LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

AUXILIARY SPACE PRECONDITIONING OF FINITE ELEMENT EQUATIONS USING A NONCONFORMING INTERIOR PENALTY REFORMULATION AND STATIC CONDENSATION

D. Z. Kalchev, P. S. Vassilevski

September 3, 2019

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

AUXILIARY SPACE PRECONDITIONING OF FINITE ELEMENT EQUATIONS USING A NONCONFORMING INTERIOR PENALTY REFORMULATION AND STATIC CONDENSATION

DELYAN Z. KALCHEV AND PANAYOT S. VASSILEVSKI

ABSTRACT. We modify the well-known interior penalty finite element discretization method so that it allows for element-by-element assembly. This is possible due to the introduction of additional unknowns associated with the interfaces between neighboring elements. The resulting bilinear form, and a Schur complement (reduced) version of it, are utilized in a number of auxiliary space preconditioners for the original conforming finite element discretization problem. These preconditioners are analyzed on the fine scale and their performance is illustrated on model second order scalar elliptic problems discretized with high order elements.

Key words. finite element method, auxiliary space, fictitious space, preconditioning, interior penalty, static condensation, algebraic multigrid, element-by-element assembly, high order

1. INTRODUCTION

The well-known interior penalty (IP) discretization method [4, 13, 22, 17] couples degrees of freedom across two neighboring elements, so it does not possess the element-by-element assembly property, which is inherent to conforming finite element discretization methods. The element-by-element assembly property is useful, for example, in "matrix-free" computations, since it minimizes the coupling across element interfaces. Also, certain element-based coarsening as in the AMGe methods, [21], can be employed then, such that it maintains the element-by-element assembly property on coarse levels. The modification we propose consists of the following simple idea. The IP bilinear form originally contains a jump term $\int_{\mathfrak{f}} [\![u]\!][\![v]\!] \, d\rho$, over each interface $\mathfrak{f}$ between any two neighboring elements τ_- and τ_+. Here, this term is replaced by two other terms, involving a new unknown, $u_{\mathfrak{f}}$, associated with $\mathfrak{f}$. We have then $\int_{\mathfrak{f}} (u_- - u_{\mathfrak{f}})(v_- - v_{\mathfrak{f}}) \, d\rho + \int_{\mathfrak{f}} (u_+ - u_{\mathfrak{f}})(v_+ - v_{\mathfrak{f}}) \, d\rho$, where u_- and v_- come from the element τ_-, while u_+ and v_+ come from the other element, τ_+. Clearly, letting $u_{\mathfrak{f}} = \frac{1}{2}(u_- + u_+)$ (and similarly, $v_{\mathfrak{f}} = \frac{1}{2}(v_- + v_+)$), recovers the original jump term, scaled by $\frac{1}{2}$. It is also clear that introducing the additional space of functions $v_b = (v_{\mathfrak{f}})$, associated with the set of interfaces $\{\mathfrak{f}\}$, allows associating each of the two new terms with a unique neighboring element, i.e., $\int_{\mathfrak{f}} (u_- - u_{\mathfrak{f}})(v_- - v_{\mathfrak{f}}) \, d\rho$ with τ_- and $\int_{\mathfrak{f}} (u_+ - u_{\mathfrak{f}})(v_+ - v_{\mathfrak{f}}) \, d\rho$ with τ_+. Consequently, the coupling occurs only through these interface unknowns. We consider $v_b = (v_{\mathfrak{f}})$ discontinuous from face to face.

The introduction of interface unknowns is a simple and basic idea to decouple neighboring elements. For example, similar idea is utilized for hybridization (see, e.g., [12, 6]) of finite element methods. In the context of hybridization, the interface spaces are used for Lagrangian multipliers associated with constraints on jump

conditions across elements, where those Lagrangian multipliers can also be interpreted as solution traces. Here, the approach is more direct and somewhat simpler – the interface spaces are explicitly built and interpreted as trace spaces, the jump conditions are between the interfaces and the elements (on both sides of each interface) instead of between neighboring elements, and these conditions are introduced as part of the minimization functional rather than as constraints. Moreover, this paper is devoted to using the reformulations for building preconditioners rather than alternative discretization methods.

Clearly, the modification can be employed with τ_- and τ_+ replaced by subdomains T_- and T_+ (for example, being unions of finite elements) and $\mathfrak{f}$ replaced by the interface F between T_- and T_+. This is the approach we exploit, when building preconditioners for an original conforming discretization (i.e., no interior penalty terms to begin with). One may view the set of subdomains, T, as a coarse triangulation $\mathcal{T}^H$. In that case, each $T \in \mathcal{T}^H$ is a union of fine elements from an initial fine triangulation $\mathcal{T}^h$. We refer to T as an *agglomerated element*, or simply an *agglomerate*. The modified IP method employs two sets of discontinuous spaces: one space of functions $u_e = (u_T)$ associated with the agglomerates $T \in \mathcal{T}^H$ and the space of functions $u_b = (u_F)$ associated with the interfaces $\{F\}$ between any two neighboring T_- and T_+ from $\mathcal{T}^H$. The resulting bilinear form consists of the local bilinear forms $a_T(u_T, v_T) + \frac{1}{\delta} \sum_{F \subset \partial T} \int_F (u_T - u_F)(v_T - v_F) \, d\rho$ associated with each T ($\delta > 0$ is the penalty parameter). The trial functions are $u_e = (u_T)$ and $u_b = (u_F)$, whereas, similarly, the test functions are $v_e = (v_T)$ and $v_b = (v_F)$. Here, $a_T(\cdot, \cdot)$ is a local, on T, version of the bilinear form in the original conforming discretization.

The goal of the present paper is to study the modified IP bilinear form as a tool for building preconditioners for the original conforming bilinear form utilizing the auxiliary space technique which goes back to S. Nepomnyaschikh (see [18]) and studied in detail by J. Xu [23]. We analyze both additive and multiplicative versions of the auxiliary space preconditioners for general smoothers, following [21, Theorem 7.18], by verifying the assumptions needed there. Another use of the element-by-element assembly property of the modified IP method is the application of the spectral AMGe technique to construct algebraic multigrid (AMG) preconditioners for the IP bilinear form and for its reduced Schur complement form. We note that the modified IP form allows for static condensation, i.e., we can eliminate the u_T unknowns (they are decoupled from each other) and, thus, end up with a problem for the interface unknowns, u_F, only.

We have implemented these auxiliary space preconditioners and tested their theoretically proven mesh independent performance on model 2D and 3D scalar second order elliptic problems, including high order elements.

The remainder of the paper is structured as follows. Section 2 introduces basic concepts, spaces, notation, and a problem of interest. The IP formulation and the respective auxiliary space preconditioners are presented in Section 3. Their properties are studied, showing (Theorem 3.4) the general optimality of a fine-scale auxiliary space preconditioning strategy via the IP reformulation. Section 4 describes a polynomial smoother employed in the preconditioner. A generic AMGe approach that can be utilized for solving the IP auxiliary space problem is outlined in Section 5. Numerical examples are demonstrated in Section 6, while conclusions and future work are left for the final Section 7.

2. Basic setting

This section is devoted to providing foundations by addressing generic notions and basic procedures. Notation and abbreviated designations are introduced to facilitate a more convenient presentation in the rest of the paper.

2.1. **Mesh and agglomeration.** A domain $\Omega \subset \mathbb{R}^d$ (d is the space dimension) with a Lipschitz-continuous boundary, a (fine-scale) triangulation $\mathcal{T}^h = \{\,\tau\,\}$ of Ω, and a finite element space $\mathcal{U}^h$ on $\mathcal{T}^h$ are given. The mesh $\mathcal{T}^h$ is seen as a set of elements and respective associated faces, where a *face* is the interface, of dimension $d-1$, between two adjacent elements. In general, the degrees of freedom (dofs) of $\mathcal{U}^h$ are split into dofs associated with the interiors of the elements and dofs associated with the faces (cf. Fig. 3), where a face dof belongs to multiple elements and can also belong to multiple faces, whereas an interior dof always belongs to a single element. The focus of this paper is on $\mathcal{U}^h$ consisting of continuous piece-wise polynomial functions, equipped with the usual nodal dofs.

Let $\mathcal{T}^H = \{\,T\,\}$ be a partitioning of $\mathcal{T}^h$ into non-overlapping connected sets of fine-scale elements, called *agglomerate elements* or, simply, *agglomerates*; see Fig. 1. In general, $\mathcal{T}^H$ can be obtained by partitioning the dual graph of $\mathcal{T}^h$, which is a graph with nodes the elements in $\mathcal{T}^h$, where two elements are connected by an edge if they share a face. It can be expressed as a relation element_element. That is, all $T \in \mathcal{T}^H$ are described in terms of the elements $\tau \in \mathcal{T}^h$ via a relation ELEMENT_element. Note that capitalization indicates agglomerate entities in $\mathcal{T}^H$, like ELEMENT (short for "agglomerate element"), FACE, or ENTITY, whereas regular letters indicate fine-scale entities in $\mathcal{T}^h$, like element, face, or entity. This convention, unless otherwise specified, is used in the rest of the paper. Using element_face next, provides

$$\text{ELEMENT\_face} = \text{ELEMENT\_element} \times \text{element\_face},$$

where element_face represents the relation, in $\mathcal{T}^h$, between elements and their adjacent faces. Then, an intersection procedure over the sets described by ELEMENT_-face constructs the agglomerate FACES in $\mathcal{T}^H$ as sets of fine-scale faces, expressed via a relation FACE_face, and related to ELEMENTS in the form of ELEMENT_FACE; cf. Fig. 2. Consequently, each FACE can be consistently recognized as the $(d-1)$-dimensional surface that serves as an interface between two adjacent ELEMENTS in $\mathcal{T}^H$. The set of obtained FACES in $\mathcal{T}^H$ is denoted by $\Phi^H = \{\,F\,\}$. Considering the dofs in $\mathcal{U}^h$, the relations element_dof, face_dof, and

$$\text{ELEMENT\_dof} = \text{ELEMENT\_element} \times \text{element\_dof},$$
$$\text{FACE\_dof} = \text{FACE\_face} \times \text{face\_dof}$$

are determined.

For additional information on relation tables (matrices) and their utilization, see [21, 20].

2.2. **Pairs of nonconforming spaces.** A main idea in this work is to obtain discontinuous (nonconforming) finite element spaces and formulations on $\mathcal{T}^H$ and Φ^H. To that purpose, define the finite element spaces $\mathcal{E}^h$ and $\mathcal{F}^h$ via restrictions (or traces) of functions in $\mathcal{U}^h$ onto $T \in \mathcal{T}^H$ and $F \in \Phi^H$, respectively. Namely,

$$\mathcal{E}^h = \left\{\, v_e^h \in L^\infty(\Omega); \qquad \forall\, T \in \mathcal{T}^H,\ \exists\, v^h \in \mathcal{U}^h\colon\ v_e^h\Big|_T = v^h\Big|_T \,\right\},$$
$$\mathcal{F}^h = \left\{\, v_b^h \in L^\infty(\cup_{F \in \Phi^H} F);\quad \forall\, F \in \Phi^H,\ \exists\, v^h \in \mathcal{U}^h\colon\ v_b^h\Big|_F = v^h\Big|_F \,\right\}.$$

(A) ELEMENTS in 3D   (B) ELEMENTS in 2D
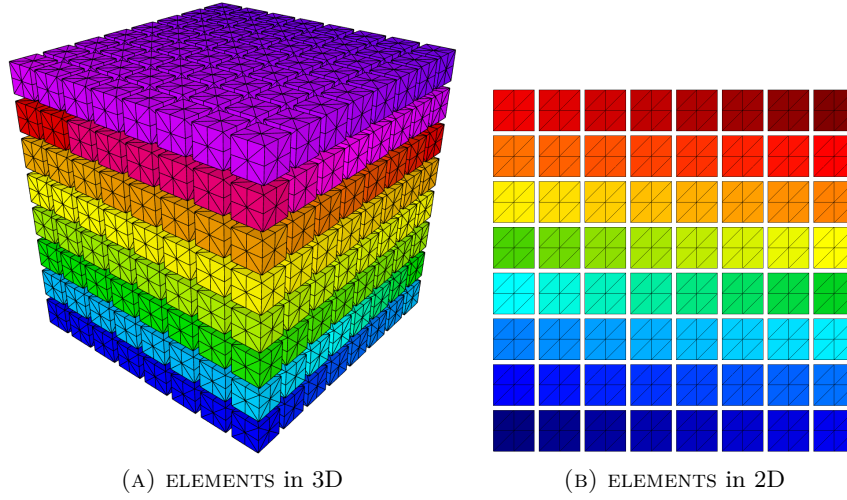
FIGURE 1. Examples of agglomerates (designated as ELEMENTS) of (fine-scale) elements, utilized in the IP reformulation.
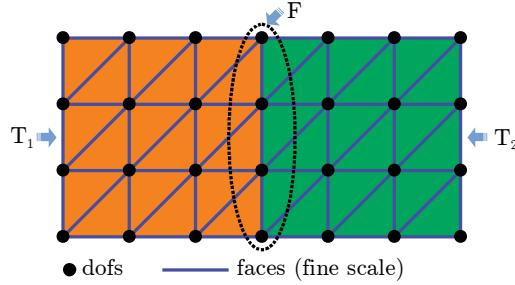

• dofs   —— faces (fine scale)

FIGURE 2. An illustration of the designation of a FACE as a set of (fine-scale) faces, serving as an interface between ELEMENTS. Here, the ELEMENTS T_1 and T_2 are related to the FACE F in the relation table ELEMENT_FACE.

