_{\mathrm{m}}N_{\mathrm{Q}}} \to \mathbb{R}^{d_{\mathrm{m}}N_{\mathrm{Q}}}$ stands for the discretized periodic Green's operator. Leute et al. [29] showed that by setting $\mathbf{C}_{(i)}^{\mathrm{ref}} = \mathbf{I_s}$, $\boldsymbol{\Gamma}_{(i)}^{0}$ projects an arbitrary field from $\mathbb{R}^{d_{\mathrm{m}}N_{\mathrm{Q}}}$ to its closest compatible part in the least square sense with respect to the L^2-norm.

Therefore, this section demonstrates that schemes (18) and (19) are equivalent and generate equivalent solutions in every step of the CG in exact arithmetic. If the corresponding stopping criteria are used, CG yields the same approximate solutions. Thus, DB and SB formulations converge equivalently, and the only decision-making argument is the efficiency of the implementation.

6.2. The connection with FEM-FFT approaches

To the best of our knowledge, our method is most similar to the linear hexahedral elements (FFT-Q_1 Hex) formulation from Schneider at al. [1] and the Fourier-Accelerated Nodal Solver (FANS) from Leuschner and Fritzen [2]. The novelty of our approach lies in the following:

- *The gradient operator.* Similar to FFT-Q_1 Hex and FANS, the gradient field is derived with respect to the FE approximation. However, we do not express the discrete gradient operator $\mathbf{D}$ in the Fourier space, but keep it in real space. The direct convolution with a short gradient kernel is cheaper than the Fourier convolution via forward and inverse FFTs. We use the Fourier representation only for the efficient inverse of the preconditioner $\mathbf{K}_{(i)}^{\mathrm{ref}}$ as discussed in Section 4.

- *Preconditioner and reference material.* Our preconditioner (11) has the same form as the fundamental solution $\mathbf{G}^0$ contained in the discretized periodic Green's operator $\boldsymbol{\Gamma}_{(i)}^{0}$ of the FFT-Q_1 Hex scheme (equation (16) of [1]), and the fundamental solution $\widehat{\phi}$ in FANS (equation (49) of [2]). Therefore we expect similar conditioning of all three schemes. We provide detailed insight from the perspective of linear algebra. The direct correspondence between the reference material $\mathbf{C}_{(i)}^{\mathrm{ref}}$, material $\mathbf{C}_{(i)}(\boldsymbol{x})$ and the resulting eigenvalues renders the optimization of $\mathbf{C}_{(i)}^{\mathrm{ref}}$ more accessible. The closer the reference material is to the real material of the sample, the better the conditioning of the discretized problem.

- *Discretization grid.* Both FFT-Q_1 Hex and FANS were developed for bi/trilinear FE basis and quadrilateral/hexahedral elements. Their authors mentioned their possible extension to more complex elements, which we present in this paper. In addition, the discretization grid in our method does not have to follow a pixel/voxel structure. We allow for an arbitrary space-filling pattern of elements to be used; recall the patterns in Fig. 2. Further extension of our formulation to FE with higher-order polynomial basis functions is therefore straightforward.

- *Computational complexity.* Computational complexity of FFT-based methods is governed by $\mathcal{O}(n\log n)$ complexity of the FFT. However, in our scheme, we compute two FFTs on dN_{n} displacement fields of size N_{I}, instead of d_{m} strain fields of size N_{Q} in FFT-Q_1. Because the number of strain components d_{m} exceeds the number of displacement components dN_{n} per stencil and the number of quadrature points N_{Q} exceeds the number of discretization nodes N_{I}, our method has smaller computational overhead than the DB methods that evaluate the gradient in the Fourier space and perform FFT on the strain-sized fields. For example, in the case of trilinear hexahedral FEs with 8 quadrature points per element, the savings factor in the number of FFTs is 24.
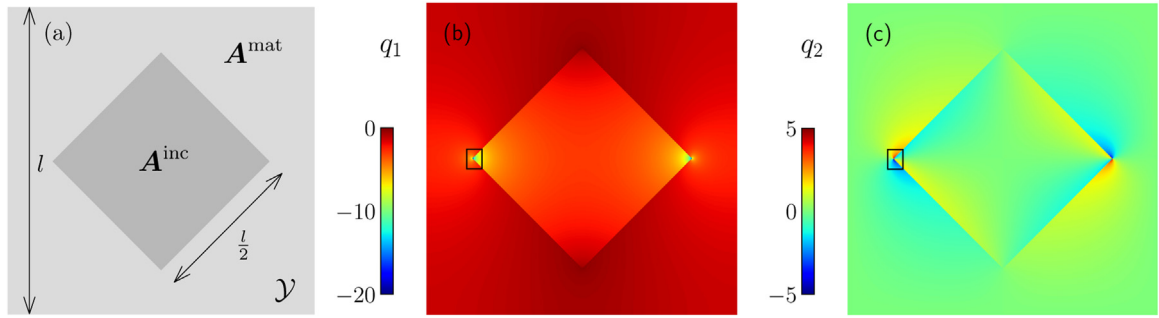
Fig. 5. A linear heat transfer problem from Section 7.1. The square periodic unit cell $\mathcal{Y}$ with a square inclusion (a). The flux field component q_1 (b) and q_2 (c) arising from average temperature gradient $\boldsymbol{e} = [0.01, 0.0]^\mathsf{T}$. Results were obtained with a one-node FE stencil ($N_\mathrm{n} = 1$) with two linear triangular elements discretization and 815 nodes in both directions ($N_\mathrm{l} = 815^2$).

7. Numerical experiments

In the previous sections, we explained the procedure for the problem of small-strain elasticity. However, the application range of this solution strategy is much broader. To highlight this versatility, we demonstrate the numerical behavior of the approach with examples of thermal conductivity, small-strain elasticity, and nonlinear finite-strain problems. The first example demonstrates the oscillation-free character of FE solutions. The second example illustrates the robustness of the approach with respect to material contrast and the effect of the reference material. The third one shows the equivalence of the SB and DB implementations. We compare the (fully- and under-integrated) FE schemes, described in the previous sections, with the (under-integrated) Fourier–Galerkin method taken from Vondřejc et al. [20], Zeman et al. [23], de Geus et al. [24]. Units employed in all examples are consistent.

Our results were obtained using μSpectre software, an open source platform for efficient FFT-based continuum mesoscale modeling, which is freely available at https://gitlab.com/muspectre/muspectre. The software package includes the examples which are described in the following sections.

