Journal of Computational Physics 508 (2024) 113013

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Scalable multiscale-spectral GFEM with an application to composite aero-structures

Jean Bénézech [a,*], Linus Seelinger [b], Peter Bastian [c], Richard Butler [a], Timothy Dodwell [d,e], Chupeng Ma [f], Robert Scheichl [c,g]

[a] Centre for Integrated Materials, Processes and Structures, University of Bath, Bath, UK
[b] Scientific Computing Center, Karlsruhe Institute of Technology, Karlsruhe, Germany
[c] Interdisciplinary Center for Scientific Computing, Heidelberg University, Heidelberg, Germany
[d] Institute of Data Science and AI, University of Exeter, Exeter, UK
[e] digiLab, Exeter, UK
[f] School of Sciences, Great Bay University, Dongguan, China
[g] Institute for Mathematics, Heidelberg University, Heidelberg, Germany

ARTICLE INFO

Dataset link: https://gitlab.dune-project.org/anne.reinarz/dune-composites/

Dataset link: https://github.com/jeanbenezech/CompositesFEMesh

Keywords:
Multiscale-spectral generalized FE methods
GenEO coarse space
A-harmonic subspace
Large-scale composite structure
Parallel scalability

ABSTRACT

In this paper, the first large-scale application of multiscale-spectral generalized finite element methods (MS-GFEM) to composite aero-structures is presented. The crucial novelty lies in the introduction of A-harmonicity in the local approximation spaces, which in contrast to Babuška and Lipton (2011) [30] is enforced more efficiently via a constraint in the local eigenproblems. This significant modification leads to excellent approximation properties, which turn out to be essential to capture accurately material strains and stresses with a low dimensional approximation space, hence maximizing model order reduction. The implementation of the framework in the Distributed and Unified Numerics Environment (DUNE) software package, as well as a detailed description of all components of the method are presented and exemplified on a composite laminated beam under compressive loading. The excellent parallel scalability of the method, as well as its superior performance compared to the related, previously introduced GenEO method are demonstrated on two realistic application cases, including a C-shaped wing spar with complex geometry. Further, by allowing low-cost approximate solves for closely related models or geometries this efficient, novel technology provides the basis for future applications in optimization or uncertainty quantification on challenging problems in composite aero-structures.

1. Introduction

Ill-conditioned and multiscale partial differential equations (PDEs) arise in many fields, where the computation of a resolved, fine-scale solution or a robust low-dimensional approximation can be challenging. Due to the interaction of their mesoscopic structure (ply level; sub-millimeter scale) and their geometric macroscopic features (structural level; meter scale), composite aero-structures are naturally, inherently multiscale. To model the behavior of a large scale composite structure, the numerical model needs to accurately represent the meso-scale configuration of the material as well as the macro-scale geometry of the part. Full mesoscopic descriptions

* Corresponding author.
E-mail address: jb3285@bath.ac.uk (J. Bénézech).

https://doi.org/10.1016/j.jcp.2024.113013
Received 3 March 2023; Received in revised form 9 April 2024; Accepted 12 April 2024

Available online 16 April 2024

of large components naturally lead to models with huge numbers of degrees of freedom. This makes the computations prohibitively expensive, particularly in contexts where many evaluations are required, e.g. optimization or uncertainty quantification. As a result, composite aero-structures provide an ideal test bed for the new multiscale method proposed in this paper, which allows for the interaction of fine and coarse scale behavior to be captured without becoming excessive in cost.

1.1. High performance solvers for composite applications

Due to strongly varying parameters across the simulation domain, elasticity problems arising in composite materials lead to Finite Element (FE) matrices that are extremely ill-conditioned [1,2]. This poor conditioning is due to the contrast in stiffness between the carbon fibbers and the surrounding resin matrix, as well as the complex anisotropy arising from the inclusion of long directional fibbers. As a result, composite laminates have both low-energy modes, whereby the stiff fibbers act as rigid body inclusions and the complement resin deforms easily; and also very high-energy modes, in deformation regimes with stretched, stiff fibbers. This contrast between low and high energy modes is at the heart of the ill-conditioning (high condition number) in composite applications.

Whilst sparse direct solvers, like UMFPACK [3] or the ones provided in the *Abaqus* package [4], can reliably solve such systems, they are inherently limited in their scalability. This immediately restricts the physical scale of composites that can be simulated. Iterative solvers such as Conjugate Gradient (CG) or GMRES [5] in turn promise massive parallel scalability for modern High Performance Computing (HPC) systems, but their efficiency (i.e. number of iterations) strongly depends on the condition number of the matrix.

In order to render such iterative solvers robust, preconditioners are essential. Whilst Algebraic Multigrid (AMG) preconditioners are in general a promising choice regarding robustness and scalability, tests with two AMG implementations, in dune-istl [6] and BoomerAMG [7], have demonstrated poor performance in composite applications [1,2]. The reason is that, without a problem-specific local aggregation strategy, coarse grids in AMG do not capture the low-energy modes in composites structures. Domain decomposition (DD) methods are another popular class of methods for constructing preconditioners for iterative solvers. The key to obtaining scalability for DD preconditioners is to add a suitable coarse space into the methods. For problems with highly varying coefficients arising in composite applications, standard coarse spaces were shown to be not robust, and spectral coarse spaces have been intensively developed to overcome this issue; see [8–16] for spectral coarse spaces developed in the context of overlapping Schwarz methods. These coarse spaces are built from selected eigenfunctions of some tailored, local generalized eigenvalue problems, leading to preconditioners that are provably robust to coefficient jumps. In particular, we mention the GenEO (Generalized Eigenproblems in the Overlaps) method [16], which provides a framework for constructing robust spectral coarse spaces for overlapping Schwarz methods. Robust scale-up to several thousands of processor cores for composites applications has been shown in [2,17]. While the iterative solver now only needs few iterations, a considerable cost is expended in solving the independent local eigenproblems.

To avoid these tremendous computational costs of direct simulations of multiscale problems, computational homogenization methods [18–20] have been well developed and widely used in the engineering community. Moreover, in practical engineering applications, multiscale problems are typically solved multiple times with different source terms and possibly local changes in model parameters, such that the higher setup cost of such methods can be offset. Most of those methods, however, are based on scale separation hypotheses, and may fail for typical problems in realistic applications that do not exhibit such a scale separation.

1.2. Multiscale methods in composite analysis

To efficiently solve multiscale problems without scale separation for repeated analysis required in practical engineering applications, various multiscale model order reduction methods have been developed, such as the multiscale finite element method (MsFEM) [21,22], the generalized multiscale finite element method (GMsFEM) [23,24], localized orthogonal decomposition (LOD) [25,26], flux norm homogenization [27], game-theoretical approaches [28,29], and the generalized finite element method (GFEM) [30–32], to cite a few; see [24,33] for more comprehensive reviews. Most of these methods were developed in the context of numerical multiscale methods and are based on approximating the solution space on a coarse grid by a low dimensional space that is spanned by some pre-computed local basis functions. These local bases are typically obtained by solving local boundary value problems or local eigenproblems tailored to the underlying PDE, and thus encode the structural meso-scale information contained in the material parameters. Whereas many multiscale methods (e.g., [21,25,26,28]) use only one basis function per local patch, it has been shown [10,23] that for high-contrast problems, it is essential to add local degrees of freedom adaptively. Moreover, existing multiscale methods were mainly developed for a single PDE and related studies for coupled multiphysics problems are much fewer. Recently, an efficient regularized multiscale method [34] was developed to solve coupled thermomechanical problems, which overcomes the limitations of classical approaches for coupled problems with asymmetric operators and low regularity.

The multiscale-spectral generalized finite element method (MS-GFEM), the focus of this work, was first proposed by Babuška and Lipton [30] for solving heterogeneous diffusion problems, but motivated also by problems in linear elasticity and in particular fibber-reinforced composites. The approach builds optimal local approximation spaces from eigenvectors of local eigenproblems posed on A-harmonic spaces defined for oversampling subdomains. Crucially, the global approximation error is fully controlled by the local approximation errors, which are rigorously proved to decay nearly exponentially – a feature not shared by most other ad-hoc constructed numerical multiscale approaches. For diffusion problems the implementation details were discussed in [32] together with numerical results on a two-dimensional toy example. We note that the A-harmonic subspaces were also used in many other related multiscale approaches [23,35,26], as well as in DD methods [8,12,13,15] to construct spectral coarse spaces. However, without using

the notion of the Kolmogrov n-width [36] and a suitable oversampling technique in the selection of the A-harmonic basis functions, the resulting spaces are not guaranteed to be optimal or to have exponential approximation accuracy.

In a recent paper [37], Ma, Scheichl and Dodwell proposed new local eigenproblems involving the partition of unity to construct new optimal local approximation spaces for the MS-GFEM method, resulting in a GenEO-type coarse-space approximation. Significant advantages of the new local approximation spaces were demonstrated and sharper decay rates for the local approximation errors were proved. In [38], the MS-GFEM method was then also formulated and analyzed for the first time in the discrete setting, as a non-iterative DD type method for solving linear systems resulting from FE discretizations of the fine-scale problem. Very similar local and global error estimates as in the continuous setting were derived. Furthermore, an efficient method to solve the (discrete) local eigenproblems was proposed, where the A-harmonic condition is directly incorporated into the eigenproblem. More recently, this approach was applied to other multiscale PDEs, such as Helmholtz [39], parabolic [40] and singularly perturbed [41] problems. Although numerical results of various simple, two-dimensional examples have demonstrated the efficiency of the MS-GFEM method, up to now, there is no study available in the literature on the application of the MS-GFEM method to realistic, large-scale, three-dimensional multiscale problems and on the implementation and performance of the method on massively parallel computers.

1.3. Contributions of this paper

This paper represents the first large-scale application of GenEO as a multiscale-spectral GFEM, providing good approximations to fine-scale solutions with a very low number of basis functions. The reformulation of GenEO as a GFEM method in local A-harmonic subspaces distinguishes our methodology from the one proposed by Babuška and Lipton [31]. Most notably, in contrast to [31] our method is inherently adaptive — we can control the error a posteriori by simply setting a threshold on the eigenvalues to decide which eigenvectors need to be included into the local spaces; see Theorem 3.1. The paper constitutes an extension of the work proposed by Ma and co-authors [37,38] by generalizing the method to three dimensional elasticity problems, demonstrated with two real-world application cases. The theoretical background of our formulation as well as its implementation are described in detail and its excellent performance and scalability are demonstrated. The resulting coarse space turns out to have significantly better approximation properties than GenEO in the elasticity problems considered here. In particular, local A-harmonicity is crucial for accurate strain approximation with significant practical improvements demonstrated in numerical experiments.

Our method provides efficient reduced order models for large-scale problems that exhibit strong dependence on local details. The efficiency arises from a decomposition of the global problem into independent sub-problems that can be treated fully in parallel. A scalability test is presented in this paper that demonstrates this high efficiency even for very large structures. The accuracy of the approximation space is fully adjustable: a single threshold parameter on the local eigenvalues allows for an optimization of the amount of model order of reduction. The approach does not rely on a scale separation hypothesis between material scales (meso- and macro-scale for the examples illustrated in the paper), the proposed multi-scale framework is particularly well suited for composite aero-structures. This is demonstrated via an application on a realistic C-spar model – a demonstrator application in the UK-EPSRC-funded CerTest project on composite structural design and certification (see the acknowledgments).

A key motivation for the approach presented in this paper is the development of an offline-online framework, where the costly local model order reductions are reused across multiple similar simulation runs, reducing the overall cost to a fraction, as suggested in [23] but with significantly smaller coarse spaces. This promises to accelerate uncertainty quantification or optimization tasks on challenging composites models and will be described in detail in a subsequent paper.

In addition, this paper reports on other technical improvements and recent developments of the `dune-composites` module within DUNE [42]. Support for GenEO is extended to unstructured DUNE grids, which is crucial for engineering applications such as composite parts with complex geometry (e.g., T-joint stiffeners) that cannot be discretized solely using a structured grid. The FE matrices in GenEO are defined on overlapping subdomains, which are however not natively supported by unstructured grids in DUNE. We therefore assemble suitable matrices on a parallelized non-overlapping grid partition using standard DUNE procedures, and then construct the desired GenEO matrices from those using an appropriate communication mechanism.

2. Problem formulation

Let us start by formulating the anisotropic, linear elasticity equations for composite structures and their finite element discretization. The composite structure is assumed to occupy a bounded and (for simplicity) polyhedral domain $\Omega \subset \mathbb{R}^3$ with boundary Γ and unit, outward normal $\mathbf{n} \in \mathbb{R}^3$. At each point $\mathbf{x} \in \Omega$ we define a vector-valued displacement $\mathbf{u}(\mathbf{x}) : \Omega \to \mathbb{R}^3$ and denote by $\mathbf{f}(\mathbf{x}) : \Omega \to \mathbb{R}^3$ the body force per unit volume. The infinitesimal strain tensor, is defined as the symmetric part of the displacement gradients:

$$\epsilon_{ij}(\mathbf{u}) = \frac{1}{2}\left(u_{i,j} + u_{j,i}\right), \tag{1}$$

where $u_{i,j} = \partial u_i / \partial x_j$. The strain tensor is connected to the Cauchy stress tensor σ_{ij} via the generalized Hooke's law:

$$\sigma_{ij}(\mathbf{u}) = C_{ijkl}(\mathbf{x})\epsilon_{kl}(\mathbf{u}), \tag{2}$$

where the material tensor $C_{ijkl}(\mathbf{x})$ is a symmetric, positive definite fourth order tensor. The studied material will be further described in Section 5.1.