Note that $\mathcal{E}^h$ and $\mathcal{F}^h$ are fine-scale spaces despite the utilization of agglomerate mesh structures like $\mathcal{T}^H$ and Φ^H, which justifies the parameter "h" (versus "H"). Accordingly, the bases in $\mathcal{E}^h$ and $\mathcal{F}^h$ are derived via respective restrictions (or traces) of the basis in $\mathcal{U}^h$. The degrees of freedom in $\mathcal{E}^h$ and $\mathcal{F}^h$ are obtained in a corresponding manner from the dofs in $\mathcal{U}^h$, as illustrated in Fig. 3. For simplicity, "dofs" is reserved for the degrees of freedom in $\mathcal{U}^h$, whereas "edofs" and "bdofs" are reserved for $\mathcal{E}^h$ and $\mathcal{F}^h$, respectively. Moreover, "adofs" designates the edofs and bdofs collectively and is associated with the product space $\mathcal{E}^h \times \mathcal{F}^h$. In more detail, edofs and bdofs are obtained by "cloning" all respective dofs for every agglomerate ENTITY that contains the dofs, in accordance with ELEMENT_dof and FACE_dof. For example, in Fig. 3: the dof in the center is cloned into four edofs (belonging to four separate ELEMENTS) and four bdofs (belonging to four separate FACES), thus, cloning that dof into eight adofs total; the dof in the interior of the FACE is cloned into three adofs – one bdof (belonging to the FACE itself) and two edofs (belonging to two separate ELEMENTS); and the dofs in the interiors of ELEMENTS are simply copied
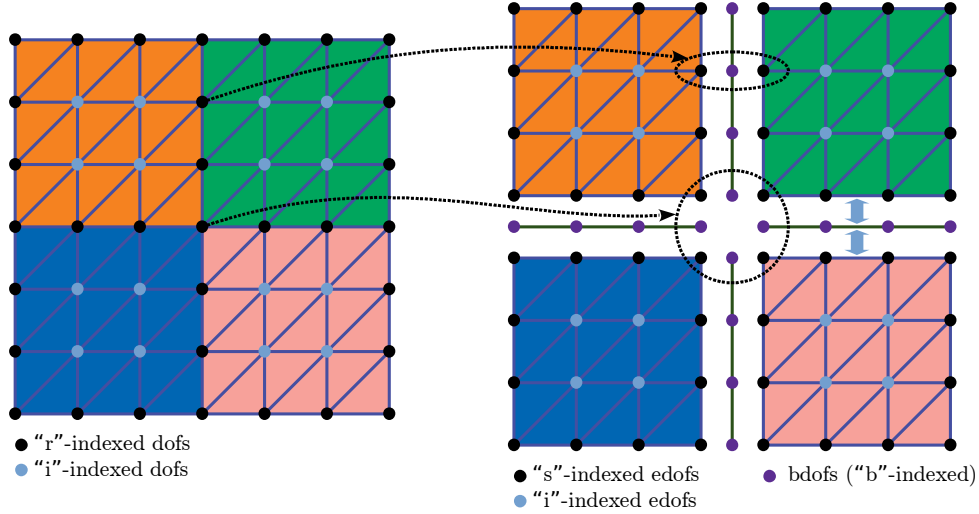
"r"-indexed dofs
"i"-indexed dofs

"s"-indexed edofs   bdofs ("b"-indexed)
"i"-indexed edofs

FIGURE 3. An illustration of the construction of (fine-scale) noncon-forming finite element spaces $\mathcal{E}^h$ and $\mathcal{F}^h$ from a conforming space $\mathcal{U}^h$, utilizing agglomeration that provides ELEMENTS and FACES. Note that the only inter-ENTITY coupling in the IP reformulation is between ELEMENTS and their respective FACES.

as single edofs. Hence, each ENTITY receives and it is the sole owner of a copy of all dofs it contains and there is no intersection between ENTITIES in terms of ed-ofs and bdofs, i.e., they are completely separated without any sharing, making $\mathcal{E}^h$ and $\mathcal{F}^h$ spaces of discontinuous functions. Nevertheless, dofs, edofs, and bdofs are connected via their common "ancestry" founded on the above "cloning" procedure. Thus, the restrictions (or traces) of finite element functions in one of the spaces and their representations as functions in some of the other spaces is seen and performed in purely algebraic context. For example, if $\boldsymbol{q}$ is a vector in terms of the edofs of some $T \in \mathcal{T}^H$, that represents locally, on T, a function in $\mathcal{E}^h$, and $F \subset \partial T$, then $\boldsymbol{q}|_F$ denotes a vector in terms of the bdofs of $F \in \Phi^H$, that represents locally, on F, a function in $\mathcal{F}^h$. This is unambiguous and should lead to no confusion as it only involves a subvector and appropriate index mapping from edofs to bdofs, without any actual transformations, since $\mathcal{E}^h$ and $\mathcal{F}^h$ are trace (restriction) spaces for $\mathcal{U}^h$. This constitutes an "algebraic" procedure, mimicking the behavior of restrictions (or traces) of finite element functions, on T, that vectors like $\boldsymbol{q}$ represent, which are, generally, also supported on F. In what follows, finite element functions are identified with vectors on the degrees of freedom in the respective space.

Coarse subspaces $\mathcal{E}^H \subset \mathcal{E}^h$ and $\mathcal{F}^H \subset \mathcal{F}^h$ can be constructed by, respectively, selecting linearly independent vectors $\{\boldsymbol{q}_{T,i}\}_{i=1}^{m_T}$, for every $T \in \mathcal{T}^H$, and $\{\boldsymbol{q}_{F,i}\}_{i=1}^{m_F}$, for every $F \in \Phi^H$, forming the bases for the coarse spaces. This is an "algebraic" procedure, formulating the coarse basis functions as linear combinations of fine-level basis functions, i.e., as vectors in terms of the fine-level degrees of freedom. The basis vectors are organized appropriately as columns of *prolongation* (or *interpolation*) matrices $\mathcal{P}_e \colon \mathcal{E}^H \mapsto \mathcal{E}^h$ and $\mathcal{P}_b \colon \mathcal{F}^H \mapsto \mathcal{F}^h$, forming $\mathcal{P} \colon \mathcal{E}^H \times \mathcal{F}^H \mapsto \mathcal{E}^h \times \mathcal{F}^h$ as $\mathcal{P} = \mathrm{diag}(\mathcal{P}_e, \mathcal{P}_b)$. For consistency, the corresponding degrees of freedom (associated

with the respective coarse basis vectors) in $\mathcal{E}^H$ and $\mathcal{F}^H$ are respectively called "Edofs" and "Bdofs", and the collective term "Adofs" is associated with $\mathcal{E}^H \times \mathcal{F}^H$. In essence, this is based on the ideas in AMGe (element-based algebraic multigrid) methods [21, 20, 10, 11, 16, 9]. A particular way to obtain coarse basis vectors and construct coarse spaces, $\mathcal{E}^H$ and $\mathcal{F}^H$, is described in Section 3.4.

2.3. **Model problem.** The model problem considered in this paper is the second order scalar elliptic partial differential equation (PDE)

$$(2.1) \qquad\qquad -\operatorname{div}(\kappa \nabla u) = f(\boldsymbol{x}) \text{ in } \Omega,$$

where $\kappa \in L^\infty(\Omega)$, $\kappa > 0$, is a given permeability field, $f \in L^2(\Omega)$ is a given source, and $u \in H^1(\Omega)$ is the unknown function. For simplicity of exposition, the boundary condition $u = 0$ on $\partial\Omega$, the boundary of Ω, is considered, i.e., $u \in H_0^1(\Omega)$. The ubiquitous variational formulation $\min_{v \in H_0^1(\Omega)}[(\kappa \nabla v, \nabla v) - 2(f, v)]$ of (2.1) is utilized, providing the weak form

$$(2.2) \qquad \text{Find } u \in H_0^1(\Omega) \colon\ (\kappa \nabla u, \nabla v) = (f, v), \quad \forall v \in H_0^1(\Omega),$$

where $(\cdot, \cdot)$ denotes the inner products in both $L^2(\Omega)$ and $[L^2(\Omega)]^d$. Consider the fine-scale finite element space $\mathcal{U}^h \subset H_0^1(\Omega)$ defined on $\mathcal{T}^h$. Using the finite element basis in $\mathcal{U}^h$, (2.2) induces the following linear system of algebraic equations:

$$(2.3) \qquad\qquad\qquad A\boldsymbol{u} = \boldsymbol{f},$$

for the global symmetric positive definite (SPD) stiffness matrix A. Moreover, the local, on agglomerates, symmetric positive semidefinite (SPSD) stiffness matrices A_T, for $T \in \mathcal{T}^H$, are obtainable, such that $A = \sum_{T \in \mathcal{T}^H} A_T$ (the summation involves an implicit local-to-global mapping).

3. Interior penalty approach

In this section, an approach for obtaining preconditioners, based on the ideas of an *interior penalty* (IP) method (see [17]) is described and studied. A main idea is that, instead of using penalty terms on the jumps directly between ELEMENTS, the interface space $\mathcal{F}^h$ is employed to avoid direct coupling between ELEMENTS.

3.1. **Formulation.** Consider the nonconforming discrete quadratic minimization formulation of (2.1)

$$\min \sum_{T \in \mathcal{T}^H} \left[\left(\kappa \nabla v_e^h,\ \nabla v_e^h\right)_T + \frac{1}{\delta} \sum_{F \subset \partial T} \left\langle v_e^h\Big|_F - v_b^h,\ v_e^h\Big|_F - v_b^h \right\rangle_F \right] - 2(f, v_e^h),$$

for $[v_e^h, v_b^h] \in \mathcal{E}^h \times \mathcal{F}^h$, providing, in lieu of (2.2), the weak form: Find $[u_e^h, u_b^h] \in \mathcal{E}^h \times \mathcal{F}^h$, such that

$$(3.1) \qquad \sum_{T \in \mathcal{T}^H} \left[\left(\kappa \nabla u_e^h,\ \nabla v_e^h\right)_T + \frac{1}{\delta} \sum_{F \subset \partial T} \left\langle u_e^h\Big|_F - u_b^h,\ v_e^h\Big|_F - v_b^h \right\rangle_F \right] = (f, v_e^h),$$

for all $[v_e^h, v_b^h] \in \mathcal{E}^h \times \mathcal{F}^h$. Here, $(\cdot, \cdot)_T$ and $\langle \cdot, \cdot \rangle_F$ denote the inner products in $[L^2(T)]^d$ and $L^2(F)$, respectively, and $\delta \in (0, 1)$ is a small penalty parameter that can, generally, depend on h. The notation in (3.1) requires an explanation. Namely, (3.1) involves an implicit restriction of u_e^h and v_e^h, which are generally discontinuous across ∂T, onto T. This removes the ambiguity from $u_e^h|_F$ and $v_e^h|_F$ by considering "one-sided" traces, involving the implicit restrictions to T prior to further restricting to F. Therefore, the only inter-ENTITY coupling, introduced by the bilinear form in

(3.1), is between the edofs in $T \in \mathcal{T}^H$ and the bdofs on ∂T. Thus, the edofs are not directly coupled across the ELEMENTS and an assembly procedure for the matrix, corresponding to the bilinear form in (3.1), would only need to add (accumulate) contributions associated with bdofs.

Equation (3.1) is slightly modified to obtain a formulation that is easier to implement algebraically and to facilitate the results in Section 3.3 below. Particularly, only the interface penalty term is changed – instead of the more standard $L^2(F)$ inner product in (3.1), $\langle \cdot, \cdot \rangle_F$, the inner product induced by the (restricted) diagonal of A is utilized for the interface term. Consider $\mathcal{A}_T$ – the local, on $T \in \mathcal{T}^H$, and corresponding to a summand in (3.1), matrix associated with the quadratic form

$$(3.2) \qquad \begin{bmatrix} \boldsymbol{v}_e \\ \boldsymbol{v}_b \end{bmatrix}^T \mathcal{A}_T \begin{bmatrix} \boldsymbol{v}_e \\ \boldsymbol{v}_b \end{bmatrix} = \boldsymbol{v}_e^T A_T \boldsymbol{v}_e + \frac{1}{\delta} \sum_{F \subset \partial T} \left(\boldsymbol{v}_e|_F - \boldsymbol{v}_b|_F \right)^T D_F \left(\boldsymbol{v}_e|_F - \boldsymbol{v}_b|_F \right),$$

where $\boldsymbol{v}_e$ is defined on the edofs of T, $\boldsymbol{v}_b$ – on the bdofs on ∂T, D_F is the restriction of the diagonal, D, of the global A onto the bdofs of F, and $\mathcal{A}_T$ takes the block form

$$(3.3) \qquad \mathcal{A}_T = \begin{bmatrix} \mathcal{A}_{T,ee} & \mathcal{A}_{T,eb} \\ \mathcal{A}_{T,be} & \mathcal{A}_{T,bb} \end{bmatrix}.$$

Here, the discussion in Section 2.2, concerning index mapping as part of the restriction, applies. Respective consistent, with the construction of $\mathcal{A}_T$, relations ELEMENT_edof, ELEMENT_bdof, and, consequently, ELEMENT_adof are also available. The global (SPD) IP matrix is obtainable via standard accumulation (assembly) – $\mathcal{A} = \sum_{T \in \mathcal{T}^H} \mathcal{A}_T$ – which, as mentioned above, would involve only accumulation (addition) of the $\mathcal{A}_{T,bb}$ blocks, while the remaining portions of the local matrices are simply copied into the global matrix.

Both A_T and $\mathcal{A}_T$ are SPSD with null spaces spanned by respective constant vectors (excluding essential boundary conditions). However, both $\mathcal{A}_{T,bb}$ and $\mathcal{A}_{T,ee}$ are positive definite due to the interface terms. In general, using that A has strictly positive diagonal entries, $\mathcal{A}_{T,ee}$ is positive definite if and only if every nonzero vector in the null space of A_T has a nonzero value on some FACE $F \subset \partial T$. This is the case here, since the null space of A_T is spanned by the constant vector.

