Uniform Subspace Correction Preconditioners for Discontinuous Galerkin Methods with \(hp\)-Refinement

Will Pazner\(^1\) · Tzanio Kolev\(^1\)

Received: 31 August 2020 / Revised: 6 February 2021 / Accepted: 5 April 2021 / Published online: 9 July 2021
© Shanghai University 2021

Abstract
In this paper, we develop subspace correction preconditioners for discontinuous Galerkin (DG) discretizations of elliptic problems with \(hp\)-refinement. These preconditioners are based on the decomposition of the DG finite element space into a conforming subspace, and a set of small nonconforming edge spaces. The conforming subspace is preconditioned using a matrix-free low-order refined technique, which in this work, we extend to the \(hp\)-refinement context using a variational restriction approach. The condition number of the resulting linear system is independent of the granularity of the mesh \(h\), and the degree of the polynomial approximation \(p\). The method is amenable to use with meshes of any degree of irregularity and arbitrary distribution of polynomial degrees. Numerical examples are shown on several test cases involving adaptively and randomly refined meshes, using both the symmetric interior penalty method and the second method of Bassi and Rebay (BR2).

Keywords Discontinuous Galerkin · Preconditioners · Domain decomposition · \(hp\)-refinement

Mathematics Subject Classification 65N55 · 65N30 · 65N22

1 Introduction
High-order discontinuous Galerkin (DG) methods have seen significant recent interest in a wide range of application areas [12, 27, 54]. One advantageous feature of the DG method is its flexibility, allowing for the natural handling of irregular (nonconforming) meshes and variable polynomial degrees, thus making it well-suited for \(hp\)-adaptive refinement [48, 55]. However, one challenging aspect of \(hp\)-refined DG methods is the iterative solution of the resulting ill-conditioned system of linear equations. Many past works have developed and analyzed solvers and preconditioners for DG discretizations of elliptic...
problems, including multigrid methods [29, 33], and domain decomposition methods [2, 9, 25], among others. However, much of the work on solvers for DG has focused on low-order elements, and some of the results hold only in the case of $h$-refinement alone. In the case of general $hp$-refinement, the presence of hanging vertices, highly graded meshes, and variable polynomial degrees makes the construction of efficient preconditioners more challenging.

Efficient solvers for DG discretizations of elliptic problems with $hp$-refinement are relevant to a large number of applications. Examples include diffusion-based preconditioning of high-order transport problems [35] and the pressure Poisson equation arising from projection methods for the incompressible Navier-Stokes equations [31, 50]. The condition number of the resulting system scales like $p^4/h^2$ (cf. [4]), thus necessitating the development of efficient preconditioners for DG discretizations of diffusion problems. In addition, these applications, and others, are highly amenable to adaptive mesh refinement and the use variable polynomial degrees, and so it is important that the preconditioners remain effective on spaces with nonconforming refinements and variable polynomial degrees.

In this work, we take a subspace correction approach, where an effective preconditioner on the conforming subspace of the DG finite element space is combined with a simple smoother, also based on a space decomposition, defined in terms of Gauss-Lobatto nodal points. Similar ideas were developed for low-order DG methods in [29], and for high-order DG methods on conforming meshes in [9, 45]. In the case of conforming meshes or $h$-refinement only, a simple point Jacobi smoother is sufficient to obtain uniform bounds on the preconditioned system. However, in the case of $hp$-refinement, the conforming subspace does not provide as good of an approximation to the DG finite element space (as quantified by a Jackson-type estimate), necessitating more powerful block smoothers, which take the form of block Jacobi methods. These block smoothers correspond to a subspace decomposition that is generated using a simple algorithm, and which depends on the irregularity of the mesh and the distribution of polynomial degrees. The resulting preconditioned system has condition number independent of the mesh size, polynomial degree, and penalty parameter.

Past works on solvers for $hp$-DG methods include multilevel methods and domain decomposition methods. In [4], a class of domain decomposition preconditioners for $hp$-DG discretizations was analyzed, where the condition number of the resulting system is bounded by $\eta p^2 H/h$, where $H$ is the coarse mesh size, and $\eta$ is the DG penalty parameter. Subsequently, in [6], optimal spectral bounds of $p^2 H/(qh)$ were established, where $q$ is the coarse grid polynomial degree. In addition, domain decomposition methods for DG methods on complicated domains featuring small features and microstructures were developed in [3], and non-overlapping additive Schwarz methods for DG discretizations on polytopic meshes were considered in [5]. A multilevel approach related to the current work was developed in [17], where a minimal conforming subspace (corresponding to a low-order $H^1$ finite element space) was used on graded (1-irregular) meshes to obtain uniform bounds on the condition number of the preconditioned system. This work was extended in [16] using a low-order preconditioner that is defined on nested dyadic grids. This is in contrast to the present work, where a low-order preconditioner is formed using a variational restriction strategy. Multigrid methods for DG discretizations were developed in [33]. The extension to meshes with local refinement was considered in [39]. In [8], multigrid algorithms for $hp$-DG methods were considered, and the dependence of the convergence factor on the type of multigrid cycle and number of smoothing steps was studied. Algebraic multigrid methods for DG methods were developed in [13] and [7], and were shown to converge uniformly with respect to the mesh size and polynomial degree.
The structure of this paper is as follows. In Sect. 2, we introduce the symmetric interior penalty DG discretization for the model problem, and enumerate several useful known results. In Sect. 3, we define a novel space decomposition that is analyzed within the abstract framework of additive Schwarz methods. This analysis depends on approximation results for an Oswald-type averaging operator that maps from the DG finite element space to its conforming subspace. An algorithm to generate the subspace decomposition based on the mesh irregularity is described. At the end of Sect. 3, we also discuss a matrix-free preconditioners for the conforming problem based on a low-order refined methodology. Finally, in Sect. 4, we present numerical results on both adaptively and randomly refined meshes, conforming the theoretical properties of the preconditioner. We end with conclusions in Sect. 5.

2 DG Discretization

Consider the model Poisson problem

\[
\begin{cases}
-\Delta u = f & \text{in } \Omega, \\
u = 0 & \text{on } \partial \Omega
\end{cases}
\] (1)

on the spatial domain $\Omega \subseteq \mathbb{R}^2$. Let $\mathcal{T}$ denote a mesh of the domain $\Omega$ consisting of non-overlapping quadrilateral elements denoted by $\kappa$. We will consider the cases of both regular (conforming) meshes (i.e., those without hanging nodes) and irregular (nonconforming) meshes. Let $\widehat{\kappa}$ denote the reference element $\widehat{\kappa} = [0,1]^2$. For given $p$, let $Q_p(\widehat{\kappa})$ denote the space of bivariate polynomials of degree at most $p$ in each variable. For each element $\kappa \in \mathcal{T}$, we associate a mapping $T_\kappa : \widehat{\kappa} \mapsto \kappa$, and a polynomial degree $p_\kappa$. Define the DG finite element space by

\[
V_h = \{ v_h \in L^2(\Omega) : v_h|_{\kappa} \circ T_\kappa \in Q_{p_\kappa}(\widehat{\kappa}) \text{ for all } \kappa \in \mathcal{T} \}. 
\] (2)

In what follows, we will assume that the mesh is quasi-uniform, and that the ratio of polynomial degrees in elements sharing a common edge remains bounded.

Remark 1 In this work, we assume that the mesh $\mathcal{T}$ consists only of quadrilateral elements. Much of the analysis depends on properties of the Gauss-Lobatto points (cf. Sect. 2.2), which are defined for meshes consisting of tensor-product elements. Many of the results are independent of the dimension, and also hold for hexahedral elements. However, the generalization of the preconditioners to meshes consisting of simplex elements remains nontrivial.

Let $\Gamma$ denote the mesh skeleton, given by the set of all mesh faces. Let $e \in \Gamma$ denote an interface bordering elements $\kappa^-$ and $\kappa^+$, and let $n^-_e$ (resp. $n^+_e$) denote the unit vector normal to $e$ pointing outward from $\kappa^-$ (resp. $\kappa^+$). Let $\phi \in V_h$ be given. Then, $\phi^-_e$ (resp. $\phi^+_e$) is used to denote the trace of $\phi$ on $e$ from within $\kappa^-$ (resp. $\kappa^+$). We define the average and jump of $\phi$ at this interface by

\[
\{ \phi \}_e = \frac{1}{2} (\phi^-_e + \phi^+_e), \quad \llbracket \phi \rrbracket_e = \phi^-_e n^-_e + \phi^+_e n^+_e.
\]
Typically, the edge $e$ in question will be clear from the context, and so we will often drop the subscript $e$.

Given these definitions, we discretize (1) using the symmetric interior penalty DG method [10, 11]. The finite element problem is: find $u_h \in V_h$, such that, for all test functions $v_h \in V_h$,

$$\mathcal{A}(u_h, v_h) = \int_{\Omega} f v_h \, dx,$$

where the bilinear form $\mathcal{A}(\cdot, \cdot)$ is defined by

$$\mathcal{A}(u_h, v_h) = \int_{\Omega} \nabla_h u_h \cdot \nabla_h v_h \, dx - \int_{\Gamma} \{\nabla_h u_h\} \cdot \{v_h\} \, ds$$

$$- \int_{\Gamma} \{\nabla_h v_h\} \cdot \{u_h\} \, ds + \int_{\Gamma} \sigma ||u_h|| \cdot ||v_h|| \, ds.$$  \hfill (4)

Here, $\nabla_h \phi$ denotes the broken gradient of $\phi$, defined elementwise over $T$. The parameter $\sigma$ must be chosen sufficiently large to ensure that the resulting discretization is stable [11]. In particular, at each interface, we take $\sigma = \eta p_e^2 / h_e$, where $p_e = \max\{p_{k-1}, p_{k+1}\}$, and $h_e = \min\{h_{k-1}, h_{k+1}\}$, where $h_k$ denotes the mesh size of the element $\kappa$ [55]. The parameter $\eta$ is known as the penalty parameter, and is chosen to satisfy $\eta \geq \eta_0 > 0$ for some $\eta_0$. An explicit formula for $\eta_0$ was given in [49], where it was also noted that the conditioning of the resulting linear system degrades for larger values of $\eta$. One of the goals of this work is to develop preconditioners for this system whose convergence is independent of the choice of $\eta$.

**Remark 2** There are many alternative DG discretizations for elliptic problems, including the local DG (LDG) method [26], the compact DG (CDG) method [46], the method of Bassi and Rebay (BR2) [12], among others. For simplicity, we focus on the symmetric interior penalty DG (SIPDG) method, but the preconditioners developed here are also applicable to these alternative methods. Numerical examples using the BR2 method are presented in Sect. 4.4.

**Remark 3** In what follows, we will use the notation $a \lesssim b$ to mean that $a \leq Cb$, where $C$ is a constant that is independent of the mesh size, the polynomial degree, and the penalty parameter, but which may depend on the level of irregularity of the mesh, and on the ratio of polynomial degrees in neighboring elements. We will write $a \gtrsim b$ to mean $b \lesssim a$, and $a \approx b$ to mean that both $a \lesssim b$ and $b \lesssim a$.