Now, let Γ_D and Γ_N be disjoint open subsets of Γ such that $\overline{\Gamma_D} \cup \overline{\Gamma_N} = \Gamma$ and consider the function space

$$V := \{\mathbf{v} \in H^1(\Omega;\mathbb{R}^3) : \mathbf{v} = \mathbf{0} \quad \text{on } \Gamma_D\}. \tag{3}$$

Given functions $\mathbf{h} : \Gamma_D \to \mathbb{R}^3$ and $\mathbf{g} : \Gamma_N \to \mathbb{R}^3$, prescribing the Dirichlet and Neumann boundary data, the weak formulation of the problem to be considered consists in seeking the unknown displacement field $\mathbf{u} \in H^1(\Omega;\mathbb{R}^3)$ with $\mathbf{u} = \mathbf{h}$ on Γ_D such that

$$a(\mathbf{u},\mathbf{v}) = b(\mathbf{v}) \quad \forall \mathbf{v} \in V, \tag{4}$$

where the bilinear form $a(\cdot,\cdot)$ and the functional $b(\cdot)$ are defined by

$$a(\mathbf{u},\mathbf{v}) = \int_\Omega \sum_{i,j} \sigma_{ij}(\mathbf{u})\epsilon_{ij}(\mathbf{v})\,dx \quad \text{and} \quad b(\mathbf{v}) = \int_{\Gamma_N} \mathbf{g}\cdot\mathbf{v}\,ds + \int_\Omega \mathbf{f}\cdot\mathbf{v}\,dx. \tag{5}$$

The variational problem (4) is discretized with conforming FEs on a mesh $\mathcal{T}_h$ on Ω by introducing the FE space $V_h \subset V$ as the tensor product $V_h := V_h^1 \times V_h^2 \times V_h^3$ of the spaces $V_h^j := \text{span}\{\phi_j^{(i)}\}_{i=1}^N$, spanned by the usual Lagrange bases on $\mathcal{T}_h$. We find a function $\mathbf{u}^p \in H^1(\Omega;\mathbb{R}^3)$ such that $\mathbf{u}^p = \mathbf{h}$ on Γ_D and then seek an approximation $\mathbf{u}_h = \mathbf{u}^p + \widetilde{\mathbf{u}}_h$, where $\widetilde{\mathbf{u}}_h \in V_h$, such that

$$a(\widetilde{\mathbf{u}}_h,\mathbf{v}_h) = \widetilde{b}(\mathbf{v}_h) := b(\mathbf{v}_h) - a(\mathbf{u}^p,\mathbf{v}_h) \quad \forall \mathbf{v}_h \in V_h. \tag{6}$$

We block together displacements from all three space dimensions, so that $\mathbf{u}_h^{(i)} \in \mathbb{B} := \mathbb{R}^3$ denotes the vector of displacement coefficients containing all space components associated with the i^{th} basis function. The displacement vector at a point $\mathbf{x}$ is then given by $\mathbf{u}_{h,j}(\mathbf{x}) = \sum_{i=1}^N u_{h,j}^{(i)} \boldsymbol{\phi}_j^{(i)}(\mathbf{x})$, $j \in \{1,2,3\}$. The system (6) is equivalent to a symmetric positive-definite (spd) system of algebraic equations:

$$\mathbf{A}\tilde{\mathbf{u}} = \tilde{\mathbf{b}} \quad \text{where} \quad \mathbf{A} \in \mathbb{R}^{N\times N} \quad \text{and} \quad \tilde{\mathbf{b}} \in \mathbb{R}^N. \tag{7}$$

The blocks in the global stiffness matrix and in the load vector are given by $\mathbf{A}_{ij} = a(\boldsymbol{\phi}^{(i)},\boldsymbol{\phi}^{(j)})$ and $\tilde{\mathbf{b}}_i = \widetilde{b}(\boldsymbol{\phi}_i)$, for any $i,j = 1,\dots,N$, while $\tilde{\mathbf{u}} = [\widetilde{\mathbf{u}}_h^{(1)},\dots,\widetilde{\mathbf{u}}_h^{(N)}]^T \in \mathbb{B}^N$ is the block vector of unknown FE coefficients.

System (7) can be assembled elementwise using Gaussian integration:

$$a(v,w) = \sum_{e\in\mathcal{T}_h} a_e(v|_e, w|_e) \quad \forall v,w \in V. \tag{8}$$

The elementwise bilinear form a_e is trivial to obtain here by restricting the integrals in eq. (4) to e. Later, restrictions of $a(\cdot,\cdot)$ to mesh-resolved subdomains will be crucial in defining coarse space approximations. For any mesh-resolved subdomain $D \subset \Omega$, the restriction of $a(\cdot,\cdot)$ to D is denoted by

$$a_D(v,w) := \sum_{e\in D} a_e(v|_e, w|_e) \quad \forall v,w \in V. \tag{9}$$

3. Multiscale-spectral generalized finite element method

In this section, the methods employed in this paper are defined alongside some theoretical results. In addition to generalized finite element methods (GFEM), we also describe the Generalized Eigenproblems in the Overlaps (GenEO) space, which was originally designed as a coarse space for two-level additive Schwarz methods in [16]. Here, it is for the first time applied as a multiscale method for stand-alone coarse approximation to a realistic three-dimensional multiscale problem in composites, in the form of the multiscale-spectral generalized finite element method (MS-GFEM) method [30,37,38].

3.1. Domain decomposition

Both the two-level additive Schwarz method and GFEM are based on a decomposition of the domain Ω into non-overlapping subdomains $\{\hat{\Omega}_j\}_{j=1}^N$ that are resolved by the mesh $\mathcal{T}_h$; see Fig. 1 for an example. Each non-overlapping subdomain $\hat{\Omega}_j$ is extended by adding layers of neighboring elements to create an overlapping partition $\{\Omega_j\}_{j=1}^N$ of Ω. We define here neighbors as elements having a sub-entity in common, e.g. a face, an edge or a node. In the present case, the neighboring elements have at least one node in common.

Next, the local finite element (FE) spaces

$$V_h(\Omega_j) := \{v|_{\Omega_j} : v \in V_h\} \quad \text{and} \quad V_{h,0}(\Omega_j) := \{v|_{\Omega_j} : v \in V_h, \text{supp}(v) \subset \Omega_j\},$$

are defined, where the former restricts V_h to the subdomain Ω_j while the latter restricts this space further to functions whose support is contained entirely in Ω_j.

A key ingredient of GenEO-type coarse spaces and of GFEM is a partition of unity (PoU) subordinate to the overlapping decomposition $\{\Omega_j\}_{j=1}^N$. A particular partition of unity specific to the FE setting was constructed in [16]. For each $1 \le j \le N$, let

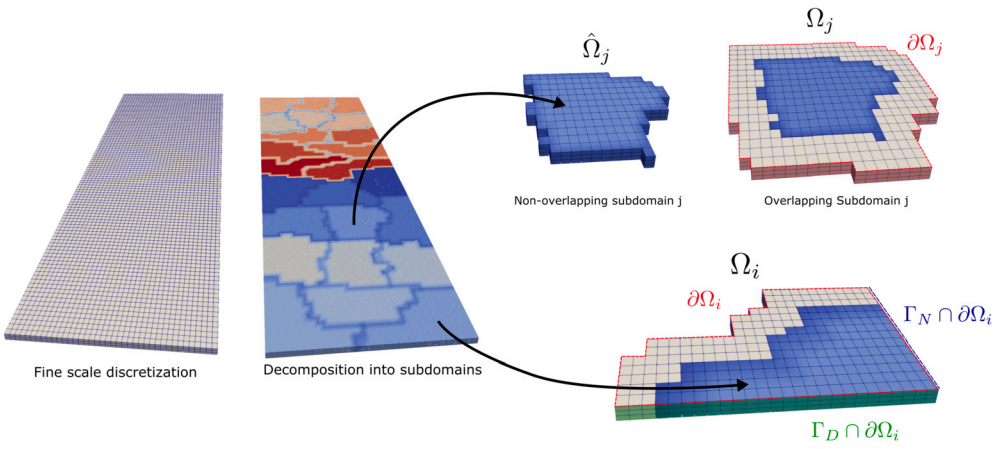

Fig. 1. Domain decomposition into 24 subdomains, illustrating the various components.

$$\mathrm{dof}(\Omega_j) = \{k : 1 \le k \le n,\ \mathrm{supp}(\boldsymbol{\phi}^{(k)}) \subset \overline{\Omega_j}\} \tag{10}$$

denote the set of internal degrees of freedom in Ω_j, and define for each degree of freedom $k \in \mathrm{dof}(\Omega_j)$ a weight $\mu_{j,k} \in [0,1]$ such that

$$\sum_{\{j:1\le j\le N, k\in\mathrm{dof}(\Omega_j)\}} \mu_{j,k} = 1.$$

With these weights we can define a family of local partition of unity operators $\Xi_j : V_h(\Omega_j) \to V_{h,0}(\Omega_j)$, $1 \le j \le N$, such that

$$\Xi_j(v) := \sum_{k\in\mathrm{dof}(\Omega_j)} \mu_{j,k} v_k \boldsymbol{\phi}^{(k)}|_{\Omega_j}, \text{ for any } v = \sum_{k\in\mathrm{dof}(\Omega_j)} v_k \boldsymbol{\phi}^{(k)} \in V_h(\Omega_j). \tag{11}$$

It follows immediately from the definition that the operators $\{\Xi_j\}_{j=1}^N$ satisfy

$$\sum_{j=1}^{N} R_j^\top \Xi_j(v|_{\Omega_j}) = v \qquad \text{for any } v \in V_h. \tag{12}$$

Here $R_j^\top : V_{h,0}(\Omega_j) \to V_h$ denotes the prolongation operator defined as the extension of a function in $V_{h,0}(\Omega_j)$ by zero. In [16], it was suggested that the weights can be set as

$$\mu_{j,k} := \frac{1}{\#\{i : 1 \le i \le N,\ k \in \mathrm{dof}(\Omega_i)\}},$$

that is, one over the number of subdomains that contain k as an internal degree of Freedom (DoF). Other partitions of unity can also be used. In the numerical experiments below, we will use a different (smoother) partition of unity.

3.2. GenEO coarse space

The GenEO space – designed in the context of additive Schwarz preconditioning methods, as a robust coarse space correction for multiscale variational problems – is based on the following generalized eigenvalue problem (GEVP) on each subdomain Ω_j: Find $\lambda^j \in \mathbb{R}$, $\varphi_h^j \in V_h(\Omega_j)$ such that

$$a_{\Omega_j}(\varphi_h^j, v_h) = \lambda^j a_{\Omega_j}(\Xi_j(\varphi_h^j), \Xi_j(v_h)), \quad \text{for all} \quad v_h \in V_h(\Omega_j). \tag{13}$$

In the original publication [16], the bilinear form on the right hand side of (13) was restricted to the overlap $\Omega_j \backslash \hat{\Omega}_j$, but as shown in subsequent publications, such as [43], the GenEO space defined by the GEVP in (13) has very similar coarse space correction properties.

Only the lowest-energy eigenfunctions in (13), i.e., the ones corresponding to the smallest eigenvalues, are used to define the GenEO coarse space. Denote by $\lambda^{j,k}$ and $\varphi_h^{j,k}$ the k-th smallest eigenvalue and the corresponding eigenfunction on subdomain Ω_j. Then, the GenEO coarse space is defined by

$$V_H := \mathrm{span}\left\{ R_j^\top \Xi_j(\varphi_h^{j,k}) : k = 1, \dots, m_j, \quad j = 1, \dots, N \right\}, \tag{14}$$

where the partition of unity operators are used to "stitch" the local approximation spaces on the subdomains Ω_j together and to guarantee that $V_H \subset V_h$. This definition of V_H still leaves open the number of eigenfunctions m_j to be included.

In the context of two-level additive Schwarz methods, where the GenEO coarse space is combined (additively) with local solves on the overlapping subdomains Ω_j to obtain a preconditioning matrix $\mathbf{M}$ for $\mathbf{A}$, it is then possible to bound the condition number κ of the preconditioned system independently of the mesh size h, the subdomain size H or the heterogeneity in the coefficient. In particular, [16, Corollary 3.23] states

$$\kappa(\mathbf{M}^{-1}\mathbf{A}) \leq C(k_0) \max_{1\leq j\leq N} \left(1 + \frac{1}{\lambda^{j,m_j+1}}\right), \tag{15}$$

where $C(k_0)$ is a typically small constant depending only on k_0, the maximum number of subdomains overlapping at any point. Thus, the condition number can be controlled by choosing the number m_j of eigenfunctions per subdomain such that $1/\lambda^{j,m_j+1}$ is bounded uniformly across all subdomains. The eigenvalues λ^{j,m_j+1} converge to 1 as m_j increases, but no theoretical results on the rate of convergence exist in general.

3.3. Generalized FE methods with GenEO-type local approximation

The GenEO space V_H in (14) is in essence a global approximation space of generalized FEM type [44], where a fairly arbitrary family of local approximation spaces can be "stitched" together via a partition of unity to build the global space. As such, for H (the subdomain size) sufficiently small or for m_j sufficiently large, it is possible to solve the FE problem directly to a required accuracy in V_H, in the spirit of the GFEM. However, as we will see below in the numerical experiments, the rate of convergence with respect to m_j is rather poor when the local bases are computed as in (13). A significantly more efficient GFEM can be designed by slightly modifying the GEVP (13) as shown in the following.

In this subsection, a particular family of GenEO-type local approximation spaces is constructed and used within the framework of the GFEM as a stand-alone coarse approximation. Compared with the original version, there are two key ingredients in this GenEO-type coarse space that provide a better accuracy for coarse approximation. The first is oversampling. Similarly to the construction of the overlapping subdomains, we extend each overlapping subdomain Ω_j further by adding more layers of fine-mesh elements to create an oversampling subdomain Ω_j^*. The local eigenproblems used for constructing the coarse space will be defined on Ω_j^* instead of Ω_j. The second ingredient is A-harmonicity. To make this notion precise, we first introduce the following local FE spaces defined on the oversampling domains:

$$V_{h,D}(\Omega_j^*) = \{v \in V_h(\Omega_j^*) : v = 0 \text{ on } \partial\Omega_j^* \cap \Gamma_D\}, \tag{16}$$

$$V_{h,DI}(\Omega_j^*) = \{v \in V_h(\Omega_j^*) : v = 0 \text{ on } \partial\Omega_j^* \cap (\Gamma_D \cup \Omega)\}. \tag{17}$$

The space $V_{h,D}(\Omega_j^*)$ consists of FE functions restricted to Ω_j^* that vanish on the external Dirichlet boundary of Ω_j^*, whereas $V_{h,DI}(\Omega_j)$ consists of FE functions that vanish on both the external Dirichlet boundary and the interior boundary of Ω_j^*. The A-harmonic local FE space on Ω_j^* is then defined as

$$V_A(\Omega_j^*) = \{u \in V_{h,D}(\Omega_j^*) : a_{\Omega_j^*}(u, v) = 0 \quad \forall v \in V_{h,DI}(\Omega_j^*)\}. \tag{18}$$

Functions in $V_A(\Omega_j^*)$ are referred to as A-harmonic FE functions. As we will see below, the local eigenvectors used for building the coarse space are A-harmonic FE functions instead of general FE functions.