3.2. **Auxiliary space preconditioners.** Here, preconditioners based on the IP formulation are derived. Consider the finite element space $\mathcal{E}^h \times \mathcal{F}^h$, which in the context here is regarded as an "auxiliary" space. Let $\Pi_h \colon \mathcal{E}^h \times \mathcal{F}^h \mapsto \mathcal{U}^h$ be a linear transfer operator, to be defined momentarily. Identifying finite element functions with vectors allows to consider Π_h as a matrix and obtain $\Pi_h^T \colon \mathcal{U}^h \mapsto \mathcal{E}^h \times \mathcal{F}^h$. In order to define the action of Π_h, recall that the relation between dofs, on one side, and adofs, on the other, is known and unambiguous. Therefore, it is reasonable to define the action of Π_h as taking the arithmetic average, formulated in terms of adofs that correspond to a particular dof, of the entries of a given (auxiliary space) vector and obtaining the respective entries of a mapped vector defined on dofs. That is, for any dof l, let J_l be the set of corresponding adofs (the respective "cloned" degrees of freedom) and consider a vector $\hat{\boldsymbol{v}}$, defined in terms of adofs. Then,

$$(\Pi_h \hat{\boldsymbol{v}})_l = \frac{1}{|J_l|} \sum_{j \in J_l} (\hat{\boldsymbol{v}})_j.$$

Clearly, all row sums of Π_h equal 1 and each column of Π_h has exactly one nonzero entry. Assuming that $\mathcal{T}^h$ is a regular (non-degenerate) mesh [7], $|J_l|$ is bounded,

independently of h. That is,

$$(3.4) \qquad\qquad 1 \leq |J_l| \leq \varkappa,$$

for a constant $1 \leq \varkappa < \infty$, which depends only on the regularity of $\mathcal{T}^h$, but not on h. Indeed, let $\varkappa$ be the global maximum number of ELEMENTS and FACES, in $\mathcal{T}^H$, that a dof can belong to, which, in turn, is bounded by the global maximum number of elements and faces, in $\mathcal{T}^h$, that a dof can belong to.

Let M be a "smoother" for A, such that $M^T + M - A$ is SPD, and $\mathcal{B}$ – a symmetric preconditioner for $\mathcal{A}$. Define the *additive auxiliary space preconditioner* for A

$$(3.5) \qquad\qquad B_{\mathrm{add}}^{-1} = \overline{M}^{-1} + \Pi_h \mathcal{B}^{-1} \Pi_h^T,$$

and the *multiplicative auxiliary space preconditioner* for A

$$(3.6) \qquad B_{\mathrm{mult}}^{-1} = \overline{M}^{-1} + (I - M^{-T}A)\Pi_h \mathcal{B}^{-1} \Pi_h^T (I - AM^{-1}),$$

where $\overline{M} = M(M + M^T - A)^{-1}M^T$ is the symmetrized (in fact, SPD) version of M. In case M is symmetric, $\overline{M}^{-1}$ in B_{add}^{-1} can be replaced by M^{-1}. The action of B_{mult}^{-1} is obtained via a standard "two-level" procedure [21]:

Given $\boldsymbol{v}_0 \in \mathbb{R}^{\dim(\mathcal{U}^h)}$, $\boldsymbol{v}_{\mathrm{m}} = B_{\mathrm{mult}}^{-1}\boldsymbol{v}_0$ is computed by the following steps:

(i) "pre-smoothing": $\boldsymbol{v}_1 = M^{-1}\boldsymbol{v}_0$;
(ii) residual transfer to the auxiliary space: $\hat{\boldsymbol{r}} = \Pi_h^T(\boldsymbol{v}_0 - A\boldsymbol{v}_1)$;
(iii) auxiliary space correction: $\hat{\boldsymbol{v}} = \mathcal{B}^{-1}\hat{\boldsymbol{r}}$;
(iv) correction transfer from the auxiliary space: $\boldsymbol{v}_2 = \boldsymbol{v}_1 + \Pi_h\hat{\boldsymbol{v}}$;
(v) "post-smoothing": $\boldsymbol{v}_{\mathrm{m}} = \boldsymbol{v}_2 + M^{-T}(\boldsymbol{v}_0 - A\boldsymbol{v}_2)$.

3.3. **Analysis.** Properties of the preconditioners, showing their optimality, are studied next. Define the operator $\mathcal{I}_h\colon \mathcal{U}^h \mapsto \mathcal{E}^h \times \mathcal{F}^h$, for[1] $\boldsymbol{v} \in \mathcal{U}^h$, via

$$(\mathcal{I}_h\boldsymbol{v})_j = (\boldsymbol{v})_l,$$

for each adof j, where $j \in J_l$, for the corresponding dof l. This describes a procedure that appropriately copies the entries of $\boldsymbol{v}$, so that the respective finite element functions, corresponding to $\boldsymbol{v}$ and $\mathcal{I}_h\boldsymbol{v}$, can be viewed as coinciding in $H^1(\Omega)$. That is, in a sense, $\mathcal{I}_h$ is an injection (embedding) of $\mathcal{U}^h$ into $\mathcal{E}^h \times \mathcal{F}^h$. Considering the respective matrices, $\mathcal{I}_h$ has the fill-in pattern of Π_h^T, but all nonzero entries are replaced by 1.

It is easy to see that $\Pi_h\mathcal{I}_h = I$, the identity operator on $\mathcal{U}^h$, implying that Π_h is surjective, i.e., it has a full row rank. Moreover, for any $\boldsymbol{v} \in \mathcal{U}^h$, $\mathcal{I}_h\boldsymbol{v} \in \mathcal{E}^h \times \mathcal{F}^h$ (exactly) approximates $\boldsymbol{v}$ in the sense

$$(3.7) \qquad\qquad \boldsymbol{v} - \Pi_h\mathcal{I}_h\boldsymbol{v} = \boldsymbol{0},$$

and $\mathcal{I}_h\boldsymbol{v}$ is "energy" stable, since

$$(3.8) \qquad\qquad (\mathcal{I}_h\boldsymbol{v})^T\mathcal{A}\mathcal{I}_h\boldsymbol{v} = \boldsymbol{v}^TA\boldsymbol{v}.$$

This is to be expected, since, in a sense, $\mathcal{E}^h \times \mathcal{F}^h$ "includes" $\mathcal{U}^h$.

The more challenging task is to show the continuity of $\Pi_h\colon \mathcal{E}^h \times \mathcal{F}^h \mapsto \mathcal{U}^h$ in terms of the respective "energy" norms. This is addressed next.

Lemma 3.1. *The operator* $E_h\colon \mathcal{E}^h \times \mathcal{F}^h \mapsto \mathcal{E}^h \times \mathcal{F}^h$, *defined as* $E_h = \mathcal{I}_h\Pi_h - I$, *where* I *is the identity on* $\mathcal{E}^h \times \mathcal{F}^h$, *is bounded in the sense*

$$(\mathcal{I}_h\Pi_h\hat{\boldsymbol{v}} - \hat{\boldsymbol{v}})^T\mathcal{A}(\mathcal{I}_h\Pi_h\hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}) \leq (1 + \Lambda\delta\varkappa^2)\,\hat{\boldsymbol{v}}^T\mathcal{A}\hat{\boldsymbol{v}},$$

---

[1]Finite element functions and algebraic vectors are identified, which should cause no confusion.

for all $\hat{\boldsymbol{v}} \in \mathcal{E}^h \times \mathcal{F}^h$, where δ is the one in (3.2), $\varkappa$ is from (3.4), and $\Lambda > 0$ is a constant, independent of h, H, the coefficient κ in (2.1), and the regularity of $\mathcal{T}^h$ (see Remark 3.3).

Proof. The portions of auxiliary space vectors corresponding to edofs and bdofs are respectively indexed by "e" and "b", leading to the notation, for $\hat{\boldsymbol{v}} \in \mathcal{E}^h \times \mathcal{F}^h$, $\hat{\boldsymbol{v}}^T = [\hat{\boldsymbol{v}}_e^T, \hat{\boldsymbol{v}}_b^T]^T$, where $\hat{\boldsymbol{v}}_e \in \mathcal{E}^h$ and $\hat{\boldsymbol{v}}_b \in \mathcal{F}^h$. By further splitting $\hat{\boldsymbol{v}}_e^T = [\hat{\boldsymbol{v}}_i^T, \hat{\boldsymbol{v}}_s^T]^T$, it is obtained $\hat{\boldsymbol{v}}^T = [\hat{\boldsymbol{v}}_e^T, \hat{\boldsymbol{v}}_b^T]^T = [\hat{\boldsymbol{v}}_i^T, \hat{\boldsymbol{v}}_s^T, \hat{\boldsymbol{v}}_b^T]^T$, where "i" denotes the edofs in the interiors of all $T \in \mathcal{T}^H$ and "s" are the edofs that can be mapped to some bdofs, based on the previously-described procedure of "cloning" dofs into edofs and bdofs. Analogously, the splitting $\boldsymbol{v}^T = [\boldsymbol{v}_i^T, \boldsymbol{v}_r^T]^T$, for $\boldsymbol{v} \in \mathcal{U}^h$, is introduced in terms of dofs, indexed "i", in the interiors of all $T \in \mathcal{T}^H$ and dofs, indexed "r", related to bdofs. Figure 3 illustrates the utilized indexing. Note that there is a clear difference between "r" and "s", but also a clear relation. Locally, on $T \in \mathcal{T}^H$, "r" and "s" can, in fact, be equated, but in a global setting, it is necessary to distinguish between "r" and "s" indices. This should not cause any ambiguity below. Based on the splittings, define

(3.9)
$$A = \begin{bmatrix} A_{ii} & A_{ir} \\ A_{ri} & A_{rr} \end{bmatrix},$$

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_{ee} & \mathcal{A}_{eb} \\ \mathcal{A}_{be} & \mathcal{A}_{bb} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{ii} & \mathcal{A}_{is} & \\ \mathcal{A}_{si} & \mathcal{A}_{ss} & \mathcal{A}_{sb} \\ & \mathcal{A}_{bs} & \mathcal{A}_{bb} \end{bmatrix},$$

$$\Pi_h = \begin{bmatrix} \Pi_{ii} & & \\ & \Pi_{rs} & \Pi_{rb} \end{bmatrix},$$

$$\mathcal{I}_h = \begin{bmatrix} I & \\ & \mathcal{I}_{sr} \\ & \mathcal{I}_{br} \end{bmatrix},$$

where $\mathcal{I}_{sr}$ is the map from "r" to "s" indices and $\mathcal{I}_{br}$ – from "r" to "b" indices. Note that $\mathcal{I}_{sr}$ is a matrix with the fill-in structure of Π_{rs}^T, where all nonzero entries are replaced by 1, each row has exactly one nonzero entry and each column has at least two nonzero entries (exactly two when the respective dofs are in the interior of a FACE). Similarly, $\mathcal{I}_{br}$ is a matrix with the fill-in of Π_{rb}^T, where all nonzero entries are replaced by 1, each row has exactly one nonzero entry and each column has at least one nonzero entry (exactly one when the respective dofs are in the interior of a FACE). Observe that, assuming everything is consistently numbered and ordered, $\mathcal{A}_{ii} = A_{ii}$ and $\Pi_{ii} = I$. It holds that $A_{ir} = \mathcal{A}_{is}\mathcal{I}_{sr}$. This is due to the fact that A_{ir} and $\mathcal{A}_{is}$ represent the same "one-sided connections" (meaning that the "connections" are on one side of the FACES $F \in \Phi^H$), but in terms of different indices.

Let $\boldsymbol{v} = \Pi_h \hat{\boldsymbol{v}} = \left[\hat{\boldsymbol{v}}_i^T, (\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b)^T \right]^T$. Then,

$$\mathcal{I}_h \boldsymbol{v} = \mathcal{I}_h \Pi_h \hat{\boldsymbol{v}} = \left[\hat{\boldsymbol{v}}_i^T, [\mathcal{I}_{sr}(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b)]^T, [\mathcal{I}_{br}(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b)]^T \right]^T,$$

$$\tilde{\boldsymbol{v}} = \mathcal{I}_h \boldsymbol{v} - \hat{\boldsymbol{v}} = \left[\mathbf{0}^T, [\mathcal{I}_{sr}(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b) - \hat{\boldsymbol{v}}_s]^T, [\mathcal{I}_{br}(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b) - \hat{\boldsymbol{v}}_b]^T \right]^T,$$

and

$$\tilde{\boldsymbol{v}}^T \mathcal{A} \tilde{\boldsymbol{v}} = \frac{1}{\delta} \sum_{T \in \mathcal{T}^H} \sum_{F \subset \partial T} (\hat{\boldsymbol{v}}_s|_F - \hat{\boldsymbol{v}}_b|_F)^T D_F (\hat{\boldsymbol{v}}_s|_F - \hat{\boldsymbol{v}}_b|_F)$$

$$+ \sum_{T \in \mathcal{T}^H} [(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b)|_T - \hat{\boldsymbol{v}}_s|_T]^T A_{T,rr}[(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b)|_T - \hat{\boldsymbol{v}}_s|_T],$$

where $A_{T,rr}$ is the local version of A_{rr} and it is utilized that locally, as noted above, "r" and "s" can be identified. Using that the stiffness matrices can be bounded from above by their diagonals, with some constant $\Lambda > 0$, it follows:

(3.10)
$$\tilde{\boldsymbol{v}}^T \mathcal{A} \tilde{\boldsymbol{v}} \leq \frac{1}{\delta} \sum_{T \in \mathcal{T}^H} \sum_{F \subset \partial T} (\hat{\boldsymbol{v}}_s|_F - \hat{\boldsymbol{v}}_b|_F)^T D_F (\hat{\boldsymbol{v}}_s|_F - \hat{\boldsymbol{v}}_b|_F)$$
$$+ \Lambda \sum_{T \in \mathcal{T}^H} \sum_{F \subset \partial T} [(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b)|_F - \hat{\boldsymbol{v}}_s|_F]^T D_F[(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b)|_F - \hat{\boldsymbol{v}}_s|_F].$$