Termination criteria To obtain comparable results, we have to choose the corresponding termination criteria for both SB and DB schemes. Newton's method stops when the relative norm of the strain increment drops below the tolerance η^NW, $\left\| \delta \boldsymbol{\partial} \tilde{\mathbf{u}}_{(i+1)} \right\| \leq \eta^\mathrm{NW} \left\| \boldsymbol{\partial} \tilde{\mathbf{u}}_{(i+1)} \right\|$, in analogy to Zeman et al. [23], de Geus et al. [24].[7] The PCG solver is stopped when the relative $\left(\mathbf{K}^\mathrm{ref}_{(i)} \right)^{-1}$-norm of the residual drops below the tolerance η^CG,

$$\left\| \mathbf{r}^k_{(i+1)} \right\|_{\left(\mathbf{K}^\mathrm{ref}_{(i)}\right)^{-1}} \leq \eta^\mathrm{CG} \left\| \mathbf{r}^0_{(i+1)} \right\|_{\left(\mathbf{K}^\mathrm{ref}_{(i)}\right)^{-1}}. \tag{20}$$

This choice is motivated by the *optimal* property of PCG to minimize the error energy norm

$$\left\| \mathbf{e}^k_{(i+1)} \right\|_{\mathbf{K}_{(i)}} = \left\| \delta \tilde{\mathbf{u}}_{(i+1)} - \delta \tilde{\mathbf{u}}^k_{(i+1)} \right\|_{\mathbf{K}_{(i)}},$$

where $\delta \tilde{\mathbf{u}}^k_{(i+1)}$ is the approximation of the solution $\delta \tilde{\mathbf{u}}_{(i+1)}$ in kth PCG step. If $\left(\mathbf{K}^\mathrm{ref}_{(i)} \right)^{-1} \mathbf{K}_{(i)} \approx \mathbf{I}$, then the $\left(\mathbf{K}^\mathrm{ref}_{(i)} \right)^{-1}$-norm of the residual $\mathbf{r}^k_{(i+1)} = \mathbf{b}_{(i)} - \mathbf{K}_{(i)} \delta \tilde{\mathbf{u}}^k_{(i+1)}$ approximates the error energy norm,

$$\left\| \mathbf{r}^k_{(i+1)} \right\|_{\left(\mathbf{K}^\mathrm{ref}_{(i)}\right)^{-1}} = {\mathbf{e}^k_{(i+1)}}^\mathsf{T} \mathbf{K}^\mathsf{T}_{(i)} \left(\mathbf{K}^\mathrm{ref}_{(i)}\right)^{-1} \mathbf{K}_{(i)} {\mathbf{e}^k_{(i+1)}}^\mathsf{T} = \left\| \mathbf{e}^k_{(i+1)} \right\|_{\mathbf{K}^\mathsf{T}_{(i)} \left(\mathbf{K}^\mathrm{ref}_{(i)}\right)^{-1} \mathbf{K}_{(i)}}.$$

Additionally, the $\left(\mathbf{K}^\mathrm{ref}_{(i)} \right)^{-1}$-norm of the residual naturally appears in the PCG algorithm, therefore it can be freely obtained.

7.1. Linear steady-state thermal conduction problem

In the first example, we demonstrate the oscillation-free character of gradient fields arising from FE discretization. For this purpose, we reconstruct the benchmark problem from Leuschner and Fritzen [2, Section 3.7.1] or Brisard and Dormieux [17, Section 3.2], where the Fourier–Galerkin methods exhibit significant discretization artifacts.

We consider a scalar problem of linear heat transfer, where we look for the flux field $\boldsymbol{q}$ satisfying the weak balance condition (A.1); see Appendix A for more details. The microstructure is defined by the square periodic unit cell $\mathcal{Y}$, as sketched on the left-hand side of Fig. 5. The composite microstructure consists of an insulating matrix with the conductivity $\mathbf{A}^\mathrm{mat} = 100\,\mathbf{I}$, and a conducting inclusion with the conductivity $\mathbf{A}^\mathrm{inc} = 100\,\mathbf{A}^\mathrm{mat}$. An average temperature gradient $\boldsymbol{e} = [0.01, 0.0]^\mathsf{T}$ is applied. The number of pixels is 815^2, and the material coefficients are constant per pixel. The choice of reference material $\mathbf{A}^\mathrm{ref}$ has

---

[7] Please note that other stopping criterion for Newton's method, such as the divergence norm [44], are also possible. However, in the SB implementation, we have direct access only to the strains and stresses. Therefore, we choose the relative norm of the strain increment.
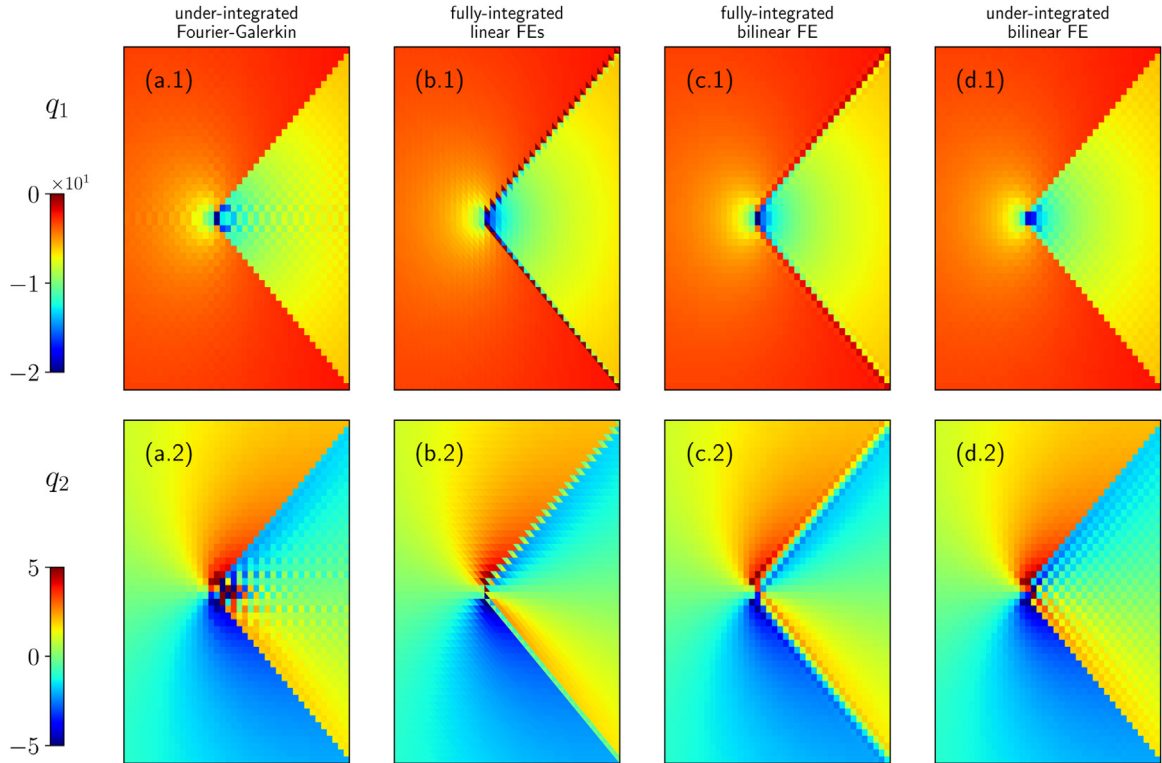
Fig. 6. Local heat flux field components q_1 **(1)** and q_2 **(2)** from the experiment descibed in Section 7.1, obtained with the under-integrated Fourier–Galerkin method is shown in column (a), a one-node FE stencil ($N_n = 1$) with two linear triangular elements and two quadrature points is shown in column (b), a one-node FE stencil ($N_n = 1$) with one bilinear rectangular element and four quadrature points is shown in column (c) and a one-node FE stencil ($N_n = 1$) with one bilinear rectangular element and one quadrature point is shown in column (d).

no effect on the discretization artifacts; thus, we set $\mathbf{A}^{\mathrm{ref}} = \mathbf{I}$ for simplicity. Components of the global flux field $\boldsymbol{q}$ are shown in Fig. 5; q_1 in the middle and q_2 on the right-hand side. The regions of details depicted in Figs. 6 and 7 are highlighted by the black rectangles in Fig. 5.