With the above notations, we now define a local eigenproblem similar to (13) on each oversampling subdomain: Find $\lambda^j \in \mathbb{R}$, $\varphi_h^j \in V_A(\Omega_j^*)$ such that

$$a_{\Omega_j^*}(\varphi_h^j, v_h) = \lambda^j a_{\Omega_j^*}\left(\Xi_j(\varphi_h^j|_{\Omega_j}), \Xi_j(v_h|_{\Omega_j})\right), \quad \text{for all} \quad v_h \in V_A(\Omega_j^*). \tag{19}$$

Note that since $\Xi_j(\varphi_h^j|_{\Omega_j})$ and $\Xi_j(v_h|_{\Omega_j})$ can be identified with FE functions in $V_h(\Omega_j^*)$, the right-hand side of the above GEVP is well-defined.

Let $(\lambda^{j,k}, \varphi_h^{j,k})$ denote the k-th eigenpair of the GEVP (19) with eigenvalues enumerated in increasing order. The desired GenEO-type GFEM coarse space is defined almost identically to the standard version (14):

$$V_H := \operatorname{span}\left\{ R_j^\top \Xi_j(\varphi_h^{j,k}|_{\Omega_j}) : k = 1, \ldots, m_j, \quad j = 1, \ldots, N \right\}. \tag{20}$$

The last ingredient of the MS-GFEM method is a global particular function built from local particular functions. On each oversampling subdomain Ω_j^*, we first define a local particular function $\mathbf{u}_{h,j}^p = \psi_{h,j}^r + \psi_{h,j}^d$, where $\psi_{h,j}^r \in V_{h,DI}(\Omega_j^*)$ satisfies

$$a_{\Omega_j^*}(\psi_{h,j}^r, v_h) = b_{\Omega_j^*}(v_h) \quad \forall v_h \in V_{h,DI}(\Omega_j^*) \tag{21}$$

with $b_{\Omega_j^*}(\cdot)$ being the restriction of $b(\cdot)$ to Ω_j^*, and $\psi_{h,j}^d \in V_h(\Omega_j^*)$ satisfies $\psi_{h,j}^d = \mathbf{h}$ on $\Gamma_D \cap \partial\Omega_j^*$ and

$$a_{\Omega_j^*}(\psi_{h,j}^d, v_h) = 0 \quad \forall v_h \in V_{h,D}(\Omega_j^*). \tag{22}$$

Note that $\psi_{h,j}^d$ vanishes on all interior subdomains where $\Gamma_D \cap \partial\Omega_j^* = \emptyset$ or whenever $\mathbf{h} = \mathbf{0}$ on $\Gamma_D \cap \partial\Omega_j^*$. On subdomains intersecting Γ_D it would in fact be possible to combine problems (21) and (22) into one local problem, but this leads to a slightly larger constant

C in Theorem 3.1 below. Therefore, we work with local particular functions defined via (21) and (22) in this paper. The global particular function is then defined by "stitching" together the local functions using the partition of unity:

$$\mathbf{u}_h^p = \sum_{j=1}^{N} R_j^\top \Xi_j (\mathbf{u}_{h,j}^p|_{\Omega_j}). \tag{23}$$

Having defined the coarse space V_H and the global particular function $\mathbf{u}_h^p$, we are now ready to give the MS-GFEM method for solving the fine-scale FE problem (6): Find $\mathbf{u}_h^G = \mathbf{u}_h^p + \mathbf{u}_H$, where $\mathbf{u}_H \in V_H$, such that

$$a(\mathbf{u}_h^G, \mathbf{v}) = b(\mathbf{v}) \quad \forall \mathbf{v} \in V_H. \tag{24}$$

To assess the quality of the MS-GFEM approximation, we estimate the energy norm $\|\mathbf{v}\|_a := \sqrt{a(\mathbf{v},\mathbf{v})}$ of the error $\mathbf{u}_h - \mathbf{u}_h^G$. The following result is proved in [45, Theorem 3.17 and Remark 3.18] in an abstract setting.

Theorem 3.1. *Let λ^{j,m_j+1} be the smallest eigenvalue corresponding to any eigenvector not included ($\notin$) in the local basis on Ω_j. Then*

$$\|\mathbf{u}_h - \mathbf{u}_h^G\|_a \le C \left(\lambda_{\min}^{\notin}\right)^{-1/2} \|\mathbf{u}_h\|_a \quad \textit{where} \quad \lambda_{\min}^{\notin} := \min_{1\le j\le N} \lambda^{j,m_j+1}, \tag{25}$$

where C is an explicitly known constant that is bounded by the maximum number of oversampling domains that overlap at any given point in Ω.

Thus, the efficiency of the MS-GFEM method is controlled by the speed at which the eigenvalues in (19) grow. To estimate this growth rate, let H_j and H_j^* denote the diameter of Ω_j and Ω_j^*, respectively, and let d denote the spatial dimension of the domain Ω. The following (informal) theorem in d dimensions, which is proved rigorously in [45, Theorem 3.7 and Section 7.3], provides an exponential bound on the eigenvalues. It relies on two important properties of the A-harmonic subspace: a weak approximation estimate, and in particular, a Caccioppoli-type inequality, which is key to achieving exponential convergence rates for multiscale problems with low regularity solutions. We refer to [45, Section 3.1] for a precise statement of those two properties and for a detailed proof.

Theorem 3.2. *Let h be sufficiently small. Then there exist $k_j, b_j, C_j > 0$ independent of h, such that for $k > k_j$*

$$\left(\lambda^{j,k}\right)^{-1} \le C_j e^{-b_j k^{1/d}}. \tag{26}$$

The constant C_j is proportional to the norm of the partition of unity operator, and k_j and b_j can be derived explicitly. The value of b_j and thus the convergence rate grows with the amount of oversampling, i.e., with decreasing H_j/H_j^*.

Combining the exponential bound (26) on the local eigenvalues and the global error estimate (25) provides a rigorous, exponential error bound for the MS-GFEM method. It is important to note that the exponential growth rate of the local eigenvalues critically relies on the two aforementioned ingredients of the new coarse space, i.e., oversampling and A-harmonicity. The global error estimate (25) also holds for the standard GenEO coarse space when used as a coarse approximation. However, without the two key ingredients, the eigenvalues of the local GEVP (13) do not grow exponentially fast, making the standard GenEO coarse space significantly less efficient; see Subsection 5.4.

Apart from the exponential decay rate of the error with respect to the number of local basis functions, Theorem 3.2 also provides an explicit decay rate of the error with respect to the oversampling size, which offers a second handle to control the error of the method, in addition to a change in the size of the local approximation spaces. This turns out to be of great importance in reducing the size of the global coarse problem; see Subsection 5.6.

We end this subsection by discussing ways to solve the local GEVP (19). To take into account the A-harmonic condition, a straightforward yet time-consuming way for solving (19) is to first construct the basis functions of the A-harmonic FE space by solving many local boundary value problems [30,32,46]. Instead, we use a different and more efficient method proposed in [38], where the A-harmonic condition is directly incorporated into the local GEVP. To this end, a Lagrange multiplier is introduced and the local GEVP (19) is rewritten in an equivalent mixed formulation: Find $\lambda \in \mathbb{R}$, $\varphi_h \in V_{h,D}(\Omega_j^*)$ and $p_h \in V_{h,DI}(\Omega_j^*)$ such that

$$\begin{aligned} a_{\Omega_j^*}(\varphi_h, v_h) + a_{\Omega_j^*}(v_h, p_h) &= \lambda\, a_{\Omega_j^*}(\Xi_j(\varphi_h|_{\Omega_j}), \Xi_j(v_h|_{\Omega_j})), && \forall v_h \in V_{h,D}(\Omega_j^*), \\ a_{\Omega_j^*}(\varphi_h, \xi_h) &= 0, && \forall \xi_h \in V_{h,DI}(\Omega_j^*). \end{aligned} \tag{27}$$

The implementation details of how the augmented system (27) is solved are presented in Subsection 4.3.

4. Implementation

The DUNE [42] package is an open-source, modular toolbox for the numerical solution of PDE problems. It leverages advanced C++ programming techniques in order to provide modularity from the ground up while producing highly efficient applications. As

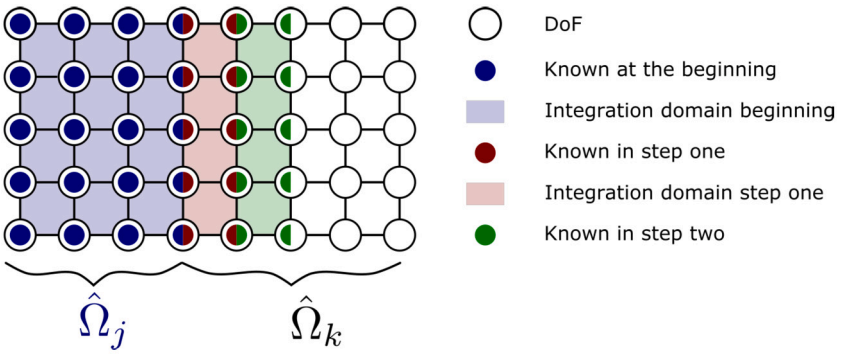

Fig. 2. The degrees of freedom known to process j after each step of recursive extension of the matrix connectivity graph.

such, it allows the reuse of many existing components when implementing our new mathematical methods for HPC applications. The new methods were integrated in the `dune-composites` module [1], which facilitates setting up elasticity models and provides access to efficient solvers that scale to thousands of cores on modern HPC systems, despite the typically bad conditioning of composites problems. This was achieved through an HPC-scale GenEO implementation [17] developed as part of `dune-composites` and later moved into the lower-level discretization module `dune-pdelab` [47] within DUNE. The `dune-pdelab` module and several lower-level DUNE modules are used within `dune-composites` to obtain the finite element discretizations on the fine level.

The DUNE framework contains a number of grid implementations for various purposes. Among those, only `YASPGrid`, a structured cube grid, provides full support for overlapping subdomains: Each process holds a copy of elements in its subdomain's overlap region and may exchange data attached to associated DoFs with neighboring processes. Since the methods considered here are constructed using overlapping subdomains, we need this kind of communication mechanism. In previous work within `dune-composites` [1], the restriction to cuboid domains induced by `YASPGrid` was overcome by applying a geometric transformation to the grid. However, for many relevant engineering applications, a smooth transformation from a cuboid geometry to the actual geometry is not available. Thus, DoF-based communication was extended to support unstructured grids as well. In order to handle more general model geometries, the grid creation was first shifted to a pre-processing step using *Gmsh* [48]. The grid import is facilitated by the IO functionality of `dune-grid`.

4.1. Algebraic overlap construction

In order to provide overlaps and communication across overlaps on unstructured grids, we use an algebraic approach. Based on the connectivity graph (i.e., sparsity pattern) of the finite element matrices $\hat{A}_j$, assembled on non-overlapping subdomains $\hat{\Omega}_j$, we construct matrices A_j defined on overlapping subdomains Ω_j. Fig. 1 illustrates our notation for non-overlapping and overlapping subdomains. We begin with a matrix $\hat{A}_j$ on a non-overlapping subdomain $\hat{\Omega}_j$. We assume that $\hat{\Omega}_j$ has neighboring subdomains $\hat{\Omega}_k$, $k \in E_j$, with corresponding matrices $\hat{A}_k$.

As a starting point, we identify coinciding DoFs in $\hat{A}_j$ and $\hat{A}_k$ from global indices. Even non-overlapping DUNE grids provide the latter on subdomain boundaries. On subdomain k, we now identify all interior DoFs directly connected to DoFs in subdomain $\hat{\Omega}_j$. Process k holds unique indices to this newly identified layer of DoFs as well as their connectivity, which we pass to process j. Since connected DoFs in a FE discretization are either associated with the same or adjacent elements, we now are in a position where each process is aware of its neighbors' DoFs within the first layer of elements along the respective subdomain boundary.

Through recursive application of this algorithm, as shown in Fig. 2, we can now grow the algebraic overlap by an arbitrary number of layers of elements, without having to rely on any overlap support in the grid implementation. In the following, we will define the number of layers of elements added to a subdomain as o (for overlap) – even though the resulting overlap of a domain with its neighbor is in fact $2o$ elements. Finally, using the extended connectivity graphs above we can employ the communication mechanisms in DUNE [49,50] to exchange overlap data of vectors between neighbors. We now have the underlying communication infrastructure to generate a GenEO space in parallel.

The method above allows us to extend connectivity graphs of non-overlapping matrices into neighboring subdomains, constructing the connectivity graphs of the overlapping matrices A_j. However, GenEO obviously requires the actual matrix entries. Using communication across the algebraic overlap, matrix entries in the overlap can be computed by the process that holds the respective DOFs and then communicated to its neighbors. As shown below, simply using neighboring matrix entries from the readily available $\hat{A}_k$ is insufficient due to edge cases.

The FE matrix entry $(A_j)_{qr}$ associated with DoFs q and r can be expressed as

$$(A_j)_{qr} = a_{\hat{\Omega}_j}(\boldsymbol{\phi}^{(q)}, \boldsymbol{\phi}^{(r)}) + \sum_{k \in E_j} \underbrace{a_{\Omega_j \cap \hat{\Omega}_k}(\boldsymbol{\phi}^{(q)}, \boldsymbol{\phi}^{(r)})}_{:=(\tilde{A}_{j,k})_{qr}} = (\hat{A}_j)_{qr} + \sum_{k \in E_j} (\tilde{A}_{j,k})_{qr},$$

where we exploit the elementwise decomposition of a in Equation (8) and the decomposition

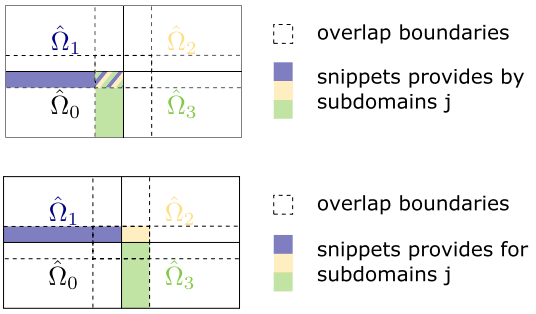

Fig. 3. Subdomain snippets used to assemble overlapping matrices from those assembled on a non-overlapping grid partition. Top: Snippets $\hat{\Omega}_0 \cap \Omega_k$ used to assemble matrices $\tilde{A}_{k,0}$ on process 0, to be sent to neighboring processes $k = 1, 2, 3$ for assembly of A_k. Bottom: Snippets $\Omega_0 \cap \hat{\Omega}_k$ used to assemble matrices $\tilde{A}_{0,k}$ on processes $k = 1, 2, 3$, to be sent to process 0 for assembly of A_0.