Let, for a dof l from the "r" dofs, J_l^s denote the set of related "s" edofs and J_l^b – the set of related bdofs; illustrated in Fig. 4. That is, $J_l = J_l^s \cup J_l^b$ with J_l^s corresponding to the l-th column of $\mathcal{I}_{sr}$ and J_l^b – to the l-th column of $\mathcal{I}_{br}$. Furthermore, let d_l be the respective diagonal entry in A, N_j^b denote, for all $j \in J_l^s$, the "b" neighbors (see Fig. 5) of j, according to the connectivity of $\mathcal{A}_{sb}$, and m_l, M_l are respectively the minimum and maximum values of $\hat{\boldsymbol{v}}$ on the adofs in J_l. Clearly,

$$d_l \sum_{j \in J_l^s} \left[\frac{1}{|J_l|} \left(\sum_{k \in J_l^s} (\hat{\boldsymbol{v}}_s)_k + \sum_{p \in J_l^b} (\hat{\boldsymbol{v}}_b)_p \right) - (\hat{\boldsymbol{v}}_s)_j \right]^2 \leq |J_l^s| d_l (M_l - m_l)^2.$$

Notice that the adofs corresponding to M_l and m_l are connected via a path, with respect to the connectivity of $\mathcal{A}_{sb}$, whose length is bounded by $|J_l|$. Following along this path, applying the triangle inequality and (3.4), it holds

$$|J_l^s| d_l (M_l - m_l)^2 \leq \varkappa^2 d_l \sum_{j \in J_l^s} \sum_{k \in N_j^b} [(\hat{\boldsymbol{v}}_s)_j - (\hat{\boldsymbol{v}}_b)_k]^2.$$

Hence,

$$d_l \sum_{j \in J_l^s} \left[\frac{1}{|J_l|} \left(\sum_{k \in J_l^s} (\hat{\boldsymbol{v}}_s)_k + \sum_{p \in J_l^b} (\hat{\boldsymbol{v}}_b)_p \right) - (\hat{\boldsymbol{v}}_s)_j \right]^2 \leq \varkappa^2 d_l \sum_{j \in J_l^s} \sum_{k \in N_j^b} [(\hat{\boldsymbol{v}}_s)_j - (\hat{\boldsymbol{v}}_b)_k]^2.$$

Summing over l in the last inequality provides

$$\sum_{T \in \mathcal{T}^H} \sum_{F \subset \partial T} [(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b)|_F - \hat{\boldsymbol{v}}_s|_F]^T D_F[(\Pi_{rs}\hat{\boldsymbol{v}}_s + \Pi_{rb}\hat{\boldsymbol{v}}_b)|_F - \hat{\boldsymbol{v}}_s|_F]$$
$$\leq \varkappa^2 \sum_{T \in \mathcal{T}^H} \sum_{F \subset \partial T} (\hat{\boldsymbol{v}}_s|_F - \hat{\boldsymbol{v}}_b|_F)^T D_F (\hat{\boldsymbol{v}}_s|_F - \hat{\boldsymbol{v}}_b|_F).$$

Combining the last estimate and (3.10) implies

$$\tilde{\boldsymbol{v}}^T \mathcal{A} \tilde{\boldsymbol{v}} \leq (1 + \Lambda \delta \varkappa^2) \frac{1}{\delta} \sum_{T \in \mathcal{T}^H} \sum_{F \subset \partial T} (\hat{\boldsymbol{v}}_s|_F - \hat{\boldsymbol{v}}_b|_F)^T D_F (\hat{\boldsymbol{v}}_s|_F - \hat{\boldsymbol{v}}_b|_F)$$
$$\leq (1 + \Lambda \delta \varkappa^2) \hat{\boldsymbol{v}}^T \mathcal{A} \hat{\boldsymbol{v}}. \qquad \square$$

Corollary 3.2. *The operator* $\Pi_h \colon \mathcal{E}^h \times \mathcal{F}^h \mapsto \mathcal{U}^h$ *is continuous in the sense*

$$(\Pi_h \hat{\boldsymbol{v}})^T A \, \Pi_h \hat{\boldsymbol{v}} \leq 2(2 + \Lambda \delta \varkappa^2) \, \hat{\boldsymbol{v}}^T \mathcal{A} \hat{\boldsymbol{v}},$$

for all $\hat{\boldsymbol{v}} \in \mathcal{E}^h \times \mathcal{F}^h$, *where the constants are the same as in Lemma 3.1.*
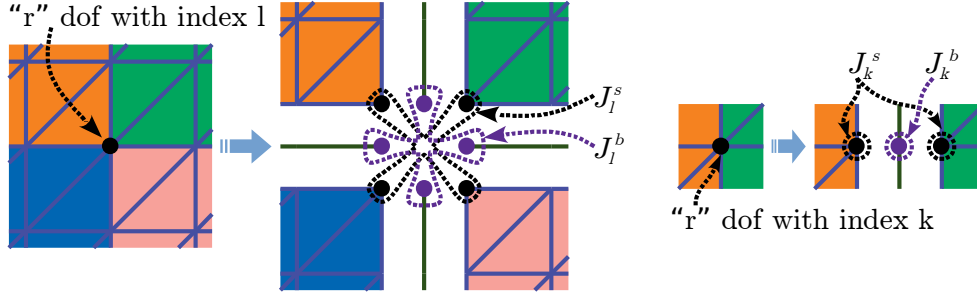
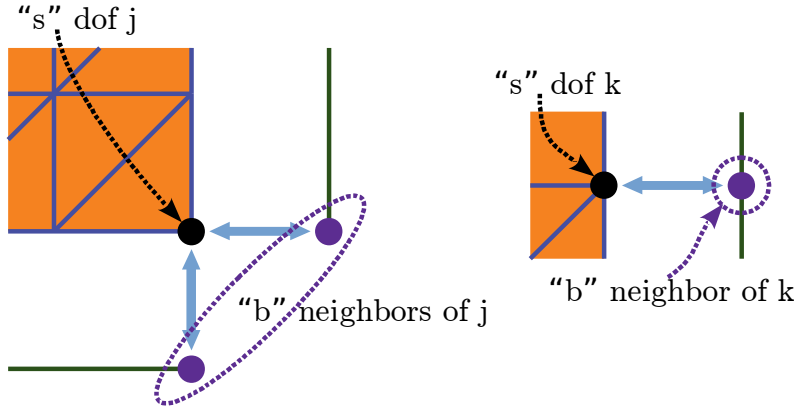FIGURE 4. An illustration of the sets of adofs, associated with an "r" dof, used in the proof of Lemma 3.1.

FIGURE 5. An illustration of the "b" neighbors, according to the connectivity of $\mathcal{A}_{sb}$ in (3.9), of an "s" dof. These are designated as N_j^b and N_k^b and used in the proof of Lemma 3.1.

Proof. Owing to (3.8), the reverse triangle inequality, and Lemma 3.1, it follows:

$$
\begin{aligned}
[(\Pi_h \hat{\boldsymbol{v}})^T A \, \Pi_h \hat{\boldsymbol{v}}]^{\frac{1}{2}} - [\hat{\boldsymbol{v}}^T \mathcal{A} \hat{\boldsymbol{v}}]^{\frac{1}{2}} &= [(\mathcal{I}_h \Pi_h \hat{\boldsymbol{v}})^T \mathcal{A} \, \mathcal{I}_h \Pi_h \hat{\boldsymbol{v}}]^{\frac{1}{2}} - [\hat{\boldsymbol{v}}^T \mathcal{A} \hat{\boldsymbol{v}}]^{\frac{1}{2}} \\
&\leq [(\mathcal{I}_h \Pi_h \hat{\boldsymbol{v}} - \hat{\boldsymbol{v}})^T \mathcal{A} (\mathcal{I}_h \Pi_h \hat{\boldsymbol{v}} - \hat{\boldsymbol{v}})]^{\frac{1}{2}} \\
&\leq \sqrt{1 + \Lambda \delta \varkappa^2} [\hat{\boldsymbol{v}}^T \mathcal{A} \hat{\boldsymbol{v}}]^{\frac{1}{2}}. \qquad \square
\end{aligned}
$$

Remark 3.3. The independence of the constants, in the bounds in Lemma 3.1 and Corollary 3.2, on the coefficient κ in (2.1) holds, since the coefficient information is contained in D_F in (3.2). In fact, the use of the diagonal entries of the global A is crucial. However, the particular argument does not exclude dependence of the constants on the order of the finite element spaces. This is due to Λ in (3.10) depending on the maximum number of dofs in an element. This can be mitigated in the analysis, if necessary, via "strengthening" the interface terms in (3.2) by using a properly scaled version of the diagonal of A (e.g., adjusting the value of δ), or employing a so called weighted ℓ_1-smoother (see [5]) in lieu of the diagonal of A for the interface term in (3.2), which would lead to $\Lambda = 1$. Note that, interestingly, the

numerical results in Section 6.3 (where the unweighted diagonal of A is used) do not demonstrate a practical necessity for adjusting the interface terms, when increasing the polynomial order. This indicates that a more intricate proof, that does not depend on the constant Λ (i.e., on the polynomial order), may be possible. Currently, this is an open question.

Finally, based on the above properties, the optimality of the auxiliary space preconditioners can be established.

Theorem 3.4 (spectral equivalence). *Let the smoother M satisfy the property that $M + M^T - A$ is SPD. Assume that $\mathcal{B}$ in (3.5) and (3.6) is a spectrally equivalent preconditioner for the IP matrix $\mathcal{A}$, in the sense that there exist positive constants α and β, such that*

$$(3.11) \qquad \alpha^{-1}\,\hat{\boldsymbol{v}}^T\mathcal{A}\hat{\boldsymbol{v}} \le \hat{\boldsymbol{v}}^T\mathcal{B}\hat{\boldsymbol{v}} \le \beta\,\hat{\boldsymbol{v}}^T\mathcal{A}\hat{\boldsymbol{v}}, \quad \forall \hat{\boldsymbol{v}} \in \mathcal{E}^h \times \mathcal{F}^h.$$

Then, the additive and multiplicative auxiliary space preconditioners, B_{add} in (3.5) and B_{mult} in (3.6), are spectrally equivalent to A in (2.3).

Proof. Owing to (3.11), (3.8) and Corollary 3.2 provide, respectively,

$$(3.12) \qquad (\mathcal{I}_h\boldsymbol{v})^T\mathcal{B}\,\mathcal{I}_h\boldsymbol{v} \le \beta\,\boldsymbol{v}^T A\boldsymbol{v}, \qquad\qquad \forall \boldsymbol{v} \in \mathcal{U}^h,$$

$$(3.13) \qquad (\Pi_h\hat{\boldsymbol{v}})^T A\,\Pi_h\hat{\boldsymbol{v}} \le 2\alpha(2 + \Lambda\delta\varkappa^2)\,\hat{\boldsymbol{v}}^T\mathcal{B}\hat{\boldsymbol{v}}, \quad \forall \hat{\boldsymbol{v}} \in \mathcal{E}^h \times \mathcal{F}^h.$$

First, consider the "fictitious space preconditioner" $\widetilde{B}^{-1} = \Pi_h\mathcal{B}^{-1}\Pi_h^T$ for A. It is clearly SPD, when $\mathcal{B}$ is SPD (implied by (3.11)), due to the full row rank of Π_h. By (3.7), the Cauchy–Schwarz inequality, and (3.12), it follows:

$$\boldsymbol{v}^T A\boldsymbol{v} = (\mathcal{I}_h\boldsymbol{v})^T\mathcal{B}\mathcal{B}^{-1}\Pi_h^T A\boldsymbol{v} \le [(\mathcal{I}_h\boldsymbol{v})^T\mathcal{B}\,\mathcal{I}_h\boldsymbol{v}]^{1/2}[\boldsymbol{v}^T A\Pi_h\mathcal{B}^{-1}\mathcal{B}\mathcal{B}^{-1}\Pi_h^T A\boldsymbol{v}]^{1/2}$$
$$\le \sqrt{\beta}[\boldsymbol{v}^T A\boldsymbol{v}]^{1/2}[\boldsymbol{v}^T A\widetilde{B}^{-1}A\boldsymbol{v}]^{1/2}.$$

Thus, $\boldsymbol{v}^T A\boldsymbol{v} \le \beta\,\boldsymbol{v}^T A\widetilde{B}^{-1}A\boldsymbol{v}$, showing

$$(3.14) \qquad \boldsymbol{v}^T A^{-1}\boldsymbol{v} \le \beta\,\boldsymbol{v}^T\widetilde{B}^{-1}\boldsymbol{v}, \quad \forall \boldsymbol{v} \in \mathcal{U}^h,$$

which is equivalent to $\boldsymbol{v}^T\widetilde{B}\boldsymbol{v} \le \beta\,\boldsymbol{v}^T A\boldsymbol{v}$. Conversely, owing to the Cauchy–Schwarz inequality and (3.13), (denoting $\gamma = 2\alpha(2 + \Lambda\delta\varkappa^2)$) it follows:

$$\boldsymbol{v}^T A\widetilde{B}^{-1}A\boldsymbol{v} \le [\boldsymbol{v}^T A\boldsymbol{v}]^{1/2}[\boldsymbol{v}^T A\widetilde{B}^{-1}A\widetilde{B}^{-1}A\boldsymbol{v}]^{1/2}$$
$$= [\boldsymbol{v}^T A\boldsymbol{v}]^{1/2}[(\Pi_h\mathcal{B}^{-1}\Pi_h^T A\boldsymbol{v})^T A\Pi_h\mathcal{B}^{-1}\Pi_h^T A\boldsymbol{v}]^{1/2}$$
$$\le \sqrt{\gamma}[\boldsymbol{v}^T A\boldsymbol{v}]^{1/2}[\boldsymbol{v}^T A\Pi_h\mathcal{B}^{-1}\mathcal{B}\mathcal{B}^{-1}\Pi_h^T A\boldsymbol{v}]^{1/2} = \sqrt{\gamma}[\boldsymbol{v}^T A\boldsymbol{v}]^{1/2}[\boldsymbol{v}^T A\widetilde{B}^{-1}A\boldsymbol{v}]^{1/2}.$$

Whence, $\boldsymbol{v}^T A\widetilde{B}^{-1}A\boldsymbol{v} \le 2\alpha(2 + \Lambda\delta\varkappa^2)\,\boldsymbol{v}^T A\boldsymbol{v}$ and

$$(3.15) \qquad \boldsymbol{v}^T\widetilde{B}^{-1}\boldsymbol{v} \le 2\alpha(2 + \Lambda\delta\varkappa^2)\,\boldsymbol{v}^T A^{-1}\boldsymbol{v}, \quad \forall \boldsymbol{v} \in \mathcal{U}^h,$$

or, equivalently, $\boldsymbol{v}^T A\boldsymbol{v} \le 2\alpha(2 + \Lambda\delta\varkappa^2)\,\boldsymbol{v}^T\widetilde{B}\boldsymbol{v}$.