In Fig. 6, we show the details of heat fluxes for various discretizations: the Fourier–Galerkin method in column (a), the one-node FE stencil ($N_n = 1$) with two linear triangular elements and two quadrature points (Fig. 2(b)) in column (b), the one-node FE stencil ($N_n = 1$) with one bilinear rectangular element and four quadrature points (Fig. 2(a)) in column (c) and the one-node FE stencil ($N_n = 1$) with one bilinear rectangular element and one quadrature point in column (d).

The Fourier–Galerkin method exhibits strong oscillations through the region. The under-integrated FE solution shows checkerboard patterns, while FE solutions of fully-integrated schemes are devoid of oscillations in the interior of the domains occupied by a single phase because the discretization discrepancies remain confined to the vicinity of the phase boundaries. For instance, triangular discretizations reduce the phase boundary discretization artifacts to the two pixel-wide layer around the phase boundary.

The zigzag patterns on the phase boundary arise from the pixel-based geometry. If the elements can capture the interface of the two phases exactly, we do not get any discretization artifacts, as can be seen in Figs. 6 and 7 in column (b). This speaks in favor of using FE over Fourier–Galerkin discretization.

7.2. Small-strain elasticity problem

The second example focuses on the effect of the preconditioner on the number of PCG iterations with respect to the number of discretization nodes N_I and phase contrast ρ. For this purpose, we use Hashin's coated sphere construction adapted from Schneider et al. [28, Section 4.1] and the references therein.