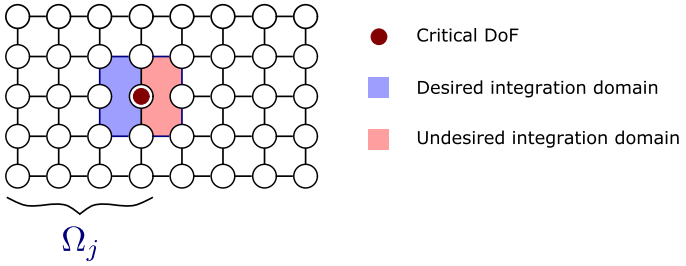

Fig. 4. Integration domain $\text{supp}(\boldsymbol{\phi}^{(q)})$ of DoF q, which is located on $\partial\Omega_j \cap \hat{\Omega}_k$. Since q is in the interior of $\hat{\Omega}_k$, process k takes the entire integration domain around the DoF into account when computing its local matrix $\hat{A}_k$. To assemble A_j process j requires integration only up to the overlapping subdomain boundary $\partial\Omega_j$, so that process k's matrix entries from $\hat{A}_k$ can not be used.

$$\Omega_j = \hat{\Omega}_j \cup \bigcup_{k \in E_j} \left(\Omega_j \cap \hat{\Omega}_k \right).$$

The matrix $\hat{A}_j$ is already known on process j. We call $\tilde{A}_{j,k}$ the snippet matrix from non-overlapping subdomain $\hat{\Omega}_k$ for subdomain Ω_j. Process k has all necessary information to compute $\tilde{A}_{j,k}$, since it holds the non-overlapping grid partition $\hat{\Omega}_k$, and it will send the relevant entries to process j. Fig. 3 illustrates the snippets needed to construct the overlapping FE matrix A_j. Importantly, the snippet matrix $\tilde{A}_{j,k}$ may differ from the readily available $\hat{A}_k$ for any DoF q on $\partial\Omega_j \cap \hat{\Omega}_k$ when $\text{supp}(\boldsymbol{\phi}^{(q)}) \not\subset \Omega_j$. Fig. 4 illustrates this. In such a case, we have

$$(\tilde{A}_{j,k})_{qr} = a_{\Omega_j \cap \hat{\Omega}_k}(\boldsymbol{\phi}^{(q)}, \boldsymbol{\phi}^{(r)}) \neq a_{\text{supp}(\boldsymbol{\phi}^{(q)}) \cap \hat{\Omega}_k}(\boldsymbol{\phi}^{(q)}, \boldsymbol{\phi}^{(r)}) = a_{\hat{\Omega}_k}(\boldsymbol{\phi}^{(q)}, \boldsymbol{\phi}^{(r)}) = (\hat{A}_k)_{qr}.$$

In practice, we compute the partition of unity Ξ_j on process j (see section 4.2), communicate it to all neighbors, and use $\Omega_j \cap \hat{\Omega}_k = \text{supp}(\Xi_j) \cap \hat{\Omega}_k$ as a convenient proxy to determine which DoFs are part of the desired snippet.

An extension operator is finally defined such that

$$\text{ext}\left(\hat{A}_j|_{\hat{\Omega}_j} \right) = A_j|_{\Omega_j} \quad \text{and} \quad \text{ext}\left(\hat{b}_j|_{\hat{\Omega}_j} \right) = b_j|_{\Omega_j}$$

to transfer matrices and vectors defined on the non-overlapping partition $(\hat{\Omega}_j)$ towards the overlapping partition (Ω_j).

4.2. Partition of unity and oversampling

As explained in section 3.1, various partition of unity operators can be used in the GenEO theory, as long as they fulfill condition (11). One suitable choice is a smooth transition from zero on $\partial\Omega_j$ to one on $\partial(\Omega_j \setminus \cup_{k \neq j} \Omega_k)$, as illustrated in Fig. 5 (left) and used in the context of the GenEO preconditioner in [1].

To construct this partition of unity, it is enough to compute Ξ_j on process j. We attribute a weight w_p to each DoF $p \in \text{dof}(\Omega_j)$. Initially, w_p is set to 0 on $\partial\Omega_j$ and to $2 \times o$ on the remaining DoFs in Ω_j, where o is the number of layers added. For each DoF p, using the extended connectivity graph of Ω_j in its vicinity and comparing with neighboring weights, the weight w_p is then reduced incrementally in the subdomain interior such that

$$w_p = \min\left(\min_{q_1 \leq q_i \leq q_n} (w_p, w_{q_i} + 1), 2o \right),$$

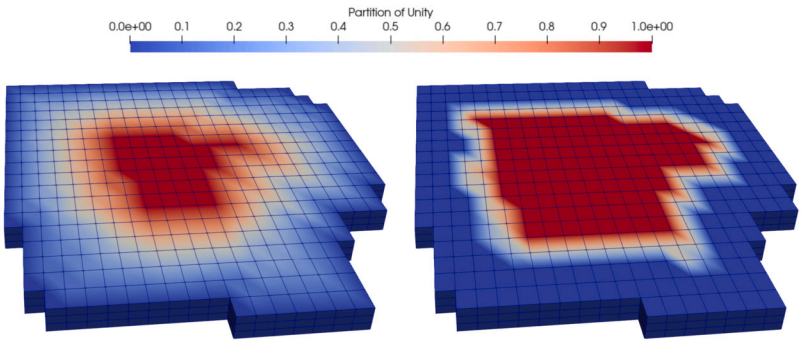

Fig. 5. Partition of unity operators Ξ_j for an overlap size of $o = 3$ (left) and $o = 1$ (right). On the right we also plot the oversampling domain Ω_j^* where $o^* = 3$, as well as the partition of unity on Ω_j where $o = 1$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

where $\{q_1, ..., q_n\}$ are the neighboring DoFs of p. It suffices to iterate this $2o - 1$ times. As an example, the partition of unity in Fig. 5 (left) has been constructed with $o = 3$. The communication mechanism between subdomains is used to associate to each DoF $p \in \text{dof}(\Omega_j)$ their corresponding counterpart in the neighboring subdomain Ω_k. The partition of unity is then simply defined by specifying the coefficients $\mu_{j,p}$ in Equation (11):

$$\mu_{j,p} \in [0,1]: \qquad \mu_{j,p} = \frac{w_{j,p}}{\sum_{\{1 \le i \le N\}} w_{i,p}}, \qquad p \in \text{dof}(\Omega_j)$$

where N is the number of subdomains sharing the DoF p.

The handling of oversampling subdomains is controlled via the choice of the partition of unity. The implementation of the partition of unity for the oversampled subdomains Ω_j^* is the same as the one described above with a different initialization of w_p. In particular, w_p is initialized to 0 not only on $\partial\Omega_j$, but also for a further $o^* - o$ layers of DoFs before applying the iterative process above. Thus, the partition of unity is a vector defined on the full oversampling subdomain Ω_j^* that takes the value zero on $\Omega_j^* \setminus \Omega_j$. In the following, we will choose a variable oversampling size o^* and keep one layer of non-zero partition of unity overlap, i.e., $o = 1$, as shown in Fig. 5 (right).

4.3. Enforcing A-harmonicity within the GEVP in DUNE

As described in [38], the eq. (27) can be formulated as a matrix eigenvalue problem. To achieve that, the DoF associate with Ω_j^* are partition into three sets:

$$\mathcal{B}_1 = \text{dof}\Big((\Omega_j^* \setminus \partial\Omega_j^*) \cup (\partial\Omega_j^* \cap \Gamma_N)\Big), \quad \mathcal{B}_2 = \text{dof}\Big(\partial\Omega_j^* \setminus (\Gamma_D \cup \Gamma_N)\Big), \quad \mathcal{B}_3 = \text{dof}\Big(\partial\Omega_j^* \cap \Gamma_D\Big).$$

Those sets of DoF are depicted in Fig. 1. We also define n_i as the sizes of the corresponding set $\mathcal{B}_i$. The GEVP in matrix form is defined as follows: Find $\lambda_h \in \mathbb{R}$, $\tilde{\phi}_j = (\phi_{j,1}, \phi_{j,2}) \in \mathbb{R}^{n_1+n_2}$ and $p \in \mathbb{R}^{n_1}$:

$$\begin{pmatrix} A_{j,11} & A_{j,12} & A_{j,11} \\ A_{j,21} & A_{j,22} & A_{j,21} \\ A_{j,11} & A_{j,12} & 0 \end{pmatrix} \begin{pmatrix} \phi_{j,1} \\ \phi_{j,2} \\ p \end{pmatrix} = \lambda \begin{pmatrix} B_{j,11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \phi_{j,1} \\ \phi_{j,2} \\ p \end{pmatrix}, \tag{28}$$

where $A_{j,mn} = a_{\Omega_j^*}(\varphi_k, \varphi_l)_{k \in \mathcal{B}_m, l \in \mathcal{B}_n}$ and $B_{j,11} = a_{\Omega_j^*}(\Xi_j \varphi_k, \Xi_j \varphi_l)_{(k,l) \in \mathcal{B}_1}$. The blocks B_{12}, B_{21}, B_{22} are zero since the partition of unity vanishes on $\mathcal{B}_2$. The oversampling, created via the choice of partition of unity, will also affect the right hand side of eq. (28), such that not only the blocks B_{12}, B_{21}, B_{22} but also all entries in $B_{j,11}$ corresponding to the oversampling region $\Omega_j^* \setminus \Omega_j$ will be zero (see Fig. 5). The GEVP solution to construct the local basis is then $\phi_j = (\phi_{j,1}, \phi_{j,2}, \phi_{j,3}) \in \mathbb{R}^{n_1+n_2+n_3}$, where $\phi_{j,3}$ is a zero-vector corresponding to the DoFs in $\mathcal{B}_3$, which combined with $(\phi_{j,1}, \phi_{j,2})$ provides an A-harmonic coefficient vector on all of Ω_j^*. Instead of generating each block $A_{j,nm}$ individually, it is extracted from A_j using the sets of DoFs $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$. In practice, the matrix

$$\begin{pmatrix} A_{j,11} & A_{j,12} & A_{j,11} \\ A_{j,21} & A_{j,22} & A_{j,21} \\ A_{j,11} & A_{j,12} & 0 \end{pmatrix} = \begin{pmatrix} M_1 & M_2^T \\ M_2 & 0 \end{pmatrix}$$

is built by block. The top left block M_1 is obtained by removing the rows and columns corresponding to $\mathcal{B}_3$ from A_j. In the elasticity case, Dirichlet boundary conditions are imposed in DUNE by altering rows in the matrix A_j – one on the diagonal, zero elsewhere – so the set $\mathcal{B}_3$ can be easily detected. Then finally M_1 is obtained by removing the rows and columns corresponding to $\mathcal{B}_2$. The block that is removed corresponds to M_2. DoFs belonging to $\mathcal{B}_2$ are detected and saved during the overlap creation phase.

Once the eigenvectors are computed, the coarse space V_H is finally obtained as shown in eq. (14) by multiplying the eigenvectors by the partition of unity. To improve the conditioning of the coarse problem, following this multiplication a re-orthogonalization

step via a Gram-Schmidt process is carried out. Contrary to [38], we directly solve the eigenproblem in eq. (28). The simplification proposed in [38] for diffusion problems requires a special handling of subdomains that touch the global Neumann boundary Γ_N that is more involved for linear elasticity in 3D.

4.4. Boundary conditions at the coarse level

The eigenvectors $(\phi_j^k)_{k\in(1,m_j)}$ of the local GEVPs on Ω_j^* have to be combined with a local particular solution ψ_j if the body force $\mathbf{f}$ is nonzero or if Ω_j^* touches the global Dirichlet boundary and the boundary displacement $\mathbf{h}$ is nonzero. Let N_j be the number of DoFs in Ω_j^*. Then, the particular solution ψ_j in (21) and (22) can be obtained by solving the local linear system

$$A_j\psi_j = b_j \quad \text{where} \quad A_j \in \mathbb{R}^{N_j}\times\mathbb{R}^{N_j} \quad \text{and} \quad b_j = \text{ext}(\hat{b}_j) \in \mathbb{R}^{N_j}. \tag{29}$$

Finally, the solution ψ_j is multiplied by the partition of unity. The resulting vector is denoted by $\hat{\psi}_j$ and appended to the basis.

Dirichlet boundary conditions are then imposed at the coarse level by altering row ℓ_j of the coarse matrix A^H and the coarse vector b^H, where ℓ_j denotes the index corresponding to $\hat{\psi}_j$ in the coarse system. Since the local Dirichlet data is taken into account in ψ_j, the index ℓ_j corresponds to the boundary DoF on Ω_j^*. As usual in DUNE a Dirichlet boundary condition is enforced at that DoF by setting:

$$A^H_{\ell_j\ell_j} = 1, \quad A^H_{\ell_j k} = 0 \ \text{ for all } \quad k \neq \ell_j, \ \text{ and } \quad b^H_{\ell_j} = 1.$$

4.5. Hardware and software

For solving the local GEVP, we use Arpack [51] through the Arpack++ wrapper in (iterative) symmetric shift-invert mode and invert the arising sparse linear systems using UMFPack [3]. The domain partition of Ω into non-overlapping subdomains $\hat{\Omega}_j$ is carried out by the graph partitioner *ParMetis* [52].

The numerical results have been carried out on the Hamilton HPC Service at Durham University. Its last version, called *Hamilton8*, provides a total of 15,616 CPU cores, 36TB RAM and 1.9PB disk space. *Hamilton8* is composed of 120 standard compute nodes, each with 128 CPU cores (2x AMD EPYC 7702), 256 GB RAM and 400 GB local SSD storage.

5. Performance experiments with MS-GFEM on composite structures

Throughout this paper, we assume a linear elastic behavior of the composite material. For aerospace applications, considering damage behavior is essential. Ultimately, our framework will be applied to model large displacement effects and the onset of failure. To achieve this, two requirements need to be validated. Firstly, it is essential to have an efficient linear elastic solver for a large-scale model to be used within any non-linear iteration. Secondly, the approximate solution needs to accurately represent relevant damage criteria, especially local extrema.

In this section, after a brief description of the composite models, we first analyze the output from all components of the MS-GFEM when applied to a composite beam. The method is then applied to a complex, aerospace part using a composite failure criterion to assess the accuracy. Finally, the parallel scaling of the method is investigated, demonstrating the efficiency of the proposed MS-GFEM on large composite structures.