Next, consider B_{add}^{-1} in (3.5). In view of the positive definiteness of the smoother in (3.5), (3.14) implies

$$\boldsymbol{v}^T A^{-1}\boldsymbol{v} \le \beta\,\boldsymbol{v}^T B_{\mathrm{add}}^{-1}\boldsymbol{v}, \quad \forall \boldsymbol{v} \in \mathcal{U}^h,$$

equivalently, $\boldsymbol{v}^T B_{\mathrm{add}}\boldsymbol{v} \le \beta\,\boldsymbol{v}^T A\boldsymbol{v}$. Conversely, using that $M + M^T - A$ is SPD and the fact that this is equivalent to $2\overline{M} - A$ being SPD [21], it holds that

$$(3.16) \qquad \boldsymbol{v}^T A\boldsymbol{v} \le 2\,\boldsymbol{v}^T\overline{M}\boldsymbol{v}, \quad \forall \boldsymbol{v} \in \mathcal{U}^h.$$

Similarly, in case M is symmetric and M^{-1} is used in (3.5), it holds $\boldsymbol{v}^T A \boldsymbol{v} \leq 2\,\boldsymbol{v}^T M \boldsymbol{v}$. This, combined with (3.15), in view of (3.5), implies

$$\boldsymbol{v}^T B_{\text{add}}^{-1} \boldsymbol{v} \leq 2[\alpha(2 + \Lambda \delta \varkappa^2) + 1]\,\boldsymbol{v}^T A^{-1} \boldsymbol{v}, \quad \forall \boldsymbol{v} \in \mathcal{U}^h,$$

equivalently, $\boldsymbol{v}^T A \boldsymbol{v} \leq 2[\alpha(2 + \Lambda \delta \varkappa^2) + 1]\,\boldsymbol{v}^T B_{\text{add}} \boldsymbol{v}$.

Finally, consider B_{mult}^{-1} in (3.6). Note that $M + M^T - A$ being SPD is equivalent to the stationary iteration with M (and M^T) being A-convergent (convergent in the norm $\|\cdot\|_A$, induced by the matrix A), i.e., $|I - A^{1/2} M^{-1} A^{1/2}| = \|I - M^{-1} A\|_A = |I - A^{1/2} M^{-T} A^{1/2}| = \|I - M^{-T} A\|_A < 1$, where $|\cdot|$ is the Euclidean norm; see, e.g., [21]. Then, using (3.15),

$$\sup_{\boldsymbol{v} \neq \boldsymbol{0}} \frac{\boldsymbol{v}^T (I - M^{-T} A) \widetilde{B}^{-1} (I - A M^{-1}) \boldsymbol{v}}{\boldsymbol{v}^T A^{-1} \boldsymbol{v}} = \sup_{\boldsymbol{v} \neq \boldsymbol{0}} \frac{\boldsymbol{v}^T A^{\frac{1}{2}} (I - M^{-T} A) \widetilde{B}^{-1} (I - A M^{-1}) A^{\frac{1}{2}} \boldsymbol{v}}{\boldsymbol{v}^T \boldsymbol{v}}$$

$$= \sup_{\boldsymbol{v} \neq \boldsymbol{0}} \frac{\boldsymbol{v}^T (I - A^{\frac{1}{2}} M^{-T} A^{\frac{1}{2}}) A^{\frac{1}{2}} \widetilde{B}^{-1} A^{\frac{1}{2}} (I - A^{\frac{1}{2}} M^{-1} A^{\frac{1}{2}}) \boldsymbol{v}}{\boldsymbol{v}^T \boldsymbol{v}}$$

$$= |(I - A^{\frac{1}{2}} M^{-T} A^{\frac{1}{2}}) A^{\frac{1}{2}} \widetilde{B}^{-1} A^{\frac{1}{2}} (I - A^{\frac{1}{2}} M^{-1} A^{\frac{1}{2}})| \leq |I - A^{\frac{1}{2}} M^{-1} A^{\frac{1}{2}}|^2 |A^{\frac{1}{2}} \widetilde{B}^{-1} A^{\frac{1}{2}}|$$

$$\leq |A^{\frac{1}{2}} \widetilde{B}^{-1} A^{\frac{1}{2}}| = \sup_{\boldsymbol{v} \neq \boldsymbol{0}} \frac{\boldsymbol{v}^T A^{\frac{1}{2}} \widetilde{B}^{-1} A^{\frac{1}{2}} \boldsymbol{v}}{\boldsymbol{v}^T \boldsymbol{v}} = \sup_{\boldsymbol{v} \neq \boldsymbol{0}} \frac{\boldsymbol{v}^T \widetilde{B}^{-1} \boldsymbol{v}}{\boldsymbol{v}^T A^{-1} \boldsymbol{v}} \leq 2\alpha(2 + \Lambda \delta \varkappa^2).$$

This and (3.16), in view of (3.6), show

$$\boldsymbol{v}^T B_{\text{mult}}^{-1} \boldsymbol{v} \leq 2[\alpha(2 + \Lambda \delta \varkappa^2) + 1]\,\boldsymbol{v}^T A^{-1} \boldsymbol{v}, \quad \forall \boldsymbol{v} \in \mathcal{U}^h,$$

equivalently, $\boldsymbol{v}^T A \boldsymbol{v} \leq 2[\alpha(2 + \Lambda \delta \varkappa^2) + 1]\,\boldsymbol{v}^T B_{\text{mult}} \boldsymbol{v}$. Conversely, using the definition (3.6) of B_{mult}^{-1}, (3.14), and the positive definiteness of $\overline{M}$, and taking $\xi = \max\{\beta, 1\}$, it holds

$$\boldsymbol{v}^T B_{\text{mult}}^{-1} \boldsymbol{v} = \boldsymbol{v}^T \overline{M}^{-1} \boldsymbol{v} + \boldsymbol{v}^T (I - M^{-T} A) \widetilde{B}^{-1} (I - A M^{-1}) \boldsymbol{v}$$

$$\geq \boldsymbol{v}^T \overline{M}^{-1} \boldsymbol{v} + \xi^{-1}\,\boldsymbol{v}^T (I - M^{-T} A) A^{-1} (I - A M^{-1}) \boldsymbol{v}$$

$$= \boldsymbol{v}^T \overline{M}^{-1} \boldsymbol{v} + \xi^{-1}\,\boldsymbol{v}^T A^{-\frac{1}{2}} (I - A^{\frac{1}{2}} M^{-T} A^{\frac{1}{2}})(I - A^{\frac{1}{2}} M^{-1} A^{\frac{1}{2}}) A^{-\frac{1}{2}} \boldsymbol{v}$$

$$= \boldsymbol{v}^T \overline{M}^{-1} \boldsymbol{v} + \xi^{-1}\,\boldsymbol{v}^T A^{-\frac{1}{2}} (I - A^{\frac{1}{2}} \overline{M}^{-1} A^{\frac{1}{2}}) A^{-\frac{1}{2}} \boldsymbol{v}$$

$$= (1 - \xi^{-1})\,\boldsymbol{v}^T \overline{M}^{-1} \boldsymbol{v} + \xi^{-1}\,\boldsymbol{v}^T A^{-1} \boldsymbol{v} \geq \xi^{-1}\,\boldsymbol{v}^T A^{-1} \boldsymbol{v},$$

or, equivalently, $\boldsymbol{v}^T B_{\text{mult}} \boldsymbol{v} \leq \max\{\beta, 1\}\,\boldsymbol{v}^T A \boldsymbol{v}$. $\qquad\square$

Remark 3.5. There is a couple of additional assumptions on the smoother, M, in [21, Theorem 7.18]. However, they are not necessary in Theorem 3.4, due to the exactness of the approximation (3.7). Only the basic property of A-convergence of the iteration with M (i.e., $M + M^T - A$ being SPD) is assumed. In fact, (3.14) and (3.15) show, counting on (3.7), that $\widetilde{B}$ alone is spectrally equivalent to A, without the necessity of additional smoothing in the fine-scale setting here. Nevertheless, when coarse auxiliary spaces (described in Section 3.4) are utilized smoothing is necessary and a smoother, M, is outlined in Section 4. Note that (3.14) and (3.15) essentially represent, based on the surjectivity of Π_h, the so called "fictitious space lemma"; cf. [18, Theorem 10.1], [23, Theorem 2.1]

Remark 3.6. Notice that the construction of the space $\mathcal{E}^h \times \mathcal{F}^h$ from $\mathcal{U}^h$ and the proofs of Lemma 3.1, Corollary 3.2, and Theorem 3.4 are abstract, algebraic, and general in nature, using only generic properties of SPD and SPSD matrices. The degrees of freedom can be viewed in a general abstract sense, using only that they can be related to mesh entities (like elements, faces, edges, and vertices). A careful

inspection of the proofs shows that the particular form of the model problem (2.1), or (2.2), and its properties (particularly, that it is an elliptic PDE) are not utilized. Thus, the IP reformulation is applicable and its spectral equivalence properties are maintained for quite general SPD systems (i.e., convex quadratic minimization problems) that can be associated with appropriate local SPSD versions. In a general setting, the IP reformulation can be constructed as in (3.2), providing the element-by-element assembly property for a given problem, via replacing A_T and D_F with the corresponding matrices associated with the respective convex quadratic minimization of interest. Consequently, the approach is not limited to preconditioning systems necessarily coming from an H^1-conforming, per se, finite element method and potentially it can even be applicable to systems that may not be associated with any finite element method. For example, in the case of Raviart-Thomas spaces [6], constructing $\mathcal{E}^h \times \mathcal{F}^h$ involves cloning all dofs associated with the fluxes on the faces constituting a FACE and, for Nédélec spaces, the dofs associated with faces and their respective edges (e.g., in 3D) are cloned. Essentially, the overall procedure remains the same and the theoretical results (together with their proofs) carry over.

3.4. **Coarse formulation.** Combining the considerations in Section 3.1 and the end of Section 2.2, an IP formulation on a coarse pair of spaces, $\mathcal{E}^H \times \mathcal{F}^H$, can be obtained. Namely, having the prolongation matrix $\mathcal{P} \colon \mathcal{E}^H \times \mathcal{F}^H \mapsto \mathcal{E}^h \times \mathcal{F}^h$, the coarse IP matrix can be obtained via a standard "RAP" procedure. That is, $\mathcal{A}^H = \mathcal{P}^T \mathcal{A} \mathcal{P}$ and, locally, $\mathcal{A}_T^H = \mathcal{P}_T^T \mathcal{A}_T \mathcal{P}_T$, where $\mathcal{P}_T$ is a local version of $\mathcal{P}$, with columns – the basis vectors associated with T and all $F \subset \partial T$.

Alternatively, $\mathcal{A}_T^H$ can be computed directly, using (3.2) with $\boldsymbol{v}_e$ in the span of $\{\boldsymbol{q}_{T,i}\}_{i=1}^{m_T}$ and $\boldsymbol{v}_b$ in the subspace spanned by $\{\boldsymbol{q}_{F,i}\}_{i=1}^{m_F}$, for all $F \subset \partial T$, while $\mathcal{A}^H$ can be assembled from the local matrices, $\mathcal{A}_T^H$. Respective consistent, with $\mathcal{A}_T^H$, relations ELEMENT_Edof, ELEMENT_Bdof, and, consequently, ELEMENT_Adof are obtainable. This practically removes the necessity to construct the fine-scale space, $\mathcal{E}^h \times \mathcal{F}^h$, and matrices – $\mathcal{A}$ and $\mathcal{A}_T$. Furthermore, the operator $\Pi_H \colon \mathcal{E}^H \times \mathcal{F}^H \mapsto \mathcal{U}^h$ is defined as $\Pi_H = \Pi_h \mathcal{P}$ and can be constructed directly. Its continuity,

$$(\Pi_H \hat{\boldsymbol{v}}_c)^T A \, \Pi_H \hat{\boldsymbol{v}}_c \le 2(2 + \Lambda \delta \varkappa^2) \hat{\boldsymbol{v}}_c^T \mathcal{A}^H \hat{\boldsymbol{v}}_c, \quad \forall \hat{\boldsymbol{v}}_c \in \mathcal{E}^H \times \mathcal{F}^H,$$

follows immediately from Corollary 3.2. The coarse-space versions of the auxiliary space preconditioners for A, where $\mathcal{B}^H$ is a symmetric preconditioner for $\mathcal{A}^H$, are

$$(3.17) \quad \begin{aligned} (B_{\mathrm{add}}^H)^{-1} &= \overline{M}^{-1} + \Pi_H (\mathcal{B}^H)^{-1} \Pi_H^T, \\ (B_{\mathrm{mult}}^H)^{-1} &= \overline{M}^{-1} + (I - M^{-T} A) \Pi_H (\mathcal{B}^H)^{-1} \Pi_H^T (I - A M^{-1}). \end{aligned}$$

A particular approach, considered here, for obtaining a coarse basis and constructing $\mathcal{E}^H \times \mathcal{F}^H$ is by solving local generalized eigenvalue problems, cf. [9], of the type

$$(3.18) \quad A_T \boldsymbol{q} = \lambda D_T \boldsymbol{q},$$

where D_T is the diagonal of A_T and the eigenvalues are ordered $\lambda_1 \le \cdots \le \lambda_{n_T}$. The first m_T ($1 \le m_T < n_T$) eigenvectors of (3.18), constitute the D_T-orthogonal basis $\{\boldsymbol{q}_{T,i}\}_{i=1}^{m_T}$ associated with T. Particularly, let $\theta \in (0,1)$ be given and m_T satisfy $\lambda_{m_T} \le \theta \lambda_{n_T}$ and $\lambda_{m_T+1} > \theta \lambda_{n_T}$. For each $F \in \Phi^H$, collect $\{\boldsymbol{q}_{T_+,i}|_F\}_{i=1}^{m_{T_+}}$ and $\{\boldsymbol{q}_{T_-,i}|_F\}_{i=1}^{m_{T_-}}$ from its adjacent ELEMENTS T_+ and T_-. After performing SVD to filter out any linear dependence, the $\ell_2(F)$-orthogonal basis $\{\boldsymbol{q}_{F,i}\}_{i=1}^{m_F}$, associated with F, is obtained. Notice that this procedure maintains the relation, similarly to the fine-scale setting, that $\mathcal{F}^H$ is the trace-space (on FACES) for the functions in $\mathcal{E}^H$.

