COMMUNICATIONS IN COMPUTATIONAL PHYSICS
Vol. **5**, No. 2-4, pp. 296-325

Commun. Comput. Phys.
February 2009

Multilevel Preconditioners for the Interior Penalty Discontinuous Galerkin Method II — Quantitative Studies

Kolja Brix[1], Martin Campos Pinto[2], Wolfgang Dahmen[1,*] and Ralf Massjung[1]

[1] *Institut für Geometrie und Praktische Mathematik, RWTH Aachen, 52056 Aachen, Germany.*

[2] *Université Louis Pasteur, Institut de Recherche Mathématique Avancée, CNRS UMR 7501, 67084 Strasbourg, France.*

Received 28 October 2007; Accepted (in revised version) 29 April 2008

Available online 1 August 2008

Abstract. This paper is concerned with preconditioners for interior penalty discontinuous Galerkin discretizations of second-order elliptic boundary value problems. We extend earlier related results in [7] in the following sense. Several concrete realizations of splitting the nonconforming trial spaces into a conforming and (remaining) nonconforming part are identified and shown to give rise to uniformly bounded condition numbers. These asymptotic results are complemented by numerical tests that shed some light on their respective quantitative behavior.

AMS subject classifications: 65F10, 65N55, 65N30

Key words: Interior penalty method, energy-stable splittings, admissible averaging operators, frames, multilevel Schwarz preconditioners, discontinuous Galerkin methods.

---

1 Introduction

An attractive feature of Discontinuous Galerkin (DG) Finite Element schemes is that this concept offers a unified and versatile discretization platform for various types of partial differential equations. The locality of the trial functions not only supports local mesh refinements but offers also a framework for comfortably varying the order of the discretization. While the error analysis has reached a fairly mature state, less appears to be known about the rigorous foundation of efficient solvers for the linear systems of equations that arise when applying the DG concept to elliptic boundary value problems e.g.

*Corresponding author. *Email addresses:* `brix@igpm.rwth-aachen.de` (K. Brix), `campos@math.u-strasbg.fr` (M. Campos Pinto), `dahmen@igpm.rwth-aachen.de` (W. Dahmen), `massjung@igpm.rwth-aachen.de` (R. Massjung)

 ©2009 Global-Science Press

for locally refined meshes with hanging nodes. In [13] a multigrid scheme was presented and shown to exhibit typical multigrid performance when the solution is sufficiently regular and when the underlying mesh is quasi-uniform. This scheme is extended in [16] to locally refined meshes. While numerical tests indicate that essentially the same efficiency is retained, a theoretical underpinning still seems to be missing. Domain decomposition preconditioners investigated in [1,2], give rise to only moderately growing condition numbers. A two-level scheme in the sense of the "auxiliary space method" (see, e.g., [6,19,24]) has been proposed in [11] and shown to exhibit mesh-independent convergence again on quasi-uniform conforming meshes.

Since the DG concept lends itself to problems whose solutions may exhibit singular behavior we have analyzed in [7] a family of multilevel preconditioners that are shown to give indeed rise to uniformly bounded condition numbers without any additional regularity assumptions and for arbitrary locally refined meshes with hanging nodes (under certain mild grading conditions). While the approach in [7] is primarily based directly on the concept of stable splittings for additive Schwarz schemes [14,18], it can also be interpreted as an "auxiliary space method" in the sense of [6,19,24], a connection that will be detailed below. [7] was primarily concerned with the principal ingredients of a general framework and proposed only a few numerical tests. Aside from commenting on the above mentioned conceptual links the central objective of this paper is to gain additional quantitative information regarding the following issues. A crucial ingredient of our approach is the suitable splitting of the trial space V_h into a conforming and (remaining) nonconforming part, and the key requirements on such splittings where shown to be satisfied in [7] only for the specific case that the conforming part consists of piecewise linear finite elements, representing in some sense the *smallest* conforming subspace contained in V_h. Here we shall consider also the *largest* conforming subspace and compare the performance of the respective preconditioners. Moreover, we shall test the robustness of the preconditioners with respect to local mesh refinements. In this context we explore a strategy for adaptive mesh refinements based on [15].

In the remainder of the introduction, we give the precise formulation of the problem, briefly highlight the typical obstructions encountered with the DG method and relate our approach to earlier more abstract results that offer remedies to such obstructions.

1.1 Problem formulation

For simplicity we shall confine the discussion to second-order elliptic boundary value problems on polygonal domains $\Omega \subset \mathbb{R}^2$. Our model problem then reads:

$$\text{find } u \in H_0^1(\Omega) \text{ such that } \quad a(u,v) := \langle A\nabla u, \nabla v\rangle + \langle bu, v\rangle = \langle f, v\rangle \quad \forall v \in H_0^1(\Omega), \quad (1.1)$$

where $\langle\cdot,\cdot\rangle$ is the canonical L_2-inner product on Ω, A is a (piecewise constant) symmetric positive definite 2×2 matrix, and b a nonnegative (piecewise constant) bounded function on Ω.

For simplicity the piecewise constant nature of the coefficients will always refer to some fixed coarse (conforming) shape regular triangulation $\mathcal{T}^0$ of Ω while our discretization will be based on refinements $\mathcal{T}_h$ of $\mathcal{T}^0$ that are allowed to be local and thus exhibit hanging vertices, see Fig. 1. However, these triangulations will be assumed to have some *grading* property that will be specified later. In particular, the edges of the triangles will contain at most one hanging vertex. By $\mathcal{E}_h$ we denote the edges of $\mathcal{T}_h$ with the convention that whenever an edge e of a triangle contains a hanging vertex, $\mathcal{E}_h$ contains the two halves of e, but not e itself.

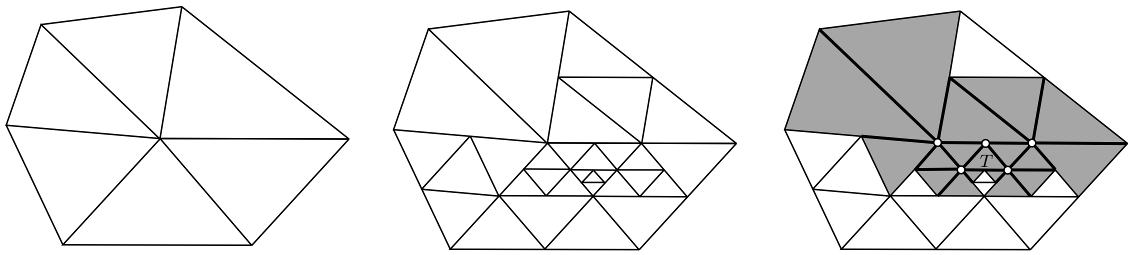

Figure 1: Example of coarse $\mathcal{T}^0$ (left), adaptive $\mathcal{T}_h$ (center) and local sets $\mathcal{N}^*_{h,1}(T)$, $\mathcal{E}^*_h(T)$ and $\mathcal{T}^*_h(T)$ (right).

Remark 1.1. Due to the nonconformity of $\mathcal{T}_h$, every vertex mentioned in the sequel should be thought as *possibly hanging*. Vertices that are not hanging will be referred to as *regular*.

We shall work with trial spaces of the form

$$V_h := \mathbb{P}_k(\mathcal{T}_h) = \{v \in L_2 : v|_T \in \mathbb{P}_k(T), \forall T \in \mathcal{T}_h\}. \tag{1.2}$$

Furthermore, it will be convenient to denote by $v^T := \chi_T v$ the element in V_h that agrees with v on T and vanishes outside T. Note that $k = k(T)$ respectively $h = h(T)$, $T \in \mathcal{T}_h$, should be understood as a piecewise constant function on Ω representing the maximal degree of the elements on the triangle T, respectively the diameter of T.

Given an edge $e = T \cap T'$, shared by two adjacent triangles T, T' we denote by n_e, n'_e the outer normals of T, T', respectively, and for $v \in V_h$ we define as usual by

$$\{v\} = \{v\}_e := \frac{1}{2}\left(v^T|_e + v^{T'}|_e\right), \quad [v] = [v]_e := n_e v^T|_e + n'_e v^{T'}|_e,$$

the averages, respectively jumps of v on $e \in \mathcal{E}_h$. Also by $\langle \cdot,\cdot \rangle_T$ we denote the standard L_2-inner product on T, and accordingly define $a(\cdot,\cdot)_T$. The Symmetric Interior Penalty Galerkin method introduced in the early 1970s reads then as follows, see e.g. [4]:

$$\text{find } u_h \in V_h \quad \text{such that} \quad a_h(u_h, v) = \langle f, v \rangle \quad \forall v \in V_h, \tag{1.3}$$

where the mesh dependent, symmetric bilinear form a_h is given by

$$a_h(v,w) := \sum_{T\in\mathcal{T}_h} a(v,w)_T - \sum_{e\in\mathcal{E}_h}\int_e \big(\{A\nabla w\}\cdot[v] + \{A\nabla v\}\cdot[w]\big) + \sum_{e\in\mathcal{E}_h}\frac{\gamma}{|e|}\int_e [w]\cdot[v].$$

This can be reformulated as an operator equation:

$$\text{find } u_h\in V_h \quad \text{such that} \quad A_h u_h = f_h, \quad \text{where} \quad \langle A_h v,w\rangle = a_h(v,w) \quad \forall v,w\in V_h \tag{1.4}$$

and f_h is the L_2-orthogonal projection of f on V_h. For sufficiently large γ, that may vary from element to element in $\mathcal{T}^0$ depending on the coefficients A and b (see the discussion in [7]), this method is known to be well posed in V_h when equipped with the mesh dependent norm

$$|||v|||_h^2 := \sum_{T\in\mathcal{T}_h} a(v,v)_T + \sum_{e\in\mathcal{E}_h}\frac{\gamma}{|e|}\|[v]\|_{L_2(e)}^2, \tag{1.5}$$

i.e., for any v and w in V_h we have

$$c_a|||v|||_h^2 \le a_h(v,v) \quad \text{and} \quad a_h(v,w) \le C_a|||v|||_h|||w|||_h, \tag{1.6}$$

where c_a and C_a are independent of the mesh sizes h.

As in the conforming case the efficient iterative solution of the linear systems (1.3) is severely hampered by the fact that the condition number $\kappa(\boldsymbol{A}) := \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\|$ of the stiffness matrix $\boldsymbol{A} := (a_h(\phi_i,\phi_k))_{i,k\in\mathcal{I}_h}$ grows like h^{-2} with $h = \inf\{h(T) : T\in\mathcal{T}_h\}$, when $\Phi_h = \{\phi_i : i\in\mathcal{I}_h\}$ is a standard nodal basis of V_h. It is fair to say that preconditioning is a bit more delicate for nonconforming discretizations and it is perhaps instructive to briefly address this issue and recall next some relevant background information.

1.2 Some background

Optimal preconditioners for conforming discretizations – optimal in the sense that they give rise to even uniformly bounded condition numbers – greatly exploit *nestedness* of hierarchies of discretizations, see, e.g., [5,10,14,18]. A very flexible framework hinges on the notion of *stable splittings*. In the present context of the DG method this amounts to looking for a collection $\mathcal{S}_h = \{V_i : i\in\mathcal{I}_h\}$ of subspaces spanning V_h in such a way that

$$c_s|||v|||_h^2 \le \inf_{\substack{v_i\in V_i\\ v=\sum_{i\in\mathcal{I}_h} v_i}}\left\{\sum_{i\in\mathcal{I}_h}|||v_i|||_h^2\right\} \le C_s|||v|||_h^2 \tag{1.7}$$

holds for any $v\in V_h$ with constants independent of h. In this case an optimal preconditioner is obtained e.g. through an additive Schwarz scheme based on the splitting, [14,18]. In the special case where each V_i is spanned by a single function and the collection of these functions forms a (uniformly) stable basis for V_h this amounts to a *change-of-basis* preconditioner whose first forerunner is perhaps the hierarchical basis preconditioner [25], while wavelet preconditioners fall into the same category and are optimal in the above sense. Therefore, one might think of using *multiwavelets* based on discontinuous piecewise polynomials in the DG case. The fact that this does not work is a consequence of the following observation from [7].

Theorem 1.1. *If Ψ_h is a multilevel basis of $V_h := \mathbb{P}_k(\mathcal{T}_h)$ that is (energy-)stable in the sense that*

$$\|\{d_\psi\}_{\psi\in\Psi_h}\|_{\ell_2} \sim \Big|\Big|\Big| \sum_{\psi\in\Psi_h} d_\psi \psi \Big|\Big|\Big|_h \quad \forall d\in\mathbb{R}^{\#(\Psi_h)},$$

then Ψ_h must contain a subset that is a stable basis of the conforming part $V_h \cap H_0^1(\Omega)$.

In particular, this means that a multilevel, stable Ψ_h must contain continuous basis functions at any level, which is for instance not the case with multiwavelets and therefore rules out this simple option. It rather suggests looking for splittings that consist of two parts, namely one working for a conforming part of V_h (which should have multilevel nature) and one that works for the (remaining) nonconforming part.