### 2.1 DG Norms and Estimates

We define the mesh-dependent DG norm $\| \cdot \|_{DG}$ by

$$\|v_h\|_{DG}^2 = \|\nabla_h v_h\|_0^2 + \sum_{e \in \Gamma} \|\sigma^{1/2} v_h\|_{0,e}^2.$$  \hfill (5)

The bilinear form $\mathcal{A}(\cdot, \cdot)$ satisfies the following continuity and coercivity bounds [4].

**Lemma 1** For all $v_h, u_h \in V_h$,
We additionally have the following useful eigenvalue estimates [4].

**Lemma 2** For any \( u_h \in V_h \),

\[
\| u_h \|_0^2 \lesssim \mathcal{A}(u_h, v_h) \lesssim \sum_{\kappa \in \mathcal{T}} \frac{p^4_{\kappa}}{h_{\kappa}} \| u_h \|_{0,\kappa}^2.
\]

The following simple result is a slight refinement of Lemma 2.

**Lemma 3** For any \( u_h \in V_h \),

\[
\mathcal{A}(u_h, u_h) \lesssim \sum_{\kappa \in \mathcal{T}} \frac{p^4_{\kappa}}{h_{\kappa}} \| u_h \|_{0,\kappa}^2 + \sum_{e \in \mathcal{E}} \frac{p^2_{e}}{h_{e}} \| \| u_h \| \|_{0,e}^2.
\]

Furthermore, for any \( w_h \in V_h \) with \( \| w_h \| = 0 \),

\[
\mathcal{A}(u_h - w_h, u_h - w_h) \lesssim \sum_{\kappa \in \mathcal{T}} \frac{p^4_{\kappa}}{h_{\kappa}} \| u_h - w_h \|_{0,\kappa}^2 + \sum_{e \in \mathcal{E}} \frac{p^2_{e}}{h_{e}} \| \| u_h \| \|_{0,e}^2.
\]

**Proof** The first statement follows from Poincaré’s inequality, and the second statement follows from the linearity of the jump.

### 2.2 Gauss-Lobatto Basis

A key ingredient in the construction of preconditioners for the DG discretization will be the use of the Gauss-Lobatto nodal basis. Given the polynomial degree \( p_\kappa \), let \( \bar{\xi}_i \in \bar{\mathcal{S}} \), \( 1 \leq i \leq (p + 1)^2 \) denote the tensor-product Gauss-Lobatto points. For each element \( \kappa \), the points \( \xi_{k,i} \) are defined as the image of \( \bar{\xi}_i \) under the mapping \( T_\kappa \), i.e., \( \xi_{k,i} = T_\kappa(\bar{\xi}_i) \). For each such point, we define the basis function \( \phi_{k,i} \) as the unique element of \( V_h \), such that

\[
\phi_{k,i}(\xi_{k',i'}) = \begin{cases} 
1, & \text{if } (k',i') = (k,i), \\
0, & \text{otherwise}.
\end{cases}
\]

The set of functions \( \{ \phi_{k,i} \} \) forms a basis for the space \( V_h \). For a given element \( \kappa \in T \), let \( B(\kappa) \) denote the set of Gauss-Lobatto nodes lying on \( \partial \kappa \), and let \( I(\kappa) \) denote the set of nodes lying in the interior of \( \kappa \). It is well-known that the discrete \( L^2 \) norm associated with the Gauss-Lobatto quadrature nodes and weights is equivalent to the (exactly integrated) \( L^2 \) norm [19, 21, 22]. As a consequence, we have the following closely related result from [9].

**Lemma 4** [9, Lemma 3] Take any \( v_h \in V_h \). Let \( v_{k,i}(x) = v_h(\xi_{k,i})\phi_{k,i}(x) \). Then,

\[
\| v_h \|_0^2 \lesssim \sum_{k,i} \| v_{k,i} \|_0^2.
\]
We can easily show the following result.

**Corollary 1** Let \( v_h \in V_h \), and write \( v_h = \sum_j v_j \), such that for any node \( \xi_{k,j} \), the value \( v_j(\xi_{k,j}) \) is nonzero for at most one function \( v_j \). Then,

\[
\|v_h\|_0^2 \approx \sum_j \|v_j\|_0^2.
\]

**Proof** By Lemma 4,

\[
\sum_j \|v_j\|_0^2 \approx \sum_j \sum_{k,i} \|v_j(\xi_{k,i})\phi_{k,i}\|_0^2 = \sum_{k,i} \|v_{k,i}\|_0^2 \approx \|v_h\|_0^2.
\]

We will also make use of the following trace and inverse trace inequalities from [18].

**Lemma 5** ([18, Lemma 3.2]) Let \( v_h \in V_h \) be given. Then,

\[
\|v_h\|_{0,\partial\kappa}^2 \leq \frac{p_k^2}{h_k} \|v_h\|_{0,k}^2. \tag{8}
\]

Now, suppose that \( v_h \) vanishes at all interior Gauss-Lobatto nodes of an element \( \kappa \), i.e., \( v_h(\xi) = 0 \) for all \( \xi \in \mathcal{I}(\kappa) \). Then,

\[
\|v_h\|_{0,k}^2 \leq \frac{h_k}{p_k^2} \|v_h\|_{0,\partial\kappa}^2. \tag{9}
\]

## 3 Preconditioning and Space Decomposition

### 3.1 Parallel Subspace Corrections

The preconditioner for the DG discretization will be constructed using the framework of parallel subspace corrections (additive Schwarz methods), which has been studied in great detail in numerous works [30, 34, 53, 56–58]. In this section, we briefly describe the abstract framework, and enumerate some results that will be useful in what follows.

Let \( W \) be a vector space with the inner product \( (\cdot, \cdot) \). Let \( A : W \rightarrow W \) be a symmetric positive-definite linear operator, and let \( \mathcal{A}(\cdot, \cdot) \) denote the induced inner product, i.e., \( \mathcal{A}(u, v) = (Au, v) \). The space \( W \) is decomposed into a sum of subspaces, \( W = \sum_{i=0}^{J} W_i \). For each \( i \), define the \( L^2 \) projection \( Q_i \) by

\[
(Q_iw, w_i) = (w, w_i) \quad \text{for all } w_i \in W_i,
\]

and define the elliptic projection \( P_i \) by

\[
\mathcal{A}(P_iw, w_i) = \mathcal{A}(w, w_i) \quad \text{for all } w_i \in W_i.
\]

Let \( A_i \) denote the restriction of \( A \) to \( W_i \). We note the useful identity

\[
A_iP_i = Q_iA. \tag{10}
\]
The preconditioned system $P$ is defined by
\[
P = \sum_{i=0}^{J} P_i = \left( \sum_{i=0}^{J} A_i^{-1} Q_i \right) A.
\]

In some circumstances, the subspace $W_i$ may be sufficiently large that $A_i^{-1}$ (and hence $P_i$) is impractical to compute. In this case, we may replace each $A_i^{-1}$ with an approximation $R_i$ to obtain the preconditioner:
\[
B = \sum_{i=0}^{J} R_i Q_i.
\]

The corresponding approximate projections are denoted by $T_i = R_i Q_i A$, and the precon-}

ditioned operator is written as $T = \sum_{i=0}^{J} T_i = BA$. Our main goal will be to estimate the iterative condition number of the preconditioned operator, defined as the ratio of extremal eigenvalues,
\[
\kappa(T) = \frac{\lambda_{\text{max}}(T)}{\lambda_{\text{min}}(T)}
\]

which will determine the speed of convergence when using the conjugate gradient method. The main tool in our analysis of the subspace correction preconditioner will be the following useful identity (cf. [53, 56, 58]).

**Lemma 6** For any $w \in W$ we have the identity
\[
A(T^{-1} w, w) = \inf_{w_i \in W_i} \sum_{i=0}^{J} A(T^{-1} w_i, w_i),
\]

and, since $P_i$ restricted to the subspace $W_i$ equals the identity operator, we have the special case for exact projections
\[
A(P^{-1} w, w) = \inf_{w_i \in W_i} \sum_{i=0}^{J} A(w_i, w_i).
\]

### 3.2 Conforming and Boundary Subspaces

In this section, we describe the subspace decomposition for the DG finite element space $V_h$. Let $V_C$ denote the conforming subspace of $V_h$, i.e., $V_C = V_h \cap H^1(\Omega)$. Let $V_B$ (where $B$ here stands for boundary) denote the set of functions $v_B$, such that $v_B(\xi_{k,i}) = 0$ for all $\xi_{k,i} \in B(\kappa)$ for all elements $\kappa \in T$. That is to say, a function $v_B \in V_B$ vanishes at every interior Gauss-Lobatto node of every element in the mesh. It is clear that $V_h = V_B + V_C$.

#### 3.2.1 Conforming Approximation

Let $u_h \in V_h$ be given. We are interested in approximating $u_h$ by a conforming function $u_c \in V_C$. We will make use of an interpolation operator $Q_h : V_h \rightarrow V_C$ that is often referred to as the Oswald operator (cf. [18, 44], among others). The operator $Q_h$ is defined as follows. Consider a nodal Gauss-Lobatto basis for the conforming space $V_C$. In the case of the $h$-refinement, the nodal points on a nonconforming edge are chosen...
to be the corresponding Gauss-Lobatto nodes of the coarse element. In the case of the \( p \)-refinement, the nodal points on an element interface are chosen to be the corresponding Gauss-Lobatto nodes of the element with the lower polynomial degree. Note that in the case of the \( hp \)-refinement, the conforming nodal points are no longer a subset of the DG nodal points. Instead, the nodes on an \( hp \)-interface are chosen to be the coarse element Gauss-Lobatto nodes corresponding to the lowest polynomial degree of any element containing the given edge. Then, any conforming function \( u_h \in V_h \) is well-defined given its value at all conforming nodal points \( \xi \). Let \( u_h \in V_h \) be given. We define \( Q_h u_h(\xi) \) to be the average value of \( u_h(\xi) \) over all elements \( \kappa \) containing the node \( \xi \):

\[
Q_h u_h(\xi) = \frac{1}{\text{card}\{\kappa \in \mathcal{T} : \xi \in \kappa\}} \sum_{\kappa \ni \xi} u_h|_{\kappa}(\xi).
\]

If the mesh \( \mathcal{T} \) is conforming, then it is possible to show the following important \( hp \) approximation property of \( Q_h \):

\[
\|v_h - Q_h v_h\|_0^2 \leq C \sum_{e \in \Gamma} h_e^k P_{e,k}^2 \|\nabla v_h\|_{0,e}^2.
\]  

(13)

This result was shown for conforming meshes and uniform polynomial degree in [18]. The case of nonconforming meshes and the uniform polynomial degree was considered in [40]. The case of conforming meshes and the variable polynomial degree was considered in [37]. In addition, similar results were shown for one-irregular meshes with the variable polynomial degree using an auxiliary mesh technique in [38, 59, 60].

However, in the case of general nonconforming meshes with variable polynomial degrees, estimates of the form (13) are not satisfactory and so one of the contributions of this paper is to propose a more precise estimate, see (14) below, that will be critical in the preconditioner analysis.