5.1. Specification of the considered composite model problems

In the following three subsections, first a laminated beam under compression is considered in order to evaluate the performance of the method proposed in this study. It is illustrated in Fig. 6 (a). The laminated beam has a length of 500 mm, a width of 140 mm and a thickness of 6 mm. The laminate is made up of a stack of three layers (or plies) of the same thickness. Each layer represents a uni-directional composite made up of carbon fibbers embedded into resin. Plies are modeled as homogeneous orthotropic elastic materials, characterized by nine parameters and a vector of orientations θ. The elastic properties of AS4/8552 ([53]) have been chosen for this example. In the global coordinate system, the material tensor is orientated using standard tensor rotations, following the stacking sequence $[0°, 45°, -45°]$. For more details see, e.g., [54]. The study considers the elastic behavior of the laminated beam under compression. The uni-axial compression, illustrated in Fig. 6 (a), is modeled via a displacement imposed through Dirichlet boundary conditions.

Then for the remainder of the paper, a realistic aerospace part is used to demonstrate the high quality of the achieved coarse approximation. The aerospace part in question is a 500 mm long C-shaped wing spar section (C-spar) with a joggle region in its center, creating a geometric feature in the structure, see Fig. 6 (b). The material is a laminated composite, composed of 24 uni-directional layers (carbon fibers and resin) of $0.2mm$ each, which are orientated as follows:

$$\left[(45°, -45°)_3, (0°, 90°)_3\right]_s$$

The behavior of the C-spar under compression will be investigated using the same boundary conditions as the ones used in the beam example above.

In both examples, as visualized in Fig. 6, for the local approximation we use FE grids with piecewise linear elements, reduced integration to avoid shear-locking and one element through thickness per layer.

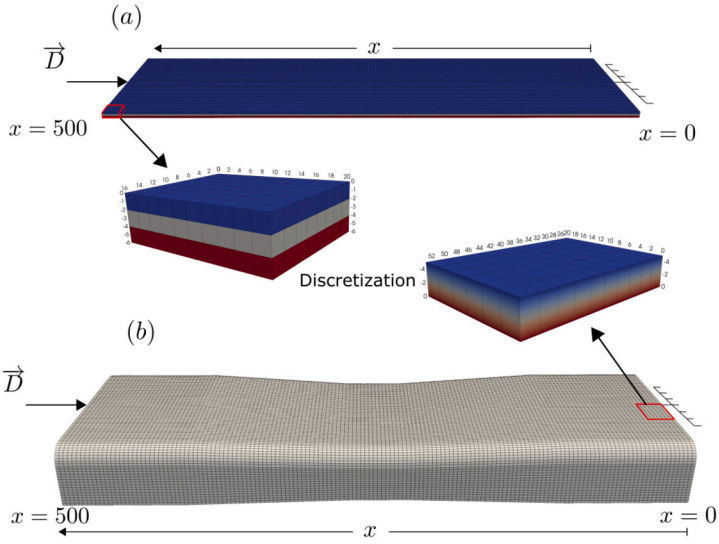

Fig. 6. Models used in the numerical experiments: laminated beam (a) and laminated C-spar (b). In both cases, Dirichlet boundary conditions are applied on all DoF belonging to the end surfaces $x = 0mm$ and $x = 500mm$.

5.2. Local GEVP outputs

We start by analyzing the local generalized eigenvalue problems (GEVP), in particular the behavior of the smallest eigenvalue $\lambda_j^{m_j+1}$ corresponding to any eigenvector not included in the local approximation space, which is the main parameter driving the method accuracy (cf. the error bound in eq. (25)). First, the decay of $1/\lambda_j^i$ for representative subdomains is analyzed, as well as the shape of the associated eigenvectors. Then, the effect of the oversampling size o^* is discussed. A particular focus will be on the local finite element aspect ratio which severely affects the accuracy of the computed eigenvectors/-values and thus also the observed actual decay of the error bound.

The beam domain in Fig. 6 (a) is decomposed into 64 subdomains for the first experiment; see Fig. 7 (b). In each subdomain, a local eigenvalue problem is solved to construct the local approximation space with a fixed number of $o^* = 8$ oversampling layers. In Fig. 7 (c,d) a selection of exemplary, local GEVP solutions on two representative subdomains are presented together with a semi-logarithmic plot of $1/\lambda_j^i$ in Fig. 7(a). For the eigenvalue plot two further subdomains are added. In total, there is one subdomain intersecting Γ_D ($j = 29$), one intersecting $\partial\Omega$ ($j = 64$) and two interior subdomains ($j = 55, 56$), with the first one of the two having a higher surface-to-volume ratio.

The semi-logarithmic plot of $1/\lambda_j^i$ in Fig. 7(a) demonstrates the predicted, nearly exponential decay of the local approximation error with respect to the basis size (m_j) in all cases. The decay is faster for subdomains intersecting $\partial\Omega$; the more the subdomain intersects $\partial\Omega$ the faster is the decay. A further factor is the surface-to-volume ratio. The higher this ratio the slower the exponential decay of $1/\lambda_j^i$, as exemplified by the relative decays of subdomains $j = 29, 64, 55$ and 56. The partitioning of the domain could be optimized to unify the surface-to-volume ratio over all subdomains. In fact, for simple, laminated composites a regular domain decomposition could be chosen, e.g., into rectangular subdomains. Here, however, we do not make this choice in order to show the robustness of the approach to rather general subdomain partitionings, as provided by automatic graph partitioners such as *ParMetis* [52], and thus to show the potential of the approach for simulating very complex structures.

In the interior subdomain $\Omega^*_{j=56}$, the first six eigenvectors (indices 0 to 5) depicted in Fig. 7 (d) correspond to the zero energy modes (or rigid body modes) representing shifts and rotations of the structure. The following modes for $\Omega^*_{j=56}$ and the first few modes for $\Omega^*_{j=29}$ in Fig. 7 (c) correspond to low-energy deformations of the subdomain: bending and shearing. The higher-energy modes correspond to higher frequency deformations ($i > 50$).

As explained in Section 3.3, the oversampling parameter o^* is key to the decay rate of the local approximation errors. In Fig. 8, the reciprocal eigenvalues $1/\lambda_j^i$ for subdomain $\Omega^*_{j=56}$ are presented for a range of oversampling sizes. The decay of $1/\lambda_j^i$ clearly accelerates as the oversampling size is increased. Thus, for a desired error bound (i.e. $1/\lambda_j^i < 10^{-7}$), subdomains with larger amounts of oversampling lead to significantly smaller local bases and consequently to an overall smaller coarse space. The trade-off, however, is that the dimensions of the local GEVPs grow significantly with the amount of oversampling and this will be discussed later. In general, the objective is to pick the minimal number of modes m_j to approximate the solution (displacement) and its derivative (strain and stress) sufficiently well. The accuracy of the coarse approximation and the effect of changing the main tuning parameters (basis / oversampling size) will be analyzed in Section 5.3. The efficiency study will be performed for the more complex C-spar structure and is presented in Section 5.6.

The results in Figs. 7 and 8 are computed using square elements with an aspect ratio of one between horizontal (through-thickness) and vertical (in-plane) edges. However, such a choice leads to an unreasonably high number of elements in larger composite structures

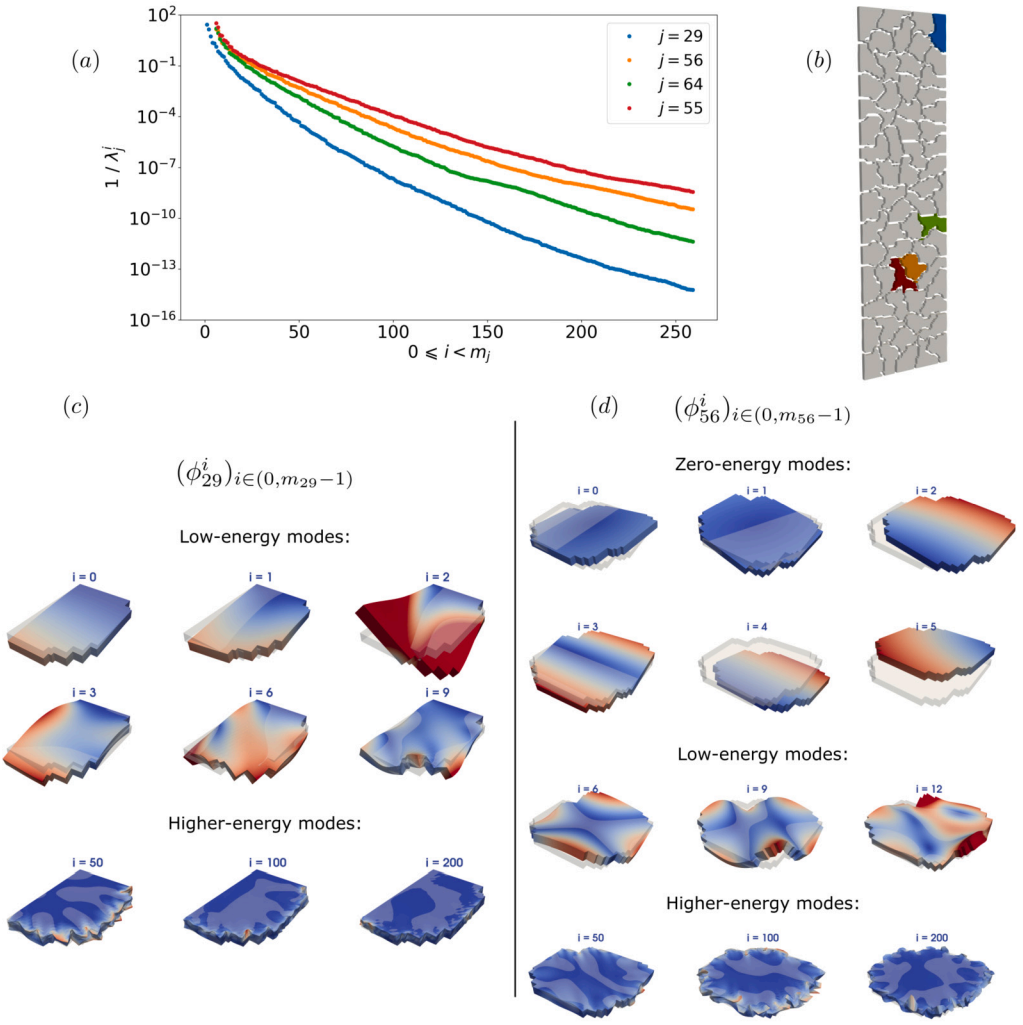

Fig. 7. Visualization of eigenpairs of the A-harmonic local GEVP with oversampling $o^* = 8$, i.e., eight layers of elements around subdomains. The reciprocal of the i^{th} eigenvalue $1/\lambda_j^i$ for four representative subdomains Ω_j is plotted in (a). The non-overlapping decomposition, with the four chosen subdomains highlighted, is presented in (b). In (c) and (d), a sample of eigenvectors ϕ_j^i for the subdomains 29 and 56 is illustrated, respectively. (For clarity, the first six eigenvalues for $\Omega_{j=55,56,64}^*$ are not plotted in (a), since by definition $\lambda_j^0 = \ldots = \lambda_j^5 = 0$.) The reader is advised to use the PDF version of the paper to fully appreciate this figure.

with a larger number of more realistic, thinner plies, such as the C-spar described in Section 5.1 and depicted at the bottom of Fig. 6. To build a reasonably sized model, the in-plane discretization has to be reduced, leading to flat elements with a larger aspect ratio. Unfortunately this reduces the accuracy significantly, as presented for four element aspect ratios and $o^* = 5$ in Fig. 9 (b), where the relative L_2-error for each eigenpair is shown. As a consequence, after a similar initial decay (up to $i = 60$) we observe a change in the slope of $1/\lambda_j^i$ for aspect ratios bigger than one in Fig. 9 (a). For very large aspect ratios of 15 and above, it even leads to a plateau in $1/\lambda_j^i$ (see Fig. 7 (a), red curve). This is to be expected, due to the larger condition numbers of the stiffness matrices in the GEVP leading to more unstable eigensolves.

Increasing the oversampling size does not alleviate this problem, as seen in Fig. 9 (c,d); the relative L_2-errors in the eigenpairs are independent of o^* and the slope of $1/\lambda_j^i$ changes roughly at the same value of i. We also tested a different type of higher-order finite element, namely a 20-DoF quadratic serendipity element, but the loss of accuracy due to high element aspect ratios persists. However, the slope change and the plateau of the eigenvalues for larger aspect ratios do not have a strong impact, since they occur only at values well below what is needed for good practial approximation, especially for higher oversampling sizes.

5.3. Coarse approximation accuracy

In this section, we investigate the accuracy of the MS-GFEM approximation. The fine-scale reference solution is computed using an iterative CG method with GenEO as preconditioner. By setting a sufficiently low tolerance, the error due to the iterative solution via CG can be neglected.

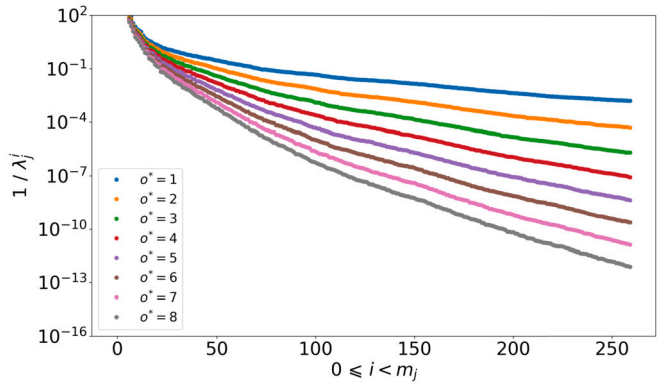

Fig. 8. Effect of the amount of oversampling on the decay of the eigenvalues, and thus on the local error bound, depicted for the interior subdomain $\Omega^*_{j=56}$.