This has been the viewpoint in [7]. But before taking this up, let us note that in most other cases of nonconforming discretizations, unlike the DG case, the lack of nestedness is the typical obstruction. This has motivated systematic attempts to reduce the task of preconditioning systems that stem from nonconforming discretizations to preconditioning conforming systems through a suitable auxiliary space. In abstract terms this leads to a two-level method, running for instance under the flag of "auxiliary space method", see e.g. [6, 19, 24]. The essence of such techniques can be summarized in abstract terms following [19] which also allows one to tie these concepts into the setting of stable splittings.

To this end, suppose that $\tilde{V}_h \subset H_0^1(\Omega)$ is an auxiliary (conforming) space for which

$$\text{find } \tilde{u}_h \in \tilde{V}_h \quad \text{such that} \quad a(\tilde{u}_h,\tilde{v}) = \langle f,\tilde{v}\rangle \quad \forall \tilde{v}\in\tilde{V}_h \tag{1.8}$$

makes sense, and suppose that the following properties hold: first, setting $\hat{V}_h = V_h + \tilde{V}_h$, there must be two symmetric positive definite bilinear forms $\hat{a}_h,\hat{b}_h : \hat{V}_h \times \hat{V}_h \to \mathbb{R}$ such that $\hat{a}_h$ is a spectrally equivalent extension of both a_h and a, i.e.

$$\hat{a}_h(v,v) \sim a_h(v,v) \quad \forall v\in V_h \quad \text{and} \quad \hat{a}_h(\tilde{v},\tilde{v}) \sim a(\tilde{v},\tilde{v}) \qquad \forall \tilde{v}\in\tilde{V}_h, \tag{1.9}$$

and $\hat{b}_h$ is an auxiliary scalar product (typically defined as an appropriately scaled L_2 inner product) satisfying the inverse estimate

$$\hat{a}_h(v,v) \lesssim \hat{b}_h(v,v) \qquad \forall v\in\hat{V}_h. \tag{1.10}$$

The next ingredients are two linear operators $\tilde{Q}: V_h \to \tilde{V}_h$, $Q:\tilde{V}_h \to V_h$ satisfying Jackson-type direct estimates, namely

$$\hat{b}_h((I-Q)\tilde{v},(I-Q)\tilde{v}) \lesssim \hat{a}_h(\tilde{v},\tilde{v}) \qquad \forall \tilde{v}\in\tilde{V}_h, \tag{1.11}$$

$$\hat{b}_h((I-\tilde{Q})v,(I-\tilde{Q})v) \lesssim \hat{a}_h(v,v) \qquad \forall v\in V_h. \tag{1.12}$$

Now the point is that, whenever (1.9)-(1.12) hold, one has the following norm equivalence [19]

$$\hat{c}a_h(v,v) \le \inf_{\substack{w\in V_h,\tilde{v}\in\tilde{V}_h:\\ v=w+Q\tilde{v}}} \left\{\hat{b}_h(w,w)+a(\tilde{v},\tilde{v})\right\} \le \hat{C}a_h(v,v) \qquad \forall v\in V_h. \tag{1.13}$$

To formulate the main consequence of this fact, let $\tilde{\boldsymbol{A}}$, $\boldsymbol{A}$ and $\boldsymbol{B}$ denote the stiffness matrices of a, a_h and $\hat{b}_h$ (restricted to $V_h\times V_h$) in the standard nodal bases of $\tilde{V}_h$, V_h and again V_h respectively. Also let $\boldsymbol{S}$ denote the $\dim(\tilde{V}_h)\times\dim(V_h)$ matrix describing the action of Q in the respective nodal bases.

Theorem 1.2 (Oswald [19])**.** *Assume that* (1.9)-(1.12) *holds, and that* $\boldsymbol{C_B}$ *and* $\boldsymbol{C_{\tilde{A}}}$ *are symmetric preconditioners for* $\boldsymbol{B}$ *and* $\tilde{\boldsymbol{A}}$*, respectively, satisfying the following spectral bounds*

$$\lambda_{\max}(\boldsymbol{C_B B}),\ \lambda_{\max}(\boldsymbol{C_{\tilde{A}}}\tilde{\boldsymbol{A}}) \le \Lambda_{\max}, \quad \lambda_{\min}(\boldsymbol{C_B B}),\ \lambda_{\min}(\boldsymbol{C_{\tilde{A}}}\tilde{\boldsymbol{A}}) \ge \Lambda_{\min}. \tag{1.14}$$

Then $\boldsymbol{C_A} := \boldsymbol{C_B}+\boldsymbol{S}^T\boldsymbol{C_{\tilde{A}}}\boldsymbol{S}$ *is a symmetric preconditioner for* $\boldsymbol{A}$*, with a bound for the spectral condition number of* $\boldsymbol{C_A A}$ *depending on the constants in* (1.14) *and* (1.13)*:*

$$\kappa(\boldsymbol{C_A A}) \le \frac{\hat{C}\Lambda_{\max}}{\hat{c}\Lambda_{\min}}. \tag{1.15}$$

While originally motivated by nonconforming schemes on non-nested meshes it should give no surprise that the abstract framework applies as well to DG methods (see [11] for the case of quasi-uniform meshes without hanging nodes). Here we shall point out that, based on two essential ingredients, namely a certain grading property of locally refined meshes and a local Jackson-type estimate, the above framework allows one to identify stable splittings for V_h also in the general case where the mesh is locally refined and the polynomial order k is arbitrary.

1.3 Layout of the paper

In Section 2 we shall reduce the applicability of Theorem 1.2 to the validity of a certain *Jackson-type estimate* provided that the underlying hierarchy of triangulations satisfies some mild grading constraints, see [7]. The key are suitable splittings of the trial spaces in a conforming and nonconforming part that are induced by suitable averaging operators. In Section 3 we shall identify two extreme cases of such splittings, both leading to asymptotically optimal preconditioners. Section 4 is devoted to numerical experiments that are to shed some light on the quantitative performance of the different versions, regarding also robustness with respect to local mesh refinements. Moreover, we discuss a simple refinement strategy based on [15].

2 Additive Schwarz preconditioner

As mentioned before, we are interested in adaptively refined triangulations. To be specific, we shall confine the discussion to subdividing a triangle T into four congruent subtriangles referred to as children of the parent T. Analogous results could be formulated for bisections as well. Given $\mathcal{T}^0$ we denote by $\mathcal{T}^j$ the jth fold uniform refinement of $\mathcal{T}^0$ according to the above rule while $\tilde{\mathcal{T}}^j$ denotes the corresponding tree representing this refinement history, i.e. $\mathcal{T}^j$ is the set of leaves of $\tilde{\mathcal{T}}^j$. We shall be concerned with triangulations $\mathcal{T}_h$ that form the set of leaves of a subtree $\tilde{\mathcal{T}}_h$ of $\tilde{\mathcal{T}}^{j_h}$, where j_h is the maximum refinement level appearing in $\mathcal{T}_h$.

Such local refinements will always be assumed to satisfy a mild *grading condition*: a hanging vertex is always the midpoint of two regular vertices (vertices that are not hanging), see [7] for an algorithmic characterization of this property. This also shows how to realize this property and the fact that it does not inflate the computational complexity in an essential way. The point is that for such graded meshes the trial space V_h contains a conforming subspace with contributions on all levels present in $\mathcal{T}_h$.

In order to link the present setting to the auxiliary space method, we will need the following localized energy norms (for certain neighborhoods $\omega = \omega(T) \subset \Omega$ of any $T \in \mathcal{T}_h$):

$$|||v|||_{h,\omega}^2 := \sum_{T\in\mathcal{T}_h: T\subseteq\omega} a(v,v)_T + \sum_{e\in\mathcal{E}_h: e\subseteq\omega} \frac{\gamma}{|e|} \|[v]\|_{L_2(e)}^2.$$

This will be used in connection with the following special neighborhoods of mesh elements that are affected by hanging vertices. To describe this it will be convenient to set

$$\mathcal{T}_h(D) := \{T\in\mathcal{T}_h : T\cap D \neq \emptyset\}, \text{ and similarly define } \mathcal{N}_{h,1}(D) \text{ and } \mathcal{E}_h(D)$$

for mesh elements (always considered as closed sets) touching a closed domain D. For instance, $\mathcal{T}_h(n)$ consists of the triangles that share the vertex n (either as a regular vertex, or as a hanging node). Unfortunately, due to hanging vertices, straightforward neighborhoods based on these notions will not suffice and we shall have to employ *extended* sets. To this end, let (with the notation of Fig. 2, right)

$$\mathcal{N}_{h,1}^*(n) := \begin{cases} \{n,n',n''\} & \text{if } n \text{ is a hanging vertex and the midpoint of } n' \text{ and } n'', \\ \{n\} & \text{otherwise, i.e. if } n \text{ is a regular vertex.} \end{cases} \tag{2.1}$$

Recall that when n is hanging, our grading property implies that both n' and n'' are regular. For any triangle $T\in\mathcal{T}_h$ we then set (denoting by $\mathcal{N}_1(T)$ the 3 vertices of T)

$$\mathcal{N}_{h,1}^*(T) := \cup_{n\in\mathcal{N}_1(T)} \mathcal{N}_{h,1}^*(n), \quad \mathcal{E}_h^*(T) := \cup_{n\in\mathcal{N}_{h,1}^*(T)} \mathcal{E}_h(n), \quad \mathcal{T}_h^*(T) := \cup_{n\in\mathcal{N}_{h,1}^*(T)} \mathcal{T}_h(n). \tag{2.2}$$

Finally we define the domain $\omega(T) := \cup_{T'\in\mathcal{T}_h^*(T)} T'$ as the union of triangles that are in contact with the extended set of vertices $\mathcal{N}_{h,1}^*(T)$. An illustration is given in Fig. 1, right, where the sets $\mathcal{N}^*_{h,1}(T)$, $\mathcal{E}^*_h(T)$ and $\mathcal{T}^*_h(T)$ are represented by white vertices, bold edges and gray triangles, respectively. The grading property implies also that the number of triangles involved in $\omega(T)$ remains uniformly bounded. Moreover, note that

$$\sum_{T'\in\mathcal{T}^*_h(T)} a(v,v)_{T'} + \sum_{e\in\mathcal{E}^*_h(T)} \frac{\gamma}{|e|}\|[v]\|^2_{L_2(e)} \le |||v|||^2_{h,\omega(T)}, \quad \forall T\in\mathcal{T}_h. \tag{2.3}$$

Proposition 2.1. Defining $\hat{b}_h(v,w) := \sum_{T\in\mathcal{T}_h}|T|^{-1}\langle v,w\rangle_T$ we have the following *Bernstein estimate*

$$|||v|||^2_h \lesssim \hat{b}_h(v,v), \quad \forall v\in V_h. \tag{2.4}$$

Now let $\mathcal{A}: V_h \to V_h\cap H^1_0(\Omega)$ be a linear projector satisfying the *local Jackson estimate*

$$\|(I-\mathcal{A})v\|_{L_2(T)} \le C^*|T|^{1/2}|||v|||_{h,\omega(T)} \qquad \forall T\in\mathcal{T}_h, \tag{2.5}$$

where $\omega(T)$ is the local patch defined above. Then, (1.9)-(1.12) hold if we set

$$\tilde{V}_h := \mathcal{A}V_h, \quad \tilde{Q} := \mathcal{A}, \quad \hat{a}_h(\cdot,\cdot) = a_h(\cdot,\cdot), \quad \text{and} \quad Q = I \ \text{ the canonical injection into } \ V_h.$$

Proof. First note that the relations in (1.9) are obviously valid. Now, (1.10) follows from (2.4) which can be proven by classical inverse inequalities and the fact that for any edge $e = T\cap T'$ one has

$$\|[w]\|^2_{L_2(e)} \le \|w^T\|^2_{L_2(e)} + \|w^{T'}\|^2_{L_2(e)}.$$

Finally (1.11) is trivial since $Q = I$, while (1.12) is a consequence of (2.5), due to the bounded overlapping of the $\omega(T)$. ☐

We could use now Theorem 1.2 to devise an optimal preconditioner for (1.3). Alternatively, in view of (1.6), we can use the consequence (1.13) of (1.9)-(1.12) to identify stable splittings for V_h in the sense of (1.7). This is the path we opt to pursue for $\tilde{V}_h = \mathcal{A}V_h$, where $\mathcal{A}$ is a suitable *admissible* projector from V_h into $V_h\cap H^1_0(\Omega)$. This means that we have to construct concrete projectors $\mathcal{A}: V_h \to V_h\cap H^1_0(\Omega)$ that are admissible – in the sense that an estimate of the form (2.5) holds – and then find stable splittings for each subspace

$$V^{\mathrm{c}}_h := \mathcal{A}V_h, \qquad V^{\mathrm{nc}}_h := (I-\mathcal{A})V_h$$

separately. In fact the completion of conforming stable splittings will be greatly simplified by the fact that any local basis yields a stable frame for the remainder (nonconforming) part. This can be formulated as follows.