We choose the linear small-strain elastic problem described in Section 2. The three-phase microstructure representing a coated sphere in the matrix with effective material properties is depicted in Fig. 8, with a core radius $r_1 = 0.2$, annulus-shaped coating outer radius $r_2 = 0.4$, and a cubic domain edge length $l = 1$. An average macroscopic strain $\boldsymbol{e} = [1, 0, 0, 0, 0, 0]^{\mathsf{T}}$ is applied. We assume isotropic phases with bulk and shear moduli K_1, G_1 in the core, K_2, G_2 in the coating and $K_{\mathrm{eff}}, G_{\mathrm{eff}}$ in the surrounding matrix. The bulk moduli K_1, K_2 are chosen in a way that the resulting response of the unit cell is equivalent to the response of a homogeneous material with K_{eff}. As a consequence, the bulk moduli K_1, K_2 must be balanced for the particular phase contrast $\rho = K_2/K_1$.
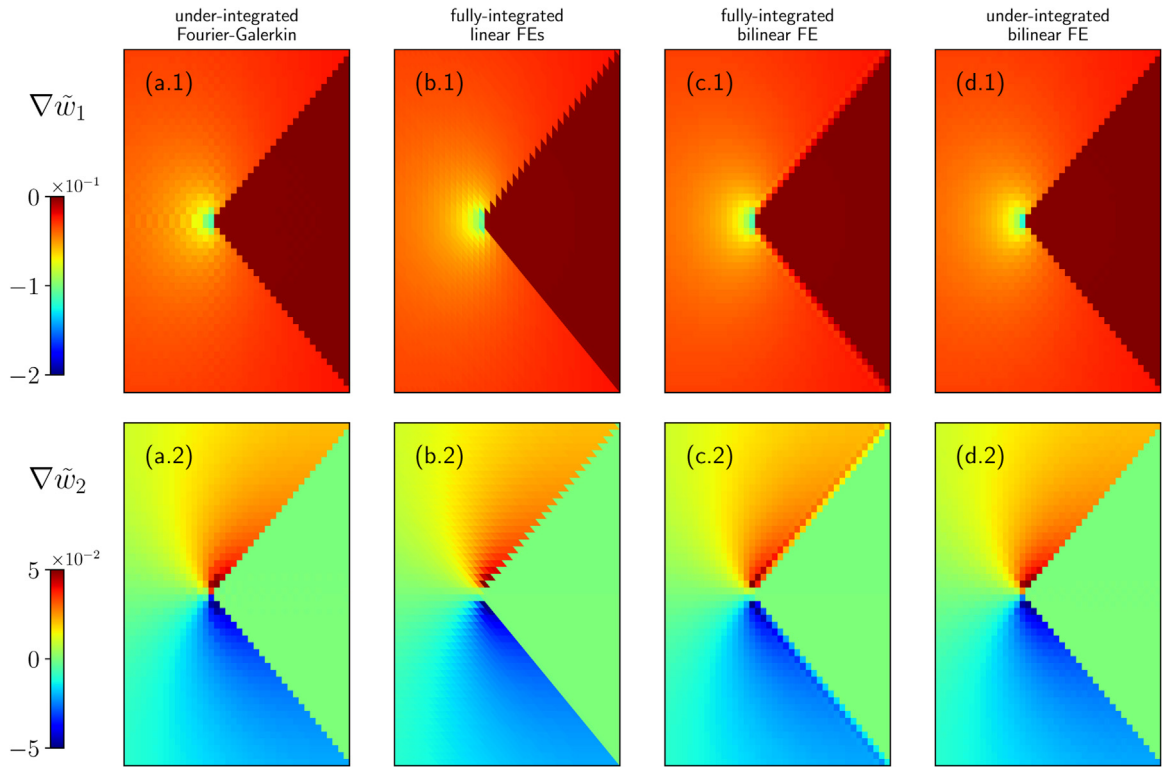
Fig. 7. Local temperature gradient field components $\nabla \tilde{w}_1$ **(1)** and $\nabla \tilde{w}_2$ **(2)** from experiment Section 7.1, obtained with the under-integrated Fourier–Galerkin method, is shown in column (a), a one-node FE stencil ($N_n = 1$) with two linear triangular elements and two quadrature points is shown in column (b), a one-node FE stencil ($N_n = 1$) with one bilinear rectangular element and four quadrature points is shown in column (c) and a one-node FE stencil ($N_n = 1$) with one bilinear rectangular element and one quadrature point is shown in column (d).
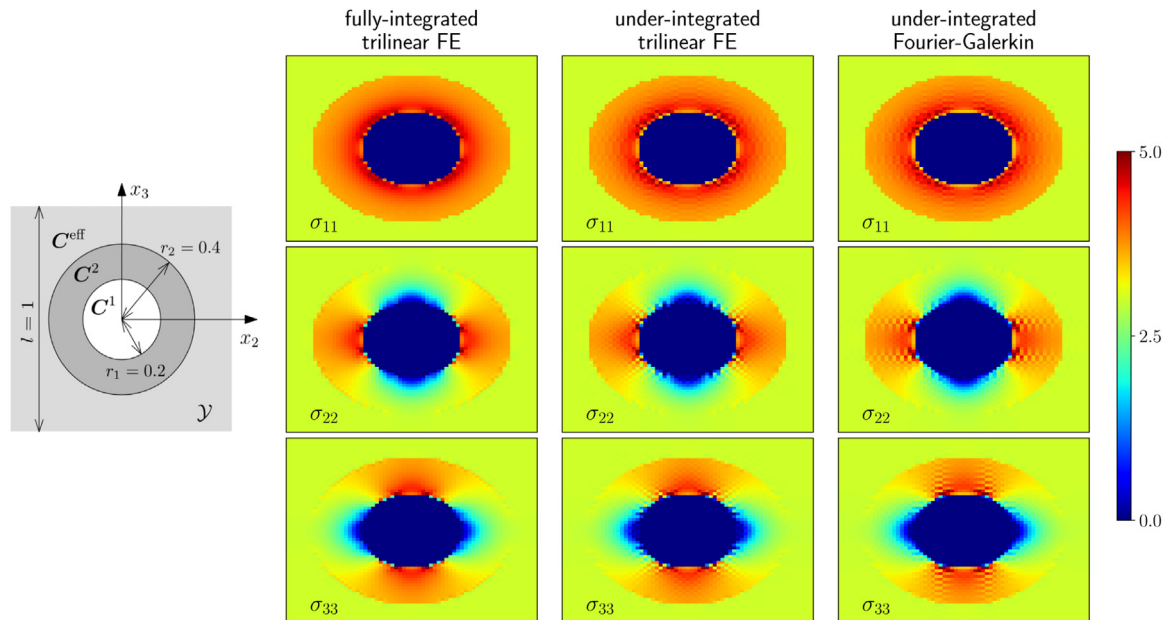
Fig. 8. Two-dimensional sections at $x_1 = 0.5$ of the 3-dimensional cubic periodic unit cell $\mathcal{Y}$ with a coated sphere inclusion. Radii $r_1 = 0.2$, $r_2 = 0.4$ with a domain size $l = 1$. Components of the local stress fields $\sigma_{\alpha\alpha}$ for trilinear FE discretization with eight quadrature points (left column), trilinear FE discretization with one quadrature point (middle column) and Fourier–Galerkin discretization (right column) with the number discretization nodes $N_l = 65^3$.
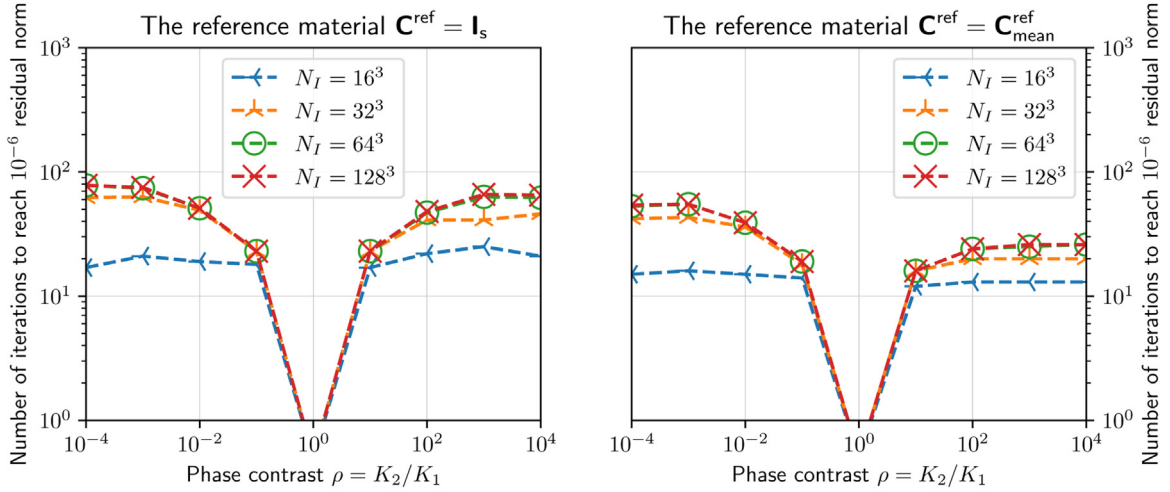
Fig. 9. The number of PCG iterations for trilinear FE discretization for different phase contrasts ρ and number of discretization nodes N_I. Termination parameter for linear solver $\eta^{CG} = 10^{-6}$.

First, in accordance with Schneider et al. [28, Section 4.1.3], we set $\rho = 10^3$ and the remaining parameters to

$$K_1 \doteq 0.00132060, \quad K_2 \doteq 1.3206033, \quad K_{\text{eff}} \doteq 1.0,$$
$$G_1 \doteq 0.00079236, \quad G_2 \doteq 0.7923620, \quad G_{\text{eff}} \doteq 0.6.$$

Two-dimensional sections at $x_1 = 0.5$ of global stress field components are shown in Fig. 8 right. Fully-integrated trilinear FE discretization (left column in Fig. 8) generates oscillation-free results, unlike under-integrated trilinear FE discretization (middle column in Fig. 8), and Fourier–Galerkin discretization (right column in Fig. 8).

Second, we are interested in how our preconditioned scheme behaves with respect to the number of discretization nodes N_I and varying phase contrast ρ. The convergence of PCG depends on the choice of the reference material $\mathbf{C}^{\text{ref}}$. We compare two cases: the first $\mathbf{C}^{\text{ref}}_{\mathbf{I_s}} = \mathbf{I_s}$, with $\mathbf{I_s} \in \mathbb{R}^{d_m \times d_m}$ being the symmetrized identity tensor $(I_s)_{\alpha\beta\iota\kappa} = \frac{1}{2}(\delta_{\alpha\iota}\delta_{\beta\kappa} + \delta_{\alpha\kappa}\delta_{\beta\iota})$ in the Mandel notation, and secondly $\mathbf{C}^{\text{ref}}_{\text{mean}} = \frac{1}{|\mathcal{Y}|}\sum_{Q=1}^{N_Q} \mathbf{C}(\boldsymbol{x}_q^Q)w^Q$, where $\mathbf{C}^{\text{ref}}_{\text{mean}}$ is the mean stiffness matrices over $\mathcal{Y}$.

The preconditioner with mean reference material $\mathbf{C}^{\text{ref}}_{\text{mean}}$ exhibits better performance in all studied cases, see Fig. 9. The numbers of iterations slowly increases as N_I grows, until the number of iterations stabilizes for sufficiently fine discretizations. In addition, $\mathbf{C}^{\text{ref}}_{\text{mean}}$ significantly reduced the phase contrast sensibility, especially for $\rho > 1$ (softer sphere core).