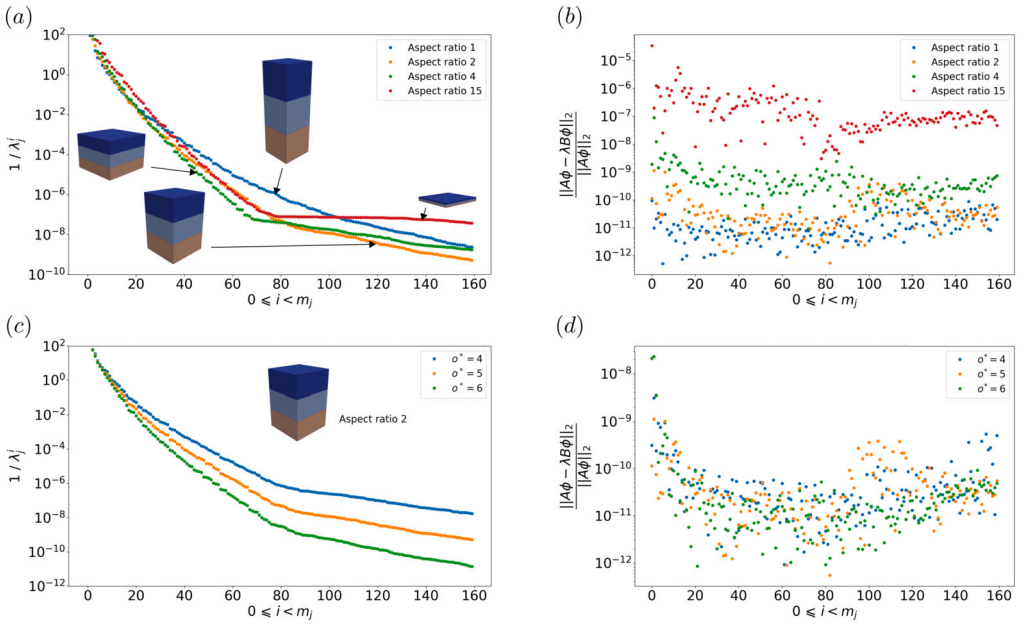

Fig. 9. The reciprocals $1/\lambda_j^i$ of the eigenvalues of the local GEVP (a,c) and the relative L_2-error in computing the associated eigenvectors (b,d); in particular, studying the effect of bad element aspect ratio (for $o^* = 5$) in (a,b), as well as oversampling size for fixed aspect ratio in (c,d).

We consider in the following the relative errors in L_2-norm between the coarse approximations of displacement and strain fields, u_H and ϵ_H, and their fine-scale counterparts, u_h and ϵ_h, i.e.,

$$e_d = \frac{\|u_h - u_H\|_2}{\|u_h\|_2}, \qquad e_\epsilon = \frac{\|\epsilon_h - \epsilon_H\|_2}{\|\epsilon_h\|_2}.$$

Fig. 10 (a) shows the fine-scale approximation of the displacement field u_h of the beam under compressive loading. The unit displacement ($\mathbf{x}(z) = -10$ mm) applied here causes the characteristic out-of-plane deformation of the structure related to the non-symmetric stacking sequence chosen for this example. A maximum displacement of 20 mm is observed on the two sides of the beam. Fig. 10 (b) depicts the fine-scale strain approximation in the direction of compression, denoted by x here. The order of magnitude of strains is $\mathcal{O}(10^{-2})$.

The main parameter driving the accuracy of the coarse approximation is the smallest eigenvalue $\min_j \lambda_j^{m_j+1}$ corresponding to any local eigenvector not included in the basis. As explained above, the size m_j of each of the local bases and the oversampling size o^* will be decisive to control this parameter. For our analysis we vary these two key parameters up to maxima of $m_j = 250$ and $o^* = 8$, for which the error bound lies below 10^{-15}. The relative L_2-errors for displacement and strain are shown in Fig. 10, resp. (c) and (d), using a logarithmic scale. Both e_d and e_ϵ decay exponentially with respect to m_j for all amounts of oversampling. The decay rate of the errors is higher for larger oversampling sizes, which is in agreement with the behavior of the reciprocal eigenvalues depicted in Fig. 8.

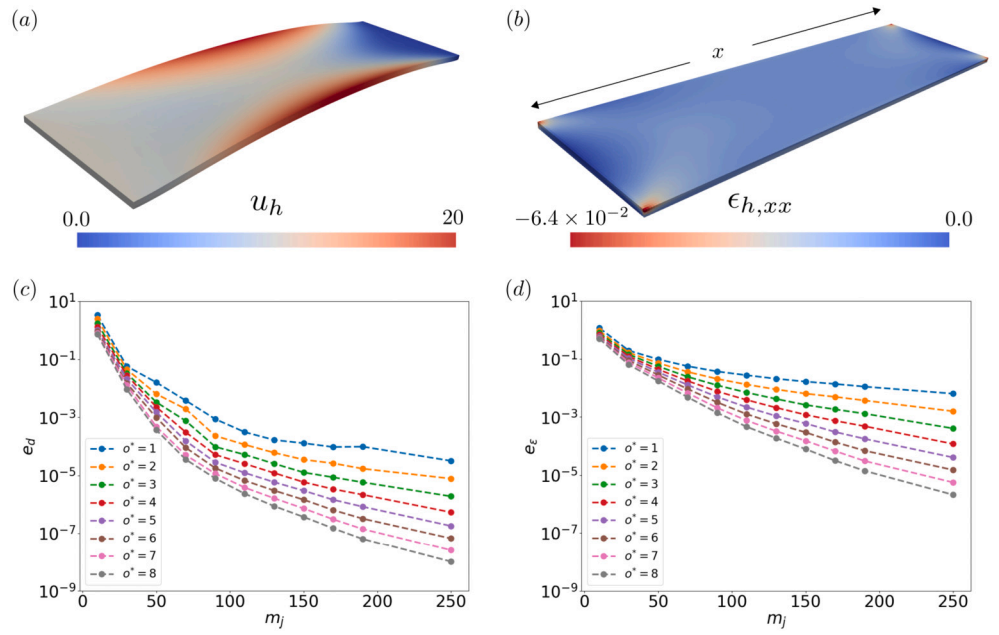

Fig. 10. L_2-errors versus local basis size for the displacement field (c) and the strain field (d). The reference fine-scale approximations are displayed above in (a) and (b).

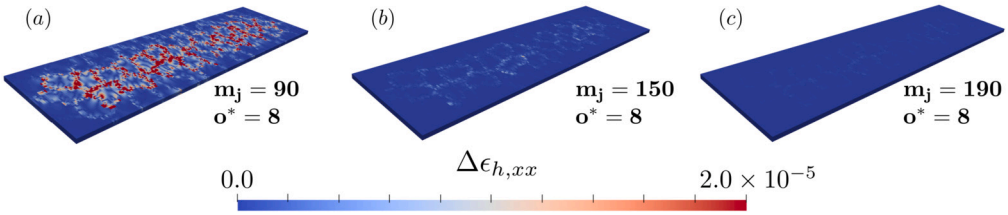

Fig. 11. Localization of the error depending on the local space size m_j. The absolute difference $\Delta\epsilon_{h,xx}$ between the coarse approximation and the fine scale approximation is plotted.

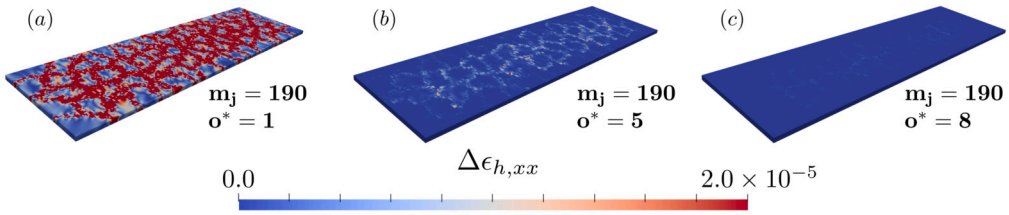

Fig. 12. Localization of the error depending on the oversampling size o^*. The absolute difference $\Delta\epsilon_{h,xx}$ between the coarse approximation and the fine scale approximation is plotted.

We further investigate the location of the largest strain error, in order to explain in more detail its dependence on m_j and o^*. The coarse approximation of the strain field has been projected onto the fine-scale space to allow a direct comparison element-by-element. The absolute differences $\Delta\epsilon_{h,xx}$ between the fine-scale solution and the coarse approximation of the strain field are depicted in Fig. 11 for three different values of the local basis size per subdomain, $m_j = \{90, 150, 190\}$, and a fixed value of $o^* = 8$. The aim is to investigate the error behavior alongside the convergence curve for $o^* = 8$ in Fig. 10 (d). The three chosen basis sizes produce coarse-space solutions of high accuracy with relative L_2-errors bellow 10^{-3}. The error maxima are observed on subdomains located in the center of the beam. Indeed, the local error on each subdomain follows the decrease of $1/\lambda_j^i$. As shown in Fig. 7 (a), interior subdomains (e.g., $\Omega^*_{j=56}$) need more eigenvectors for the same local error. Eventually, a sufficiently large local basis size ensures an accurate approximation for all subdomains.

Conversely, for fixed $m_j = 190$ Fig. 12 assesses the impact of the oversampling size, varying $o^* \in \{1, 5, 8\}$. As predicted by the theory, a smaller oversampling size leads to a larger error, which is localized mainly at subdomain interfaces. Increasing the oversampling area decreases the interface errors, and, eventually a smooth and very low error is observed across the domain for

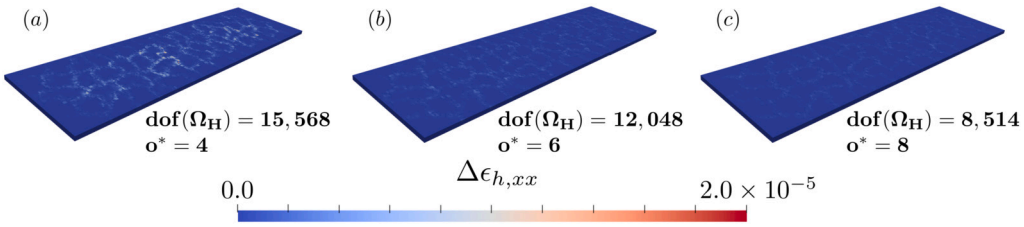

Fig. 13. Localization of the error depending on the oversampling size o^* for a given targeted accuracy. The absolute difference $\Delta\epsilon_{h,xx}$ between the coarse approximation and the fine scale approximation is plotted.

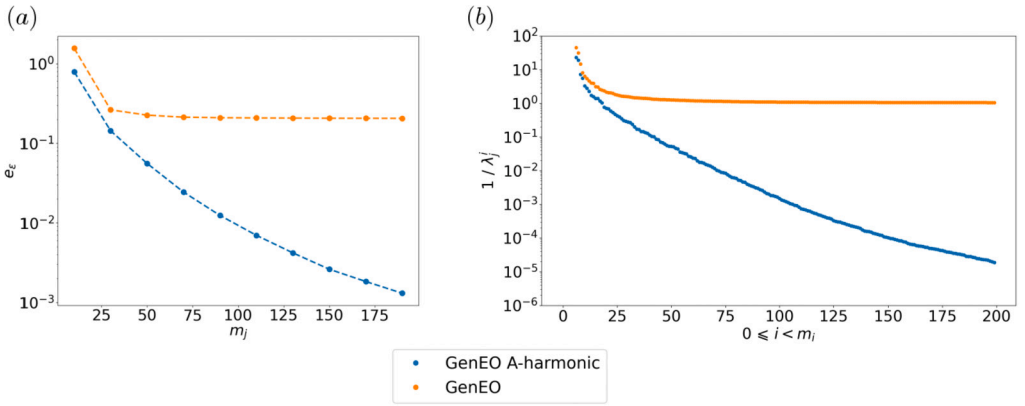

Fig. 14. Accuracy of the GenEO-type coarse spaces. Absolute relative L_2-error on the strain field for both the classical GenEO and the A-harmonic GenEO (MS-GFEM) approximations (a). The quotient $1/\lambda_j^i$ for one particular subdomain for each of the approaches (b).

$o^* > 5$. This example showcases one of the main advantages of the MS-GFEM: the interface problem at subdomain boundaries is handled by oversampling. The coarse approximation can then resolve the displacement, strain, and stress distribution at the ply level, independently of the model choices for the coarse space construction, in particular the choice of domain decomposition. The scales represented by the artificial subdomain partitioning on the macroscale and the mesoscopic geometry and orientations of the plies within each subdomain are mixed as if the entire solution were computed on the finer scale, proving that there is no scale separation between the ply scale (mesoscopic) and the structural scale (macroscopic).

In Fig. 13, the strain error is evaluated for a fixed eigenvalue threshold $1/\lambda_j^{m_j+1} = 10^{-7}$ for various oversampling sizes, to study the impact on the overall coarse space size. For fixed accuracy, the size of the coarse space is reduced by 22.6% when increasing $o^* = 4$ to $o^* = 6$ and by 45.3% from $o^* = 4$ to $o^* = 8$. This improved model order reduction obviously reduces the cost of the coarse problem solve, but it does increase the size of the local problems and thus the cost of solving the local GEVPs. This cost trade-off will be studied more carefully in Section 5.6 for the C-spar.

5.4. *Importance of the A-harmonic condition for composite problems*

To emphasize the necessity of the A-harmonic condition in our framework, we compare the new MS-GFEM space to the classic GenEO coarse space with the GEVPs formulated directly in the FE space V_h, i.e., without enforcing A-harmonicity. The theoretical error bound in Theorem 3.1 is the same for both GFEM spaces, with and without the A-harmonic condition in the GEVP. Crucially, in both cases the approximation error is bounded by the reciprocal $1/\lambda_j^i$ of the smallest eigenvalue corresponding to any eigenvector that is not included in the coarse space V_H.

Fig. 14 (right) shows that the decay of $1/\lambda_j^i$ for $i < 10$ is comparable between classic GenEO and the A-harmonic formulation, since the first modes (rigid body, shear and bending) appear in both. In fact, the classic GenEO coarse space with a small number of lowest-eigenvalue modes was shown in Reinarz et al. [2] to provide a robust preconditioner within CG that reduces the condition number effectively and leads to a low number of CG iterations for composites problems. Beyond that, we observe in Fig. 14 (right) that the spectrum for the classic GenEO formulation decays much slower and eventually almost stagnates, when compared to MS-GFEM. This is reflected in the coarse approximation error in Fig. 14 (left), as predicted in Theorem 3.1. Even including 190 eigenvectors in classic GenEO is not sufficient to accurately represent the strain field in this application (with $e_\epsilon \approx 1$). MS-GFEM with A-harmonic GEVPs, on the other hand, achieves an excellent coarse approximation at much lower basis sizes.

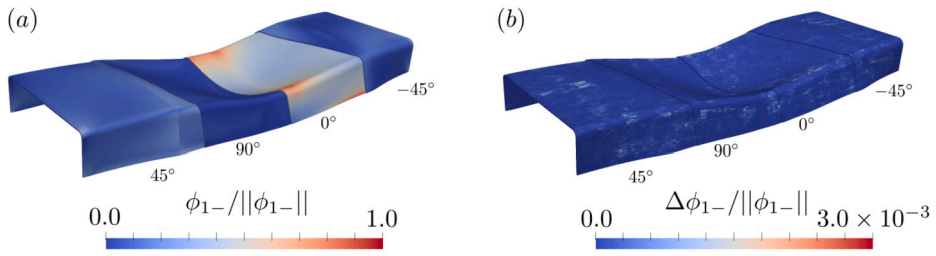

Fig. 15. Application of the method to the C-spar. The elementwise, normalized compressive failure criterion is shown in (a) with a focused view of four different plies oriented at 45°, 90°, 0° and −45°, clipped through the thickness of the laminate, successively along the length of the part. It shows the influence of local material orientation on ϕ_{1-} and the ability of our method to deal with strong through-thickness heterogeneities. The error $\Delta\phi_{1-}$ of the MS-GFEM approximation of the failure criterion with respect to a fine scale computation is shown in (b) with the corresponding visualization of individual plies through thickness.