To see why (13) is not sufficient in the general \( hp \) case, consider the simple mesh shown in Fig. 1. Let \( f \) be a given biquadratic function, and define \( u_h \in V_h \) by pointwise interpolation of \( f \) at the Gauss-Lobatto nodes. Since the left element has degree \( p = 2 \),

![Fig. 1](image_url)

**Fig. 1** Left: example of problematic case for estimates of type (13). A large high-order element \( (p = 2) \) borders \( n \) low-order \( (p = 1) \) elements. The value of any conforming function at the nonconforming interface \( e \) is determined by its values at the two indicated nodes. Right: numerically computed best value of the constants in estimates (13) and (14) for this case.
we have \( u_h|_{e_{k_i}} = f \). On the remaining elements, \( u_h \) is given by piecewise bilinear interpolation of this function, and so the pointwise error will scale as \( 1/n^2 \), where \( n \) is the number of refined elements on the right. As a result, we see that

\[
\|\|u_h\|\|^2_{0,e} \lesssim 1/n^4.
\]

Let \( u_c = Q_h u_h \in V_C \). Notice that if we refine the elements on the right (i.e., increase \( n \)), the error on the interface \( \|u_h - u_c\|_{0,e} \) remains unchanged, because \( u_c \) is always determined by its values at the two black nodes indicated in the diagram. We attempt to estimate the constant in the inequality (13). We have

\[
C \geq p_e^2 \frac{\|u_h - u_c\|_0}{h_e \|\|u_h\|\|_{0,e}} \\
\geq p_e^2 \frac{n^4}{h_e} \|\|u_h - u_c\|_0^2 \\
\geq n^4 \|\|u_h - u_c\|_0^2 \approx n^4,
\]

where the second-to-last step follows from the inverse trace inequality. We see that the constant in this inequality degrades very quickly with the number of nonconforming refinements in the \( hp \) case. This estimate is verified numerically by computing the minimum value of \( C \), such that that the inequality (13) holds for this particular configuration. The values are shown in Fig. 1, illustrating the degradation of the constant \( C \) with increasing refinements.

To address this issue, we bound the difference \( \|u_h - u_c\|_0^2 \) by both \( ||\nabla u_h||_0^2 \) and \( \sum_{e \in I} h_e^2 ||u_h||_{0,e}^2 \). The numerical computations in Fig. 1 indicate that the resulting modified estimates remain constant with increasing refinements. This is confirmed by the following lemma, which is closely related to the Jackson-type estimates of [17].

**Lemma 7** Let \( u_h \in V_h \) and \( u_c = Q_h u_h \in V_C \). Then,

\[
\|u_h - u_c\|_0^2 \lesssim \sum_{e \in I} h_e^2 \|\nabla u_h\|_{0,e}^2 + \sum_{e \in I} \|u_h\|_{0,e}^2.
\]

**Proof** Let \( u_h \in V_h \) be given. We construct a function \( \tilde{u} \in V_h \) element-by-element as follows. If a nodal point \( \xi_{k,i,j} \) does not lie on a nonconforming edge (i.e., it is an interior node, \( \xi_{k,j} \in \mathcal{I}(\kappa) \), or it lies on a conforming edge), then we set \( \tilde{u}(\xi_{k,j}) = u_h(\xi_{k,j}) \). It remains to define \( \tilde{u} \) on the nonconforming edges of the mesh. Let \( e \) denote a nonconforming edge, bordering coarse element \( \kappa_0 \), and fine elements \( \kappa_1, \ldots, \kappa_n \). Let \( p_e = \min(p_{\kappa_i}) \), and let \( \xi_j \) denote the \( p_e + 1 \) Gauss-Lobatto nodes on \( e \) (these are the conforming nodal points defined at the beginning of Sect. 3.2.1). Then, let \( \tilde{u}_{k_{i,e}} \) be chosen to be the degree-\( p_e \) polynomial that interpolates \( u_h|_{\kappa_0} \) at the points \( \xi_j \). For \( i > 0 \), if \( p_{\kappa_i} = p_e \), then \( \tilde{u}|_{k_{i,e}} = u_h|_{k_{i,e}} \). However, if \( p_{\kappa_i} > p_e \), we choose \( \tilde{u}|_{k_{i,e}} \) to be a degree-\( p_e \) interpolant as follows. Note that fewer than \( p_e + 1 \) of the points \( \xi_j \) lie within \( e \cap \kappa_i \). We select interpolation points \( \xi_{ij} \) that consist of those nodal points \( \xi_j \) that lie within \( e \cap \kappa_i \), supplemented with additional Gauss-Lobatto points to obtain \( p_e + 1 \) distinct points lying within \( e \cap \kappa_i \). Then, \( \tilde{u}|_{k_{i,e}} \) is chosen to interpolate \( u_h|_{k_{i}} \) at the \( p_e + 1 \) points \( \xi_{ij} \).

By this definition, we have \( \tilde{u}|_e(\xi) = u|_e(\xi) \) for all conforming nodal points \( \xi \). Consequently, \( Q_h \tilde{u} = Q_h u_h = u_c \). However, for any nonconforming interface, \( \tilde{u}|_e \) has degree no
higher than $p_e$. Therefore, we can apply the arguments of [37, Proposition 5.2] and [40, Theorem 2.3] to show that

$$\|\tilde{u} - u_c\|_0^2 \lesssim \sum_{e \in E} \frac{h_e}{p_e^2} \|\tilde{u}\|_{0,e}^2.$$ 

In addition, we have

$$\|u_h - u_c\|_0^2 \lesssim \|u_h - \tilde{u}\|_0^2 + \|\tilde{u} - u_c\|_0^2 \lesssim \|u_h - \tilde{u}\|_0^2 + \sum_{e \in E} \frac{h_e}{p_e^2} \|\tilde{u}\|_{0,e}^2.$$ 

By accuracy of the polynomial interpolant

$$\|u_h - \tilde{u}\|_0^2 = \sum_k \|u_h - \tilde{u}\|_{0,k}^2 \lesssim \sum_k \frac{h_k^2}{p_{\kappa}^4} \|\nabla u_h\|_{0,k}^2,$$

where $p_{\kappa}$ is the minimum polynomial degree of all elements $\kappa$ neighboring $\kappa$ (note that we make that assumption that the ratio of polynomial degrees on neighboring elements is bounded, and so $p_{\kappa'} \approx p_{\kappa}$).

It remains to estimate the term $\sum_{e \in E} \frac{h_e}{p_e} \|\tilde{u}\|_{0,e}$. On a given edge $e$, we have

$$\|\tilde{u}\|_{0,e} = \|\tilde{u}\| - \|u_h\| + \|u_h\|_{0,e} = \|\tilde{u}^- - \tilde{u}^+ - u_h^- + u_h^+ + u_h^- - u_h^+\|_{0,e}$$

$$\leq \|\tilde{u}^- - u_h^-\|_{0,e} + \|\tilde{u}^- - u_h^+\|_{0,e} + \|\tilde{u}^+ - u_h^+\|_{0,e} + \|\tilde{u}^+ - u_h^-\|_{0,e} + \|\tilde{u}^- - u_h^-\|_{0,e}.$$ 

Again, using accuracy of the interpolant and the trace inequality (8), we obtain

$$\frac{h_e}{p_e^2} \|\tilde{u}^- - u_h^+\|_{0,e} \lesssim \|\tilde{u}^- - u_h^+\|_{0,k}^2 \lesssim \frac{h_e^2}{p_e^2} \|\nabla u_h\|,$$

and so

$$\sum_{e \in E} \frac{h_e}{p_e} \|\tilde{u}\|_{0,e} \lesssim \sum_k \frac{h_k^2}{p_{\kappa}^4} \|\nabla u_h\|_{0,k}^2 + \sum_{e \in E} \frac{h_e}{p_e} \|\tilde{u}\|_{0,e}^2,$$

from which the conclusion (14) follows.

Now, we consider some special cases in which the stronger estimates of the form (13) hold. The following lemmas are generalizations of the $hp$ estimates from [18] to the case of nonconforming meshes. Let $v$ denote a vertex of the mesh $T$, and let $K_v$ denote the set of all elements $\kappa \in T$ containing $v$ as a vertex. A vertex is called hanging if it is contained in an element of which it is not a vertex (i.e., $v \in \kappa$, where $\kappa \notin K_v$). A vertex is called regular if it is not hanging. That is, a vertex $v$ is regular if for each $\kappa$, such that $v \in \kappa$, we have $\kappa \in K_v$. For a given vertex $v$, for each element $\kappa \in K_v$, there exists a Gauss-Lobatto node $\xi_{v,\kappa}$ that is coincident with $v$. Let $\Xi_v$ denote the set of all such Gauss-Lobatto nodes coincident with $v$.

**Lemma 8** Let $u_h \in V_h$ be given and let $v_b = u_h - Q_h u_h$. Let $v$ be a regular vertex. Then,
\[
\sum_{\kappa \in \mathcal{K}_T} \sum_{\xi_{\kappa,i} \in \Xi} \| v_b(\xi_{\kappa,i}) \phi_{\kappa,i} \|_{0,\kappa}^2 \lesssim \sum_{e \in V} \frac{h_e}{p_e^2} \| u_h \|_{0,e}^2.
\]

**Proof** Applying the inverse trace inequality (9) twice, we obtain
\[
\| v_b(\xi_{\kappa,i}) \phi_{\kappa,i} \|_{0,\kappa}^2 \lesssim \frac{h_e^2}{p_e^2} \| v_b(\xi_{\kappa,i}) \phi_{\kappa,i} \|_{0,v}^2.
\]

Recall that \( Q_b u_h \) is defined by
\[
Q_b u_h(v) = \frac{1}{\text{card}(\mathcal{K}_v)} \sum_{\kappa \in \mathcal{K}_v} u_h|_{\kappa'}(v),
\]
and so
\[
v_b|_{\kappa}(v) = u_h|_{\kappa}(v) - \frac{1}{\text{card}(\mathcal{K}_v)} \sum_{\kappa \in \mathcal{K}_v} u_h|_{\kappa'}(v),
\]
which can be written as, for appropriate choice of coefficients \( \alpha_e \),
\[
v_b|_{\kappa}(v) = \sum_{e \ni \kappa} \alpha_e \| u_h \|_e.
\]
By the trace inequality (8), we have
\[
\| \| u_h \|_e \|_{0,v}^2 \lesssim \frac{p_e^2}{h_e} \| \| u_h \|_e \|_{0,e}^2,
\]
and the desired result follows.