Theorem 2.1 ([7]). *Assume that $\mathcal{A}$ satisfies the Jackson estimate (2.5), and let $\{V^{\mathrm{c}}_i : i\in\mathcal{I}^{\mathrm{c}}_h\}$ be any energy stable splitting for the conforming subspace $V^{\mathrm{c}}_h := \mathcal{A}V_h$. Moreover, assume that $V^{\mathrm{nc}}_{T,p}$, $T\in\mathcal{T}_h$, $p\in\mathcal{I}^{\mathrm{nc}}_T$ are subspaces of V_h with the property*

$$\mathrm{supp}\, v \subseteq T \quad \forall v\in V^{\mathrm{nc}}_{T,p}, \quad \text{and} \quad (I-\mathcal{A})V_h =: V^{\mathrm{nc}}_h \subseteq \sum_{T\in\mathcal{T}_h}\sum_{p\in\mathcal{I}^{\mathrm{nc}}_T} V^{\mathrm{nc}}_{T,p}. \tag{2.6}$$

Then, setting $\mathcal{I}_h^{\rm nc} := \{i=(T,p) : p\in\mathcal{I}_T^{\rm nc}, T\in\mathcal{T}_h\}$, *the collection*

$$\mathcal{S}_h = \{V_i : i\in\mathcal{I}_h\} := \{V_i^{\rm nc} : i\in\mathcal{I}_h^{\rm nc}\}\cup\{V_i^{\rm c} : i\in\mathcal{I}_h^{\rm c}\} \tag{2.7}$$

is an energy stable splitting for V_h *in the sense of* (1.7).

The actual construction of the preconditioner follows then standard lines, summarized as follows. Given a collection of subspaces $\mathcal{S}_h=\{V_i : i\in\mathcal{I}_h\}$ satisfying (1.7), consider auxiliary inner products on the spaces V_i, namely $b_i(\cdot,\cdot): V_i\times V_i\to\mathbb{R}$, $i\in\mathcal{I}_h$ yielding norms that are equivalent to $|||\cdot|||_h$ on V_i, i.e. satisfying

$$c_b|||v_i|||_h^2 \le b_i(v_i,v_i) \le C_b|||v_i|||_h^2 \quad \forall v_i\in V_i,\ i\in\mathcal{I}_h, \tag{2.8}$$

for constants c_b,C_b depending only on the degree k, the shape properties of $\mathcal{T}^0$ and possibly on the coefficients in the bilinear form $a(\cdot,\cdot)$. From (1.6) we see that one possible strategy is to set $b_i=a_h$ for all i, but other choices are conceivable, see (2.12). Next define the operators $P_i: V_h\to V_i$ and elements $f_i\in V_i$ by

$$b_i(P_iw,v_i)=a_h(w,v_i) \quad\text{and}\quad b_i(f_i,v_i)=\langle f,v_i\rangle, \quad \forall v_i\in V_i,\ i\in\mathcal{I}_h. \tag{2.9}$$

Note that whenever V_i is a one-dimensional space, the application of P_i just amounts to solving a linear equation with a single unknown. The central result we shall use reads then as follows, see, e.g., [14, 18].

Theorem 2.2. *Define* $P_h: V_h\to V_h$ *and* $\bar{f}_h\in V_h$ *by* $P_h:=\sum_{i\in\mathcal{I}_h}P_i$ *and* $\bar{f}_h:=\sum_{i\in\mathcal{I}_h}f_i$. *Then* P_h *is* a_h*-symmetric, positive definite, and the problem (1.3) is equivalent to the operator equation*

$$P_hu=\bar{f}_h. \tag{2.10}$$

Moreover, if (1.7) and (2.8) *hold, then the spectral values of the symmetric positive definite (with respect to the inner product* $a_h(\cdot,\cdot)$*) operator* P_h *can be bounded by constants independent of* $\mathcal{T}_h$*, namely*

$$\frac{c_a}{C_bC_S}\le\lambda_{\min}(P_h):=\inf_{\substack{v\in V_h\\ v\neq 0}}\frac{a_h(P_hv,v)}{a_h(v,v)}, \qquad \sup_{\substack{v\in V_h\\ v\neq 0}}\frac{a_h(P_hv,v)}{a_h(v,v)}=:\lambda_{\max}(P_h)\le\frac{C_a}{c_bc_S},$$

remember c_a,C_a *are the constants from* (1.6)*. In particular, the spectral condition number of* P_h *is bounded independently of* $\mathcal{T}_h$*,*

$$\kappa(P_h):=\frac{\lambda_{\max}(P_h)}{\lambda_{\min}(P_h)}\le\frac{C_aC_bC_S}{c_ac_bc_S}. \tag{2.11}$$

Let us end this section by mentioning the following admissible choices for the $b_i(\cdot,\cdot)$:

$$\begin{cases} a(v,w)_T + \displaystyle\sum_{e\in\mathcal{E}_h, e\subset\partial T} \frac{\gamma}{|e|}\int_e [v][w] & \text{or} \\ |T|^{-1}\langle v,w\rangle_T + \displaystyle\sum_{e\in\mathcal{E}_h, e\subset\partial T} \frac{\gamma}{|e|}\int_e [v][w] & \text{when } i=(T,p)\in\mathcal{I}_h^{\mathrm{nc}}, \\ a(v,w)_{\omega(i)} & \text{or} \\ |\omega(i)|^{-1}\langle v,w\rangle_{\omega(i)} & \text{when } i\in\mathcal{I}_h^{\mathrm{c}}, \end{cases} \tag{2.12}$$

where $\omega(i)$ denotes the union of all the supports of the nodal basis functions spanning the subspace V_i^c (typically a coarse grid space).

3 Averaging projectors

We shall now turn to the construction of admissible projectors $\mathcal{A}$. We will focus on two extreme cases, namely a *low order* and a *high order* averaging operator:

$$\mathcal{A}_1 : V_h \to V_{h,1}^{\mathrm{c}} := H_0^1(\Omega)\cap\mathbb{P}_1(\mathcal{T}_h) \quad \text{and} \quad \mathcal{A}_k : V_h \to V_{h,k}^{\mathrm{c}} := H_0^1(\Omega)\cap V_h.$$

Thus, the auxiliary space $\tilde{V}_h := V_{h,1}^{\mathrm{c}}$ consists just of globally continuous piecewise linear finite elements and is therefore the *minimal* conforming subspace of V_h, while the auxiliary space $\tilde{V}_h := V_{h,k}^{\mathrm{c}}$ is the largest one that contains continuous finite elements of the actual degree present in V_h.

3.1 Minimal conforming subspaces

The case $V_{h,1}^{\mathrm{c}}$ has been already analyzed in [7]. For the convenience of the reader we briefly recall the main facts. To this end, let $\mathcal{T}_h^j := \mathcal{T}_h \cap \mathcal{T}^j$, $j=1,2,\cdots,j_h$, be the subsets of $\mathcal{T}_h$ comprised of level-j triangles, $\Omega_h^j := \cup\{T : T\in\mathcal{T}_h^j\}$ and denote by

$$\mathcal{N}_{h,1}^{j,\mathrm{c}} := \{n\in\mathcal{N}_1(T) : T\in\mathcal{T}_h^j,\ n\notin\partial\Omega_h^j\}$$

all the vertices of $\mathcal{T}_h^j$ that lie in the interior of Ω_h^j. Clearly the union of the sets consists of the regular vertices that lie in the interior of Ω. Next, for every $n\in\mathcal{N}_{h,1}^{j,\mathrm{c}}$ let $\phi_{j,n}^c$ denote the standard nodal (piecewise affine) continuous hat function at vertex n supported on the star of triangles in $\mathcal{T}_h^j$ sharing n. Then the multilevel collection

$$\mathcal{S}_{h,1}^{\mathrm{c}} := \{\mathrm{Span}(\phi_i^c) : i\in\mathcal{I}_{h,1}^{\mathrm{c}}\} \quad \text{with} \quad \mathcal{I}_{h,1}^{\mathrm{c}} := \{i=(j,n) : j\in\{1,2,\cdots,j_h\},\ n\in\mathcal{N}_{h,1}^{j,\mathrm{c}}\} \tag{3.1}$$

is known to be a H_0^1-stable splitting of $V_{h,1}^{\mathrm{c}} = H_0^1(\Omega)\cap\mathbb{P}_1(\mathcal{T}_h)$, see [14], while

$$\Phi_{h,1}^{\mathrm{c}} := \{\phi_n^c := \phi_{j(n),n}^c :\ n\in\mathcal{N}_{h,1}^{\mathrm{c}} := \cup_{j=1}^{j_h}\mathcal{N}_{h,1}^{j,\mathrm{c}}\}, \quad \text{where} \quad j(n) := \max\{j : n\in\mathcal{N}_{h,1}^{j,\mathrm{c}}\}, \tag{3.2}$$

which clearly yields a nodal basis for $V^{c}_{h,1}$, is unstable. Now, since boundary vertices are subject to a zero boundary condition in $H^1_0(\Omega)$, $\mathcal{A}_1 v$ is simply defined by prescribing its values at every interior regular vertex $n\in\mathcal{N}^{c}_{h,1}$. In [7] it has been shown that setting

$$(\mathcal{A}_1 v)(n) := \frac{1}{\#(\mathcal{T}_h(n))}\sum_{T\in\mathcal{T}_h(n)} v^T(n) \qquad \forall n\in\mathcal{N}^c_{h,1} \tag{3.3}$$

defines an admissible projector, i.e. $\mathcal{A}_1$ satisfies (2.5). Hence, as a viable stable splitting one could take the union of a nodal basis for V_h and the collection $\mathcal{S}^c_{h,1}$, see Theorem 3.1. The following remark links this case directly to Theorem 1.2.

Remark 3.1. Clearly, Theorem 2.2 applied to the above collection $\mathcal{S}^{c}_{h,1}$ yields an optimal preconditioner for the auxiliary problem (1.8) when $\tilde{V}_h = V^{c}_{h,1}$. In order to write the matrix $\tilde{\boldsymbol{P}}$ describing the action of the resulting operator $\tilde{P}_h := \sum_{i\in\mathcal{I}^{c}_{h,1}} P_i$ in the nodal basis $\Phi^{c}_{h,1}$ (3.2) of $\tilde{V}_h$, let $\tilde{\boldsymbol{A}}$ and $\boldsymbol{F}$ denote respectively the stiffness matrix of a, and the matrix representation of the frame functions $\{\phi^{c}_i : i\in\mathcal{I}^{c}_{h,1}\}$ in the nodal basis $\Phi^{c}_{h,1}$. Since every element of $\mathcal{S}^{c}_{h,1}$ is spanned by one single ϕ^{c}_i, it can easily be checked that

$$P_i\tilde{v} = \frac{a(\tilde{v},\phi^{c}_i)}{b_i(\phi^{c}_i,\phi^{c}_i)}\phi^{c}_i \quad \forall \tilde{v}\in\tilde{V}_h,\ i\in\mathcal{I}^{c}_{h,1}, \quad \text{and} \quad \tilde{\boldsymbol{A}}\tilde{\boldsymbol{P}} = \Big(a\left(\tilde{P}_h\phi^{c}_n,\phi^{c}_m\right)\Big)_{n,m\in\mathcal{N}^{c}_{h,1}} = \tilde{\boldsymbol{A}}\boldsymbol{F}^T\tilde{\boldsymbol{D}}\boldsymbol{F}\tilde{\boldsymbol{A}},$$

where $\tilde{\boldsymbol{D}} := \mathrm{diag}(b_i(\phi^{c}_i,\phi^{c}_i) : i\in\mathcal{I}^{c}_{h,1})$. Therefore it follows from Theorem 2.2 that the spectral condition number of $\tilde{\boldsymbol{P}} = \boldsymbol{F}^T\boldsymbol{D}\tilde{\boldsymbol{F}}\tilde{\boldsymbol{A}}$ is uniformly bounded, i.e. $\boldsymbol{C}_{\tilde{\boldsymbol{A}}} := \boldsymbol{F}^T\tilde{\boldsymbol{D}}\boldsymbol{F}$ is an optimal preconditioner for $\tilde{\boldsymbol{A}}$. Now let $\boldsymbol{A}$ and $\boldsymbol{P}$ respectively denote the stiffness matrix of a_h and the matrix describing the action of $P_h := \tilde{P}_h + \sum_{i\in\mathcal{I}_h} P_i$ in a given nodal basis $\Phi_h = \{\phi_i : i\in\mathcal{I}_h\}$ of V_h (see Theorem 3.1). Also let $\boldsymbol{S}$ denote the matrix describing the action of the canonical injection from $\tilde{V}_h$ into V_h in the respective bases $\Phi^{c}_{h,1}$, Φ_h. Computing as above, we find

$$P_i v = \frac{a(v,\phi_i)}{b_i(\phi_i,\phi_i)}\phi_i \quad \forall v\in V_h,\ i\in\mathcal{I}_h, \quad \text{and} \quad \boldsymbol{A}\boldsymbol{P} = \Big(a_h\left(P_h\phi_i,\phi_l\right)\Big)_{i,l\in\mathcal{I}_h} = \boldsymbol{A}(\boldsymbol{D}+\boldsymbol{S}^T\boldsymbol{F}^T\tilde{\boldsymbol{D}}\boldsymbol{F}\boldsymbol{S})\boldsymbol{A}$$

with $\boldsymbol{D} := \mathrm{diag}(b_i(\phi_i,\phi_i) : i\in\mathcal{I}_h)$, therefore Theorem 2.2 yields that $\boldsymbol{C}_{\boldsymbol{A}} := (\boldsymbol{D}+\boldsymbol{S}^T\boldsymbol{F}^T\boldsymbol{D}\boldsymbol{F}\boldsymbol{S})$ is an optimal preconditioner for the matrix $\boldsymbol{A}$. Note that since normalized nodal bases are L_2-stable, $\boldsymbol{D}$ is an optimal preconditioner for the stiffness matrix of $\hat{b}$ defined in Proposition 2.1, hence we identify indeed the preconditioner given in Theorem 1.2.