7.3. Finite-strain elasto-plastic problem

The purpose of the last example is two-fold. First, we demonstrate the applicability of the approach to real-world problems in the finite-strain setting and the effect of nonphysical oscillations on the results. Second, we point out the equivalence of the DB FE scheme to the SB FE scheme with the projection operator recently proposed by Leute et al. [29]. The equivalence of these two approaches is briefly explained later in Section 6.1.

For this purpose, we adapt the example from Section 5.5 of de Geus et al. [24]. This example studies a sample of dual-phase steel obtained using a scanning electron microscope. Responses of the material phases are elastic and homogeneous in the elastic part of deformation with Young's moduli $E = 200\,\text{GPa}$ and Poisson's ratios $\nu = 0.3$, and differ in the parameters of linear hardening in the plastic regions, see de Geus et al. [24, Section 5] for more details about the material model.

The yield stress τ_y evolves with respect to plastic strain ε_p, initial yield stresses τ_{y0}^{hard}, and hardening moduli H_0^{hard} such that $\tau_y = \tau_{y0} + H\varepsilon_p$. We set these parameters to

$$\tau_{y0}^{\text{hard}} = 2\tau_{y0}^{\text{soft}} = 0.003E, \quad \text{and} \quad H_0^{\text{hard}} = 2H_0^{\text{soft}} = 0.01E.$$

Total macroscopic deformation gradient is applied in 5 load increments:

$$\mathbf{F} = \frac{\sqrt{3}}{2}\begin{bmatrix} 0.995 & 0 \\ 0 & -0.995 \end{bmatrix}. \tag{21}$$

We solved this problem with the following schemes: an under-integrated SB Fourier–Galerkin scheme with the Fourier projection operator from [23,24], an SB scheme with two linear triangular FEs and the FE projection operator from Leute et al. [29], a DB FE scheme with two linear triangular FEs, and a DB FE scheme with one bilinear rectangular FE. We set the
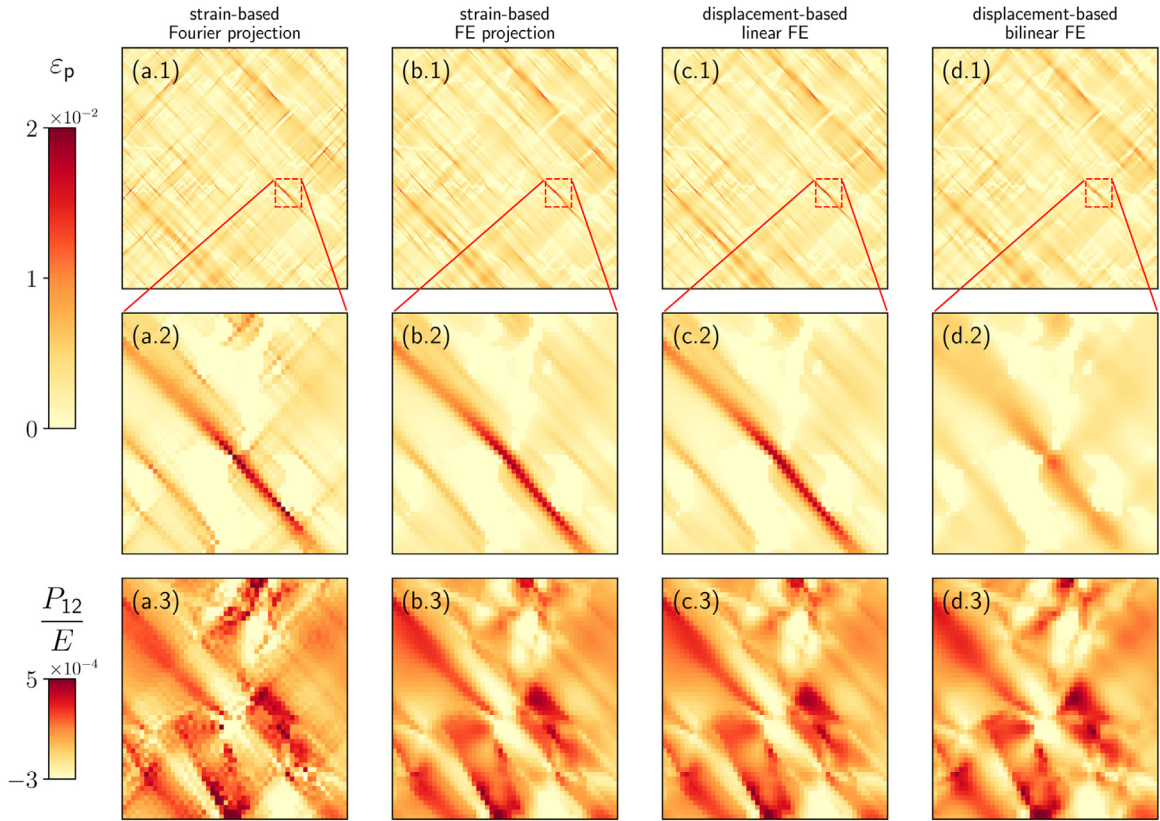
Fig. 10. Global plastic strains ε_p in dual-phase steel with applied deformation gradient (21) in row **(1)** with local details in row **(2)**. Row **(3)** shows accompanying normalized shear stresses P_{12} in the detailed area. Discretization schemes in columns: (a) the standard SB scheme with a Fourier projection operator, (b) the SB scheme with an FE projection operator with two linear triangular elements, (c) the DB FE scheme with two linear triangular elements, and (d) the DB FE scheme with one bilinear rectangular elements. All quantities are averaged per pixel.

Table 1
The number of Newton's method steps and the total number of PCG steps required to solve the finite-strain elasto-plastic problem of Section 7.3 for three selected reference materials, with Newton tolerance $\eta^{NW} = 10^{-5}$ and PCG tolerance $\eta^{CG} = 10^{-5}$. Discretization approaches from left to right: the under-integrated SB Fourier–Galerkin scheme with Fourier projection operator, the fully-integrated SB scheme with FE projection operator with two linear triangular elements, the fully-integrated DB FE scheme with two linear triangular elements, the fully-integrated DB FE scheme with one bilinear rectangular element, and the under-integrated DB FE scheme with one bilinear rectangular element per pixel. Numbers in boldface highlight the equivalence of our DB FE scheme and the SB FE scheme presented by Leute et al. [29].

		Strain-based (SB)		Displacement-based (DB)		
	$\mathbf{C}^{ref}$	Under-integrated Fourier projection	Fully-integrated linear FE projection	Fully-integrated linear FE	Fully-integrated bilinear FE	Under-integrated bilinear FE
Newton steps		11	**9**	**9**	10	11
	$\mathbf{I}$	1012	**861**	**861**	761	1014
PCG steps	$\mathbf{I_s}$	781	**609**	**609**	540	779
	$\mathbf{C}^{ref}_{mean}$	585	**457**	**457**	407	586

Newton tolerance to $\eta^{NW} = 10^{-5}$ and PCG tolerance to $\eta^{CG} = 10^{-5}$. We solved three cases with identity $\mathbf{C}^{ref} = \mathbf{I}$, symmetrized identity $\mathbf{C}^{ref} = \mathbf{I_s}$ and mean value $\mathbf{C}^{ref} = \mathbf{C}^{ref}_{mean}$ reference materials, in analogy to Section 7.2.