**Lemma 9** Let \( e \) be a conforming edge (i.e., no hanging vertex lies in the interior of \( e \), denoted by \( e \overset{\circ}{\ni} \)), and suppose that \( p_{\kappa'} = p_{\kappa} \), where \( \kappa' \) and \( \kappa \) are the two elements containing \( e \). Let \( u_h \in V_h \) be given and let \( v_b = u_h - Q_b u_h \). Then,
\[
\sum_{\kappa \in \{\kappa'\}} \sum_{\xi_{\kappa,i} \in e} \| v_b(\xi_{\kappa,i}) \phi_{\kappa,i} \|_{0,\kappa}^2 \lesssim \frac{h_e}{p_e^2} \| \| u_h \|_e \|_{0,e}^2.
\]

**Proof** Let \( \xi_i \in e \) denote a Gauss-Lobatto node lying in the interior of \( e \), and let \( \phi^\pm \) denote the corresponding basis functions. Then, \( Q_b u_h(\xi_i) = \frac{1}{2}(u^-_h(\xi_i) + u^+_h(\xi_i)) \) and so \( v^\pm_b(\xi_i) = \pm \frac{1}{2} \| u_h \|_e(\xi_i) \). By Lemma 4 and the inverse trace inequality,
\[
\sum_{\kappa \in \{\kappa'\}} \sum_{\xi_{\kappa,i} \in e} \| v_b(\xi_{\kappa,i}) \phi_{\kappa,i} \|_{0,\kappa}^2 \approx \sum_{\kappa \in \{\kappa'\}} \sum_{\xi_{\kappa,i} \in e} \| u_h \|_e(\xi_{\kappa,i}) \phi_{\kappa,i} \|_{0,\kappa}^2 \\
\approx \sum_{\kappa \in \{\kappa'\}} \sum_{\xi_{\kappa,i} \in e} \frac{h_e}{p_e^2} \| u_h \|_e(\xi_{\kappa,i}) \phi_{\kappa,i} \|_{0,e}^2 \approx \frac{h_e}{p_e^2} \| \| u_h \|_e \|_{0,e}^2.
\]

We will call a vertex \( v \) of the mesh \( T \) an \( hp \)-vertex if \( v \) is a hanging vertex, and the elements containing \( v \) do not all have the same polynomial degree. In the case that an edge does not contain any \( hp \)-vertices, we can apply the result of Karakashian and Pascal [40] to obtain the following result.
Lemma 10 Let \( e \) denote an edge that does not contain any hp-vertices. Let \( u_h \in V_h \) be given and let \( v_b = u_h - Q_h u_h \). Then,

\[
\sum_{\kappa \in e} \| v_b \|_{0,k}^2 \lesssim \frac{h_e}{P_e^2} \| \| u_h \| \|_{0,e}^2.
\]

Remark 4 In practice, we observe that certain nonconforming interfaces not included in the above three cases also satisfy estimate (13). In particular, this is observed for interfaces for which the coarse element has polynomial degree no higher than any of the fine elements containing the given face. This is corroborated by numerical examples shown in Sect. 4; however, it is not implied as a consequence of the above estimates.

3.3 Preconditioning

Let \( P_C \) denote elliptic projection onto \( V_C \), and likewise let \( P_B \) denote elliptic projection onto \( V_B \). We will show that the additive Schwarz system

\[
P = P_B + P_C
\]

is uniformly well-conditioned with respect to \( h, p, \) and \( \eta \). First, we show that the decomposition \( V_h = V_B + V_C \) is stable.

Lemma 11 Let \( u_h \in V_h \), and write \( u_h = u_b + u_c \), where \( u_c = Q_h u_h \). Then,

\[
A(u_b, u_b) \lesssim A(u_h, u_h)
\]

and

\[
A(u_c, u_c) \lesssim A(u_h, u_h).
\]

Proof By the estimates in Lemmas 7 and 3, we have

\[
A(u_h - u_c, u_h - u_c) \lesssim \sum_{\kappa \in \Gamma} \frac{p^4_{\Gamma}}{h^2_{\kappa}} \| u_h - Q_h u_h \|_{0,k}^2 + \sum_{e \in \Gamma} \eta \frac{p^2_{\Gamma}}{h_{\kappa}} \| \| u_h \| \|_{0,e}^2
\]

\[
\lesssim \| \nabla_h u_h \|_{0,k}^2 + \sum_{e \in \Gamma} \eta \frac{h}{p^2} \| \| u_h \| \|_{0,e}^2
\]

\[
\approx \| u_h \|_{DG}^2 \lesssim A(u_h, u_h),
\]

proving the first assertion. The second assertion follows by writing \( u_c = (u_h - u_c) + u_c \) and using the triangle inequality.

Theorem 1 The iterative condition number of \( P \) satisfies

\[
\kappa(P) = \lambda_{\text{max}}(P) / \lambda_{\text{min}}(P) \approx 1.
\]

Proof Since both \( P_B \) and \( P_C \) are projections, we have \( \lambda_{\text{max}}(P) \leq 2 \). To establish the bound for the minimum eigenvalue of \( P \), we use the identity (12) from Lemma 6:

\[
A(P^{-1} v_h, v_h) = \inf_{v_h = v_b + v_c} \left( A(v_b, v_h) + A(v_c, v_c) \right).
\]
Lemma 11 shows that $V_B + V_C$ is a stable decomposition of $V_h$, i.e., for all $v_h \in V_h$, there exist $v_b \in V_B, v_c \in V_C$, such that $v_h = v_b + v_c$ and

$$\mathcal{A}(v_b, v_b) + \mathcal{A}(v_c, v_c) \lesssim \mathcal{A}(v_h, v_h).$$

Consequently, we have

$$\mathcal{A}(P^{-1}v_h, v_h) \lesssim \mathcal{A}(v_h, v_h),$$

and the desired result follows.

The spaces $V_B$ and $V_C$ are themselves large, and the computation of $P_B$ and $P_C$ requires the inversion of the bilinear forms obtained by restricting $\mathcal{A}$ to the subspaces, which are denoted by $\mathcal{A}_B$ and $\mathcal{A}_C$, respectively. This cost is clearly prohibitive, and so we seek to replace $P_B$ and $P_C$ with approximations $T_B$ and $T_C$. On the conforming space $V_C$, note that $\mathcal{A}_C$ corresponds to a standard $H^1$-conforming discretization, and so we may replace $\mathcal{A}_C^{-1}$ with a uniform preconditioner for the conforming problem. In this work, we use a low-order refined matrix-free preconditioner, which is described in greater detail in Sect. 3.4. In principle, any uniform preconditioner for $\mathcal{A}_C$ may be used, and in the remainder of this section we will assume that $\mathcal{A}(T_C^{-1}u_c, u_c) \approx \mathcal{A}(u_c, u_c)$ for all $u_c \in V_C$.

The projection $P_B$ onto the boundary space $V_B$ is approximated using a further space decomposition. We decompose the space $V_B$ as the sum of yet-to-be-defined subspaces:

$$V_B = V_E + V_J, \quad V_E = \sum_e V_{e}, \quad V_J = \sum_j V_{j},$$

and define the corresponding approximate projection $T_B$ by

$$T_B = T_E + T_J, \quad T_E = \sum_e P_{e}, \quad T_J = \sum_j P_{j}.$$

We begin by defining the space $V_J$. The subscript $J$ is used to indicate that the approximate projection $T_J$ onto $V_J$ will be a simple point Jacobi method. For this reason, it is advantageous to choose $V_J$ to be as large as possible while still obtaining a stable decomposition. Let $V_J$ consist of any degree of freedom coincident with a regular mesh vertex, lying on the interior of a conforming edge with uniform polynomial degrees, or lying on an edge which contains no $hp$-vertices (see the definitions in Sect. 3.2.1). In other words, $V_J$ is designed to consist of those degrees of freedom for which the stronger approximation estimate (13) holds for the Oswald operator $Q_h$. As noted in Remark 4, the conditions above are not exhaustive, and in practice $V_J$ can be chosen to also include those degrees of freedom lying on edges for which the coarse element has polynomial degree no higher than any of the neighboring fine elements. For the $j$th degree of freedom in the space $V_J$, let $V_j$ denote the corresponding one-dimensional subspace, so that $V_J$ can be written as the direct sum $V_J = \sum_j V_j$.

Now, we define the spaces $V_e$. For every edge $e \in \Gamma$ define a (potentially empty) subspace $V_e$ according to the following process:

- $V_e^{(0)}$ is given as the span of all basis functions $\phi_{x,i} \notin V_J$ whose corresponding Gauss-Lobatto node $\xi_{x,i}$ lies on the edge $e$.
- $V_e^{(i+1)}$ is defined as the span of all basis functions $\phi_{x,i} \in V_B, \phi_{x,i} \notin V_J$, such that $\|\phi_{x,i}\| \cdot \|v\| \neq 0$ for some $v \in V_e^{(i)}$. 

 Springer
• The iteration terminates when $V_e^{(i+1)} = V_e^{(i)}$, at which point we set $V_e = V_e^{(i)}$.

An example of this process is illustrated in Fig. 2. For any pair of spaces $V_e$ and $V_e'$, either $V_e = V_e'$ or $V_e \cap V_e' = \{0\}$, and in such a case we may simply omit one of the two spaces. Therefore, without loss of generality, we may assume that the spaces $V_e$ are disjoint. Furthermore, $V_e \cap V_j = \{0\}$ for all $e$ by construction, and so the decomposition $V_B = \sum_e V_e + \sum_j V_j$ is in fact a direct sum.

**Remark 5** The process described above is guaranteed to terminate; however, the spaces $V_e$ could, in principle, be quite large. In practice, this occurs only in pathological cases, and in realistic cases these spaces remain relatively small. Furthermore, we observe in the numerical results in Sect. 4.3 that for 1-irregular meshes, the dimension of the spaces $V_e$ remain bounded with increasing refinements.

As a consequence of this choice of subspaces, we have the following result concerning the stability of the decomposition.

**Lemma 12** Let $u_h \in V_h$ be given, and let $v_b = u_h - Q_h u_h \in V_B$. Let $v_b = \sum_e v_e + \sum_j v_j$ be the unique representation of $v_b$ in the subspace decomposition (15). Then,

$$\sum_{e} \frac{p^4}{h^2} \sum_j \|v_j\|_{0,k}^2 \lesssim \sum_{e} \sum_{e' \in \partial e} \frac{p^2}{h_{e'}} \|\|u_h\|\|_{0,e'}^2, \quad (17)$$

and

$$\sum_{e' \in \partial e} \frac{p^2}{h_{e'}} \left( \sum_e \|v_e\|_{0,e'}^2 + \sum_j \|v_j\|_{0,e'}^2 \right) \lesssim \sum_{e' \in \partial e} \frac{p^2}{h_{e'}} \|\|u_h\|\|_{0,e'}^2.$$

**Proof** The first assertion follows from the definition of the spaces $V_j$ and by Lemmas 8, 9, and 10.

We now prove the second assertion. By the trace inequality (8),

![Fig. 2 Example of the creation of the subspaces $V_e$ on a non-conforming mesh with variable polynomial degrees. The subspaces are spanned by the basis functions corresponding to the indicated nodes. Note that nodes that are coincident with regular mesh vertices are omitted.](image-url)
\[
\sum_{e' \in \partial \Omega} p_{e'}^2 \sum_j \| v_j \|_{0,e'}^2 \lesssim \sum_j \frac{p_j^4}{h_J^2} \| v_j \|_{0,k}^2.
\]  
(18)

Using (17),

\[
\sum_j \sum_k \frac{p_j^4}{h_J^2} \| v_j \|_{0,k}^2 \lesssim \sum_j \sum_{e' \in \partial \Omega} \frac{p_{e'}^2}{h_{e'}^2} \| u_{e'} \|_{0,e'}^2.
\]  
(19)