Remark 3.7. In general, as far as the feasibility of the approach is concerned, $\mathcal{F}^H$ does not need to be precisely the trace-space for $\mathcal{E}^H$. Particularly, it can be a proper subspace of the trace-space (on FACES) for the functions in $\mathcal{E}^H$, potentially providing additional reduction, when employing static condensation (Section 3.5). The study of the choices of $\mathcal{F}^H$ and the effect this may have on the properties of the preconditioners is a subject of future work.

Remark 3.8. An alternative choice is to utilize polynomial bases for $\mathcal{E}^H$ and $\mathcal{F}^H$ of order lower than the order of $\mathcal{E}^h$ and $\mathcal{F}^h$. This paper concentrates on coarse basis vectors obtained via the above described generalized eigenvalue problems.

3.5. **Static condensation.** Consider, to simplify the presentation, fine-scale spaces and matrices for the IP method – $\mathcal{E}^h \times \mathcal{F}^h$, $\mathcal{A}$, and $\mathcal{A}_T$. The idea is, in lieu of directly preconditioning $\mathcal{A}$ in B_{add}^{-1} and B_{mult}^{-1}, to eliminate all edofs in $\mathcal{A}$ and precondition the resulting Schur complement, expressed only on the bdofs. This procedure is referred to as *static condensation*, since it involves "condensing" the formulation on the interfaces.

Namely, using the block factorization (cf. (3.9))

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_{ee} & \\ \mathcal{A}_{be} & \mathcal{S} \end{bmatrix} \begin{bmatrix} I & \mathcal{A}_{ee}^{-1}\mathcal{A}_{eb} \\ & I \end{bmatrix},$$

where $\mathcal{S} = \mathcal{A}_{bb} - \mathcal{A}_{be}\mathcal{A}_{ee}^{-1}\mathcal{A}_{eb}$ is the respective SPD Schur complement, a preconditioner for $\mathcal{A}$ is obtained as follows:

$$(3.19) \qquad \mathcal{B}_{\text{sc}}^{-1} = \begin{bmatrix} I & -\mathcal{A}_{ee}^{-1}\mathcal{A}_{eb} \\ & I \end{bmatrix} \begin{bmatrix} \mathcal{A}_{ee}^{-1} & \\ -S^{-1}\mathcal{A}_{be}\mathcal{A}_{ee}^{-1} & S^{-1} \end{bmatrix},$$

where S is a SPD preconditioner for $\mathcal{S}$, providing a SPD $\mathcal{B}_{\text{sc}}$. Observe that computing the action of $\mathcal{B}_{\text{sc}}^{-1}$ involves applying S^{-1} once and $\mathcal{A}_{ee}^{-1}$ twice: first, during the "elimination" stage, represented by the second factor in (3.19), that provides the argument for S^{-1}; second, during the "backward substitution" stage, represented by the first factor in (3.19), that updates the edof portion of the solution. If S is spectrally equivalent to $\mathcal{S}$, then $\mathcal{B}_{\text{sc}}$ is spectrally equivalent to $\mathcal{A}$.

Clearly, obtaining $\mathcal{A}_{ee}^{-1}$, or an appropriate approximation thereof, is necessary for computing both $\mathcal{S}$ and the action of $\mathcal{B}_{\text{sc}}^{-1}$. As mentioned, no edofs are coupled across ELEMENTS. That is, $\mathcal{A}_{ee}$ is block-diagonal – $\mathcal{A}_{ee} = \text{diag}(\mathcal{A}_{T,ee})_{T\in\mathcal{T}^H}$; see (3.3). Thus, computing $\mathcal{A}_{ee}^{-1}$, for obtaining $\mathcal{S}$ and the action of $\mathcal{B}_{\text{sc}}^{-1}$, only involves local work on ELEMENTS. Moreover, consider the local, for T and the respective $\mathcal{A}_T$, cf. (3.3), SPSD Schur complements $\mathcal{S}_T = \mathcal{A}_{T,bb} - \mathcal{A}_{T,be}\mathcal{A}_{T,ee}^{-1}\mathcal{A}_{T,eb}$, formulated on the bdofs associated with T (more precisely, with ∂T). Notice that the $\mathcal{S}_T$ matrices provide a local structure associated with $\mathcal{S}$, since $\mathcal{S}$ can be assembled from $\mathcal{S}_T$. Hence, while $\mathcal{S}_T$ are generally dense, $\mathcal{S}$ possesses a typical sparsity pattern associated with matrices coming from finite element methods. The local structures of $\mathcal{A}$ and $\mathcal{S}$, respectively provided by $\mathcal{A}_T$ and $\mathcal{S}_T$, facilitate the application of AMGe methods for preconditioning $\mathcal{A}$ and $\mathcal{S}$, as described in Section 5.

Note that everything is general and also applies in the coarse-scale setting, of Section 3.4, by instead utilizing $\mathcal{E}^H \times \mathcal{F}^H$, $\mathcal{A}^H$, and $\mathcal{A}_T^H$, providing the Schur matrices $\mathcal{S}^H$ and $\mathcal{S}_T^H$, as well as the preconditioners S^H for $\mathcal{S}^H$ and $(\mathcal{B}_{\text{sc}}^H)^{-1}$ for $\mathcal{A}^H$.

4. THE SMOOTHER M

A particular smoother M, for A, which is a part of the auxiliary space preconditioners proposed in this paper, is shortly described now. Particularly, a polynomial smoother is utilized based on the Chebyshev polynomial of the first kind.

For a given integer $\nu \geq 1$, consider the polynomial of degree $3\nu + 1$ on $[0, 1]$

$$p_\nu(t) = \left(1 - T_{2\nu+1}^2(\sqrt{t})\right)(-1)^\nu \frac{1}{2\nu + 1} \frac{T_{2\nu+1}(\sqrt{t})}{\sqrt{t}},$$

satisfying $p_\nu(0) = 1$, where $T_l(t)$ is the Chebyshev polynomial of the first kind on $[-1, 1]$. Then, M is defined as $M^{-1} = [I - p_\nu(b^{-1}D^{-1}A)]A^{-1}$ or, equivalently, $I - M^{-1}A = p_\nu(b^{-1}D^{-1}A)$, where $b = \mathcal{O}(1)$ is a parameter satisfying $\boldsymbol{v}^T A \boldsymbol{v} \leq b\,\boldsymbol{v}^T D \boldsymbol{v}$, for all $\boldsymbol{v} \in \mathcal{U}^h$, and D is the diagonal of A (or another appropriate diagonal matrix). Note that M is SPD and the action of such a polynomial smoother is computed via $3\nu + 1$ Jacobi-type iterations, using the roots of the polynomial, which makes it convenient for parallel computations; see [14, Section 4.2.2].

In practice, D, in the definition of M, can be replaced by a diagonal weighted ℓ_1-smoother like $W = \text{diag}(w_i)_{i=1}^{\dim(\mathcal{U}^h)}$, where $w_i = \sum_{j=1}^{\dim(\mathcal{U}^h)} |a_{ij}| \sqrt{a_{ii}/a_{jj}}$. Such a choice allows setting $b = 1$.

For more information on the subject, consult [21, 5, 9, 8, 19, 15].

5. AMGe for the nonconforming formulations

A procedure for constructing AMGe hierarchies and preconditioners, for SPD matrices and applicable to the auxiliary space problems, is outlined in this section. The approach is related to the ideas in [11, 16, 9]. First, the method is described in a general abstract setting, which requires reusing some of the notions and notation, introduced in Section 2, in a slightly different context. Then, details concerning the specifics of applying the approach in the auxiliary space setting are discussed, including the mapping between the terminology here and in Section 2. This should avoid any confusion between the notions and they should not get mixed.

5.1. **General description.** Let a collection $\mathfrak{T}^h = \{\tau\}$ of elements and an associated set of dofs $\mathfrak{D}^h$, on $\mathfrak{T}^h$, be given. Here, $\mathfrak{T}^h$ and $\mathfrak{D}^h$ supply the respective relations element_element and element_dof. Additionally, SPSD element matrices A_τ, for $\tau \in \mathfrak{T}^h$, are obtainable, whose global assembly, in accordance with element_dof, provides a SPD matrix A. This constitutes the given data and the goal is to construct a preconditioner for A. Typically, such data comes from meshes, finite element spaces, and weak formulations, as is the case in this paper, but the approach can be evoked with data that is generally not necessarily associated with finite element methods.

Using element_element and, e.g., a graph partitioner, a collection $\mathfrak{T}^H = \{T\}$ of ELEMENTS is obtained, where the ELEMENTS, T, are non-overlapping and connected, in terms of element_element, sets of elements, τ. This is expressed via the relation ELEMENT_element. As before, compute

$$\text{ELEMENT\_dof} = \text{ELEMENT\_element} \times \text{element\_dof},$$

which represents the ELEMENTS as sets of dofs. An intersection procedure[2] over ELEMENT_dof generates so called *minimal intersection sets* (MISes); see [21, 11, 16]. In more detail, MISes are (non-overlapping) equivalence classes of dofs with respect to the relation that two dofs are equivalent if they belong to identical sets of ELEMENTS, in accordance with (ELEMENT_dof)^T. The collection of MISes partitions the set of dofs, $\mathfrak{D}^h$, and, for every T, there is a subcollection of MISes that partitions the dofs

---

[2] This is fundamentally the same procedure as the one identifying FACES in Section 2.1.

in T, as determined by ELEMENT_dof. The intersection procedure constructs the relations MIS_dof and ELEMENT_MIS, where

$$\text{ELEMENT\_MIS} = \text{ELEMENT\_dof} \times (\text{MIS\_dof})^T.$$

Next, the derivation of a coarse basis, as a linearly independent set of vectors defined on dofs, is described. Based on ELEMENT_element (also, utilizing element_-dof and ELEMENT_dof), assemble local, on T, SPSD matrices A_T from the element matrices A_τ, for $\tau \in T$. Solve local generalized eigenvalue problems

$$(5.1) \qquad\qquad A_T \boldsymbol{q} = \lambda D_T \boldsymbol{q},$$

where D_T is the diagonal of A_T and the eigenvalues are ordered $\lambda_1 \leq \cdots \leq \lambda_{n_T}$. Take the first m_T $(1 \leq m_T < n_T)$ eigenvectors of (5.1) forming the D_T-orthogonal set $\{\boldsymbol{q}_{T,i}\}_{i=1}^{m_T}$ associated with T. Particularly, let $\theta_s \in (0,1)$ be given and m_T satisfy $\lambda_{m_T} \leq \theta_s \lambda_{n_T}$ and $\lambda_{m_T+1} > \theta_s \lambda_{n_T}$. Then, for each MIS $\mathscr{M}$, collect $\{\boldsymbol{q}_{T,i}\}_{i=1}^{m_T}$ from all T associated with $\mathscr{M}$, via $(\text{ELEMENT\_MIS})^T$, and restrict them to $\mathscr{M}$ based on the relations MIS_dof and ELEMENT_dof. After performing SVD to filter out any linear dependence, the $\ell_2(\mathscr{M})$-orthogonal basis $\{\boldsymbol{q}_{\mathscr{M},i}\}_{i=1}^{m_\mathscr{M}}$, associated with $\mathscr{M}$, is obtained and organized as the columns of a local, on $\mathscr{M}$, prolongation matrix $P_\mathscr{M} = [\boldsymbol{q}_{\mathscr{M},1}, \cdots, \boldsymbol{q}_{\mathscr{M},m_\mathscr{M}}]$. The resulting global prolongation matrix takes the form

$$P = \begin{bmatrix} P_{\mathscr{M}_1} & & \\ & \ddots & \\ & & P_{\mathscr{M}_{n_a}} \end{bmatrix}.$$

Similarly, for $T \in \mathfrak{T}^H$ and all $\mathscr{M}$ related to T, via ELEMENT_MIS, collect all $P_\mathscr{M}$ to construct a local, on T, prolongation matrix P_T.

By building P, a coarse space is constructed, where the columns of P form the basis. Identify the coarse degrees of freedom, designated by "Dofs" and constituting the set $\mathfrak{D}^H$, with the columns of P, i.e., with the basis vectors, and define the relation Dof_dof as a representation of the sparsity pattern of P^T. Consequently, a Dof is related to a dof if and only if the respective coarse basis vector is supported on that dof. Similarly, an ELEMENT is related to a Dof, determining ELEMENT_Dof, whenever the respective coarse basis vector is supported on any dof of the ELEMENT. That is,

$$\text{ELEMENT\_Dof} = \text{ELEMENT\_dof} \times (\text{Dof\_dof})^T.$$

Notice that this corresponds to the columns of P_T. Hence, coarse SPSD element matrices are consistently obtained via local "RAPs" – $A_{T,c} = P_T^T A_T P_T$ – and the global SPD coarse matrix is $A_c = P^T A P$, which can be assembled from $A_{T,c}$, for $T \in \mathfrak{T}^H$, in accordance with ELEMENT_Dof. Furthermore, ELEMENTS are related via

$$\text{ELEMENT\_ELEMENT} = \text{ELEMENT\_element} \times \text{element\_element}$$
$$\times (\text{ELEMENT\_element})^T.$$

By assigning

$$\mathfrak{T}^h \leftarrow \mathfrak{T}^H, \ \mathfrak{D}^h \leftarrow \mathfrak{D}^H, \ \text{element\_element} \leftarrow \text{ELEMENT\_ELEMENT},$$
$$\text{element\_dof} \leftarrow \text{ELEMENT\_Dof}, \ \{A_\tau\} \leftarrow \{A_{T,c}\},$$

the above procedure can be executed recursively, building a hierarchy of meshes, spaces, bases, degrees of freedom, element matrices, global matrices, relations, and transition operators between the spaces. To obtain multilevel cycles as preconditioners, only remains to appoint a "relaxation" ("smoothing") method on each level. To that purpose, the polynomial smoother in Section 4 is employed with an integer

parameter $\nu_s \geq 1$. This algorithm is implemented, including parallel, in SAAMGE [3].