First, the distributions of global plastic strain ε_p obtained for these four approaches are shown in the first row of Fig. 10. The regions of details (the second row) uncover the checkerboard patterns in the plastic strain fields of the under-integrated SB Fourier–Galerkin solution (a.2) that are a direct consequence of the oscillating stress field (a.3). The other three schemes, columns (b) to (d), produce solutions without oscillations.

Second, the number of Newton's method steps and the total number of PCG iterations needed to solve the problem with these four approaches are shown in Table 1. The table highlights the equivalence of our DB scheme and the SB scheme presented by Leute et al. [29], if equivalent discretizations are used. Additionally, Table 1 shows that under-integrated schemes exhibit worse convergence compared to fully-integrated variants.

8. Conclusions

In this paper, we presented a novel and *geometrically optimal* approach for computational homogenization of nonlinear micromechanical and thermal problems in periodic media. Efficiency was achieved due to a clever interplay between the PCG solver and the geometry and physical properties of the problem [45]. Standard FE discretization on a regular grid was coupled with Newton's method to handle the nonlinear system iteratively. The linearized system was solved using the PCG method, which was enhanced with a preconditioner based on a discretized inverse (Green's) operator for a problem with data from a spatially uniform reference material. The proposed method exhibited excellent convergence properties because the number of linear solver iterations was bounded (independent of the number of discretization nodes) and showed mild phase-contrast sensitivity. Our main findings are summarized as follows:

- The condition number associated with the preconditioned linear system decreased as the data for the reference material approached the data for the material. Two-sided bounds for all eigenvalues of the preconditioned linear system were easily accessible, and thus provided valuable insight into the choice of reference material.
- The complexity of computation was governed by the FFT algorithm applied to the displacement field. The preconditioning operator was, in terms of computation, cheaply inverted and applied in Fourier space, while the gradient was evaluated through the convolution with a short kernel in real space.
- The FE bases produced oscillation-free stress and strain solution fields with marginal discretization artifacts at the phase interfaces. Additional variability of discretization patterns facilitated a reduction in mesh anisotropy and a more accurate representation of the geometry and the solution.

We consider this approach optimal because of the following reasons:

- The fully-integrated Galerkin approach provides the optimal approximation for the solution from the approximation subspace and minimizes the energy norm of the error. The energy norm of the error is equivalent to the error in homogenized properties; see Vondřejc and de Geus [46].
- For symmetric positive definite algorithmic tangents, the CG solver minimizes the energy norm of the error and provides the optimal approximation in every iteration.
- Using the stopping criterion for CG that estimates the energy norm of error is optimal.
- The Green's function preconditioner makes the CG iteration count independent of the size of the linear system. This geometrically-optimal preconditioner (recall Section 4.3) is inexpensive and applicable at $\mathcal{O}(n \log(n))$ complexity. In addition, an appropriate choice of the reference material can further improve its efficiency.

In addition, the Galerkin nature of the FE method connected with the minimization of the related energy functional allows us to use a well-built theory based on the FE method for error estimation, convergence analysis, and other useful analyses.

The major disadvantage of a fully-integrated FE scheme is its significant computational memory consumption, especially in 3D, when a single voxel requires at least 5 quadrature points (5 linear tetrahedral FEs). This significantly increases memory use and costs of constitutive model evaluations compared to single-quadrature point schemes. Our formulation allows use of the under-integrated FE as well; however, using this kind of FE leads to lower quality solutions. To obtain locally accurate solutions that are necessary for problems with localized deformations (e.g., damage or plasticity problems), fully-integrated schemes appear, to the best of our knowledge, to be necessary at present.[8]

In the future, the extension of the equivalence of DB and SB schemes to a general reference material and the fusion of Vondřejc et al.'s [48] low-rank tensor approximation technique with our FE scheme will be the areas in which we will focus our investigations.

Data availability

Data will be made available on request.

---

[8] Note that this does not necessarily hold for homogenized or averaged fields, as was pointed out recently in Schneider [47]. Schneider [47] also proposed an hourglass stabilization for under-integrated FE scheme, which improves the quality of the local solution fields. However, the solution of this promising approach yet depends on the selection of the hourglass stabilization parameter.

anonymous reviewers whose comments and suggestions helped improve the quality of the manuscript. Finally, the authors are grateful to Dr. Stephanie Krueger for her helpful English editorial suggestions and proofreading of the manuscript.

Thermal conduction The proposed preconditioned FE method can also be used for potential problems investigating thermal conduction or electrostatics. From a mathematical viewpoint, these problems are described by a scalar elliptic partial differential equation.

For the scalar thermal conduction problem, we split the overall temperature gradient $\nabla w : \mathcal{Y} \to \mathbb{R}^d$ into an average temperature gradient $\boldsymbol{e} = \frac{1}{|\mathcal{Y}|} \int_{\mathcal{Y}} \nabla w(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \in \mathbb{R}^d$ and a periodically fluctuating field $\nabla \tilde{w} : \mathcal{Y} \to \mathbb{R}^d$

$$\nabla w = \boldsymbol{e} + \nabla \tilde{w} \quad \text{for all } \boldsymbol{x} \in \mathcal{Y}.$$

Here, $\nabla \tilde{w}$ denotes the temperature gradient, and the fluctuating temperature field $\tilde{w}$ belongs to the space of admissible functions $\mathcal{V} = \{\tilde{v} : \mathcal{Y} \to \mathbb{R}, \ \tilde{v} \text{ is } \mathcal{Y}\text{-periodic }\}$. The governing equation for $\nabla \tilde{w}$ follows from the thermal equilibrium condition

$$-\nabla \cdot \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{e} + \nabla \tilde{w}(\boldsymbol{x})) = 0 \quad \text{for all } \boldsymbol{x} \in \mathcal{Y},$$

in which $\boldsymbol{q} : \mathcal{Y} \times \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^d$ is the flux field. As usual, the equilibrium equation is converted to the weak form

$$\int_{\mathcal{Y}} \nabla \tilde{v}(\boldsymbol{x})^\mathsf{T} \boldsymbol{q}(\boldsymbol{x}, \boldsymbol{e} + \nabla \tilde{w}(\boldsymbol{x})) \, \mathrm{d}\boldsymbol{x} = 0 \quad \text{for all } \tilde{v} \in \mathcal{V} \tag{A.1}$$

that serves as the starting point for the FE method. Following the discretization scheme described in Section 3, the linearization in Section 3.1 and preconditioning in Section 4 leads to a well-conditioned linear system