Note that \( \| v_j \|_{0,e'}^2 = \| \| v_j \|_{0,e'}^2 \), and so combining (18) and (19) gives

\[
\sum_{e' \in \partial \Omega} \frac{p_{e'}^2}{h_{e'}^2} \sum_j \| v_j \|_{0,e'}^2 \lesssim \sum_j \sum_{e' \in \partial \Omega} \frac{p_{e'}^2}{h_{e'}^2} \| u_{e'} \|_{0,e'}^2.
\]

Let \( v_j = \sum_j v_j \) and \( v_E = \sum_e v_e \). Note that as a consequence of Corollary 1, we have

\[
\sum_{e' \in \Omega} \frac{p_{e'}^2}{h_{e'}^2} \| v_E \|_{0,e'}^2 \lesssim \sum_{e' \in \Omega} \frac{p_{e'}^2}{h_{e'}^2} \| u_{e'} \|_{0,e'}^2.
\]

By the triangle inequality, writing \( v_E = v_b - v_j \) and noting that \( \| v_b \| = \| u_h \| \),

\[
\sum_{e' \in \Omega} \frac{p_{e'}^2}{h_{e'}^2} \| v_E \|_{0,e'}^2 \lesssim \sum_{e' \in \Omega} \frac{p_{e'}^2}{h_{e'}^2} \left( \| v_b \|_{0,e'}^2 + \| v_j \|_{0,e'}^2 \right) \lesssim \sum_{e' \in \Omega} \frac{p_{e'}^2}{h_{e'}^2} \| u_E \|_{0,e'}^2.
\]

Furthermore, by definition of the subspaces \( V_e \), we have, for \( e_1 \neq e_2 \), \( \| v_{e_1} \| \cdot \| v_{e_2} \| = 0 \). Therefore,

\[
\sum_{e' \in \Omega} \frac{p_{e'}^2}{h_{e'}^2} \sum_{e} \| v_E \|_{0,e}^2 = \sum_{e} \frac{p_{e'}^2}{h_{e'}^2} \| v_E \|_{0,e'}^2,
\]

and the second statement follows.

We also have the following lower bounds on \( T_{E}^{-1} \) and \( T_{J}^{-1} \).

**Lemma 13** For all \( v_E \in V_E \),

\[
\mathcal{A}(T_{E}^{-1} v_E, v_E) \gtrsim \mathcal{A}(v_E, v_E)
\]  
(20)

and

\[
\mathcal{A}(T_{J}^{-1} v_j, v_j) \gtrsim \mathcal{A}(v_j, v_j) \quad \text{for all } v_j \in V_J.
\]  
(21)

**Proof** To prove (20), we use the finite overlap property of the spaces \( V_e \). For each subspace \( V_e \), define the subdomain \( \Omega_e \subseteq \Omega \) as the union of all elements \( \kappa \in \mathcal{T} \), such that \( v_e |_{\kappa} \neq 0 \) for some \( v_e \in V_e \). Each element \( \kappa \) is contained in a number of subdomains \( \Omega_e \) bounded by the number of edges of \( \kappa \).

We then have

\[
\mathcal{A}(v_E, v_E) \approx \| v_E \|_{\text{DG}}^2 = \sum_{\kappa \in \mathcal{T}} \| \nabla v_E \|_{0,\kappa}^2 + \sum_{e' \in \Omega} \| \sigma^{1/2} [v_E] \|_{0,e'}^2.
\]

Note that \( \| \nabla v_e \|_{0,\kappa}^2 = 0 \) if \( \kappa \not\in \Omega_e \), and so by the finite overlap property,
\[ \| \nabla v_e \|_{0,k}^2 = \left\| \sum_{e} \nabla v_e \right\|_{0,k}^2 \lesssim \sum_{e} \|\nabla v_e\|_{0,k}^2. \]

and so
\[ \sum_{x \in T} \| \nabla v_E \|_{0,k}^2 \lesssim \sum_{x \in T} \sum_{e} \|\nabla v_e\|_{0,k}^2. \]

In addition, since \( \| v_{e_1} \| \cdot \| v_{e_2} \| = 0 \) for \( e_1 \neq e_2 \), we have
\[ \sum_{e' \in F} \| \sigma^{1/2} [v_{E}] \|_{0,e'}^2 = \sum_{e' \in F} \sum_{e} \|\sigma^{1/2} [v_e]\|_{0,e'}^2. \]

We conclude that
\[ \mathcal{A}(v_E, v_E) \lesssim \sum_{e} \left( \sum_{x \in T} \| \nabla v_e \|_{0,k}^2 + \sum_{e' \in F} \| \sigma^{1/2} [v_e] \|_{0,e'}^2 \right) = \sum_{e} \| v_e \|_{DG}^2 \lesssim \sum_{e} \mathcal{A}(v_e, v_e), \]

and so, by Lemma 6,
\[ \mathcal{A}(T_E^{-1} v_E, v_E) = \sum_{e} \mathcal{A}(v_e, v_e) \gtrsim \mathcal{A}(v_E, v_E). \]

To prove (21), we use an argument from [9]. Write \( v_j = \sum_j v_j \). From the eigenvalue estimate (Lemma 2), we have
\[ \mathcal{A}(v_j, v_j) \lesssim \sum_{x \in T} \frac{p_x^4}{h_x^2} \| v_j \|_{0,k}^2 \approx \sum_{x \in T} \frac{p_x^4}{h_x^2} \sum_j \| v_j \|_{0,k}^2. \]

Using the inverse trace inequality and noting that \( \| v_j \|_{0,e}^2 = \| [v_j] \|_{0,e'}^2 \), we see
\[ \sum_{x \in T} \frac{p_x^4}{h_x^2} \sum_j \| v_j \|_{0,k}^2 \lesssim \sum_{x \in T} \frac{p_x^2}{h_x^2} \sum_j \| v_j \|_{0,k}^2 = \sum_{x \in T} \frac{p_x^2}{h_x} \sum_j \| [v_j] \|_{0,e}^2 \lesssim \sum_j \| v_j \|_{DG}^2. \]

Then, by Lemma 6
\[ \mathcal{A}(T_j^{-1} v_j, v_j) = \sum_j \mathcal{A}(v_j, v_j) \gtrsim \sum_j \| v_j \|_{DG}^2 \gtrsim \mathcal{A}(v_j, v_j), \]

proving (21).

We are now ready to show that \( T_B \) satisfies the following bounds.

**Lemma 14** For all \( v_h \in V_h \),
\[ \mathcal{A}(T_B^{-1} v_h, v_h) \gtrsim \mathcal{A}(v_h, v_h) \quad \text{(22)} \]

and
\[ \mathcal{A}(T_B^{-1} (v_h - Q_h v_h), v_h - Q_h v_h) \lesssim \mathcal{A}(v_h - Q_h v_h, v_h - Q_h v_h) \quad \text{for all } v_h \in V_h. \quad \text{(23)} \]

**Proof** We begin by proving the lower bound (22). By Lemmas 6 and 13, and the triangle inequality, we see, for \( v_h = v_E + v_J \),

\[ \mathcal{A}(T_B^{-1} v_h, v_h) \gtrsim \mathcal{A}(v_h, v_h). \]
proving (22).

Now, we turn to the upper bound (23). This is equivalent to showing that \( \sum_e V_e + \sum_j V_j \) is a stable decomposition of \( \text{Ran}(I - Q_h) \). Let \( v_h \) be given, and let \( v_b = v_h - Q_h v_h \). By Lemma 3,

\[
\mathcal{A}(T_B^{-1} v_b, v_b) \lesssim \sum_e \left( \sum_k \frac{p_k^4}{h_k^2} \|v_e\|_{0,k}^2 + \sum_{e'} \eta \frac{p_{e'}^2}{h_{e'}} \|\|v_{e'}\|\|_{0,e'}^2 \right) + \sum_j \sum_k \eta \frac{p_k^4}{h_k^2} \|v_j\|_{0,k}^2.
\] (24)

By Lemma 12, we have

\[
\sum_e \sum_{e'} \eta \frac{p_{e'}^2}{h_{e'}} \|\|v_{e'}\|\|_{0,e'}^2 + \sum_j \sum_k \eta \frac{p_k^4}{h_k^2} \|v_j\|_{0,k}^2 \lesssim \sum_{e'} \eta \frac{p_{e'}^2}{h_{e'}} \|\|v_{e'}\|\|_{0,e'}^2.
\] (25)

and by Corollary 1,

\[
\sum_e \sum_k \frac{p_k^4}{h_k^2} \|v_e\|_{0,k}^2 \approx \sum_k \frac{p_k^4}{h_k^2} \|v_b\|_{0,k}^2.
\] (26)

Inserting (25) and (26) into (24), noting that \( (I - Q_h) v_b = v_b \) and applying Lemma 7,

\[
\mathcal{A}(T_B^{-1} v_b, v_b) \lesssim \sum_k \frac{p_k^4}{h_k^2} \|v_b\|_{0,k}^2 + \sum_{e'} \eta \frac{p_{e'}^2}{h_{e'}} \|\|v_{e'}\|\|_{0,e'}^2
\]

\[
\lesssim \sum_k \|\nabla v_b\|_{0,k}^2 + \sum_{e'} \eta \frac{p_{e'}^2}{h_{e'}} \|\|v_{e'}\|\|_{0,e'}^2 \approx \|v_b\|_{DG}^2 \lesssim \mathcal{A}(v_b, v_b),
\]

proving the claim.

Finally, by means of these results, we may prove that the resulting preconditioned system \( T = T_B + T_C \) is uniformly well conditioned with respect to the mesh size, polynomial degree, and penalty parameter \( \eta \).

**Theorem 2** The preconditioned system \( T = T_B + T_C \) is uniformly well-conditioned, that is,

\[
\kappa(T) = \lambda_{\text{max}}(T) / \lambda_{\text{min}}(T) \approx 1,
\]

independent of the mesh size, the polynomial degree, and the penalty parameter \( \eta \).

**Proof** By Lemma 6, we have
\[ A(T^{-1}u_h, u_h) = \inf_{u_h+v_h = u_h} \left( A(T_B^{-1}u_h, u_h) + A(T_C^{-1}u_c, u_c) \right) \]
\[ \leq A(T_B^{-1}(I - Q_h)u_h, (I - Q_h)u_h) + A(T_C^{-1}Q_hu_h, Q_hu_h) \]
\[ \approx A(u_h - Q_hu_h, u_h - Q_hu_h) + A(Q_hu_h, Q_hu_h) \]
\[ \leq A(u_h, u_h), \]

using Lemmas 11 and 14. This establishes the upper bound.