3.2 Maximal conforming subspaces

We now turn to the averaging operators $\mathcal{A}_k$ that produce H^1_0–conforming functions of polynomial degree according to k. In order to limit technicalities we shall greatly simplify the setting by assuming from now on that the polynomial degrees are equal to some fixed $k=\bar{k}$ throughout the mesh. This does not preclude using locally lower order polynomials (that could always be written artificially as higher order ones) but $\mathcal{A}_k$ will generally map into the full space V_h. The concrete realization of $\mathcal{A}_k$ (given in Section 3.2.4 below) requires a few auxiliary tools to be prepared next.

3.2.1 High order Bernstein-Bézier polynomial bases

It will be convenient to employ the so called *principal lattices* which makes it very convenient to ensure continuity across element faces. Since we shall make use of these concepts for triangles as well as for edges we quickly recall the relevant facts in a general d-dimensional setting, which also indicates what happens when dealing with higher spatial dimensions. We refer, for instance, to [12] for further details. Given a d-dimensional simplex S, we denote by

$$\mathcal{N}_k(S):=\Big\{p=p_\beta:=\frac{1}{k}\sum_{j=0}^{d}\beta_j n_j:\beta\in\mathbb{Z}_+^{d+1},\sum_{j=0}^{d}\beta_j=k\Big\} \tag{3.4}$$

the principal lattice of order k induced on S, see Fig. 2, left. Of course, for $k=1$ we simply have $\mathcal{N}_1(S)=\{n_0,\cdots,n_d\}$, the set of vertices of S itself, as used before. Note that the β/k are the barycentric coordinates of the mesh points $p=p_\beta$, a fact that will be used later again. Clearly we have $\dim(\mathbb{P}_k(S))=\#\mathcal{N}_k(S)$.

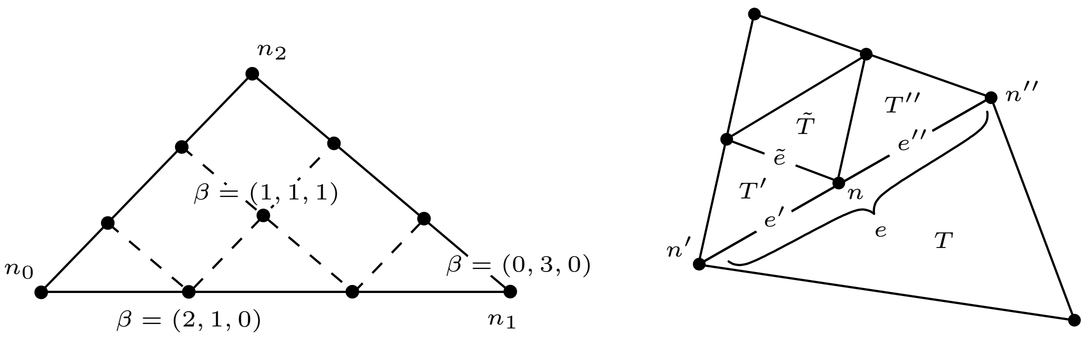

Figure 2: Principal lattice $\mathcal{N}_3(T)$ for $k=3$ (left) and local notation for nonconforming configurations (right).

Now let $\{P_p^k(\cdot;\hat{S}):p\in\mathcal{N}_k(\hat{S})\}$ be a fixed basis for $\mathbb{P}_k(\hat{S})$, where $\hat{S}$ is a reference simplex and define for any $S=F\hat{S}$ with $F=F_S$ an affine mapping, $P_p^S:=\chi_S P_p^k(\cdot;\hat{S})\circ F_S^{-1}$, $p\in\mathcal{N}_k(S)$ (the degree k being then implicit through $p\in\mathcal{N}_k(S)$). Obviously,

$$\Phi_h:=\{\phi_{T,p}:=P_p^T:p\in\mathcal{N}_k(T),T\in\mathcal{T}_h\} \tag{3.5}$$

is then a basis for V_h. Later, the P_p^k will be (essentially) normalized in L_∞. Thus one has

$$c_k\|\{a_p\}_{p\in\mathcal{N}_k(S)}\|_{\ell_\infty}\le\Big\|\sum_{p\in\mathcal{N}_k(S)}a_p P_p^S\Big\|_{L_\infty(S)}\le C_k\|\{a_p\}_{p\in\mathcal{N}_k(S)}\|_{\ell_\infty} \tag{3.6}$$

with constants depending only on k and the specific choice of the polynomial basis. For instance, one may consider Lagrange polynomials pieces defined on the local k-meshes

$$L_p^T:=\chi_T L_p^k(\cdot;T),\quad p\in\mathcal{N}_k(T),\ T\in\mathcal{T}_h,\quad \text{where}\quad L_p^k(q;T)=\delta_{p,q}\quad\forall\ p,\ q\in\mathcal{N}_k(T), \tag{3.7}$$

and (3.6) allows us to estimate function norms by coefficient norms.

A convenient alternate basis can be comprised of Bernstein-Bézier polynomials whose definition will be recalled next. Given a simplex S, and using standard multi-index notation $\lambda^\alpha := \lambda_0^{\alpha_0}\cdots\lambda_d^{\alpha_d}$, $\alpha\in\mathbb{Z}_+^{d+1}$, $\alpha! := \alpha_0!\cdots\alpha_d!$ the corresponding Bernstein-Bézier basis functions of degree $|\beta| := \beta_0+\cdots+\beta_d$ are first defined by

$$B_\beta^{|\beta|}(x;S) := \frac{|\beta|!\lambda^\beta}{\beta!}, \quad \beta\in\mathbb{Z}_+^d, \tag{3.8}$$

where the barycentric coordinates $\lambda=\lambda(x;S)=(\lambda_0,\cdots,\lambda_d)$ of $x\in\mathbb{R}^d$ are given by the nonsingular system of linear equations $x=\lambda_0 n_0+\cdots+\lambda_d n_d$, $\lambda_0+\cdots+\lambda_d=1$. Note that for $k=1$, $d=2$, the three polynomials

$$B^1_{(1,0,0)}(\cdot;T)=\lambda_0, \quad B^1_{(0,1,0)}(\cdot;T)=\lambda_1, \quad B^1_{(0,0,1)}(\cdot;T)=\lambda_2$$

correspond to the three canonical piecewise linear nodal shape functions on the underlying triangle, and more generally, the set $\{B_\beta^k(\cdot;S) : |\beta|=k\}$ is known to form a basis for $\mathbb{P}_k(\mathbb{R}^d)$, see [12]. As above, we will index the polynomials $B_\beta^k(\cdot;S)$ by the point $p=p_\beta\in\mathcal{N}_k(S)$ instead of by the multi-index β. Also, since we are interested in discontinuous bases we set

$$B_p^S(x) := \chi_S(x)B_\beta^k(x;S), \quad p=p_\beta\in\mathcal{N}_k(S), \tag{3.9}$$

the dependence of B_p^S on k being then implicit through $p\in\mathcal{N}_k(S)$. Hence, given $\mathcal{T}_h$, every $v\in V_h$ has a unique representation

$$v=\sum_{T\in\mathcal{T}_h}\sum_{p\in\mathcal{N}_k(T)} b_p^T B_p^T(x). \tag{3.10}$$

One easily verifies that on T the polynomials B_p^T, $p\in\mathcal{N}_k(T)$, form a nonnegative partition of unity on T

$$1=\sum_{p\in\mathcal{N}_k(T)} B_p^T(x), \qquad B_p^T(x)\geq 0, \quad x\in T,\, p\in\mathcal{N}_k(T), \tag{3.11}$$

so that the graph of v in (3.10) (over T) is contained in the convex hull of the points (p,b_p^T), $p\in\mathcal{N}_k(T)$. Therefore, the points (p,b_p^T) are often called *control points* and the coefficients b_p^T in the Bernstein-Bézier representation are usually referred to as *control coefficients*.

Now, another important property of the Bernstein-Bézier representations lies in the fact that the above inequality holds with $C_k=1$. The following is indeed a simple consequence of (3.11).

Remark 3.2. There exists a constant c_k depending only on the spatial dimension d and the degree k of the Bernstein polynomial pieces B_p^S, such that

$$c_k\|\{b_p\}_{p\in\mathcal{N}_k(S)}\|_{\ell_\infty} \leq \|\sum_{p\in\mathcal{N}_k(S)} b_p B_p^S\|_{L_\infty(S)} \leq \|\{b_p\}_{p\in\mathcal{N}_k(S)}\|_{\ell_\infty} \tag{3.12}$$

holds for any set of control coefficients $\{b_p\}_{p\in\mathcal{N}_k(S)}$.

We will now carefully study how such polynomial pieces can be continuously stitched together at common faces. To this end we state a few known facts concerning traces of Bernstein polynomials, and to begin with we observe that the restriction of the k-mesh of a triangle to one of its edges is again a k-mesh of that edge. Now, the same holds for Bézier bases, i.e. the trace of a Bézier basis polynomial B_p^T on an edge e of T is again a Bézier basis polynomial B_p^e on that edge, see [12]:

$$B_p^T|_e = B_p^e\,, \quad \text{for any edge } e \text{ of } T,\ p \in \mathcal{N}_k(e) = \mathcal{N}_k(T) \cap e, \tag{3.13}$$

and in fact the collection $\{B_p^e : p \in \mathcal{N}_k(e)\}$ is a basis of $\mathbb{P}_k(e)$. From this property, one readily infers the following facts.

(i) Given any (closed) edge e of T, we have

$$p \in \mathcal{N}_k(T) \setminus e \Longrightarrow B_p^T|_{\mathcal{N}_k(e)} = B_p^T|_e = 0. \tag{3.14}$$

In particular, Bernstein polynomials associated to interior nodes vanish on ∂T

$$p \in \mathcal{N}_k^0(T) := \mathcal{N}_k(T) \setminus \partial T \Longrightarrow B_p^T|_{\partial T} = 0, \tag{3.15}$$

i.e. they are globally continuous. (Note that $p = p_\beta \in \mathcal{N}_k(T)$ lies in the interior of T if and only if all the entries of β are strictly positive.)

(ii) If $e = T \cap T'$ is an edge of both T and T' in $\mathcal{T}_h$, then writing v as in (3.10) we have

$$v^T|_e = v^{T'}|_e \quad \Longleftrightarrow \quad b_p^T = b_p^{T'},\ \forall\, p \in \mathcal{N}_k(e), \tag{3.16}$$

i.e. two polynomials (with Bézier representations of the same degree) on adjacent triangles merge continuously across an edge if and only if the control coefficients coincide on that edge, see (3.13). See also Remark 3.4.

(iii) For any degree k, we have $\mathcal{N}_1(T) \subseteq \mathcal{N}_k(T)$ and

$$B_p^T(n) = \delta_{p,n} \ \text{ for any } \ n \in \mathcal{N}_1(T),\ p \in \mathcal{N}_k(T).$$

Thus if n is a vertex of both T and T', writing again v as in (3.10) we have

$$v^T(n) = \sum_{p \in \mathcal{N}_k(T)} b_p^T B_p^T(n) = b_n^T \tag{3.17}$$

i.e. the different polynomial pieces of v merge continuously on n if and only if the control coefficients coincide on that vertex. Note, however, that this argument doesn't apply to hanging vertices.

3.2.2 A reduced frame for the remainder $V_{h,k}^{\mathrm{nc}} := (I - \mathcal{A}_k)V_h$

Unlike the case of $\mathcal{A}_1$ (see Section 3.1), an energy stable frame for $V_{h,k}^{\mathrm{nc}}$ can be formed from a strict subset of the full basis for V_h. It suffices to take only edge-based basis functions. Relation (3.15) indeed means that Bernstein polynomial pieces associated to interior nodes are continuous, and it is easily checked that the same holds for Lagrange polynomial pieces, i.e.

$$\phi_{T,p}|_{\partial T} = 0 \quad \text{when} \quad p \in \mathcal{N}_k^0(T) := \mathcal{N}_k(T) \setminus \partial T \tag{3.18}$$

for $\phi_{T,p}=B_p^T$ or L_p^T. Hence in the case $\mathcal{A}_k$ whose range is all of $V_h\cap H_0^1(\Omega)$, for all $p\in\mathcal{N}_k^0(T)$ one has $(I-\mathcal{A}_k)\phi_{T,p}=0$. The following remark is an immediate consequence of this fact.