$$\underbrace{\mathbf{F}^H (\mathbf{F}\mathbf{D}^\mathsf{T}\mathbf{W}\mathbf{A}_{(i)}^{\mathrm{ref}}\mathbf{D}\mathbf{F}^H)^{-1}}_{(\mathbf{K}_{(i)}^{\mathrm{ref}})^{-1}} \underbrace{\mathbf{F}\,\mathbf{D}^\mathsf{T}\mathbf{W}\mathbf{A}_{(i)}\mathbf{D}}_{\mathbf{K}_{(i)}} \delta\tilde{\mathbf{w}}_{(i+1)} = \underbrace{\mathbf{F}^H (\mathbf{F}\mathbf{D}^\mathsf{T}\mathbf{W}\mathbf{A}_{(i)}^{\mathrm{ref}}\mathbf{D}\mathbf{F}^H)^{-1}}_{(\mathbf{K}_{(i)}^{\mathrm{ref}})^{-1}} \underbrace{\mathbf{F}\,\mathbf{D}^\mathsf{T}\mathbf{W}\boldsymbol{q}(\mathbf{e} + \mathbf{D}\tilde{\mathbf{w}}_{(i)})}_{-\mathbf{b}_{(i)}}, \tag{A.2}$$

for a finite Newton's method increment $\delta\tilde{\mathbf{w}}_{(i+1)}$. Material data matrix $\mathbf{A}_{(i)} \in \mathbb{R}^{dN_Q \times dN_Q}$ stores values of conductivity tangent matrix $\mathbf{A}_{(i)}(\boldsymbol{x}) = \frac{\partial \boldsymbol{q}}{\partial \nabla \tilde{w}}(\boldsymbol{x}, \boldsymbol{e} + \nabla \tilde{w}(\boldsymbol{x})) \in \mathbb{R}^{d \times d}$ in (i)th Newton's method step, and $\mathbf{A}_{(i)}^{\mathrm{ref}} \in \mathbb{R}^{dN_Q \times dN_Q}$ comes from spatially uniform material data $\mathbf{A}_{(i)}^{\mathrm{ref}} \in \mathbb{R}^{d \times d}$. Another small difference lies in the form of the gradient matrix $\mathbf{D}$,

$$\nabla \tilde{\mathbf{w}} = \mathbf{D}\tilde{\mathbf{w}} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{bmatrix} [\tilde{\mathbf{w}}]. \tag{A.3}$$

Here, the entries are the same as in the elasticity problem; recall Eq. (5).

References

[1] M. Schneider, D. Merkert, M. Kabel, FFT-based homogenization for microstructures discretized by linear hexahedral elements, Int. J. Numer. Methods Eng. 109 (10) (2017) 1461–1489, doi:10.1002/nme.5336.
[2] M. Leuschner, F. Fritzen, Fourier-accelerated nodal solvers (FANS) for homogenization problems, Comput. Mech. 62 (3) (2018) 359–392, doi:10.1007/s00466-017-1501-5.
[3] J. LLorca, C. González, J.M. Molina-Aldareguía, J. Segurado, R. Seltzer, F. Sket, M. Rodríguez, S. Sádaba, R. Muñoz, L.P. Canal, Multiscale modeling of composite materials: a roadmap towards virtual testing, Adv. Mater. 23 (44) (2011) 5130–5147, doi:10.1002/adma.201101683.
[4] K. Matouš, M.G.D. Geers, V.G. Kouznetsova, A. Gillman, A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials, J. Comput. Phys. 330 (2017) 192–220, doi:10.1016/j.jcp.2016.10.070.
[5] J. Fish, G.J. Wagner, S. Keten, Mesoscopic and multiscale modelling in materials, Nat. Mater. 20 (2021) 774–786, doi:10.1038/s41563-020-00913-0.
[6] K. Terada, T. Miura, N. Kikuchi, Digital image-based modeling applied to the homogenization analysis of composite materials, Comput. Mech. 20 (1997) 331–346, doi:10.1007/s004660050255.
[7] E. Maire, P.J. Withers, Quantitative X-ray tomography, Int. Mater. Rev. 59 (1) (2014) 1–43, doi:10.1179/1743280413Y.0000000023.
[8] B. Sonon, K. Ehab Moustafa Kamel, T.J. Massart, Advanced Geometry Representations and Tools for Microstructural and Multiscale Modeling, in: Advances in Applied Mechanics, vol. 54, Elsevier, 2021, pp. 1–111, doi:10.1016/bs.aams.2020.12.001.
[9] C. Johnson, Numerical Solution of Partial Differential Equations by the Finite Element Method, Cambridge University Press, 1995.
[10] M.R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, J. Res. Natl. Bur. Stand. 49 (1952) 409–435, doi:10.6028/JRES.049.044.
[11] H. Moulinec, P. Suquet, A fast numerical method for computing the linear and nonlinear mechanical properties of composites, C. R. Acad. Sci. Sér. II. Méc., Phys., Chim., Astron. 318 (1–2) (1994) 1417–1423.
[12] H. Moulinec, P. Suquet, A numerical method for computing the overall response of nonlinear composites with complex microstructure, Comput. Methods Appl. Mech. Eng. 157 (1–2) (1998) 69–94, doi:10.1016/S0045-7825(97)00218-1.
[13] G.H. Golub, C.F. Van Loan, Matrix Computations, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, 2013.
[14] M. Schneider, A review of nonlinear FFT-based computational homogenization methods, Acta Mech. 232 (6) (2021) 2051–2100, doi:10.1007/s00707-021-02962-1.
[15] S. Lucarini, M.V. Upadhyay, J. Segurado, FFT based approaches in micromechanics: fundamentals, methods and applications, Model. Simul. Mater. Sci. Eng. 30 (2) (2021) 023002, doi:10.1088/1361-651x/ac34e1.
[16] C. Gierden, J. Kochmann, J. Waimann, B. Svendsen, S. Reese, A review of FE-FFT-based two-scale methods for computational modeling of microstructure evolution and macroscopic material behavior, Arch. Comput. Methods Eng. 29 (6) (2022) 4115–4135, doi:10.1007/s11831-022-09735-6.
[17] S. Brisard, L. Dormieux, FFT-based methods for the mechanics of composites: ageneral variational framework, Comput. Mater. Sci 49 (3) (2010) 663–671, doi:10.1016/j.commatsci.2010.06.009.
[18] J. Zeman, J. Vondřejc, J. Novák, I. Marek, Accelerating a FFT-based solver for numerical homogenization of periodic media by conjugate gradients, J. Comput. Phys. 229 (21) (2010) 8065–8071, doi:10.1016/j.jcp.2010.07.010.
[19] S. Brisard, L. Dormieux, Combining Galerkin approximation techniques with the principle of Hashin and Shtrikman to derive a new FFT-based numerical method for the homogenization of composites, Comput. Methods Appl. Mech. Eng. 217–220 (2012) 197–212, doi:10.1016/j.cma.2012.01.003.

[20] J. Vondřejc, J. Zeman, I. Marek, An FFT-based Galerkin method for homogenization of periodic media, Comput. Math. Appl. 68 (3) (2014) 156–173, doi:10.1016/j.camwa.2014.05.014.