For the lower bound, by Lemma 14 and the triangle inequality, we have
\[ A(T^{-1}u_h, u_h) = \inf_{u_h+v_h = u_h} \left( A(T_B^{-1}u_h, u_h) + A(T_C^{-1}u_c, u_c) \right) \]
\[ \geq \inf_{u_h+v_h = u_h} (A(u_h, u_h) + A(u_c, u_c)) \]
\[ \geq A(u_h, u_h). \]

### 3.4 Matrix-Free Preconditioners for the Conforming Problem

In the above analysis, the approximate projection \( T_C \) onto the conforming subspace \( V_C = V_h \cap H^1(\Omega) \) corresponds to a preconditioned system for the standard \( H^1 \)-conforming finite element problem. As long as this preconditioned system is well conditioned, i.e., it satisfies
\[ A_C(T_C^{-1}u_c, u_c) \approx A_C(u_c, u_c) \]
for all \( u_c \in V_C \), where \( A_c(\cdot, \cdot) \) is the standard \( H^1 \) bilinear form \( A(u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx \), then the result of Theorem 2 holds and the proposed preconditioner is efficient.

However, constructing a good preconditioner for the general \( hp \)-conforming problem is challenging, because assembling the associated matrix can be costly at high polynomial degrees. Therefore, in this section, we propose matrix-free preconditioners for the conforming problem which are much more efficient at higher orders. While such preconditioners have been previously considered in the case of \( h \)-refinement with a fixed \( p \), the extension to the general \( hp \)-refinement case is new and is one of the contributions of this paper.

Our approach is based on a low-order refined methodology [20, 43]. It is well known that a low-order \((p = 1)\) finite element discretization on a Gauss-Lobatto refined mesh is spectrally equivalent to the high-order conforming discretization [19]. This equivalence is often also referred to as the finite element method-spectral element method (FEM-SEM) equivalence [21]. Low-order refined preconditioners with parallel subspace corrections have been used in the context of DG discretizations on conforming meshes in [45]. However, in the case of nonconforming meshes or variable polynomial degrees, the low-order refined meshes do not match at coarse element interfaces, and the refined spaces corresponding to different polynomial degrees are not nested. An illustration of one such mesh is shown in Fig. 3. For these reasons it is not immediately clear how to construct an equivalent low-order refined discretization in either the case of \( p \)-refinement or nonconforming meshes.

In this work, we make use of the variational restriction perspective for nonconforming adaptive mesh refinement [24]. Let \( P : V_C \to V_h \) denote the natural injection, which we call the conforming prolongation operator. Let \( \hat{A} \) denote the matrix corresponding to the bilinear form \( \hat{A}(u, v) = \int_\Omega \nabla_h u \cdot \nabla_h v \, dx \). Note that \( \hat{A} : V_h \to V_h \) is a block diagonal
matrix, since it is defined on the “broken” DG space \( V_h \). Then, the matrix \( A_C : V_C \rightarrow V_C \) corresponding to the conforming bilinear form \( A_C \) is given by the variational restriction:

\[
A_C = P^T \hat{A} P.
\]

Let \( \hat{A}_\kappa \) denote the diagonal block of \( \hat{A} \) corresponding to element \( \kappa \). Let \( \tilde{\kappa} \) denote the low-order refined mesh of element \( \kappa \), defined as the image under the element mapping \( T_\kappa \) of the Cartesian mesh whose vertices are the \( p_\kappa + 1 \) tensor-product Gauss-Lobatto nodes. Then, let \( \tilde{A}_\kappa \) denote the matrix corresponding to the bilinear \((p = 1)\) finite element stiffness matrix on the low-order refined element mesh \( \tilde{\kappa} \). The spectral equivalence of low-order refined discretizations (cf. [19]) implies that \( \tilde{A}_\kappa \) is spectrally equivalent to \( A_\kappa \), that is

\[
u^T \tilde{A}_\kappa u \approx \nu^T A_\kappa u \quad \text{for all } u,
\]

independent of the polynomial degree \( p_\kappa \). Let \( \tilde{A} \) be the block diagonal matrix whose diagonal blocks correspond to the low-order refined elemental matrices \( \tilde{A}_\kappa \). Then, define \( \tilde{A}_C \) by

\[
\tilde{A}_C = P^T \tilde{A} P.
\]

This definition gives us the following simple result.

**Proposition 1** The low-order refined discretization \( \tilde{A}_C \) is spectrally equivalent to the high-order conforming discretization \( A_C \).

**Proof** For any \( u \in V_C \) we have, by (27) and setting \( v = Pu \),

\[
\frac{u^T A_C u}{u^T \tilde{A}_C u} = \frac{u^T P^T \hat{A} P u}{u^T \tilde{A}_C u} = \frac{v^T \hat{A} v}{v^T \tilde{A}_C v} \approx 1.
\]

Fig. 3 Example of high-order mesh \( \mathcal{T} \) and corresponding low-order refined mesh \( \tilde{T} \), illustrating the resulting non-matching, non-nested interfaces

The advantage of the discretization \( \tilde{A}_C \) is that the elemental matrices can be assembled in constant time per degree of freedom, as opposed to the high-order discretization \( A_C \), for which naive implementations require \( \mathcal{O}(p^{2d}) \) operations per degree of freedom (optimized
implementations using sum factorizations can reduce this cost to $O(p^{d+1})$ operations per degree of freedom) [41, 43]. Then, any uniform matrix-based preconditioner for $\widetilde{A}_C$ can be used to precondition $\hat{A}$. In particular, algebraic multigrid methods such as BoomerAMG [36], which require an assembled matrix, can be applied easily. In this work, the approximate projection $\mathcal{P}_C$ is given by approximating $A_C^{-1}$ by one V-cycle of the BoomerAMG preconditioner applied to $\widetilde{A}_C$.

**Remark 6** The low-order refined discretization $\widetilde{A}_C$ is of interest in and of itself. Let $\mathcal{T} = \{\kappa : \kappa \in T\}$ denote the (non-matching) low-order defined mesh. In this context, the elements $\kappa \in \mathcal{T}$ will be referred to as macroelements, which will be refined to obtain the mesh $\mathcal{T}$. Any high-order function $v_h \in V_h$ can be identified with a low-order refined function $\tilde{v}_h$, which is given on each element $\kappa$ by the low-order ($p = 1$) interpolant at the Gauss-Lobatto nodes. Let $\mathcal{T} : v_h \mapsto \tilde{v}_h$ denote this identification, and let $\widetilde{V}_h = \mathcal{T}(V_h)$ denote the image of $\mathcal{T}$, which consists of piecewise $p = 1$ functions defined on the refined mesh $\mathcal{T}$ that are continuous within each macroelement $\kappa$, and potentially discontinuous across macroelements. We define the nonconforming finite element space $\widetilde{V}_C$ by

$$\widetilde{V}_C = \left\{ \tilde{v}_h \in \widetilde{V}_h : \mathcal{T}^{-1}(\tilde{v}_h) \in H^1(\Omega) \right\},$$

that is, $\widetilde{V}_C$ consists of all low-order refined functions $\tilde{v}_h$, whose corresponding high-order function $v_h = \mathcal{T}^{-1}(\tilde{v}_h)$ is conforming. Then, the low-order refined $\tilde{A}_C$ can be seen to correspond to the bilinear form $\mathcal{A} : \widetilde{V}_C \times \widetilde{V}_C \to \mathbb{R}$, $\mathcal{A}(\tilde{u}_h, \tilde{v}_h) = \int_\Omega \nabla \tilde{u}_h \cdot \nabla \tilde{v}_h \, dx$.

It is straightforward to see (using the norm-equivalence of low-order refined functions) that the operator $\mathcal{A}$ is bounded and coercive, with respect to the broken $H^1$ norm, denoted by $\| \cdot \|_{1,h}$. Using techniques similar to those of mortar element methods [14] together with a discrete Poincaré inequality, it is possible to bound the approximation and consistency errors of the discretization to obtain the error estimate:

$$\| \tilde{u}_h - u \|_{1,h} \lesssim h^{1/2} \| u \|_1.$$  

By the nonconforming Aubin-Nitsche lemma (cf. [15]), we can obtain the $L^2$ estimate:

$$\| \tilde{u}_h - u \|_0 \lesssim h \| u \|_2.$$  

These estimates indicate that the discretization $\widetilde{A}_C$ is of limited utility in terms of accuracy of the discrete solution; however, because of Proposition 1, it will be quite useful for preconditioning the high-order problem.

## 4 Numerical Results

### 4.1 Implementation and Algorithmic Details

The algorithms described in this paper have been implemented in the framework for the MFEM finite element library [1, 42]. The main components of the solver are as follows.

(i) Efficient matrix-free evaluation of the high-order DG bilinear form $\mathcal{A}(\cdot, \cdot)$. 

© Springer
(ii) Assembly of the diagonal of the DG matrix (corresponding to the subspaces $V_j$), and assembly of the diagonal blocks corresponding to the subspaces $V_{e}$.

(iii) Assembly of the low-order refined conforming stiffness matrix $A_C$.

(iv) Assembly of the conforming prolongation operator $P$.

(v) Application of a uniform preconditioner (e.g., BoomerAMG) approximating $A_C^{-1}$.

We now consider the number of operations required to perform these operations. In particular, we are interested in the scaling with respect to the polynomial degree. For operations which are local to a given element or edge, the polynomial degree $p$ will be used to refer to $p_e$ or $p_e$, respectively. Using matrix-free sum-factorized operator evaluation, the evaluation of the high-order DG bilinear form $A(\cdot, \cdot)$ can be performed element-by-element, requiring $O(p^{d+1}) = O(p^3)$ operations and constant memory per degree of freedom [41, 43, 47]. In addition, the diagonal of the matrix can be assembled in the same complexity. Since the spaces $V_{e}$ are composed of degrees of freedom lying on edges, the size of $V_{e}$ scales like $O(n_e p_e)$, where $n_e$ is the number of edges that are included in $V_{e}$ through the generating process. As a result, the assembly and inversion of these local blocks can be computed in $O(p^3)$ operations, which is the same scaling as operator evaluation. The low-order refined conforming stiffness matrix has $O(1)$ nonzeros per row and, therefore, can be assembled in constant time per degree of freedom (i.e., $O(p^2)$ operations). The number of nonzeros in the conforming prolongation operator scales as $O(p^2)$, and each nonzero entry can be computed in constant time. Finally, the construction and application of the BoomerAMG preconditioner also requires constant operations per degree of freedom.

In the numerical examples below, we will study the performance of the preconditioners developed in this paper applied to several problems involving nonconforming mesh refinement and variable polynomial degree. We consider the preconditioned system $T = T_B + T_C$, where $T_B$ is defined by (16) and denote the corresponding preconditioner as $B = B_B + B_C$, such that $T = BA$. In addition to $B$, we also consider a simplified preconditioner $\tilde{B} = J_B + B_C$, where $J_B$ corresponds to a simple Jacobi method applied to the subspace $V_B$ (that is, the edge spaces $V_{e}$ are not used in the simplified preconditioner). This preconditioner is known to be uniform for the case of conforming meshes with uniform polynomial degree [9, 45]; however, the failure of estimates of the form (13) to hold in the $hp$-refinement case suggests that this preconditioner will not perform well in situations similar to that shown in Fig. 1. Finally, we will compare our results to one V-cycle of the BoomerAMG algebraic multigrid method with Gauss-Seidel smoothing applied to the DG system. Conjugate gradient iteration counts are reported using a relative tolerance of $10^{-8}$.