Remark 3.3. Suppose that the basis functions $\phi_{T,p}$ satisfy (3.18). Then setting

$$\mathcal{I}_{h,k}^{\mathrm{nc}}:=\{i=(T,p):p\in\partial\mathcal{N}_k(T),T\in\mathcal{T}_h\}\quad\text{yields}\quad V_{h,k}^{\mathrm{nc}}:=(I-\mathcal{A}_k)V_h\subseteq\operatorname{span}\{\phi_i:i\in\mathcal{I}_{h,k}^{\mathrm{nc}}\}.$$

In particular, it follows from Theorem 2.1 that for any $w\in V_{h,k}^{\mathrm{nc}}$ one can write

$$|||w|||_h^2\sim\sum_{i\in\mathcal{I}_{h,k}^{\mathrm{nc}}}|w_i|^2|||\phi_i|||_h^2\quad\text{where}\quad w=\sum_{i\in\mathcal{I}_{h,k}^{\mathrm{nc}}}w_i\phi_i \tag{3.19}$$

with the previously specified dependence of the constants.

3.2.3 The conforming high order subspace $V_{h,k}^{\mathrm{c}}=V_h\cap H_0^1(\Omega)$

Before defining $\mathcal{A}_k$, it is convenient to describe the subspace $V_{h,k}^{\mathrm{c}}:=\mathcal{A}_kV_h=V_h\cap H_0^1(\Omega)$. Let us begin with the following observation.

Remark 3.4. Suppose that the edge e of T contains a hanging vertex as its midpoint so that $e=T\cap(T'\cup T'')$ for two adjacent triangles T', T'' (as in Fig. 2, right). Then both traces $(v^{T'}+v^{T''})|_e$, $v^T|_e$ of any $v\in V_{h,k}^{\mathrm{c}}$ must agree with a single polynomial of degree k.

In order to express the degrees of freedom in $V_{h,k}^{\mathrm{c}}$ we need some additional notation. So far, the set $\mathcal{E}_h$ has been defined by replacing any edge with a hanging vertex by its two halves. In the sequel we will need the other convention, so let $\mathcal{E}_h^{\mathrm{u}}$ be the set where two such half edges are replaced by their *union*. For instance in Fig. 2, right, $\mathcal{E}_h$ contains e', e'' while $\mathcal{E}_h^{\mathrm{u}}$ contains e. Moreover since the elements of $V_{h,k}^{\mathrm{c}}$ vanish on $\partial\Omega$, we don't count in $\mathcal{E}_h^{\mathrm{u}}$ edges that lie in the boundary $\partial\Omega$. Next, similarly to $\mathcal{N}_k^0(T)$ we denote

$$\mathcal{N}_k^0(e):=\mathcal{N}_k(e)\setminus\{n',n''\},\ \text{ for every } e\in\mathcal{E}_h^{\mathrm{u}} \text{ with endpoints } n',n'', \tag{3.20}$$

be the set of k-th order nodes located in the (relative) interior of e.

Since the value of a polynomial at a vertex always agrees, by (3.17), with the corresponding control coefficient (regardless of the degree of the polynomial), the above observations (3.15)-(3.17) and Remark 3.4 suggest to group the degrees of freedom of $V_{h,k}^{\mathrm{c}}$ into three disjoint sets, namely

$$\begin{aligned}&\text{1) the set } \mathcal{N}_{h,1}^{\mathrm{c}} \text{ consisting of regular vertices in the interior of } \Omega,\\ &\text{2) the nodes in the interior of the ``united'' edges } \mathcal{N}_{h,k}^{\mathcal{E}}:=\bigcup\{\mathcal{N}_k^0(e):e\in\mathcal{E}_h^{\mathrm{u}}\},\\ &\text{3) the nodes in the interior of the triangles } \mathcal{N}_{h,k}^{\mathcal{T}}:=\bigcup\{\mathcal{N}_k^0(T):T\in\mathcal{T}_h\}.\end{aligned} \tag{3.21}$$

Now, it is easily checked that the following collection

$$\Phi_h^{\mathrm{c}}:=\{\phi_n^{\mathrm{c}}:n\in\mathcal{N}_{h,1}^{\mathrm{c}}\}\cup\{\phi_p^{\mathrm{c}}:p\in\mathcal{N}_{h,k}^{\mathcal{E}}\}\cup\{\phi_p^{\mathrm{c}}:p\in\mathcal{N}_{h,k}^{\mathcal{T}}\} \tag{3.22}$$

forms a basis for $V^{c}_{h,k} = V_h \cap H^1_0(\Omega)$ when the ϕ^{c}_p are defined as follows: for $n \in \mathcal{N}^{c}_{h,1}$ we set $\phi^{c}_n|_T = B^T_n$ for all T in the *regular star* of n that consists of all triangles in the tree $\widehat{\mathcal{T}}_h$ sharing n as a vertex and having the same minimum level among those belonging to $\mathcal{T}_h$. When $p \in \mathcal{N}^{\mathcal{E}}_{h,k}$ we form ϕ^{c}_p by adjoining the corresponding two Bernstein polynomials B^T_p, $B^{T'}_p$ where either both T, T' belong to $\mathcal{T}_h$ (in which case the shared edge has no hanging node) or T' is the parent of the higher level triangles adjacent to T. Finally, for $p \in \mathcal{N}^{\mathcal{T}}_{h,k}$ we simply set $\phi^{c}_p = B^T_p$ when p is located in the interior of T.

3.2.4 Construction of $\mathcal{A}_k$

The construction of $\mathcal{A}_k$ will identify the coefficients with respect to Φ^{c}_h. According to the structure of $V^{c}_{h,k}$ explained above we shall define the coefficients of $\mathcal{A}_k v$ depending on their membership to conforming vertices, nodes belonging to the interior of a triangle or to the relative interior of an edge. Due to Remark 3.4, in the presence of a hanging vertex this latter group of nodes suggests considering the spaces of polynomials and piecewise polynomials of degree k on e that vanish at the end points of e. In other terms, using the notation of Fig. 2, right, we define

$$\mathbb{P}^0_k(e) := \mathrm{span}\{B^e_p : p \in \mathcal{N}^0_k(e) = \mathcal{N}_k(e) \setminus \{n', n''\}\}$$

and

$$\mathbb{P}^0_k(e', e'') := \mathrm{span}\{B^{e'}_p : p \in \mathcal{N}_k(e') \setminus \{n'\}\} \cup \{B^{e''}_p : p \in \mathcal{N}_k(e'') \setminus \{n''\}\}.$$

Let then $\mathbf{Q} = \mathbf{Q}(e, e', e'')$ be the matrix representation of the orthogonal projector from $\mathbb{P}^0_k(e', e'')$ into $\mathbb{P}^0_k(e)$ in the respective Bézier bases. Clearly this is a $(k-1) \times (2k)$-matrix which, due the affine invariance of the Bézier representation depends only on the degree k and not on the specific edges e, e', e'' (and in fact, any projector will do, as long as its stability constant $\|\mathbf{Q}\|$ only depends on k). The reverse procedure is commonly called *subdivision* and consists in writing any polynomial in $\mathbb{P}^0_k(e)$ as a linear combination of Bernstein pieces in $\mathbb{P}^0_k(e', e'')$. Hence is it represented by a $(2k) \times (k-1)$-matrix $\mathbf{M}$ that depends also only on the degree k, and one has

$$\mathbf{QM} = \mathbf{Id}_{k-1}. \tag{3.23}$$

In practice, subdivision is not carried out by first assembling the matrix $\mathbf{M}$ but via efficient recursive procedures, see e.g. [12].

We can now describe the averaging operator $\mathcal{A}_k$ preserving the original degrees of the elements: given $v \in V_h$ written as in (3.10), i.e. in terms of the full discontinuous basis $v = \sum_{T \in \mathcal{T}_h} \sum_{p \in \mathcal{N}_k(T)} b^T_p B^T_p$, we define $\mathcal{A}_k : V_h \to V_h \cap H^1_0(\Omega)$ by

$$\mathcal{A}_k v := \sum_{T \in \mathcal{T}_h} \sum_{p \in \mathcal{N}_k(T)} c^T_p B^T_p \tag{3.24}$$

$$:= \sum_{n \in \mathcal{N}^{c}_{h,1}} a_n \phi^{c}_n + \sum_{p \in \mathcal{N}^{\mathcal{E}}_{h,k} \cup \mathcal{N}^{\mathcal{T}}_{h,k}} a_p \phi^{c}_p\,, \tag{3.25}$$

where the coefficients c_p^T (with respect to the full discontinuous basis for V_h) and the coefficients a_n, a_p (with respect to the conforming basis $\Phi_h^{\rm c}$) will be specified next, in particular see Proposition 3.1 below. As mentioned in the previous section, we distinguish the three types of conforming nodes (3.21), namely (1) regular vertices inside Ω, i.e. $\mathcal{N}_{h,1}^{\rm c}$, (2) nodes located in the interior of the "united" edges of $\mathcal{E}_h^{\rm u}$, i.e. $\mathcal{N}_{h,k}^{\mathcal{E}}$, and (3) nodes located in the interior of the triangles, i.e. $\mathcal{N}_{h,k}^{\mathcal{T}}$. In particular, note that the control coefficients c_n^T corresponding to hanging vertices will not be specified directly, see (3.32) below.

Case (1) is the same as in the construction of $\mathcal{A}_1$, namely we simply average all values coming from the adjacent triangles

$$a_n := \frac{1}{\#\mathcal{T}_h(n)} \sum_{T\in\mathcal{T}_h(n)} b_n^T, \quad n\in\mathcal{N}_{h,1}^{\rm c}. \tag{3.26}$$

Accordingly we set

$$c_n^T := a_n, \quad n\in\mathcal{N}_{h,1}^{\rm c}, \;\; T\in\mathcal{T}_h(n). \tag{3.27}$$

Case (3) is particularly simple, because we can simply set

$$a_p := c_p^T := b_p^T, \quad p\in\mathcal{N}_k(T)\setminus\partial T, \;\; T\in\mathcal{T}_h. \tag{3.28}$$

Thus it remains to deal with case (2) where p belongs to the interior of one edge $e\in\mathcal{E}_h^{\rm u}$ (recall that this set contains no edges in $\partial\Omega$). Having (3.16) and Remark 3.4 in mind we distinguish again two subcases, depending whether (2.a) e contains no hanging vertex, or (2.b) it contains one. In the first subcase, e is an edge of two adjacent triangles T, T' in $\mathcal{T}_h$ such that $e = T\cap T'$. According to (3.13) we have then

$$v^T|_e = \sum_{p\in\mathcal{N}_k(e)} b_p^T B_p^e, \qquad v^{T'}|_e = \sum_{p\in\mathcal{N}_k(e)} b_p^{T'} B_p^e. \tag{3.29}$$

Now, it will be convenient to treat simultaneously the whole arrays of control coefficients in the interior of e, so we let $\mathbf{b}^{T,e} := \{b_p^T : p\in\mathcal{N}_k^0(e)\}$, and similarly define $\mathbf{b}^{T',e}$, $\mathbf{c}^{T,e}$, $\mathbf{c}^{T',e}$. Guided by (3.16), we then set

$$\mathbf{c}^{T,e} := \frac{1}{2}\big(\mathbf{b}^{T,e} + \mathbf{b}^{T',e}\big) =: \mathbf{c}^{T',e}. \tag{3.30}$$

As for the representation (3.25), we set, of course,

$$a_p = c_p^T \quad p\in\mathcal{N}_k^0(e). \tag{3.31}$$

Now for the subcase (2.b), i.e. when p lies in the interior of one edge e that contains a hanging vertex, we refer again to the notation in Fig. 2 (right). In the same spirit as before we denote by

$$\mathbf{b}^{T',e'} = \big\{b_p^{T'} : p\in\mathcal{N}_k(e')\setminus\{n'\}\big\}, \quad \mathbf{b}^{T'',e''} = \big\{b_p^{T''} : p\in\mathcal{N}_k(e'')\setminus\{n''\}\big\}$$

and

$$\mathbf{b}^{T,e}=\{b_p^T : p\in\mathcal{N}_k^0(e)=\mathcal{N}_k(e)\setminus\{n',n''\}\}$$

the arrays of interior control coefficients of the traces $v^{T'}|_{e'}$, $v^{T''}|_{e''}$, $v^T|_e$, respectively. For $\mathbf{c}^{T,e}$ defined as above, guided by the definitions of $\mathbf{Q}$ and $\mathbf{M}$, we then set

$$\mathbf{c}^{T,e}:=\frac{1}{2}\left(\mathbf{b}^{T,e}+\mathbf{Q}\begin{pmatrix}\mathbf{b}^{T',e'}\\ \mathbf{b}^{T'',e''}\end{pmatrix}\right),\quad \begin{pmatrix}\mathbf{c}^{T',e'}\\ \mathbf{c}^{T'',e''}\end{pmatrix}:=\mathbf{M}\mathbf{c}^{T,e} \tag{3.32}$$

(note that the latter also sets the values of control coefficients corresponding to hanging vertices), and the coefficients a_p, $p\in\mathcal{N}_k^0(e)$, are given as in subcase (2.a) by (3.31).

Proposition 3.1. The operator $\mathcal{A}_k$ defined by (3.24) with coefficients from (3.27), (3.28), (3.30) and (3.32) is a projection from V_h onto $V_h\cap H_0^1(\Omega)$. Relation (3.25) with coefficients given by (3.26), (3.28), and (3.31) is an equivalent representation in the basis (3.22).