Remark 5.1. Observe that the approach here can be invoked with given data: $\mathfrak{T}^h$, $\mathfrak{D}^h$, element_element, $\mathfrak{T}^H$ (in the form of ELEMENT_element), ELEMENT_dof, and $\{ A_T; \ T \in \mathfrak{T}^H \}$. In that case, the procedures for obtaining these structures are skipped and the precursors (element_dof and $\{ A_\tau; \ \tau \in \mathfrak{T}^h \}$), which, in fact, may not be available, are not needed.

5.2. **On the nonconforming formulations.** Specifics, concerning the application of the method outlined in Section 5.1 in the context of the nonconforming formulations, are addressed now.

First, consider the fine-scale IP setting, involving $\mathcal{E}^h \times \mathcal{F}^h$, $\mathcal{A}$, and $\mathcal{A}_T$. Since agglomeration is already performed to form the IP problem, without employing a coarsening process for the spaces, it is natural to reuse the same agglomeration and assign $\mathfrak{T}^h \leftarrow \mathcal{T}^h$ (using its element_element), the set of adofs (associated with $\mathcal{E}^h \times \mathcal{F}^h$) as $\mathfrak{D}^h$, $\mathfrak{T}^H \leftarrow \mathcal{T}^H$ with the respective ELEMENT_element and ELEMENT_dof $\leftarrow$ ELEMENT_adof; see Sections 2.1 and 3.1. In this case, the respective local, for $T \in \mathfrak{T}^H$, agglomerate matrices are the already assembled IP matrices $\{ \mathcal{A}_T; \ T \in \mathfrak{T}^H \}$. Then, in view of Remark 5.1, the general scheme in Section 5.1 produces a hierarchy of spaces and a multilevel preconditioner for $\mathcal{A}$.

Notice that the outlined methodology automatically generates a hierarchy of space pairs $\mathcal{E}^{H_l} \times \mathcal{F}^{H_l}$, for $l = 0, \ldots, n_\ell$ and $H_0 = h$, and the IP method formulated on these spaces via a variational ("RAP") procedure. This is due to the MISes' construction, using ELEMENT_adof, automatically reidentifying the FACES and separating edofs from bdofs, which eventually results in obtaining separate bases associated with ELEMENTS and interfaces. Thus, once the space pair $\mathcal{E}^h \times \mathcal{F}^h$ is constructed as discontinuous, the whole hierarchy maintains the same general complexion.

Comparing $\mathcal{E}^{H_1} \times \mathcal{F}^{H_1}$, constructed here by a single coarsening, with the space pair in Section 3.4, resemblance can be seen. While the notion of ELEMENT is the same, a major difference is that $\mathcal{E}^{H_1} \times \mathcal{F}^{H_1}$ uses basis vectors obtained via generalized eigenvalue problems for the local IP matrices $\mathcal{A}_T$, as defined by (3.2), whereas Section 3.4 utilizes A_T – the local representations of (2.2). Moreover, the final preconditioners for A have clear differences. Namely, B_{add} in (3.5) and B_{mult} in (3.6) employ a fine-scale auxiliary space, $\mathcal{E}^h \times \mathcal{F}^h$, and a multigrid preconditioner, $\mathcal{B}$, for $\mathcal{A}$ acts on the respective fine level via relaxation. In contrast, B_{add}^H and B_{mult}^H, in (3.17), directly utilize a coarse auxiliary space, $\mathcal{E}^H \times \mathcal{F}^H$, and a multigrid preconditioner, $\mathcal{B}^H$, for $\mathcal{A}^H$ performs no action on the fine level. In the latter case, the method of Section 5.1 can still be employed to precondition $\mathcal{A}^H$ from Section 3.4. This is described next.

Since one coarsening is readily accomplished in Section 3.4, the coarse level there acts as a fine one for the preconditioner here. That is, assign $\mathfrak{T}^h \leftarrow \mathcal{T}^H$, element_element $\leftarrow$ ELEMENT_ELEMENT, the set of Adofs (associated with $\mathcal{E}^H \times \mathcal{F}^H$) as $\mathfrak{D}^h$, and element_dof $\leftarrow$ ELEMENT_Adof; see Sections 2.1 and 3.4. Also, the available coarse IP matrices $\{ \mathcal{A}_T^H; \ T \in \mathcal{T}^H \}$ act as fine-level element matrices for the solver here. Then, the full procedure in Section 5.1 can be invoked, obtaining a multilevel preconditioner for $\mathcal{A}^H$.

Finally, the AMGe method of Section 5.1 can be applied to the "condensed" formulations of Section 3.5, since the approach in Section 3.5 provides a local structure. Particularly, the static condensation method eliminates edofs, leaving only bdofs. That is, the only remaining space is $\mathcal{F}^h$ and a "condensed" formulation on
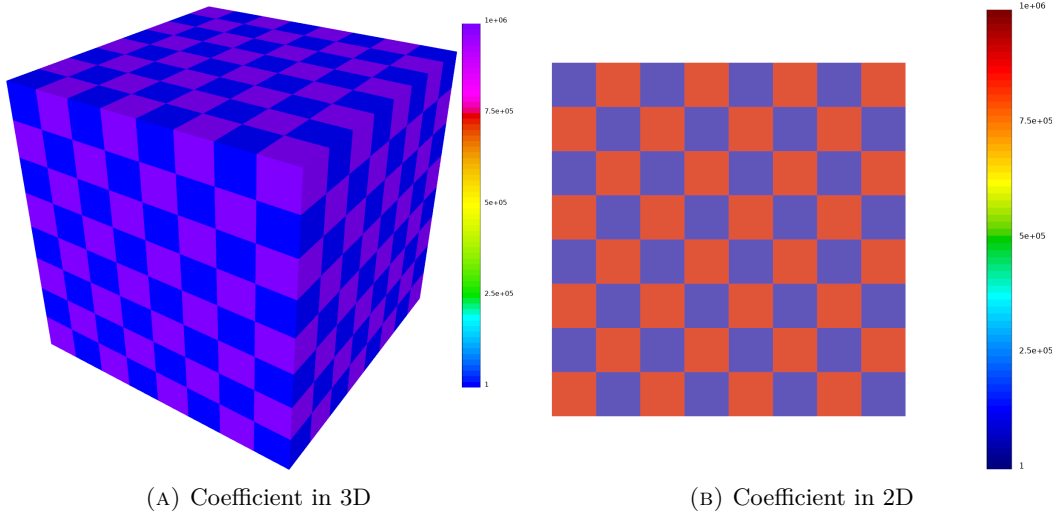
(A) Coefficient in 3D   (B) Coefficient in 2D

FIGURE 6. The coefficient κ in (2.1).

that space. The procedure is analogous and obtainable from the above discussion by simply replacing: adof by bdof, $\mathcal{A}$ by $\mathcal{S}$, and $\mathcal{A}_T$ by $\mathcal{S}_T$. Then, the agglomeration and MISes' construction coarsen the interfaces and produce basis vectors to generate a space hierarchy $\mathcal{F}^{H_l}$ and a multilevel preconditioner for $\mathcal{S}$, where "RAP" provides "condensed" formulations on each level. Similarly, having in mind Section 3.4, a multigrid preconditioner for $\mathcal{S}^H$ can be obtained starting from $\mathcal{F}^H$ by using the above procedure and replacing: adof by bdof, $\mathcal{A}^H$ by $\mathcal{S}^H$, and $\mathcal{A}_T^H$ by $\mathcal{S}_T^H$. Note that, as before, starting with $\mathcal{F}^h$ reuses the agglomeration from the IP method, whereas starting with $\mathcal{F}^H$ interprets those agglomerates as fine-level elements.

6. NUMERICAL EXAMPLES

A set of numerical results is shown. Two test cases are presented: on low and high order discretizations. First, the test setting is outlined.

6.1. **Setting.** The test problem is (2.1) with $f \equiv 1$, κ representing the high contrast coefficients shown in Figs. 6a and 6b, and $\delta = 1$ in (3.2). Both $\mathcal{T}^h$ and $\mathcal{T}^H$, for the IP formulation, are regular and of the general type depicted in Figs. 1a and 1b. Any further agglomeration, required by the procedure of Section 5, is produced using METIS [2]. Note that the ELEMENTS in Figs. 1a and 1b get respectively refined as the mesh is refined, i.e., refining $\mathcal{T}^h$ leads to a corresponding refinement of $\mathcal{T}^H$. In contrast, the coefficients in Figs. 6a and 6b remain fixed and their pattern is invariant with respect to mesh refinement.

A few measures of operator complexity (OC), representing relative sparsity in the obtained space hierarchies, are reported. Namely, the OC of the IP reformulation:

$$\mathrm{OC_{IP}} = 1 + \mathrm{NNZ}(\mathcal{A}^{H_0})/\mathrm{NNZ}(A);$$

the OC of the "auxiliary" multigrid hierarchy (e.g., generated by the method in Section 5), relative to the IP matrix:

$$\mathrm{OC_{aux}} = 1 + \sum_{l=1}^{n_\ell} \mathrm{NNZ}(\mathcal{A}^{H_l})/\mathrm{NNZ}(\mathcal{A}^{H_0});$$

Refs	# dofs	# Adofs	OC_{IP}	n_{it}
0	4913	2240	1.304	2
1	35937	17152	1.296	4
2	274625	134144	1.291	6
3	2146689	1060864	1.288	9
4	16974593	8437760	1.287	12
5	135005697	67305472	1.286	15

(A) $(\mathcal{B}^H)^{-1} = (\mathcal{A}^H)^{-1}$

Refs	$n_\ell + 1$	OC_{aux}	OC_{orig}	n_{it}
0	4	1.542	1.468	2
1	5	1.711	1.506	4
2	7	1.762	1.513	6
3	8	1.753	1.506	11
4	9	1.728	1.496	21
5	9	1.708	1.489	41

(B) One AMGe (Section 5) V-cycle as $(\mathcal{B}^H)^{-1}$

TABLE 1. Low order test results with coarsening and without static condensation. That is, requiring an (approximate) inverse, $(\mathcal{B}^H)^{-1}$, of the respective $\mathcal{A}^H$.

and the total OC, relative to the matrix A in (2.3):

$$\mathrm{OC}_{\mathrm{orig}} = 1 + \sum_{l=0}^{n_\ell} \mathrm{NNZ}(\mathcal{A}^{H_l})/\mathrm{NNZ}(A) = 1 + \mathrm{OC}_{\mathrm{aux}} \times (\mathrm{OC}_{\mathrm{IP}} - 1).$$

Here, NNZ denotes the number of nonzero entries in the sparsity pattern of a matrix, n_ℓ is the number of levels (excluding the fine one) in the "auxiliary" multigrid hierarchy (e.g., generated by the method in Section 5), and $\mathcal{A}^{H_0} = \{\mathcal{A}, \mathcal{S}, \mathcal{A}^H, \text{ or } \mathcal{S}^H\}$ (depending on the case) is the matrix of the IP reformulation, involving or not coarsening and/or static condensation. The matrices $\mathcal{A}^{H_l}$, for $l \geq 1$, represent the coarse versions of $\mathcal{A}^{H_0}$ in the solver hierarchy built by the method in Section 5.

Recall that dofs are associated with $\mathcal{U}^h$ and the matrix A in (2.3), whereas adofs, edofs, bdofs and Adofs, Edofs, Bdofs are related to $\mathcal{E}^h \times \mathcal{F}^h$ and $\mathcal{E}^H \times \mathcal{F}^H$ and the matrices $\mathcal{A}$, $\mathcal{S}$ or $\mathcal{A}^H$, $\mathcal{S}^H$.

In all cases, the preconditioned conjugate gradient (PCG) method and the respective multiplicative auxiliary space preconditioner are applied for solving the linear system (2.3) and the number of iterations, n_{it}, is reported. The relative tolerance is 10^{-8} and the measure in the stopping criterion is $\boldsymbol{r}^T B^{-1} \boldsymbol{r}$, for a current residual $\boldsymbol{r}$ and the respective preconditioner B^{-1} in the PCG method. The smoother in Section 4 is employed. Moreover, coarsening (Section 3.4) is utilized in all cases and the particular preconditoner is B_{mult}^H in (3.17). Thus, the problem is reduced to the choice of $(\mathcal{B}^H)^{-1}$ as an approximate inverse of $\mathcal{A}^H$. Particularly, when static condensation (Section 3.5) is employed, i.e., $(\mathcal{B}^H)^{-1} = (\mathcal{B}_{\mathrm{sc}}^H)^{-1}$ (a coarse version of (3.19)), the problem is reduced to the choice of $(S^H)^{-1}$ as an approximate inverse of the respective $\mathcal{S}^H$.

6.2. **Low order discretization.** Here, a 3D mesh of the type shown in Fig. 1a is sequentially refined and spaces, $\mathcal{U}^h$, of piecewise linear finite elements are obtained. Coarse auxiliary spaces of Section 3.4 are built with a single eigenvector per ELEMENT, corresponding to the smallest eigenvalue. Static condensation (Section 3.5) is very cheap in this case, involving the elimination of only one edof per ELEMENT. The smoother in Section 4 is used throughout with $\nu = 4$. Also, whenever an AMGe hierarchy (Section 5) is constructed, a single (smallest) eigenvector is taken from all local eigenvalue problems in Section 5 on all levels.