In the examples below, the penalty parameter $\eta$ is fixed to be $\eta = 100$, with the exception of Sect. 4.4, in which we vary $\eta$ to study the impact of the penalty parameter on the preconditioner performance.

### 4.2 Adaptive Refinement

We consider two examples of adaptively refined meshes and spaces. The first problem is the standard L-shaped domain test [51, 52]:

$$
\begin{cases}
-\Delta u = 0 & \text{in } \Omega, \\
u = g_D & \text{on } \partial \Omega,
\end{cases}
$$
where the Dirichlet boundary conditions are given by
\[ g_D(x) = r(x)^{2/3} \sin \left( \frac{2\theta(x)}{3} + \frac{\pi}{3} \right), \]
where \( r \) and \( \theta \) denote polar coordinates in \( \mathbb{R}^2 \). The exact solution to this problem has a singular gradient, triggering refinements near the corner of the domain. The mesh, polynomial degrees, and solution after 14 adaptive refinement steps are shown in Fig. 4.

We adaptively refine the mesh and polynomial degrees 20 times, beginning with a coarse mesh with 12 elements. The final mesh is 1-irregular, and consists of 138 elements, with a total of 5 207 degrees of freedom. The conjugate gradient iteration counts for this problem are shown in Fig. 5 and Table 1. We report the iteration counts for the subspace correction preconditioner \( B = B_B + B_C \), where \( T_B \) is defined by (16), and \( T_C \) is given by one V-cycle of BoomerAMG applied to the low-order refined discretization of the conforming problem (Fig. 6). The corresponding low-order refined mesh is illustrated in Fig. 7. In addition, we show the iteration counts for the simplified preconditioner \( \tilde{B} = J_B + B_C \), where \( J_B \) is a point Jacobi method applied to the subspace \( V_B \). This simplified preconditioner is expected to perform well on cases without problematic \( hp \)-interfaces of the type illustrated in Fig. 1. Finally, we also consider one V-cycle of BoomerAMG with Gauss-Seidel smoothing applied to the DG problem. We notice that for the L-shaped domain test case, because of the choice of \( h \)- and \( p \)-refinements, the first 18 refinements do not introduce problematic \( hp \)-interfaces, and both the subspace correction preconditioner \( T \) and the simplified preconditioner \( \tilde{T} \) result in good performance that is essentially independent of refinement level. The last two AMR steps introduce several problematic \( hp \)-interfaces, which result in degraded convergence for \( \tilde{T} \). After only a couple of AMR steps, the BoomerAMG preconditioner applied to this problem results in large iteration counts.

The second case we consider is a problem with an internal layer \([28, 51]\). We solve the problem
\[
\begin{aligned}
-\Delta u &= f \quad \text{in } \Omega, \\
u &= g_D \quad \text{on } \partial \Omega,
\end{aligned}
\]

![Fig. 4 Solution and polynomial degrees for the L-shaped domain problem after 14 adaptive refinement steps](image-url)
Table 1 Convergence results and subspace sizes for the L-shaped domain test problem. \#\(V_e\) indicates the number of nontrivial subspaces \(V_e\) in the space decomposition (15). max dim\((V_e)\) indicates the dimension of the largest such subspace.

| AMR step | # DOFs | Iters. \(B_B + B_C\) | Iters. \(J_B + B_C\) | Iters. Boomer-AMG | \# \(V_e\) | max dim\((V_e)\) |
|----------|--------|----------------------|----------------------|---------------------|--------|----------------|
| 0        | 63     | 24                   | 24                   | 25                  | 0      | –              |
| 1        | 84     | 25                   | 25                   | 28                  | 0      | –              |
| 2        | 121    | 27                   | 27                   | 36                  | 0      | –              |
| 3        | 174    | 30                   | 30                   | 60                  | 0      | –              |
| 4        | 223    | 30                   | 30                   | 72                  | 0      | –              |
| 5        | 268    | 31                   | 31                   | 69                  | 0      | –              |
| 6        | 319    | 32                   | 32                   | 87                  | 0      | –              |
| 7        | 662    | 34                   | 33                   | 184                 | 6      | 12             |
| 8        | 1,014  | 32                   | 32                   | 207                 | 6      | 12             |
| 9        | 1,359  | 33                   | 33                   | 218                 | 6      | 12             |
| 10       | 1,683  | 33                   | 33                   | 226                 | 6      | 12             |
| 11       | 2,025  | 33                   | 33                   | 227                 | 6      | 13             |
| 12       | 2,367  | 33                   | 33                   | 234                 | 6      | 13             |
| 13       | 2,709  | 33                   | 33                   | 223                 | 6      | 13             |
| 14       | 3,042  | 32                   | 32                   | 223                 | 6      | 13             |
| 15       | 3,384  | 32                   | 32                   | 223                 | 6      | 13             |
| 16       | 3,752  | 35                   | 35                   | 209                 | 2      | 13             |
| 17       | 4,098  | 33                   | 33                   | 193                 | 0      | –              |
| 18       | 4,422  | 34                   | 34                   | 194                 | 0      | –              |
| 19       | 4,805  | 33                   | 57                   | 191                 | 2      | 15             |
| 20       | 5,207  | 35                   | 71                   | 198                 | 8      | 25             |

Fig. 5 Conjugate gradient iteration counts for the L-shaped domain (left) and internal layer problem (right) adaptively refined problems. Comparison of subspace correction preconditioner \(B = B_B + B_C\), simplified preconditioner \(B = J_B + T_C\), and BoomerAMG.
where $f$ and $g_D$ are chosen to give the exact solution

$$u(x) = \atan(200(r(x) - 0.7)),$$

where $r$ here denotes the distance from the point $(-0.05, -0.05)$. This problem is characterized by a steep gradient near the circle centered at $(-0.05, -0.05)$ of radius 0.7. The results for this problem are shown in Fig. 5 and Table 2. In contrast to the previous test case, the adaptive refinement procedure in this case results in a large number of difficult $hp$-interfaces, as indicated by the large number of nontrivial $V_e$ subspaces. As a consequence, the simplified preconditioner $\vec{T} = J_B = J_C$ does not perform well for this problem. On the other hand, the subspace correction preconditioner results in iteration counts that remain bounded independent of the refinement level. We note that the number of nontrivial subspaces $V_e$ increases with refinement level, thus increasing the cost of the preconditioner. However, the majority of these subspaces are quite small, and they can be processed independently and in parallel. For example, after 20 AMR steps, there are 121 nontrivial $V_e$ subspaces, of which the majority have dimension less than 10, and only 6 of which have dimension greater than 20.

4.3 Random Refinement

To test the robustness of the new preconditioner, we now consider a sequence of nonconforming refinements made randomly. Starting with an initial mesh, each element is marked for refinement with probability 0.5. It is possible to limit the degree of irregularity of the final mesh (i.e., to ensure that an $\ell$-irregular mesh is obtained for given $\ell$) by propagating certain refinements. We consider both the case of 1-irregular meshes and meshes with no limit on the degree of irregularity. After the final mesh is obtained through this random refinement process, polynomial degrees are randomly assigned to each element. We solve the problem

$$-\nabla \cdot (a \nabla u) = f,$$

where $a(x)$ is a piecewise constant diffusion coefficient, which takes values of 1 and 20 according to a numbering of the elements the initial coarse mesh. The coefficient and an example of a randomly refined mesh are shown in Fig. 8.

We study the convergence of the preconditioner for this problem using a combination of random and uniform refinements, and considering both 1-irregular meshes, and
Table 2  Convergence results and subspace sizes for the internal layer test problem. $\# V_e$ indicates the number of nontrivial subspaces $V_e$ in the space decomposition (15). max dim($V_e$) indicates the dimension of the largest such subspace.

| AMR step | # DOFs | Iters. ($B_B + B_C$) | Iters. ($J_B + B_C$) | Iters. (Boomer-AMG) | $\# V_e$ | max dim($V_e$) |
|----------|--------|----------------------|----------------------|---------------------|----------|----------------|
| 0        | 64     | 21                   | 21                   | 23                  | 0        | –              |
| 1        | 124    | 32                   | 32                   | 38                  | 0        | –              |
| 2        | 196    | 36                   | 36                   | 46                  | 0        | –              |
| 3        | 279    | 38                   | 57                   | 72                  | 13       | 8              |
| 4        | 354    | 44                   | 57                   | 105                 | 13       | 12             |
| 5        | 448    | 46                   | 69                   | 128                 | 22       | 12             |
| 6        | 502    | 43                   | 70                   | 135                 | 22       | 12             |
| 7        | 806    | 48                   | 92                   | 140                 | 36       | 16             |
| 8        | 958    | 47                   | 101                  | 185                 | 39       | 20             |
| 9        | 1229   | 43                   | 137                  | 212                 | 42       | 22             |
| 10       | 1602   | 43                   | 203                  | 255                 | 53       | 24             |
| 11       | 1846   | 40                   | 218                  | 255                 | 58       | 22             |
| 12       | 2086   | 42                   | 205                  | 316                 | 63       | 22             |
| 13       | 2365   | 45                   | 267                  | 364                 | 72       | 28             |
| 14       | 2421   | 47                   | 281                  | 398                 | 72       | 31             |
| 15       | 2547   | 42                   | 209                  | 274                 | 76       | 24             |
| 16       | 2892   | 42                   | 310                  | 323                 | 84       | 27             |
| 17       | 3256   | 49                   | 451                  | 378                 | 93       | 27             |
| 18       | 3805   | 44                   | 279                  | 350                 | 111      | 27             |
| 19       | 4253   | 45                   | 357                  | 391                 | 121      | 33             |
| 20       | 4401   | 45                   | 274                  | 294                 | 121      | 34             |

Fig. 7  Low-order refined meshes for the L-shaped domain and internal layer problems.
arbitrary $\ell'$-irregular meshes. The results are presented in Table 3. To begin, we refine the mesh once randomly, and then twice uniformly. We note that the conjugate gradient iterations remain roughly constant with each uniform refinement. Furthermore, the maximum dimension of the edge subspaces $V_e$ does not increase with uniform refinement. We also consider increasing levels of random refinement. We note that with increased random refinements (and increased irregularity of the mesh), we observe a slight degradation in the quality of the preconditioner. In addition, the dimension of the spaces $V_e$ is seen to grow rapidly with the irregularity of the mesh. However, if we require that the mesh be 1-irregular, then the preconditioner performance and dimension of the spaces $V_e$ remains constant when performing random refinements.
4.4 Dependence on Penalty Parameter

An attractive feature of the preconditioners developed in this work is that the conditioning of the preconditioned system is independent of the value of the penalty parameter $\eta$. In general, larger values of the penalty parameter result in systems that are worse-conditioned, and more difficult to solve using standard preconditioners and multigrid methods [23, 49]. In this section, we numerically study the dependence of the preconditioner on the choice of the penalty parameter. The same mesh is used as in Sect. 4.3, with one level of random refinements. Each element of the mesh is randomly assigned a polynomial degree $5 \leq p_e \leq 9$. The symmetric interior penalty parameter $\eta$ is increased from 10 to 10 000 by factors of 10. The resulting iteration counts are shown in Fig. 9. We note that the preconditioned system $T = T_B + T_C$ remains uniformly well-conditioned, independent of the choice of $\eta$, whereas both the simplified preconditioner $\tilde{B}$ and BoomerAMG result in severely degraded convergence for large values of $\eta$.