Proof. Due to (3.17), the continuity at the vertices in $\mathcal{N}_{h,1}^{\mathrm{c}}$ is obvious from (3.26) and (3.27). As for the continuity across edges in $\mathcal{E}_h^{\mathrm{u}}$, we refer to (3.16) and note first that in case (2.a), by (3.30), the two arrays $\mathbf{c}^{T,e}$, $\mathbf{c}^{T',e}$ define the same polynomial trace on e. For case (2.b), we see by (3.32) that $\binom{\mathbf{c}^{T',e'}}{\mathbf{c}^{T'',e''}}$ results from subdividing the polynomial with coefficients $\mathbf{c}^{T,e}$ thus representing again the same function on e. The argument for the remaining cases is analogous, and to complete the proof we only have to note that Remark 3.4 gives the precise conditions for continuity which, combined with (3.23), show that any continuous function in V_h is reproduced by $\mathcal{A}_k$. ☐

3.2.5 An energy stable frame for $V_{h,k}^{\mathrm{c}}$

Since the conforming part of the splitting possesses a multilevel structure we introduce an analogous grouping into conforming vertices, edge nodes and interior nodes for the lower levels as well. Recalling that $\mathcal{T}_h^j$ and Ω_h^j were defined in Section 3.1, we let

$$\mathcal{N}_{h,k}^{j,\mathrm{c}}:=\mathcal{N}_{h,1}^{j,\mathrm{c}}\cup\mathcal{N}_{h,k}^{j,\mathcal{E}}\cup\mathcal{N}_{h,k}^{j,\mathcal{T}},\qquad j=1,\cdots,j_h, \tag{3.33}$$

where $\mathcal{N}_{h,1}^{j,\mathrm{c}}$, $\mathcal{N}_{h,k}^{j,\mathcal{E}}:=\bigcup\{\mathcal{N}_k^0(e):e$ edge of $T\in\mathcal{T}_h^j$, $e\not\subset\partial\Omega_h^j\}$ and $\mathcal{N}_{h,k}^{j,\mathcal{T}}:=\bigcup\{\mathcal{N}_k^0(T):T\in\mathcal{T}_h^j\}$ denote respectively the sets of j level regular, interior vertices, (also defined in Section 3.1), nodes in the interior of j level edges inside Ω_h^j and nodes in the interior of j level triangles. Next, setting

$$\mathcal{I}_{h,k}^{\mathrm{c}}:=\{i=(j,p):j=0,\cdots,j_h,\ p\in\mathcal{N}_{h,k}^{j,\mathrm{c}}\}, \tag{3.34}$$

and choosing P_p^T to be either a Lagrange L_p^T or a Bernstein B_p^T polynomial piece supported on T, see (3.7) and (3.9), we define ϕ_i^{c} for any $i=(j,p)\in\mathcal{I}_{h,k}^{\mathrm{c}}$ by

$$\phi_i^{\mathrm{c}}|_T:=P_p^T,\quad \text{for any } T\in\mathcal{T}_h^j \text{ such that } p\in T. \tag{3.35}$$

Actually this defines three types of nodal functions according to the decomposition (3.33), but in every case the continuity follows from the conformity of $\mathcal{T}_h^j$ and the fact that every node of $\mathcal{N}_{h,k}^{j,c}$ is inside Ω_h^j. The following fact is well-known, see e.g. [9,18].

Proposition 3.2. The multilevel collection

$$\mathcal{S}_{h,k}^{c} := \{\phi_i^c : i \in \mathcal{I}_{h,k}^c\} \tag{3.36}$$

defined by (3.34) and (3.35) is an energy stable splitting for $V_{h,k}^c := \mathcal{A}_k V_h = V_h \cap H_0^1(\Omega)$, i.e.

$$|||v|||_h^2 \sim \inf_{v=\sum_{i\in\mathcal{I}_{h,k}^c} c_i\phi_i^c} \sum_{i\in\mathcal{I}_{h,k}^c} c_i^2 |||\phi_i^c|||_h^2 \qquad \forall v \in V_{h,k}^c \tag{3.37}$$

holds with constants depending only on k, the shape properties of $\mathcal{T}^0$ and the constants from (1.6) (recall that $|||\cdot|||_h^2$ and $a(\cdot,\cdot)$ coincide on $V_h \cap H_0^1(\Omega)$).

3.2.6 Jackson Estimate for $\mathcal{A}_k$

It remains to confirm that $\mathcal{A}_k$ satisfies (2.5). Our main tool can be formulated as follows.

Lemma 3.1. *For any v in V_h, any $T \in \mathcal{T}_h$ and any edge e of T, there is some $C = C(k)$ for which*

$$\|v^T - (\mathcal{A}_k v)^T\|_{L_\infty(e)} \le C \sum_{e' \in \mathcal{E}_h^*(T)} \|[v]\|_{L_\infty(e')}. \tag{3.38}$$

Proof. Here we distinguish four possible configurations for the edge e.

(I) e (as a closed set) contains no hanging vertex,

(II) e contains a hanging vertex at its midpoint (as in Fig. 2, right). Note that the grading of $\mathcal{T}_h$ implies then that both ends of e are regular vertices, which is equivalent to saying that $e \in \mathcal{E}_h^u \setminus \mathcal{E}_h$.

(III) e contains a hanging vertex at one end, in such a way that it is strictly included in the edge of the adjacent triangle (like e' or e'' in Fig. 2). Now the grading of $\mathcal{T}_h$ implies that e contains no other hanging vertex, and $e \in \mathcal{E}_h \setminus \mathcal{E}_h^u$.

(IV) e contains (at least) one hanging vertex at one end, in such a way that it coincides with the edge of the adjacent triangle, like $\tilde{e}$ in Fig. 2.

In case (I) we know from (3.15) that the trace of $(v - \mathcal{A}_k v)^T$ on e is a univariate Bernstein polynomial whose control coefficients lie on e, see (3.13). Also recall that we denote by b_p^T and c_p^T the control coefficients of v and $\mathcal{A}_k v$, respectively. Since the end points n, m of e are by assumption regular, using (3.26), (3.27) and (3.17), we compute

$$\begin{aligned}|b_n^T - c_n^T| &= \Big|\frac{1}{\#\mathcal{T}_h(n)} \sum_{T'\in\mathcal{T}_h(n)} (b_n^T - b_n^{T'})\Big| \\ &= \Big|\frac{1}{\#\mathcal{T}_h(n)} \sum_{T'\in\mathcal{T}_h(n)} \big(v^T(n) - v^{T'}(n)\big)\Big| \le \sum_{e'\in\mathcal{E}_h(n)} \|[v]\|_{L_\infty(e')} \end{aligned} \tag{3.39}$$

(where the last step simply follows from turning around n), and clearly the same holds for the other end point m as well. As to the control coefficients in the interior of e, we adhere to the notation used in (3.29) describing this case, and from (3.30) we have

$$\mathbf{b}^{T,e}-\mathbf{c}^{T,e}=\frac{1}{2}\Big(\mathbf{b}^{T,e}-\mathbf{b}^{T',e}\Big), \tag{3.40}$$

where T' is such that $e=T\cap T'$. Using Remark 3.2 we find

$$\begin{aligned}\|\mathbf{b}^{T,e}-\mathbf{b}^{T',e}\|_{\ell_\infty}&=\|\{b_p^T-b_p^{T'}\}_{p\in\mathcal{N}_k^0(e)}\|_{\ell_\infty}\le\|\{b_p^T-b_p^{T'}\}_{p\in\mathcal{N}_k(e)}\|_{\ell_\infty}\\&\lesssim\|\sum_{p\in\mathcal{N}_k(e)}(b_p^T-b_p^{T'})B_p^e\|_{L_\infty(e)}=\|[v]\|_{L_\infty(e)}.\end{aligned} \tag{3.41}$$

Now, since $\mathbf{b}^{T,e}-\mathbf{c}^{T,e}$ is the array of Bézier coefficients of $(v-\mathcal{A}_kv)^T$ that lie in the interior of the edge e, Remark 3.2 also says that

$$\|(v-\mathcal{A}_kv)^T\|_{L_\infty(e)}\sim\max\Big\{|b_n^T-c_n^T|,|b_m^T-c_m^T|,\|\mathbf{b}^{T,e}-\mathbf{b}^{T',e}\|_{\ell_\infty}\Big\},$$

and the assertion (3.38) follows therefore in this case from (3.39), (3.40) and (3.41).

Let us consider next case (II) where e has a hanging vertex n as its midpoint separating the two edges e',e'' as illustrated in Fig. 2, right. The principle is similar as before, and since the endpoints of e (now denoted n', n'') are regular, the differences $|b_p^T-c_p^T|$, $p\in\{n',n''\}$, are estimated exactly as in (3.39). So we need to consider again only p in the interior of e. Now we are in the situation (3.32) and obtain upon using (3.23),

$$\mathbf{b}^{T,e}-\mathbf{c}^{T,e}=\frac{1}{2}\Big(\mathbf{b}^{T,e}-\mathbf{Q}\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}\Big)=\frac{1}{2}\mathbf{Q}\Big(\mathbf{M}\mathbf{b}^{T,e}-\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}\Big).$$

Applied in the same spirit than in (3.41), but now taking into account the meaning of the subdivision operator $\mathbf{M}$, Remark 3.2 gives

$$\Big\|\mathbf{M}\mathbf{b}^{T,e}-\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}\Big\|_{\ell_\infty}\lesssim\|[v]\|_{L_\infty(e)},$$

hence

$$\|(v-\mathcal{A}_kv)^T\|_{L_\infty(e)}\lesssim\sum_{\check{e}\in\mathcal{E}_h(n')\cup\mathcal{E}_h(n'')}\|[v]\|_{L_\infty(\check{e})} \tag{3.42}$$

with constants depending only on k, and this yields (3.38).

Let us consider next case (III). Since for this case we adhere to the notation in Fig. 2, we should estimate $\|(v-\mathcal{A}_kv)^{T'}\|_{L_\infty(e')}$, but in the light of the foregoing discussion we might as well estimate directly $\|(v-\mathcal{A}_kv)^{T'\cup T''}\|_{L_\infty(e)}$. Thus we compute, in view of (3.32),

$$\begin{aligned}\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}-\begin{pmatrix}\mathbf{c}^{T',e'}\\\mathbf{c}^{T'',e''}\end{pmatrix}&=\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}-\mathbf{M}\mathbf{c}^{T,e}=\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}-\frac{1}{2}\mathbf{M}\Big(\mathbf{b}^{T,e}+\mathbf{Q}\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}\Big)\\&=\frac{1}{2}\Big(\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}-\mathbf{M}\mathbf{b}^{T,e}\Big)+\frac{1}{2}\Big(\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}-\mathbf{M}\mathbf{Q}\begin{pmatrix}\mathbf{b}^{T',e'}\\\mathbf{b}^{T'',e''}\end{pmatrix}\Big)\\&=:B_1+B_2.\end{aligned}$$

Now, upon using again (3.23) we see that $B_2 = B_1 + \mathbf{MQ}B_1$, and B_1 can be bounded as in (3.42). Since for the (regular) endpoints of e, we can apply (3.39) (replacing n by n' or n''), inequality (3.38) holds in this case as well.

Finally we consider case (IV) and adhere again to the notation given in Fig. 2, i.e. we estimate the left hand side of (3.38) on the edge $\tilde{e}$. Note that two situations may occur concerning the triangle T in (3.38): either it is adjacent to one coarser triangle, like $\tilde{T}$, or not, like T'. Therefore we will estimate $\|(v-\mathcal{A}_k v)^{\check{T}}\|_{L_\infty(\tilde{e})}$ with $\check{T} = T'$ or $\tilde{T}$. Again, due to (3.14), we proceed by estimating differences of control coefficients located on the edge $\tilde{e}$. For the differences $b_p^{\check{T}} - c_p^{\check{T}}$, with $p \in \mathcal{N}_k^0(\tilde{e})$ located in the interior of $\tilde{e}$ we apply (3.40)-(3.41), which leads to a bound by $\|[v]\|_{L_\infty(\tilde{e})}$. Now for the control coefficients at the endpoints of $\tilde{e}$, we observe that (3.39) applies at a regular vertex (if any), and that at a hanging vertex like n, we have

$$\begin{aligned} |(v-\mathcal{A}_k v)^{\tilde{T}}(n)| &\le |v^{\tilde{T}}(n) - v^{T'}(n)| + |(v-\mathcal{A}_k v)^{T'}(n)| \\ &\le \|[v]\|_{L_\infty(\tilde{e})} + \|(v-\mathcal{A}_k v)^{T'}\|_{L_\infty(e')}. \end{aligned}$$

Therefore using (3.38) for (T',e') finally yields

$$\|(v-\mathcal{A}_k v)^{\check{T}}\|_{L_\infty(\tilde{e})} \lesssim \sum_{\check{e} \in \mathcal{E}_h^*(T')} \|[v]\|_{L_\infty(\check{e})}$$

for $\check{T} = T'$ or $\tilde{T}$, where we have used (twice) that $\tilde{e}$ is comprised in $\mathcal{E}_h^*(T')$, see (2.2). Now since the vertex n' has to be regular from the grading of $\mathcal{T}_h$, we observe that $\mathcal{N}_{h,1}^*(T')$ is always included in $\mathcal{N}_{h,1}^*(\tilde{T})$, so that $\mathcal{E}_h^*(T')$ is always included in $\mathcal{E}_h^*(\tilde{T})$. This establishes (3.38) in the last case, and finishes the proof. $\square$

We can now prove the main result of this section

Proposition 3.3. The projector $\mathcal{A}_k$ defined in Section 3.2.4 satisfies (2.5) with a constant that depends only on k and on the shape properties of $\mathcal{T}^0$.