5.5. Method accuracy on an aerospace part

In this section, the behavior of the C-spar, described in Section 5.1, under compression is investigated. The domain, shown in Fig. 6 (b), is divided into 256 subdomains. For an oversampling size of $o^* = 4$ and a tolerance of $t = 10^{-6}$ on $1/\lambda_j^{m_j+1}$ we obtain local basis sizes m_j varying between 60 and 170. These parameters lead to an average subdomain size of $\text{dof}(\Omega_j^*) = 30000$ and a model order reduction of 32×. With 4 GB of RAM per subdomain to compute the local GEVP this parameter choice requires 4 nodes of the HPC cluster *Hamilton8* for the 256 subdomains. The element aspect ratio is 15 for this example (see Fig. 6). Thus, as for the simple beam example, the possible coarse approximation accuracy is limited by the fine scale error.

In order to assess the coarse approximation, we compute

$$\begin{aligned}
\phi_{1-} &= \frac{\left\langle |\sigma_{12}^R| + \eta^L \sigma_{22}^R \right\rangle}{S_L} \\
\sigma_{12}^R &= \sigma_{11}\sin^2(\varphi_C) + \sigma_{22}\cos^2(\varphi_C) - 2|\sigma_{12}|\sin(\varphi_C)\cos(\varphi_C), \\
\sigma_{22}^R &= (\sigma_{22} - \sigma_{11})\sin(\varphi_C)\cos(\varphi_C) + |\sigma_{12}|(\sin^2(\varphi_C) + \cos^2(\varphi_C)),
\end{aligned} \tag{30}$$

the longitudinal compressive failure criterion from [55], which is a linear combination of relevant components of the Cauchy stress tensor (eq. (2)): longitudinal (σ_{11}), transverse (σ_{22}) and shear (σ_{12}). A detailed description of the material constants (η^L, S_L, φ_C) is available in [55]; the notation $\langle\cdot\rangle$ denotes the positive part of the argument. In fact, damage mechanisms such as fiber kinking [56] are observed for fiber reinforced composites under compression. Here, ϕ_{1-} (eq. (30)) is a good indicator of the activation of this damage mechanism within the structure and, along with other criteria (corresponding also to linear combinations of the Cauchy stress tensor and more material properties), is used for structural design of composite laminated parts. In slender composite structures, this failure mode typically appears after the buckling of the structure, but even in the prebuckling context, this criterion gives a good insight into the usefulness of the method for composite structural design. The longitudinal compressive failure criterion, computed in local coordinates, is plotted in Fig. 15 (a). To visualize it, the coarse approximation u_H is first projected onto the fine scale FE space V_h. The relative error with respect to a direct meso-scale approximation in V_h is shown in Fig. 15 (b). With the chosen parameters, the relative error on the failure criterion is below 3.0×10^{-3}. In Fig. 15 (a), four plies (ply 6, 12, 18 and 24) are highlighted in a certain part of the domain, representing each stacking orientation and showing that the coarse approximation is able to accurately represent the quick variation of the compressive failure criterion through thickness. The remaining error is small and uniformly distributed, with higher values near subdomain boundaries and on interior subdomains, in agreement with the observations on the beam example.

5.6. Method scalability

A parallel (weak) scaling test of the method is presented in Fig. 16, where the cost of the local GEVP solves and the cost of the coarse solve are assessed using C-spars of different lengths while fixing the discretization of the meso-structure: 24 elements through thickness (one per ply) and a constant element size in the other directions (aspect ratio 15). The C-spar models vary between $L = 62.5$ mm with $\text{dof}(V_h) = 1.53 \times 10^5$ (Model 1) and $L = 2$ m with $\text{dof}(V_h) = 4.6 \times 10^6$ (Model 7). The number of processors (P) employed for each model has been chosen such that $\text{dof}(V_h)/P$ remains constant, here $\text{dof}(Vh)/P \approx 10,000$. Therefore, there is no mesh refinement between models: the characteristic elementsize remains the same and the total number of elements scales with the length of the part. Keeping $\text{dof}(Vh)/P$ constant means that the size of the local eigenvalue problems also remains constant between the different models.

Three combinations of parameters have been tested, choosing $o^* = 4$ and $o^* = 6$ and selecting a threshold of $t = 10^{-6}$ and $t = 10^{-3}$ for the reciprocal of the local eigenvalues. Parameter sets #1 and #2 assess the scalability of the method for a highly accurate solution with $t = 10^{-6}$. In Set #1 $o^* = 4$, chosen to reduce RAM consumption, whereas in Set #2 a larger oversampling size of $o^* = 6$ is selected, thus requiring more RAM. This trade-off will be further discussed below. Set #3 has been selected to study how the method scales for a lower eigenvalue threshold of $t = 10^{-3}$ with smaller $o^* = 3$ and half the amount of subdomains in Sets #1 and #2, such that $N_j = \text{dof}(\Omega_j^*)$ is roughly the same.

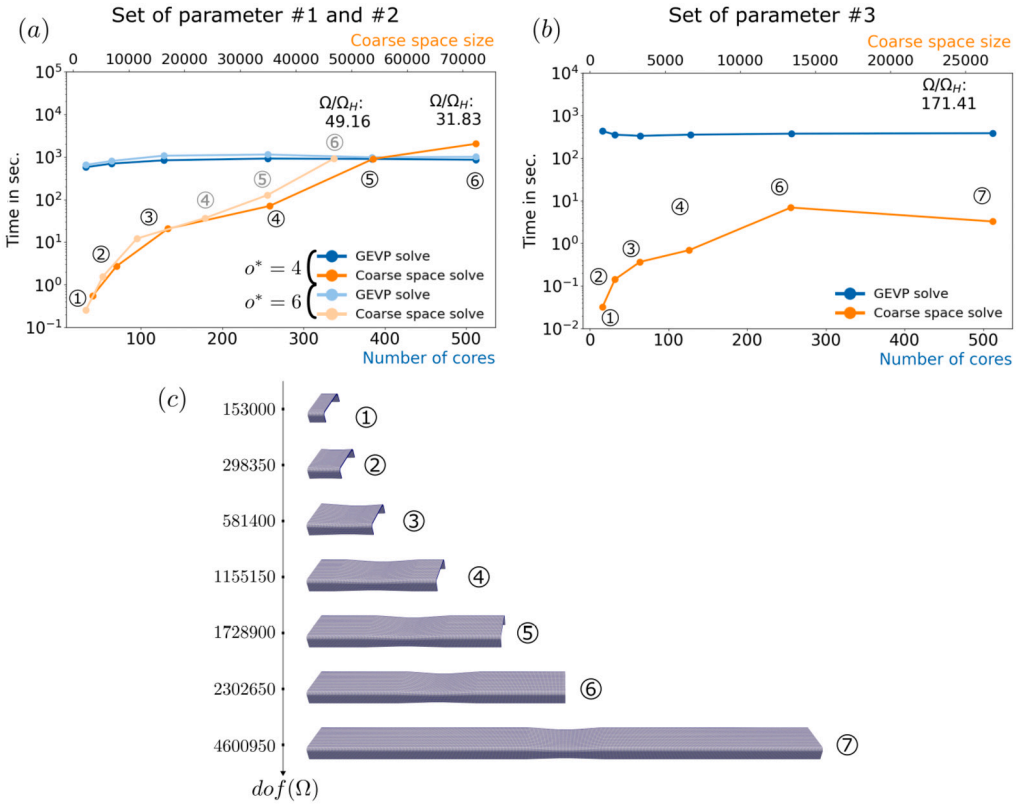

Fig. 16. Scaling of the method applied to C-spars of different lengths for three sets of parameters. In (a) the Sets #1 and #2 optimize the approximation accuracy, while in (b) Set #3 aims at a very efficient simulation with acceptable accuracy. The employed FE grids are displayed in (c). In all models, the characteristic element sizes in the fine space discretization are the same. Fig. 6 (b) represents model 4 with a detailed visualization of the mesh refinement.

As shown in Fig. 16 (a-b), the GEVP solve step scales perfectly for all three sets, since no parallel communication is required. As expected, the solve times increase slightly between Sets #1 and #2 due to the slightly bigger local problem sizes. There is also some variation in problem sizes across subdomains. However, even though the number of ARPACK iterations also varies, the cost per GEVP is essentially almost entirely due to the sparse direct solver UMFPack. For a FE discretization of a 3D elasticity problem its cost scales as expected roughly like $\mathcal{O}(N_j^{1.5})$. The overall cost for the coarse space setup for each model is essentially identical, since it is dominated by the time spent on the local GEVPs and can be carried out fully in parallel with only local data exchange. As the subdomain size N_j is kept constant during this numerical experiment, the time to solve the local particular problem (section 4.4) is constant too and does not exceed 10 seconds for any of the tested models. Therefore, to avoid cluttering we did not include that cost as a specific line in Fig. 16, similarly to the algebraic overlap construction, which required even less time per subdomain.

In contrast, the final coarse space problem (24) is (currently) solved on one processor, and thus eventually dominates the overall cost for larger models. To remain efficient, the coarse space size needs to be optimized. For all considered models, the model order reduction is around 32× for Set #1 and around 50× for Set #2. As a consequence, the increased oversampling size in Set #2 significantly reduces the cost of the coarse solve, particularly for the larger models (4–5–6). These gains clearly outweigh the higher cost for the local GEVP solves, but the amount of local memory (RAM) needed to process these GEVPs in parallel is a limit to this improvement, preventing the use of too large oversampling sizes. In this example, subdomains with $o^* = 6$ have more than $50,000$ DoF, necessitating over 8 GB of RAM per subdomain for the local GEVPs. A reasonable oversampling size thus needs to balance performance and memory consumption.

The strength of our approach is the ability to construct approximation spaces of adjustable complexity in a very simple way. A cheaper approximation can be built by choosing a lower threshold t for selecting the local eigenvectors. In particular, for Set #3, where $o^* = 3$ and $t = 10^{-3}$, the size of local bases (m_j) reduces to between 30 and 50, leading to a model order reduction of about 170×. This allows the construction and solve of the coarse-space approximation for a 2 m C-spar (Model 7) on 512 processors in 10 minutes, thus extending the range of our scalability test. As shown in Fig. 16 (b), the model order reduction is then sufficient to keep the computational times for the coarse solves under 10 seconds in all cases, even for Model 7. In fact, in contrast to the situation for parameter sets #1 and #2, the time spent on the coarse solve for set #3 is dominated by the communication cost for collecting the coarse system on one processor.

The accuracy of the solution of two sets of parameters: Sets #1 and #3 are compared in Fig. 17 for Model 6. Both qualitatively, as visible on Fig. 17, and quantitatively (relative L_2-errors) the displacement field u_H appears to be well represented for both sets, which is in agreement with the beam example observations. However, as expected, since it is derived from the stresses (eq. (30)), the

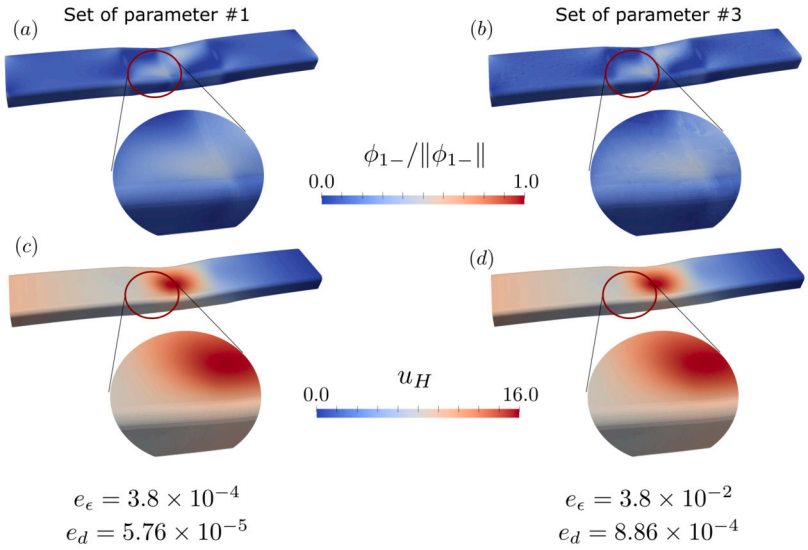

Fig. 17. Visualization of the solution accuracy for Set #3 (b,d) compared to Set #1 (a,c) for Model 6 (see Fig. 16), plotting for both sets in (a, b) the compressive criterion (eq. (30)) and in (c, d) the displacement field. The relative L_2-errors (on strain e_ϵ and displacement e_d) with respect to the fine scale approximation are associated with each parameter set.

accuracy of the compressive criterion in Set #3 is affected by the small basis size. Despite a visible noise on the compressive failure criterion of Set #3 (see the zoom-in Fig. 17), global extrema are preserved. Hence, this set of parameters is able to detect the global maximum of the criterion at very cheap cost.

6. Conclusion & future work

In this paper, we have presented the first scalable HPC implementation of a MS-GFEM method and demonstrated that it delivers high quality approximate solutions for very small coarse space sizes. As proven in previous theoretical work, this is due to the nearly exponential decay of the reciprocal eigenvalues in the local generalized eigenvalue problems. Here, we have demonstrated that this nearly exponential decay crucially relies on enforcing an A-harmonic constraint on the local eigenproblems also in composites applications, and that oversampling of the local subdomains is essential to achieve good accuracies at small local basis sizes. While the related GenEO-coarse space, which does not enforce A-harmonicity in the local eigenproblems, also leads to acceptable results in approximating displacements, we have seen that A-harmonicity is crucial to accurately approximate strains, stresses and derived failure criteria in composite applications.

We have demonstrated good parallel scalability on several hundreds of processor cores. While a single solve of the fine-scale problem is cheaper using, e.g., GenEO-preconditioned Krylov methods, the localized approach of MS-GFEM opens up new opportunities for parallel scalability. When solving large numbers of closely related problems, eigenvectors from unaffected subdomains may be retained, solving costly eigenproblems only where model parameters or geometry changes between runs. If only a few subdomains are affected, the global solution can be computed using a significantly smaller number of processors in the same time as a full run. This is especially interesting in future Uncertainty Quantification (UQ) applications, such as the impact of (meso-scale) localized wrinkles in composite structures on the strength or the failure behavior of the overall (macro-scale) structure. In such applications, the problem setup will essentially be identical in all but a few subdomains and large numbers of runs are required.