First, no static condensation is employed and the IP problem is inverted (almost) exactly, resulting in $(\mathcal{B}^H)^{-1} = (\mathcal{A}^H)^{-1}$ in the auxiliary space preconditioner in (3.17). Results are shown in Table 1a. Then, $(\mathcal{B}^H)^{-1}$ is implemented via a single

Refs	# dofs	# Bdofs	$\mathrm{OC_{IP}}$	n_{it}
0	4913	1728	1.647	2
1	35937	13056	1.619	4
2	274625	101376	1.604	6
3	2146689	798720	1.597	9
4	16974593	6340608	1.593	12
5	135005697	50528256	1.591	15

(A) $(\mathcal{B}_{\mathrm{sc}}^H)^{-1}$ with $(S^H)^{-1} = (\mathcal{S}^H)^{-1}$

Refs	$n_\ell + 1$	$\mathrm{OC_{aux}}$	$\mathrm{OC_{orig}}$	n_{it}
0	4	1.337	1.864	2
1	5	1.571	1.973	4
2	7	1.604	2.046	6
3	8	1.814	2.083	9
4	9	1.852	2.098	16
5	9	1.871	2.106	28

(B) $(\mathcal{B}_{\mathrm{sc}}^H)^{-1}$ with fixed three PCG iterations, preconditioned by an AMGe (Section 5) V-cycle, as $(S^H)^{-1}$

Refs	$n_\ell + 1$	$\mathrm{OC_{aux}}$	$\mathrm{OC_{orig}}$	n_{it}
0	4	1.061	1.686	2
1	5	1.090	1.675	4
2	5	1.117	1.675	6
3	6	1.153	1.688	9
4	7	1.166	1.692	12
5	9	1.185	1.701	17

(C) $(\mathcal{B}_{\mathrm{sc}}^H)^{-1}$ with fixed one PCG iteration, preconditioned by a BoomerAMG V-cycle, as $(S^H)^{-1}$

TABLE 2. Low order test results with coarsening and with static condensation. That is, a coarse version, $(\mathcal{B}_{\mathrm{sc}}^H)^{-1}$, of (3.19) is used, requiring an (approximate) inverse, $(S^H)^{-1}$, of the respective $\mathcal{S}^H$. Here, in two of the cases, $(S^H)^{-1}$ is obtained invoking a fixed number of conjugate gradient iterations, preconditioned by multigrid V-cycles.

V-cycle of the solver in Section 5; see Table 1b. The number of iterations does not scale very well, in the particular case shown in Table 1b, while trying to preserve reasonable cost per cycle of the method in Section 5. This can be partially remedied by using static condensation (Section 3.5) and wrapping the AMGe V-cycle in a few PCG iterations, as in Table 2b, but the cost of each application of the action of $(S^H)^{-1}$ is increased. However, invoking BoomerAMG (a well-known AMG implementation and a part of the HYPRE library[1]) wrapped in a single PCG iteration provides a much more efficient alternative in this case. Indeed, observe that the number of iterations in Table 2c is almost the same as in Table 2a, which involves the (almost) exact inversion of the respective Schur complement.

The AMGe method (Section 5) demonstrates much better behavior in the high order tests below.

6.3. **High order discretization.** Intuitively, the IP reformulation is expected to exhibit more beneficial properties in the context of increasing the polynomial order of the finite element spaces, particularly in a combination with coarsening (Section 3.4) and static condensation (Section 3.5). The available implementation is currently not matrix-free, therefore results in 2D are demonstrated using a fixed mesh of the type shown in Fig. 1b and increasing the polynomial order. Coarse auxiliary spaces described in Section 3.4 are constructed with $\theta = 0.05$ and the static condensation in Section 3.5 is employed in Table 3, while Table 4 is produced utilizing the same parameters without static condensation. Notice that, for the condensed case, the

Order	# dofs	# Bdofs	OC_{IP}	n_{it}
1	289	136	1.738	2
2	1089	144	1.117	6
3	2401	756	1.816	5
4	4225	980	1.557	4
5	6561	1204	1.408	5
6	9409	1428	1.313	7
7	12769	1652	1.248	7
8	16641	1876	1.201	4
9	21025	2100	1.167	4
10	25921	2324	1.141	3
11	31329	2548	1.120	3
12	37249	2772	1.104	4

(A) $(\mathcal{B}_{sc}^H)^{-1}$ with $(S^H)^{-1} = (\mathcal{S}^H)^{-1}$

Order	$n_\ell + 1$	OC_{aux}	OC_{orig}	n_{it}
1	2	1.685	2.243	3
2	2	1.626	1.191	6
3	3	1.254	2.023	6
4	4	1.551	1.865	6
5	5	1.758	1.717	8
6	5	1.956	1.614	9
7	6	2.185	1.541	9
8	7	2.402	1.484	8
9	7	2.618	1.437	9
10	11	3.986	1.561	12
11	12	4.318	1.520	14
12	13	4.650	1.484	15

(B) $(\mathcal{B}_{sc}^H)^{-1}$ with one AMGe (Section 5) V-cycle as $(S^H)^{-1}$

TABLE 3. High order test results with coarsening and with static condensation. That is, a coarse version, $(\mathcal{B}_{sc}^H)^{-1}$, of (3.19) is used, requiring an (approximate) inverse, $(S^H)^{-1}$, of the respective $\mathcal{S}^H$.

Order	# dofs	# Adofs	OC_{IP}	n_{it}
1	289	200	1.573	3
2	1089	208	1.090	6
3	2401	984	1.430	5
4	4225	1494	1.466	4
5	6561	2130	1.492	5
6	9409	3216	1.632	7
7	12769	4982	1.891	7
8	16641	7518	2.231	4
9	21025	10978	2.666	4
10	25921	15894	3.299	3

(A) $(\mathcal{B}^H)^{-1} = (\mathcal{A}^H)^{-1}$

Order	$n_\ell + 1$	OC_{aux}	OC_{orig}	n_{it}
1	2	1.791	2.026	3
2	2	1.736	1.155	6
3	3	1.231	1.529	10
4	4	1.505	1.701	12
5	5	1.779	1.875	16
6	5	2.161	2.366	19
7	6	2.562	3.282	25
8	7	2.841	4.499	34
9	7	3.161	6.267	46
10	11	4.766	11.958	60

(B) One AMGe (Section 5) V-cycle as $(\mathcal{B}^H)^{-1}$

TABLE 4. High order test results with coarsening and without static condensation. That is, requiring an (approximate) inverse, $(\mathcal{B}^H)^{-1}$, of the respective $\mathcal{A}^H$.

Edofs are eliminated, involving exclusively local coarse-scale work, leaving only Bdofs. The smoother in Section 4 is used throughout with $\nu = 2$.

The method in Section 5 is constructed with $\theta_s = 0.05$. Since no mesh refinement is performed, no additional agglomeration is invoked. That is, the ELEMENTS for the IP reformulation are maintained throughout the hierarchy and only the basis functions are reduced using the eigenvalue problems and imposing an additional requirement on the SVD filtering of linear dependencies on MISes. Namely, during the coarsening on each MIS, it is enforced that the SVD should remove at least 2 or 3 vectors relative to the current level being coarsened. This is reminiscent of p-multigrid but in a spectral AMGe setting.

Results, with static condensation, are shown in Table 3, demonstrating the especially good performance of the auxiliary space formulation – the case of utilizing a

coarse version, $(\mathcal{B}_{sc}^H)^{-1}$, of (3.19) with $(S^H)^{-1} = (\mathcal{S}^H)^{-1}$. The method of Section 5 also shows more reliable performance in the case of high order discretizations and static condensation. Maintaining all parameters the same and only disabling static condensation produces the results in Table 4. Observe the differences in operator complexities as the order is increased and the declined performance of the AMGe solver of Section 5. This illustrates the utility of static condensation, when high order discretizations are involved.

7. Conclusions and future work

In this paper, we have introduced a modification of the classical interior penalty (IP) discretization method by having an additional space of discontinuous functions associated with the interfaces between the subdomains (or elements) used to couple the individual local bilinear forms. This allows for element-by-element (or subdomain-by-subdomain) assembly. The resulting modified IP bilinear form and a reduced (Schur complement) form of it, obtained by static condensation, are then utilized in the construction of a number of auxiliary space preconditioners. These are analyzed and their proven mesh-independence spectral equivalenvce properties are illustrated by a set of numerical tests on 2D and 3D second order scalar elliptic equations, including the case of high order finite element discretizations. The element-by-element assembly property of the modified IP bilinear form was also beneficial in the construction of element-based algebraic multigrid (AMGe) that is used to replace the exact inverses in the auxiliary space preconditioners. A possible feasible extension of this work is to utilize the studied auxiliary space preconditioners with AMGe coarse solves in the setting of matrix-free solvers for high order finite element discretizations in combination with the polynomial smoothers presented in Section 4. Moreover, applying the approach of this paper to formulations that are conforming in $H(\text{div})$ or $H(\text{curl})$ (see Remark 3.6) is an interesting future study.

References

[1] HYPRE: Scalable Linear Solvers and Multigrid Methods.

[2] METIS: Graph Partitioning and Fill-reducing Matrix Ordering.

[3] SAAMGE: Smoothed Aggregation Element-based Algebraic Multigrid Hierarchies and Solvers.

[4] D Arnold. An Interior Penalty Finite Element Method with Discontinuous Elements. *SIAM J. Numer. Anal.*, 19(4):742–760, 1982.

[5] A Baker, R Falgout, T Kolev, and U Yang. Multigrid Smoothers for Ultraparallel Computing. *SIAM J. Sci. Comput.*, 33(5):2864–2887, 2011.

[6] Daniele Boffi, Franco Brezzi, and Michel Fortin. *Mixed Finite Element Methods and Applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Berlin, Heidelberg, 2013.

[7] Susanne C Brenner and L Ridgway Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, 3rd edition, 2008.

[8] Marian Brezina, Petr Vaněk, and Panayot S Vassilevski. An improved convergence analysis of smoothed aggregation algebraic multigrid. *Numer. Linear Algebr. with Appl.*, 19(3):441–469, 2012.

[9] Marian Brezina and Panayot S Vassilevski. Smoothed Aggregation Spectral Element Agglomeration AMG: SA-ρAMGe. In Ivan Lirkov, Svetozar Margenov, and Jerzy Waśniewski, editors, *Large-Scale Sci. Comput.*, pages 3–15, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[10] T Chartier, R Falgout, V Henson, J Jones, T Manteuffel, S McCormick, J Ruge, and P Vassilevski. Spectral AMGe (ρAMGe). *SIAM J. Sci. Comput.*, 25(1):1–26, 2003.

[11] Timothy Chartier, Robert Falgout, Van Emden Henson, Jim E Jones, Tom A Manteuffel, John W Ruge, Steve F McCormick, and Panayot S Vassilevski. Spectral Element Agglomerate AMGe. In Olof B Widlund and David E Keyes, editors, *Domain Decompos. Methods Sci. Eng. XVI*, pages 513–521, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

[12] Philippe G Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, reprint edition, 2002.

[13] Jim Douglas and Todd Dupont. Interior Penalty Procedures for Elliptic and Parabolic Galerkin Methods. In R Glowinski and J L Lions, editors, *Comput. Methods Appl. Sci.*, volume 58 of *Lecture Notes in Physics*, pages 207–216, Berlin, Heidelberg, 1976. Springer.

[14] Delyan Kalchev. Adaptive Algebraic Multigrid for Finite Element Elliptic Equations with Random Coefficients. Technical report, LLNL-TR-553254, Lawrence Livermore National Laboratory, Livermore, CA, 2012.

[15] Johannes K Kraus, Panayot S Vassilevski, and Ludmil T Zikatanov. Polynomial of best uniform approximation to $1/x$ and smoothing in two-level methods. *Comput. Methods Appl. Math.*, 12(4):448–468, 2012.

[16] Ilya Lashuk and Panayot S Vassilevski. On some versions of the element agglomeration AMGe method. *Numer. Linear Algebr. with Appl.*, 15(7):595–620, 2008.

[17] R D Lazarov, J E Pasciak, J Schöberl, and P S Vassilevski. Almost optimal interior penalty discontinuous approximations of symmetric elliptic problems on non-matching grids. *Numer. Math.*, 96(2):295–315, dec 2003.

[18] Sergey Nepomnyaschikh. Domain Decomposition Methods. In Johannes Kraus and Ulrich Langer, editors, *Lect. Adv. Comput. Methods Mech.*, volume 1 of *Radon Series on Computational and Applied Mathematics*, pages 89–159. Walter de Gruyter, Berlin, 2007.

[19] P Vanek, M Brezina, and R Tezaur. Two-grid Method for Linear Elasticity on Unstructured Meshes. *SIAM J. Sci. Comput.*, 21(3):900–923, 1999.

[20] Panayot S Vassilevski. Sparse matrix element topology with application to AMG(e) and preconditioning. *Numer. Linear Algebr. with Appl.*, 9(6-7):429–444, 2002.

[21] Panayot S Vassilevski. *Multilevel Block Factorization Preconditioners: Matrix-based Analysis and Algorithms for Solving Finite Element Equations*. Springer, New York, 2008.

[22] M Wheeler. An Elliptic Collocation-Finite Element Method with Interior Penalties. *SIAM J. Numer. Anal.*, 15(1):152–161, 1978.

[23] J Xu. The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids. *Computing*, 56(3):215–235, sep 1996.

Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O. Box 808, L-561, Livermore, CA 94551, USA.

Email address: kalchev1@llnl.gov

Department of Mathematics and Statistics, Portland State University, Portland, OR 97207, USA, and Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O. Box 808, L-561, Livermore, CA 94551, USA.

Email address: panayot@pdx.edu, vassilevski1@llnl.gov