Proof. Since by definition, see (3.28), the control coefficients in the interior of a triangle are left unchanged by $\mathcal{A}_k$, we see that $(v-\mathcal{A}_k v)^T$ has for every triangle $T \in \mathcal{T}_h$ the form

$$(v-\mathcal{A}_k v)^T =: w = \sum_{p \in \partial\mathcal{N}_k(T)} w_p^T B_p^T, \quad \text{where} \quad \partial\mathcal{N}_k(T) := \mathcal{N}_k(T) \cap \partial T.$$

Thus, keeping (3.13) in mind and applying Remark 3.2 to the Bézier representation on T as well as on the edges of T, yields

$$\|w\|_{L_\infty(T)} \sim \max_{e \subset T} \|w\|_{L_\infty(e)} \lesssim \sum_{e \in \mathcal{E}_h^*(T)} \|[v]\|_{L_\infty(e)} \sim \sum_{e \in \mathcal{E}_h^*(T)} |e|^{-1/2} \|[v]\|_{L_2(e)}, \tag{3.43}$$

where we have used Lemma 3.1 in the second but last and the equivalence between norms of polynomials in the last step. Also note that a standard scaling argument gives

$|T|^{-1/2}\|w\|_{L_2(T)} \sim \|w\|_{L_\infty(T)}$ with a constant that only depends on k and the shape properties of $\mathcal{T}^0$. Recalling then (2.3), and the fact that the sets $\mathcal{E}_h^*(T)$ have uniformly bounded cardinality, this latter estimate together with (3.43) confirms (2.5). $\square$

Now, using Theorem 2.1, the facts recalled in Section 3.1, Remark 3.3 and finally Propositions 3.2 and 3.3, we have proven the following result.

Theorem 3.1. *Let* $\mathcal{S}_{h,1}^{\mathrm{c}}$ *and* $\mathcal{S}_{h,k}^{\mathrm{c}}$ *be defined as in* (3.1) *and* (3.36)*. Then if* $\phi_{T,p}$ *is either a Lagrange polynomial piece* L_p^T*, see* (3.7)*, or a Bernstein piece* B_p^T*, see* (3.9)*, both the collections*

$$\mathcal{S}_{h,1}^{\mathrm{c}} \cup \{\mathrm{span}(\phi_{T,p}) : T \in \mathcal{T}_h,\ p \in \mathcal{N}_k(T)\} \quad \textit{and} \quad \mathcal{S}_{h,k}^{\mathrm{c}} \cup \{\mathrm{span}(\phi_{T,p}) : T \in \mathcal{T}_h,\ p \in \partial\mathcal{N}_k(T)\}$$

are stable splitting for V_h *in the sense of* (1.7)*. In particular, one can apply Theorem 1.2 on both splittings to build optimal preconditioners for the DG problem* (1.3)*.*

Let us finally remark that the above results for $\mathcal{A}_k$ could be extended to varying degrees as well under the slight restriction that the triangles sharing an edge with a hanging vertex carry the same degree. In the terms of Fig. 2 this would mean $k(T) = k(T') = k(T'')$.

4 Numerical experiments

The subsequent numerical experiments refer to the simple example of Poisson's equation, i.e. $a(v,w) = \int_\Omega \nabla v \nabla w$, on the L-shaped domain $\Omega \subset \mathbb{R}^2$ created by cutting out the upper right square $[1,2)^2$ from the square $[0,2]^2$. For simplicity, we choose $f = 1$ as the right hand side.

We use Bernstein polynomial basis functions in the non-conforming part according to Theorem 3.1 as well as in the higher order conforming part according to (3.35). By $A1$ and Ak we denote the preconditioners based on the splittings induced by the projectors $\mathcal{A}_1$ and $\mathcal{A}_k$, respectively. By the *order* of the preconditioner we mean the polynomial order of the corresponding conforming subspace. We consider only the cases where the degree function is constant with $\bar{k} = k \in \{1,2,3\}$ and apply the preconditioners $A1$ and Ak in a standard CG method. As stopping criterion, for a given ℓ-th CG iterate u_h^ℓ we calculate the residual vector $r_h^\ell = f_h - A_h u_h^\ell$ and check the norm of its coefficient vector $\|\mathbf{r}_h^\ell\|_{\mathcal{C}_h^\Psi}$, where for given basis Ψ_h of V_h the norm $\|\mathbf{v}\|_{\mathcal{C}_h^\Psi} := \langle \mathbf{v}, \mathcal{C}_h^\Psi \mathbf{v}\rangle^{\frac{1}{2}}$ for the coefficient vectors is induced by the symmetric and positive definite matrix $\mathcal{C}_h^\Psi := \big(\langle \psi_i, \mathcal{C}_h \psi_j\rangle\big)_{i,j\in\mathcal{I}}$ generated by the preconditioner $\mathcal{C}_h : V_h \to V_h$ with $P_h = \mathcal{C}_h A_h$ (see (1.4) and (2.10)). The residual norm $\|\mathbf{r}_h^\ell\|_{\mathcal{C}_h^\Psi}$ is equivalent to $|||u_h^\ell - u_h|||_h$, where u_h is the exact solution of the discrete system in V_h [7]. For $k = 3$ the preconditioner $A3$ is perhaps of special interest, because the corresponding trial spaces has the smallest order, for which $\mathcal{N}_{h,k}^{j,\mathcal{T}} \neq \emptyset$.

In a preliminary study we have inspected the dependence of the condition number on the stabilization parameter γ. It seems that the three choices $\gamma = 7.5$ for $k = 1$, $\gamma = 10$

for $k=2$ and $\gamma=15$ for $k=3$ come close to minimizing the condition numbers for the respective k in all tests.

To study the quantitative effect of the specific choice of the auxiliary bilinear forms $b_i(\cdot,\cdot)$, we refer to the case C0 for

$$b_i(v,w)=\begin{cases} a(v,w)_T+\sum\limits_{e\in\mathcal{E}_h,e\subset\partial T}\dfrac{\gamma}{|e|}\displaystyle\int_e [v][w] & \text{when } i=(T,p)\in\mathcal{I}_h^{\mathrm{nc}},\\ a(v,w)_{\omega(i)} & \text{when } i\in\mathcal{I}_h^{\mathrm{c}},\end{cases}$$

and to C1 for

$$b_i(v,w)=\begin{cases} |T|^{-1}\langle v,w\rangle_T+\sum\limits_{e\in\mathcal{E}_h,e\subset\partial T}\dfrac{\gamma}{|e|}\displaystyle\int_e [v][w] & \text{when } i=(T,p)\in\mathcal{I}_h^{\mathrm{nc}},\\ |\omega(i)|^{-1}\langle v,w\rangle_{\omega(i)} & \text{when } i\in\mathcal{I}_h^{\mathrm{c}}.\end{cases}$$

Specifically, the subsequent numerical experiments are to shed some light on the quantitative dependence of the preconditioner on:

(i) the choice of $b_i(\cdot,\cdot)$;

(ii) the degree k;

(iii) the dimension m of the subspaces V_i in the underlying stable splitting;

(iv) the type of the mesh.

As for (i), we consider only the above cases C0, C1, where C0 is close to the DG-bilinear form while C1 is a little simpler and involves local scaled L_2-inner products.

In (ii) we do not expect the scheme to be fully robust in k but wish to see the effect for the range $k\leq 3$.

(iii) addresses the question whether it is of advantage to choose subspaces of dimension larger than one in the nonconforming part of the stable splitting, e.g. the local polynomial spaces associated with the triangles. Thus $m=1$ refers to point relaxation while $m=\dim V_i>1$ boils down to solving m-dimensional linear systems and thus to block relaxation.

Finally, in (iv) we wish to compare the performance of the respective variants of preconditioners for nonuniform adaptively generated meshes with hanging nodes.

To this end, our initial study refers to the same standardized preselected hierarchy of nonuniform meshes refined around the reentrant corner for all versions of preconditioners. More precisely, we start on a uniform triangulation with 24 triangles. In each loop we refine first all cells that are in contact with the reentrant corner and apply then the algorithm from [7] to ensure that the mesh fulfills the necessary grading property. The initial guess is always the zero function. Tables 1, 2 and 3 display the number of iterations needed to reduce the residual by a factor of 10^{-8} for the degrees $k=1,2,3$. As expected the number of iterations increases with increasing k but stays bounded for each k.

Table 1: Number of CG iterations for $k=1$ and $\gamma=7.5$.

Loop	#DOF	Preconditioner none	A1, $m=1$ C0	A1, $m=1$ C1	A1, $m=3$ C0	A1, $m=3$ C1
		#its	#its	#its	#its	#its
1	72	30	25	30	23	27
5	180	52	34	47	31	44
10	315	70	37	53	35	55
15	450	87	37	54	35	58
20	585	101	37	57	35	58
25	720	116	37	59	35	60
30	855	130	37	61	35	62

Table 2: Number of CG iterations for $k=2$ and $\gamma=10.0$.

Loop	#DOF	Preconditioner none	A1, $m=1$ C0	A1, $m=1$ C1	A1, $m=3$ C0	A1, $m=3$ C1	A2, $m=1$ C0	A2, $m=1$ C1	A2, $m=6$ C0	A2, $m=6$ C1
		#its	#its	#its	#its	#its	#its	#its	#its	#its
1	144	51	47	52	33	37	38	86	30	67
5	360	78	58	76	47	67	45	117	36	99
10	630	99	64	88	55	87	46	126	36	109
15	900	117	65	91	57	96	46	134	36	115
20	1170	135	66	93	57	98	46	143	36	122
25	1440	152	66	92	57	99	46	144	36	119
31	1764	171	66	95	57	101	46	145	36	124

Table 3: Number of CG iterations for $k=3$ and $\gamma=15.0$.

Loop	#DOF	Preconditioner none	A1, $m=1$ C0	A1, $m=1$ C1	A1, $m=10$ C0	A1, $m=10$ C1	A3, $m=1$ C0	A3, $m=1$ C1	A3, $m=9$ C0	A3, $m=9$ C1
		#its	#its	#its	#its	#its	#its	#its	#its	#its
1	240	97	89	132	47	67	66	217	45	151
5	600	120	100	163	60	116	78	297	54	206
10	1050	142	103	173	70	147	83	308	57	227
15	1500	160	104	181	74	164	84	310	57	225
20	1950	176	105	184	74	164	84	313	57	228
25	2400	195	106	186	74	166	84	316	57	230
30	2850	214	106	188	74	168	84	318	57	232

The choice of the auxiliary bilinear form seems to have the strongest impact on the preconditioning effect. In fact, for all tested k and m the case C0 using the auxiliary bilinear form which corresponds to the localized DG bilinear form is superior to C1 with an increasing difference in performance between the cases C0 and C1 when the polynomial degree k grows. Moreover, we see that it is generally favorable for either auxiliary inner product to choose $m>1$, i.e. to use block relaxation, except for C1 with $A1$ when $k=2$.

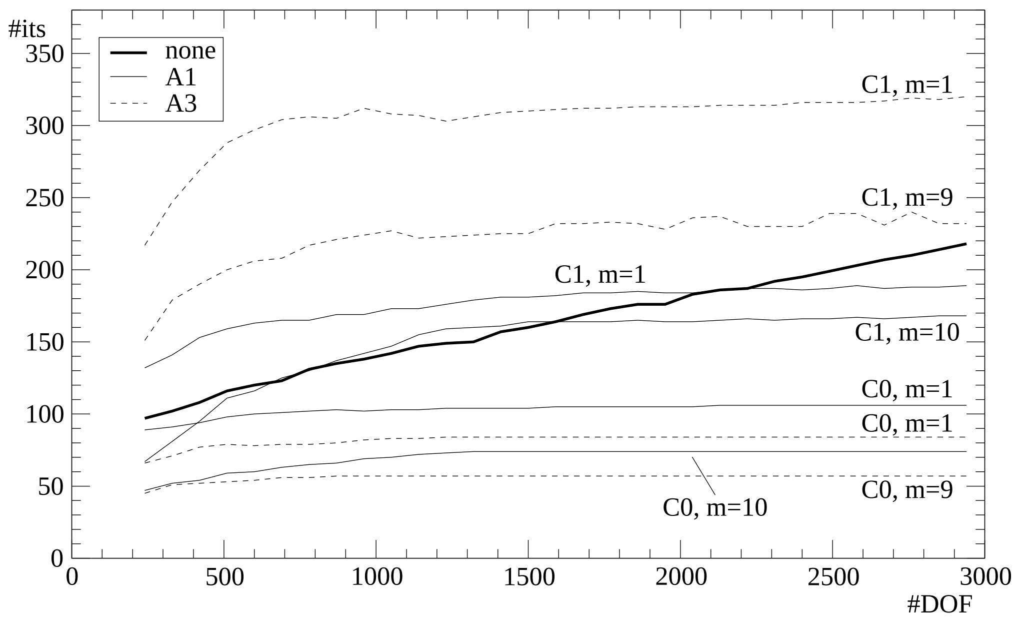

Figure 3: Number of CG iterations for $k=3$, $\gamma=15$ and different preconditioners.