The integration of our new MS-GFEM method into an offline/online framework, where local approximation spaces will only be updated in a few subdomains between runs, is currently ongoing and will form part of a subsequent publication. This will also include the application of the offline/online framework as part of Uncertainty Quantification (UQ) methods for composites; in particular, exploiting the natural hierarchy of approximate models in the MS-GFEM framework within multilevel UQ methods such as Multilevel Monte Carlo (MLMC) [57] or Multilevel Markov Chain Monte Carlo (MLMCMC) [58].

The coarse space solves have been handled by a sequential direct solver here. A way to push scalability beyond the limits imposed by coarse system size would be to parallelize the coarse solve. Since the coarse system is itself ill-conditioned, this is not trivial. Earlier experiments conducted for [2] (but not included) indicate that, in case of GenEO, straightforward application of AMG is not sufficient. This aspect and management of parallel resources will be explored in future work.

On the practical side, our approach has a largely automatic workflow, from domain decomposition to the automatic generation of a coarse space specifically tuned to the given problem. The balance between global approximation error and basis size can be controlled by setting a single threshold for the selection of eigenvectors. In multiscale applications, and specifically composites, MS-GFEM is particularly interesting since we obtain a low-dimensional approximation without assuming scale separation. Instead, the eigenproblems capture the structure of the given problem, providing better quality than hand-tuned approximations. The resulting coarse space then accurately captures fine- and coarse-scale interaction.

A very relevant aspect in the study of composite materials is material failure under load. In contrast to linear elasticity (as covered by this work), non-linear models are needed to simulate the failure of composites aero-structures (non-linear geometry, damage initiation and propagation). We are therefore also extending our methods to nonlinear solvers and implement nonlinear material behavior in `dune-composites`.

CRediT authorship contribution statement

Jean Bénézech: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Linus Seelinger:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. **Peter Bastian:** Methodology, Software, Supervision, Writing – review & editing. **Richard Butler:** Funding acquisition, Supervision, Writing – review & editing. **Timothy Dodwell:** Conceptualization, Funding acquisition, Methodology. **Chupeng Ma:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Robert Scheichl:** Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The main source code is available at 'https://gitlab.dune-project.org/anne.reinarz/dune-composites/'. FE grids generation code is available at 'https://github.com/jeanbenezech/CompositesFEMesh'.

Acknowledgement

The research was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) through the Programme Grant: 'Certification of Design: Reshaping the Testing Pyramid' EP/S017038/1 (https://www.composites-certest.com/). This multidisciplinary project aims at developing new approaches to enable the design and certification of lighter, more cost and fuel efficient composite aero-structures. The funding received is gratefully acknowledged. This work made use of the facilities of the Hamilton HPC Service of Durham University. We thank Anne Reinarz (Durham University) for her support in the HPC experiments. Butler holds a Royal Academy of Engineering-GKN Aerospace Research Chair in Composites. Seelinger, Bastian and Scheichl acknowledge support of the Deutsche Forschungsgemeinschaft (German Research Foundation) under Germany's Excellence Strategy EXC 2181/1 – 390900948 (the Heidelberg STRUCTURES Excellence Cluster).

References

[1] A. Reinarz, T. Dodwell, T. Fletcher, L. Seelinger, R. Butler, R. Scheichl, Dune-composites — a new framework for high-performance finite element modelling of laminates, Compos. Struct. 184 (2018) 269–278.
[2] R. Butler, T. Dodwell, A. Reinarz, A. Sandhu, R. Scheichl, L. Seelinger, High-performance dune modules for solving large-scale, strongly anisotropic elliptic problems with applications to aerospace composites, Comput. Phys. Commun. 249 (2020) 106997.
[3] T.A. Davis, Algorithm 832: UMFPACK v4.3 — an unsymmetric-pattern multifrontal method, ACM Trans. Math. Softw. 30 (2) (2004) 196–199.
[4] Dassault Systèmes, Abaqus analysis user's manual, Simulia Corp. Providence, RI, USA 40, 2007.
[5] Y. Saad, Iterative Methods for Sparse Linear Systems, SIAM, 2003.
[6] M. Blatt, P. Bastian, The iterative solver template library, in: International Workshop on Applied Parallel Computing, Springer, 2006, pp. 666–675.
[7] U.M. Yang, V.E. Henson, BoomerAMG: a parallel algebraic multigrid solver and preconditioner, Appl. Numer. Math. 41 (1) (2002) 155–177.
[8] V. Dolean, F. Nataf, R. Scheichl, N. Spillane, Analysis of a two-level Schwarz method with coarse spaces based on local Dirichlet-to-Neumann maps, Comput. Methods Appl. Math. 12 (4) (2012) 391–414.
[9] E. Eikeland, L. Marcinkowski, T. Rahman, Overlapping Schwarz methods with adaptive coarse spaces for multiscale problems in 3d, Numer. Math. 142 (1) (2019) 103–128.
[10] J. Galvis, Y. Efendiev, Domain decomposition preconditioners for multiscale flows in high-contrast media, Multiscale Model. Simul. 8 (4) (2010) 1461–1483.
[11] J. Galvis, Y. Efendiev, Domain decomposition preconditioners for multiscale flows in high contrast media: reduced dimension coarse spaces, Multiscale Model. Simul. 8 (5) (2010) 1621–1644.
[12] M.J. Gander, A. Loneland, T. Rahman, Analysis of a new harmonically enriched multiscale coarse space for domain decomposition methods, Preprint, arXiv: 1512.05285, 2015.
[13] A. Heinlein, A. Klawonn, J. Knepper, O. Rheinbach, Adaptive gdsw coarse spaces for overlapping Schwarz methods in three dimensions, SIAM J. Sci. Comput. 41 (5) (2019) A3045–A3072.
[14] A. Heinlein, A. Klawonn, J. Knepper, O. Rheinbach, Multiscale coarse spaces for overlapping Schwarz methods based on the acms space in 2d, Electron. Trans. Numer. Anal. 48 (2018) 156–183.
[15] A. Heinlein, K. Smetana, A fully algebraic and robust two-level Schwarz method based on optimal local approximation spaces, Preprint, arXiv:2207.05559, 2022.
[16] N. Spillane, V. Dolean, P. Hauret, F. Nataf, C. Pechstein, R. Scheichl, Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps, Numer. Math. 126 (4) (2014) 741–770.
[17] L. Seelinger, A. Reinarz, R. Scheichl, A high-performance implementation of a robust preconditioner for heterogeneous problems, in: R. Wyrzykowski, E. Deelman, J. Dongarra, K. Karczewski (Eds.), Parallel Processing and Applied Mathematics, Springer International Publishing, Cham, 2020, pp. 117–128.

[18] J. Guedes, N. Kikuchi, Preprocessing and postprocessing for materials based on the homogenization method with adaptive finite element methods, Comput. Methods Appl. Mech. Eng. 83 (2) (1990) 143–198.
[19] V. Kouznetsova, M.G. Geers, W. Brekelmans, Multi-scale second-order computational homogenization of multi-phase materials: a nested finite element solution strategy, Comput. Methods Appl. Mech. Eng. 193 (48–51) (2004) 5525–5550.
[20] V.P. Nguyen, M. Stroeven, L.J. Sluys, Multiscale continuous and discontinuous modeling of heterogeneous materials: a review on recent developments, J. Multiscale Model. 3 (04) (2011) 229–270.
[21] T.Y. Hou, X.-H. Wu, A multiscale finite element method for elliptic problems in composite materials and porous media, J. Comput. Phys. 134 (1) (1997) 169–189.
[22] Y. Efendiev, T.Y. Hou, Multiscale Finite Element Methods: Theory and Applications, vol. 4, Springer Science & Business Media, 2009.
[23] Y. Efendiev, J. Galvis, T.Y. Hou, Generalized multiscale finite element methods (GMsFEM), J. Comput. Phys. 251 (2013) 116–135.
[24] E. Chung, Y. Efendiev, T.Y. Hou, Multiscale Model Reduction: Multiscale Finite Element Methods and Their Generalizations, vol. 212, Springer Nature, 2023.
[25] A. Målqvist, D. Peterseim, Localization of elliptic multiscale problems, Math. Comput. 83 (290) (2014) 2583–2603.
[26] M. Hauck, D. Peterseim, Super-localization of elliptic multiscale problems, Math. Comput. 92 (341) (2023) 981–1003.
[27] L. Berlyand, H. Owhadi, Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast, Arch. Ration. Mech. Anal. 198 (2) (2010) 677–721.
[28] H. Owhadi, C. Scovel, Operator-Adapted Wavelets, Fast Solvers, and Numerical Homogenization: From a Game Theoretic Approach to Numerical Approximation and Algorithm Design, vol. 35, Cambridge University Press, 2019.
[29] H. Owhadi, L. Zhang, Gamblets for opening the complexity-bottleneck of implicit schemes for hyperbolic and parabolic odes/pdes with rough coefficients, J. Comput. Phys. 347 (2017) 99–128.
[30] I. Babuška, R. Lipton, Optimal local approximation spaces for generalized finite element methods with application to multiscale problems, Multiscale Model. Simul. 9 (1) (2011) 373–406.
[31] I. Babuška, X. Huang, R. Lipton, Machine computation using the exponentially convergent multiscale spectral generalized finite element method, ESAIM: Math. Model. Numer. Anal. 48 (2) (2014) 493–515.
[32] I. Babuška, R. Lipton, P. Sinz, M. Stuebner, Multiscale-spectral GFEM and optimal oversampling, Comput. Methods Appl. Mech. Eng. 364 (2020) 112960.
[33] R. Altmann, P. Henning, D. Peterseim, Numerical homogenization beyond scale separation, Acta Numer. 30 (2021) 1–86.
[34] X. Guan, L. Jiang, Y. Wang, Regularized coupling multiscale method for thermomechanical coupled problems, J. Comput. Phys. 499 (2024) 112737.
[35] V.M. Calo, Y. Efendiev, J. Galvis, G. Li, Randomized oversampling for generalized multiscale finite element methods, Multiscale Model. Simul. 14 (1) (2016) 482–501.
[36] A. Pinkus, N-Widths in Approximation Theory, vol. 7, Springer Science & Business Media, 2012.
[37] C. Ma, R. Scheichl, T. Dodwell, Novel design and analysis of generalized finite element methods based on locally optimal spectral approximations, SIAM J. Numer. Anal. 60 (1) (2022) 244–273.
[38] C. Ma, R. Scheichl, Error estimates for discrete generalized FEMs with locally optimal spectral approximations, Math. Comput. 91 (338) (2022) 2539–2569.
[39] C. Ma, C. Alber, R. Scheichl, Wavenumber explicit convergence of a multiscale generalized finite element method for heterogeneous Helmholtz problems, SIAM J. Numer. Anal. 61 (3) (2023) 1546–1584.
[40] J. Schleuß, K. Smetana, Optimal local approximation spaces for parabolic problems, Multiscale Model. Simul. 20 (1) (2022) 551–582.
[41] C. Ma, J.M. Melenk, Exponential convergence of a generalized fem for heterogeneous reaction-diffusion equations, Multiscale Model. Simul. 22 (1) (2024) 256–282.
[42] P. Bastian, M. Blatt, A. Dedner, N.-A. Dreier, C. Engwer, R. Fritze, C. Gräser, C. Grüninger, D. Kempf, R. Klöfkorn, M. Ohlberger, O. Sander, The Dune framework: basic concepts and recent developments, Comput. Math. Appl. 81 (2021) 75–112.
[43] P. Bastian, R.S. Scheichl, L. Seelinger, A. Strehlow, Multilevel spectral domain decomposition, SIAM J. Sci. Comput. (2022) S1–S26.
[44] J.M. Melenk, On generalized finite element methods, PhD Thesis, University of Maryland, 1995, https://www.asc.tuwien.ac.at/melenk/publications/diss.ps.gz.
[45] C. Ma, A unified framework for multiscale spectral generalized fems and low-rank approximations to multiscale pdes, arXiv preprint, arXiv:2311.08761, 2023.
[46] K. Chen, Q. Li, J. Lu, S.J. Wright, Randomized sampling for basis function construction in generalized finite element methods, Multiscale Model. Simul. 18 (2) (2020) 1153–1177.
[47] P. Bastian, F. Heimann, S. Marnach, Generic implementation of finite element methods in the distributed and unified numerics environment (DUNE), Kybernetika 46 (2) (2010) 294–315.
[48] C. Geuzaine, J.-F. Remacle, Gmsh: a 3-D finite element mesh generator with built-in pre- and post-processing facilities, Int. J. Numer. Methods Eng. 79 (11) (2009) 1309–1331.
[49] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, M. Ohlberger, O. Sander, A generic grid interface for parallel and adaptive scientific computing. Part I: abstract framework, Computing 82 (2008) 103–119.
[50] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, R. Kornhuber, M. Ohlberger, O. Sander, A generic grid interface for parallel and adaptive scientific computing. Part II: implementation and tests in dune, Computing 82 (2008) 121–138.
[51] R.B. Lehoucq, D.C. Sorensen, C. Yang, ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods, SIAM, 1998.
[52] G. Karypis, V. Kumar, Multilevel k-way partitioning scheme for irregular graphs, J. Parallel Distrib. Comput. 48 (1) (1998) 96–129.
[53] O. Falcó, R. Ávila, B. Tijs, C. Lopes, Modelling and simulation methodology for unidirectional composite laminates in a virtual test lab framework, Compos. Struct. 190 (2018) 137–159.
[54] C.G. Koay, On the six-dimensional orthogonal tensor representation of the rotation in three dimensions: a simplified approach, Mech. Mater. 41 (8) (2009) 951–953.
[55] C. Furtado, G. Catalanotti, A. Arteiro, P. Gray, B. Wardle, P. Camanho, Simulation of failure in laminated polymer composites: building-block validation, Compos. Struct. 226 (2019) 111168.
[56] J. Srisuriyachot, J. Bénézech, G. Couégnat, S.A. McNair, T. Maierhofer, R. Butler, A.J. Lunt, Synchrotron micro-ct in kink-band formation of ud-cfrp laminates with microdefects, Composites, Part B, Eng. 266 (2023) 111038.
[57] M.B. Giles, Multilevel Monte Carlo path simulation, Oper. Res. 56 (3) (2008) 607–617.
[58] T.J. Dodwell, C. Ketelsen, R. Scheichl, A.L. Teckentrup, Multilevel Markov chain Monte Carlo, SIAM Rev. 61 (3) (2019) 509–545.