We also consider an alternative DG formulation, known as the second method of BR2 [12]. The BR2 method proceeds by defining, for each edge $e \in \Gamma$, a lifting operator $r_e : [L^1(e)]^2 \rightarrow [V_h]^2$ given by

$$
\int_{\Omega} r_e(\varphi) \cdot \tau \, dx = - \int_e \varphi \cdot \{\tau\} \, ds \quad \text{for all } \tau \in [V_h]^2.
$$

(28)

The BR2 bilinear form $A_{BR}$ is obtained by replacing the symmetric interior penalty term $\int_{\Omega} \eta \frac{p_e^2}{h_e} \|u_h\| \cdot \|v_h\| \, ds$ with an alternative stabilization term of the form $\sum_e \int_{\Omega} \eta h_e r_e(\|u_h\|) \cdot r_e(\|v_h\|) \, dx$. This method has the advantage that the penalty parameter $\eta$ can be chosen to be $O(1)$, and the scaling by a factor of $p_e^2/h_e$ is not required. However, the presence of the lifting operators $r_e$ can result in degraded convergence for multigrid methods [32]. It can be seen that the norm $\| \cdot \|_{BR}$ induced by the bilinear form $A_{BR}$ is equivalent to the DG norm $\| \cdot \|_{DG}$, independent of the mesh size and the polynomial degree [45].
Consequently, we expect the preconditioner $B = B_B + B_C$ to result in performance independent of the magnitude of BR2 penalization. This agrees with the numerical results presented in Fig. 9.

5 Conclusions

In this work, we presented new preconditioners for DG methods applied to $hp$-refined meshes. The preconditioners are based on a subspace decomposition, using a coarse space of $H^1$-conforming functions, together with subspaces corresponding to nonconforming interfaces. For the coarse space we proposed a new matrix-free low-order refined preconditioner which is shown to be spectrally equivalent to the high-order conforming problem. The nonconforming interface subspaces are generally small in size, and can be processed independently in parallel. Analysis of the overall preconditioner shows that the condition number of the resulting system is independent of the mesh size, the polynomial degree, and the penalty parameter. Numerical examples are presented on both adaptively refined and randomly refined meshes. Comparisons to alternative preconditioners, including a simplified preconditioner with diagonal correction and an algebraic multigrid preconditioner demonstrate the utility and benefits of the current approach.

Acknowledgements This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 and was partially supported by the LLNL-LDRD Program under Project No. 20-ERD-002 (LLNL-JRNL-814157). 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.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Anderson, R., et al.: MFEM: a modular finite element methods library. Comput. Math. Appl. 81, 42–74 (2021). https://doi.org/10.1016/j.camwa.2020.06.009
2. Antonietti, P.F., Ayuso, B.: Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case. ESAIM: Mathematical Modelling and Numerical Analysis 41(1), 21–54 (2007) https://doi.org/10.1051/m2an:2007006
3. Antonietti, P.F., Giani, S., Houston, P.: Domain decomposition preconditioners for discontinuous Galerkin methods for elliptic problems on complicated domains. J. Sci. Comput. 60(1), 203–227 (2013) https://doi.org/10.1007/s10915-013-9792-y
4. Antonietti, P.F., Houston, P.: A class of domain decomposition preconditioners for $hp$-discontinuous Galerkin finite element methods. J. Sci. Comput. 46(1), 124–149 (2010) https://doi.org/10.1007/s10915-010-9390-1
29. Dobrev, V.A., Lazarov, R.D., Vassilevski, P.S., Zikatanov, L.T.: Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations. Numer. Linear Algebra Appl. 13(9), 753–770 (2006) https://doi.org/10.1002/nla.504
30. Dryja, M., Widlund, O.B.: Some domain decomposition algorithms for elliptic problems. In: Kincaid, D.R., Hayes, L.J. (eds.) Iterative Methods for Large Linear Systems, pp. 273–291. Academic Press, New York (1990). https://doi.org/10.1016/B978-0-12-407475-0.50022-X
31. Fehn, N., Wall, W.A., Kronbichler, M.: On the stability of projection methods for the incompressible Navier-Stokes equations based on high-order discontinuous Galerkin discretizations. J. Comput. Phys. 351, 392–421 (2017) https://doi.org/10.1016/j.jcp.2017.09.031
32. Fortunato, D., Rycroft, C.H., Saye, R.: Efficient operator-coarsening multigrid schemes for local discontinuous Galerkin methods. SIAM J. Sci. Comput. 41(6), A3913–A3937 (2019) https://doi.org/10.1137/18m1206357
33. Gopalakrishnan, J., Kanschat, G.: A multilevel discontinuous Galerkin method. Numer. Math. 95(3), 527–550 (2003) https://doi.org/10.1007/s002110030092
34. Griebel, M., Oswald, P.: On the abstract theory of additive and multiplicative Schwarz algorithms. Numer. Math. 70(2), 163–180 (1995) https://doi.org/10.1007/s002110050115
35. Haut, T.S., Southworth, B.S., Maginot, P.G., Tomov, V.Z.: Diffusion synthetic acceleration preconditioning for discontinuous Galerkin discretizations of SN transport on high-order curved meshes. SIAM J. Sci. Comput. 42(5), B1271–B1301 (2020) https://doi.org/10.1137/19m124993x
36. Houston, P., Schötzau, D., Wihler, T.P.: Energy norm a posteriori error estimation of hp-adaptive discontinuous Galerkin methods for elliptic problems. Math. Models Methods Appl. Sci. 17(01), 33–62 (2007) https://doi.org/10.1142/s0218202507001826
37. Houston, P., Süli, E., Wihler, T.P.: A posteriori error analysis of hp-version discontinuous Galerkin finite-element methods for second-order quasi-linear elliptic PDEs.IMA J. Numer. Anal. 28(2), 245–273 (2007) https://doi.org/10.1093/imanum/drm009
38. Haut, T.S., Southworth, B.S., Maginot, P.G., Tomov, V.Z.: Diffusion synthetic acceleration preconditioning for discontinuous Galerkin discretizations of SN transport on high-order curved meshes. SIAM J. Sci. Comput. 42(5), B1271–B1301 (2020) https://doi.org/10.1137/19m124993x
39. Hansson, V.E., Yang, U.M.: BoomerAMG: a parallel algebraic multigrid solver and preconditioner. Appl. Numer. Math. 41(1), 155–177 (2002) https://doi.org/10.1016/s0168-9274(01)00115-5
40. Kanschat, G.: Multilevel methods for discontinuous Galerkin FEM on locally refined meshes. Comput. Struct. 82(28), 2437–2445 (2004) https://doi.org/10.1016/j.compstruc.2004.04.015
41. Karakashian, O.A., Pascal, F.: A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. SIAM J. Numer. Anal. 41(6), 2374–2399 (2003) https://doi.org/10.1137/s0036142902405217
42. Kansačn, G.: Multilevel methods for discontinuous Galerkin FEM on locally refined meshes. Comput. Struct. 82(28), 2437–2445 (2004) https://doi.org/10.1016/j.compstruc.2004.04.015
43. Orszag, S.A.: Spectral methods for problems in complex geometries. J. Comput. Phys. 37(1), 70–92 (1980) https://doi.org/10.1002/9781-9991(80)90005-4
44. Oswald, P.: On a BPX-preconditioner for PI elements. Computing 51(2), 125–133 (1993) https://doi.org/10.1007/bf02243847
45. Pazner, W.: Efficient low-order refined preconditioners for high-order matrix-free continuous and discontinuous Galerkin methods. J. Sci. Comput. 42(5), A3055–A3083 (2020)
46. Peraire, J., Persson, P.-O.: The compact discontinuous Galerkin (CDG) method for elliptic problems. SIAM J. Sci. Comput. 30(4), 1806–1824 (2008) https://doi.org/10.1137/070685518
47. Pazner, W., Persson, P.-O.: Approximate tensor-product preconditioners for very high order discontinuous Galerkin methods. J. Comput. Phys. 354, 344–369 (2018) https://doi.org/10.1016/j.jcp.2017.10.030
48. Perugia, I., Schötzau, D.: An hp-analysis of the local discontinuous Galerkin method for diffusion problems. J. Sci. Comput. 17(1/2/3/4), 561–571 (2002) https://doi.org/10.1023/a:1015118613130
49. Shahbazi, K.: An explicit expression for the penalty parameter of the interior penalty method. J. Comput. Phys. 205(2), 401–407 (2005) https://doi.org/10.1016/j.jcp.2004.11.017
50. Shahbazi, K., Fischer, P.F., Ethier, C.R.: A high-order discontinuous Galerkin method for the unsteady incompressible Navier-Stokes equations. J. Comput. Phys. 222(1), 391–407 (2007). https://doi.org/10.1016/j.jcp.2006.07.029
51. Sošín, P., Červený, J., Doležel, I.: Arbitrary-level hanging nodes and automatic adaptivity in the hp-FEM. Math. Comput. Simul. 77(1), 117–132 (2008) https://doi.org/10.1016/j.matcom.2007.02.011
52. Szabo, B., Babuška, I.: Wiley Series in Computational Mechanics. Finite Element Analysis. Wiley, Amsterdam (1991)
53. Toselli, A., Widlund, O.B.: Domain Decomposition Methods-Algorithms and Theory. Springer Series in Computational Mathematics. Springer-Verlag, Berlin/Heidelberg (2005). https://doi.org/10.1007/b137868
54. Wang, Z., et al.: High-order CFD methods: current status and perspective. Int. J. Numer. Meth. Fluids 72(8), 811–845 (2013) https://doi.org/10.1002/fld.3767
55. Wihler, T., Frauenfelder, P., Schwab, C.: Exponential convergence of the $hp$-DGFEM for diffusion problems. Comput. Math. Appl. 46(1), 183–205 (2003) https://doi.org/10.1016/s0898-1221(03)90088-5
56. Xu, J.: Iterative methods by space decomposition and subspace correction. SIAM Rev. 34(4), 581–613 (1992) https://doi.org/10.1137/1034116
57. Xu, J.: The method of subspace corrections. J. Comput. Appl. Math. 128(1/2), 335–362 (2001) https://doi.org/10.1016/s0377-0427(00)00518-5
58. Xu, J., Zikatanov, L.: The method of alternating projections and the method of subspace corrections in Hilbert space. J. Am. Math. Soc. 15(03), 573–598 (2002) https://doi.org/10.1090/s0894-0347-02-00398-3
59. Zhu, L., Giani, S., Houston, P., Schötzau, D.: Energy norm a posteriori error estimation for $hp$-adaptive discontinuous Galerkin methods for elliptic problems in three dimensions. Math. Models Methods Appl. Sci. 21(02), 267–306 (2011) https://doi.org/10.1142/s0218202511005052
60. Zhu, L., Schötzau, D.: A robust a posteriori error estimate for $hp$-adaptive DG methods for convection-diffusion equations. IMA J. Numer. Anal. 31(3), 971–1005 (2010) https://doi.org/10.1093/imanum/drp038