Again, the favorable effect of block relaxation increases with growing polynomial degree k.

The results favor the choice of the auxiliary bilinear form in case C0 in combination with a higher-order preconditioner and block relaxation. It should be noted though that one iteration step in Ak for $k>1$ is generally more expensive than a step in $A1$.

An overview of results for various combinations (choices of auxiliary bilinear forms, point-, block relaxation) for $k \leq 3$ and both types $A1$ and $A3$ in the above setting is displayed in Fig. 3.

The locally refined meshes already give rise to small truncation errors for relatively small numbers of degrees of freedom and we are content here with problem sizes for which the iteration counts start to settle. Fig. 3 shows also that for the above range of problem sizes some versions do actually worse than no preconditioning with a clear indication though that this would change for larger problems.

To test also significantly larger systems we consider in a second test setting a hierarchy of *uniform meshes* for $k=3$, however, now in combination with *nested iteration* in order to properly couple solution and discretization error accuracy. That means the result on each mesh is used as an initial guess for the next refined mesh on which the error is reduced to a level that is significantly smaller than the expected associated discretization error. To estimate this, we apply the a-posteriori error estimator derived in [15] and stop the CG method, when the residual norm $\|\mathbf{r}_h^\ell\|_{\mathcal{C}_h^\Psi}$ reaches 10^{-2} of the estimated error.

In the following we concentrate on those variants that have turned out to be most effective in the first study, namely the case C0 with $A1$ and Ak with maximal m. For a

Table 4: Nested iteration on uniform mesh: Number of CG iterations and estimated error for $k=1$ and $\gamma=7.5$.

Loop	#DOF	Preconditioner C0, A1, $m=1$		C0, A1, $m=3$	
		#its	estimated error	#its	estimated error
0	72	37	5.61E-01	34	5.61E-01
1	288	7	3.79E-01	6	3.79E-01
2	1152	7	2.28E-01	6	2.29E-01
3	4608	6	1.30E-01	7	1.30E-01
4	18432	7	7.32E-02	7	7.31E-02
5	73728	7	4.15E-02	7	4.15E-02
6	294912	7	2.40E-02	7	2.40E-02
7	1179648	8	1.42E-02	8	1.42E-02

Table 5: Nested iteration on uniform mesh: Number of CG iterations and estimated error for $k=2$ and $\gamma=10.0$.

Loop	#DOF	Preconditioner C0, A1, $m=6$		C0, A2, $m=6$	
		#its	estimated error	#its	estimated error
0	144	55	2.24E-01	52	2.24E-01
1	576	7	1.05E-01	8	1.05E-01
2	2304	7	5.93E-02	9	5.99E-02
3	9216	8	3.63E-02	10	3.66E-02
4	36864	8	2.27E-02	10	2.30E-02
5	147456	8	1.43E-02	11	1.44E-02
6	589824	9	9.00E-03	11	9.08E-03
7	2359296	9	5.67E-03	12	5.71E-03

Table 6: Nested iteration on uniform mesh: Number of CG iterations and estimated error for $k=3$ and $\gamma=15.0$.

Loop	#DOF	Preconditioner C0, A1, $m=10$		C0, A3, $m=9$	
		#its	estimated error	#its	estimated error
0	240	78	1.70E-01	84	1.70E-01
1	960	6	1.03E-01	9	1.07E-01
2	3840	6	6.37E-02	10	6.75E-02
3	15360	6	3.98E-02	12	4.15E-02
4	61440	7	2.50E-02	13	2.59E-02
5	245760	7	1.57E-02	14	1.62E-02
6	983040	8	9.88E-03	14	1.03E-02
7	3932160	8	6.21E-03	15	6.43E-03

comparison, we add $A1$ with $m=1$ for $k=1$. The results are recorded in Tables 4, 5 and 6. Note that $A3$ has been used with the reduced frame for the nonconforming subspace omitting the (globally continuous) shape function associated with the interior node in each triangle, which explains $m=9$ for $A3$ in Table 6. The number of iterations in each loop stays bounded and all versions show essentially the same performance. However, in contrast to the previous test set the more economic variant $A1$ is slightly better than

Table 7: Nested iteration on adaptive mesh: Number of CG iterations and estimated error for $k=1$ and $\gamma=7.5$.

Loop	Preconditioner					
	C0, A1, $m=1$			C0, A1, $m=3$		
	#DOF	#its	estimated error	#DOF	#its	estimated error
0	72	37	5.61E-01	72	34	5.61E-01
5	513	5	3.10E-01	558	5	3.02E-01
10	2718	5	1.58E-01	2826	5	1.55E-01
15	12825	6	7.69E-02	13014	5	7.63E-02
20	54630	5	3.81E-02	54630	5	3.81E-02
25	218322	6	1.93E-02	217341	5	1.94E-02
30	846333	4	9.90E-03	838503	4	9.95E-03

Table 8: Nested iteration on adaptive mesh: Number of CG iterations and estimated error for $k=2$ and $\gamma=10.0$.

Loop	Preconditioner					
	C0, A1, $m=6$			C0, A2, $m=6$		
	#DOF	#its	estimated error	#DOF	#its	estimated error
0	144	55	2.24E-01	144	52	2.24E-01
5	648	6	6.89E-02	648	8	6.89E-02
10	2124	6	2.33E-02	2160	7	2.31E-02
15	6246	6	8.34E-03	6408	8	8.11E-03
20	17550	7	2.93E-03	18216	8	2.84E-03
25	45972	7	1.09E-03	47520	8	1.06E-03
30	126756	8	4.20E-04	130896	8	4.07E-04

Table 9: Nested iteration on adaptive mesh: Number of CG iterations and estimated error for $k=3$ and $\gamma=15.0$.

Loop	Preconditioner					
	C0, A1, $m=10$			C0, A3, $m=9$		
	#DOF	#its	estimated error	#DOF	#its	estimated error
0	240	78	1.70E-01	240	84	1.70E-01
5	570	5	6.33E-02	570	8	6.42E-02
10	1440	5	2.10E-02	1440	9	2.11E-02
15	3420	6	5.46E-03	3450	10	5.44E-03
20	7890	6	1.50E-03	8400	9	1.36E-03
25	18780	6	4.50E-04	19710	8	4.20E-04
30	39030	6	1.35E-04	41460	9	1.23E-04

Ak for $k>1$. Moreover, the number of iterations needed to realize discretization error accuracy is nearly independent of k. In view of the geometric increase of the number of degrees of freedom the computational work on the finest level dominates the overall work reflecting a performance that is actually better than one might predict from the first test set.

Finally, we address the most realistic setting, namely nested iteration in combination with adaptively refined meshes. This time, adaptation is controlled by a standard bulk chasing schemed based on the a-posteriori error bounds derived in [15, 17]. In fact, we select the cells with the largest error indicators until a ϑ-fraction of the total estimated error is captured. We fix $\vartheta = 0.55$. The absolute tolerance for the CG method is again set to 10^{-2} of the estimated error in the previous loop. Note that now different versions of preconditioners will generate slightly different meshes because the approximate solutions may differ somewhat giving rise to slightly different error indicators. The results are displayed in Tables 7, 8 and 9. The respective numbers of degrees of freedom indicate that the difference between the various meshes are quite marginal.

As expected the mesh is refined mainly at the reentrant corner which is known to cause a singularity. As soon as after some loops the estimated error near the reentrant corner has been sufficiently reduced some mild refinement (typically by just one level) takes place on very coarse triangles. As expected, the larger k the less the coarse cells away from the reentrant corner are refined.

Again, all variants exhibit a very similar performance in terms of the number of iterations. As in the case of uniform refinements, in view of the higher cost per iteration for Ak when $k > 1$, the overall most economic version seems to be $(C0, A1, m = 3)$.

The last group of tests confirms the findings for the case of uniform refinements. First, due to the geometric increase of degrees of freedom in all cases the overall work is dominated by that on the final mesh resulting in an overall much better work/accuracy rate than indicated by the first set of tests. Moreover, the dependence on the degree of the trial spaces seems to be again negligible. Note also that the error decay rate is roughly $N^{-(k-1)/2}$ when N is the number of unknowns which is optimal and much better than in the previous case of uniform refinements, due to the low Sobolev regularity of the solution.

At this point a few comments relating the above findings to previous work in the literature seems to be in order. $A1$ being the preferable choice, the realization of the preconditioner is similar to that for familiar additive Schwarz schemes in the conforming setting. In contrast to [11, 16] the primary objective of this work is a rigorous foundation for a possibly flexible scope of discretizations covering, in particular, locally refined meshes. Additive schemes offer perhaps the best chance from an analysis point of view. It should therefore not surprise that in strict quantitative terms multiplicative versions or two-level schemes where the auxiliary problem is solved exactly perform even better, also in cases not covered by the respective analysis, see the numerical tests in [11, 16]. Nevertheless, in combination with nested refinements the above variants, certainly leaving room for quantitative improvements, appear to work quite well for all admissible meshes.

Acknowledgments

This work has been supported in part by the French-German PROCOPE contract 11418YB, by the European Commission Human Potential Programme under contract HPRN-CT-2002-00286, "Breaking Complexity", by the SFB 401 and the Leibniz Programme funded by DFG.

References

[1] P. Antonietti and B. Ayuso, Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case, M2AN Math. Model. Numer. Anal., 41 (2007), 21-54.

[2] P. Antonietti and B. Ayuso, Multiplicative Schwarz methods for discontinuous Galerkin approximations of elliptic problems, Submitted (2006).

[3] D. N. Arnold, An interior penalty finite element method with discontinuous elements, SIAM J. Numer. Anal., 19 (1982), 742-760.

[4] D. N. Arnold, F. Brezzi, B. Cockburn and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal., 39 (2002), 1749-1779.

[5] J. H. Bramble, J. E. Pasciak and J. Xu, Parallel multilevel preconditioners, Math. Comput., 55 (1990), 1-22.

[6] S. C. Brenner, Two-level additive Schwarz preconditioners for nonconforming finite element methods, Math. Comput., 65 (1996), 897-921.

[7] K. Brix, M. Campos Pinto and W. Dahmen, A multilevel preconditioner for the interior penalty discontinuous Galerkin method, IGPM Report # 277, RWTH Aachen, July 2007, to appear in SIAM J. Numer. Anal.

[8] W. Dahmen, Adaptive approximation by multivariate smooth splines, J. Approx. Theory, 36 (1982), 119-140.

[9] W. Dahmen, Wavelet and Multiscale Methods for Operator Equations, Acta Numerica, Cambridge University Press, 6 (1997), 55-228.

[10] W. Dahmen and A. Kunoth, Multilevel Preconditioning, Numer. Math., 63 (1992), 315-344.

[11] V.A. Dobrev, R.D. Lazarov, P.S. Vassilevski and L.T. Zikatanov, Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations, Numer. Linear Algebra Appl., 13 (2006), 753–770.

[12] G. Farin, Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide, Academic Press, New York, 1988.

[13] J. Gopalakrishnan and G. Kanschat, A multilevel discontinuous Galerkin method, Numer. Math., 95 (2003), 527-550.

[14] M. Griebel and P. Oswald, On the abstract theory of additive and multiplicative Schwarz algorithms, Numer. Math., 70 (1995), 163-180.

[15] P. Houston, D. Schötzau and T.P. Wihler, Energy norm a-posteriori error estimation of hp-adaptive discontinuous Galerkin methods for elliptic problems, Math. Models Methods Appl. Sci., 17 (2007), 33-62.

[16] G. Kanschat, Multilevel methods for discontinuous Galerkin FEM on locally refined meshes, Comput. & Structures, 82 (2004), 2437-2445.

[17] O. Karakashian and F. Pascal, A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems, SIAM J. Numer. Anal., 41 (2003), 2374-2399.

[18] P. Oswald, Multilevel Finite Element Approximation - Theory and Applications, Teubner Skripten zur Numerik, B.G. Teubner, Stuttgart, 1994.

[19] P. Oswald, Preconditioners for nonconforming discretizations, Math. Comput., 65 (1996), 923-941.

[20] U. Rüde, Mathematical and computational techniques for multilevel adaptive methods, SIAM, Philadelphia, 1993

[21] E. Süli, C. Schwab and P. Houston, hp-DGFEM for partial differential equations with nonnegative characteristic form, in: Discontinuous Galerkin Methods: Theory, Computation

and Applications, B. Cockburn, G. E. Karniadakis and C.-W. Shu, eds., Lecture Notes in Computational Science and Engineering, Springer, Berlin, 11 (2000), 211-230.

[22] T.P. Wihler, P. Frauenfelder and C. Schwab, Exponential convergence of the hp-DGFEM for diffusion problems, Comput. Math. Appl., 46 (2003), 183-205.

[23] J. Xu, Iterative methods by space decomposition and subspace correction, SIAM Rev., 34 (1992), 581-613.

[24] J. Xu, The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids, Computing, 56 (1996), 215-235.

[25] H. Yserentant, On the multi-level splitting of finite element spaces, Numer. Math., 49 (1986), 379-412.