[21] L. Gélébart, R. Mondon-Cancel, Non-linear extension of FFT-based methods accelerated by conjugate gradients to evaluate the mechanical behavior of composite materials, Comput. Mater. Sci 77 (2013) 430–439, doi:10.1016/j.commatsci.2013.04.046.

[22] M. Kabel, T. Böhlke, M. Schneider, Efficient fixed point and Newton–Krylov solvers for FFT-based homogenization of elasticity at large deformations, Comput. Mech. 54 (6) (2014) 1497–1514, doi:10.1007/s00466-014-1071-8.

[23] J. Zeman, T.W.J. de Geus, J. Vondřejc, R.H.J. Peerlings, M.G.D. Geers, A finite element perspective on nonlinear FFT-based micromechanical simulations, Int. J. Numer. Methods Eng. 111 (10) (2017) 903–926, doi:10.1002/nme.5481.

[24] T.W.J. de Geus, J. Vondřejc, J. Zeman, R.H.J. Peerlings, M.G.D. Geers, Finite strain FFT-based non-linear solvers made simple, Comput. Methods Appl. Mech. Eng. 318 (2017) 412–430, doi:10.1016/j.cma.2016.12.032.

[25] S. Kaßbohm, W.H. Müller, R. Feßler, Improved approximations of Fourier coefficients for computing periodic structures with arbitrary stiffness distribution, Comput. Mater. Sci 37 (1–2) (2006) 90–93, doi:10.1016/j.commatsci.2005.12.010.

[26] P. Shanthraj, P. Eisenlohr, M. Diehl, F. Roters, Numerically robust spectral methods for crystal plasticity simulations of heterogeneous materials, Int. J. Plast. 66 (2015) 31–45, doi:10.1016/j.ijplas.2014.02.006.

[27] F. Willot, B. Abdallah, Y.P. Pellegrini, Fourier-based schemes with modified Green operator for computing the electrical response of heterogeneous media with accurate local fields, Int. J. Numer. Methods Eng. 98 (7) (2014) 518–533, doi:10.1002/nme.4641.

[28] M. Schneider, F. Ospald, M. Kabel, Computational homogenization of elasticity on a staggered grid, Int. J. Numer. Methods Eng. 105 (9) (2016) 693–720, doi:10.1002/nme.5008.

[29] R.J. Leute, M. Ladecký, A. Falsafi, I. Jödicke, I. Pultarová, J. Zeman, T. Junge, L. Pastewka, Elimination of ringing artifacts by finite-element projection in FFT-based homogenization, J. Comput. Phys. 453 (2022) 110931, doi:10.1016/j.jcp.2021.110931.

[30] X. Ma, M. Shakoor, D. Vasiukov, S.V. Lomov, C.H. Park, Numerical artifacts of fast Fourier transform solvers for elastic problems of multi-phase materials: their causes and reduction methods, Comput. Mech. (2021), doi:10.1007/s00466-021-02013-5.

[31] S. Lucarini, J. Segurado, An algorithm for stress and mixed control in Galerkin-based FFT homogenization, Int. J. Numer. Methods Eng. 119 (8) (2019) 797–805, doi:10.1002/nme.6069.

[32] O. Axelsson, J. Karátson, Equivalent operator preconditioning for elliptic problems, Numer. Algorithms 50 (3) (2009) 297–380, doi:10.1007/s11075-008-9233-4.

[33] I. Pultarová, M. Ladecký, Two-sided guaranteed bounds to individual eigenvalues of preconditioned finite element and finite difference problems, Numer. Linear Algebra Appl. 28 (5) (2021) e2382, doi:10.1002/nla.2382.

[34] T. Gergelits, K.-A. Mardal, B.F. Nielsen, Z. Strakoš, Laplacian preconditioning of elliptic PDEs: localization of the eigenvalues of the discretized operator, SIAM J. Numer. Anal. 57 (3) (2019) 1369–1394, doi:10.1137/18M1212458.

[35] M. Ladecký, I. Pultarová, J. Zeman, Guaranteed two-sided bounds on all eigenvalues of preconditioned diffusion and elasticity problems solved by the finite element method, Appl. Math. 66 (1) (2020) 21–42, doi:10.21136/AM.2020.0217-19.

[36] J.W. Cooley, J.W. Tukey, An algorithm for the machine calculation of complex Fourier series, Math. Comput. 19 (1965) 297–301, doi:10.1090/S0025-5718-1965-0178586-1.

[37] H. Grimm-Strele, M. Kabel, Runtime optimization of a memory efficient CG solver for FFT-based homogenization: implementation details and scaling results for linear elasticity, Comput. Mech. 64 (2019) 1339–1345, doi:10.1007/s00466-019-01713-3.

[38] S. Lucarini, J. Segurado, DBFFT: A displacement based FFT approach for non-linear homogenization of the mechanical behavior, Int. J. Eng. Sci. 144 (2019), doi:10.1016/j.ijengsci.2019.103131.

[39] Y. Saad, Iterative Methods for Sparse Linear Systems, second ed., Society for Industrial and Applied Mathematics, 2003, doi:10.1137/1.9780898718003.

[40] M.W. Frazier, An Introduction to Wavelets Through Linear Algebra, Springer New York, NY, 1999.

[41] V. Eijkhout, P. Vassilevski, The role of the strengthened Cauchy–Buniakowskii–Schwarz inequality in multilevel methods, SIAM Rev. 33 (1991) 405–419, doi:10.1137/1033098.

[42] O. Axelsson, Iterative Solution Methods, Cambridge University Press, 1996, doi:10.1017/CBO9780511624100.

[43] B.F. Nielsen, A. Tveito, W. Hackbusch, Preconditioning by inverting the Laplacian: an analysis of the eigenvalues, IMA J. Numer. Anal. 29 (1) (2009) 24–42, doi:10.1093/imanum/drm018.

[44] H. Moulinec, F. Silva, Comparison of three accelerated FFT-based schemes for computing the mechanical response of composite materials, Int. J. Numer. Methods Eng. 97 (2014) 960–985, doi:10.1002/nme.4614.

[45] J. Málek, Z. Strakoš, Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs, SIAM Spotlight Series, Society for Industrial and Applied Mathematics, 2015.

[46] J. Vondřejc, T.W.J. de Geus, Energy-based comparison between the Fourier–Galerkin method and the finite element method, J. Comput. Appl. Math. 374 (2020) 112585, doi:10.1016/j.cam.2019.112585.

[47] M. Schneider, Voxel-based finite elements with hourglass control in fast Fourier transform-based computational homogenization, Int. J. Numer. Methods Eng. 123 (24) (2022) 6286–6313, doi:10.1002/nme.7114.

[48] J. Vondřejc, D. Liu, M. Ladecký, H.G. Matthies, FFT-based homogenisation accelerated by low-rank tensor approximations, Comput. Methods Appl. Mech. Eng. 364 (2020) 112890, doi:10.1016/j.cma.2020.112890.