V-cycle multigrid algorithms for discontinuous Galerkin methods on non-nested polytopic meshes

P. F. Antonietti · G. Pennesi.

Abstract In this paper we analyse the convergence properties of V-cycle multigrid algorithms for the numerical solution of the linear system of equations arising from discontinuous Galerkin discretization of second-order elliptic partial differential equations on polytopal meshes. Here, the sequence of spaces that stands at the basis of the multigrid scheme is possibly non nested and is obtained based on employing agglomeration with possible edge/face coarsening. We prove that the method converges uniformly with respect to the granularity of the grid and the polynomial approximation degree p, provided that the number of smoothing steps, which depends on p, is chosen sufficiently large.

1 Introduction

The Discontinuous Galerkin (DG) method was introduced in 1973 by Reed and Hill for the discretization of hyperbolic equations [41]. Extensions of the method were quickly proposed to deal with elliptic and parabolic problems: some of the most relevant works include Arnold [10], Baker [12], Nitsche [40] and Wheeler

P. F. Antonietti
MOX-Laboratory for Modelling and Scientific Computing, Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy.
Tel.: (+39) 02 2399 4601
E-mail: paola.antonietti@polimi.it

G. Pennesi
MOX-Laboratory for Modelling and Scientific Computing, Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy.
Tel.: (+39) 02 2399 4604
E-mail: giorgio.pennesi@polimi.it

[45], whose contributions put the basis for the development of the interior penalty DG methods. In the last 40 years the scientific and industrial community has shown an exponentially growing interest in DG methods - see for example [28, 29, 37, 42] for an overview. On one side, the features of DG methods have been naturally enhanced by the recent development of High Performance Computing technologies as well as the growing request for high-order accuracy. In particular, as the discrete polynomial space can be defined locally on each element of the mesh, DG methods feature a high-level of intrinsic parallelism. Moreover, the local conservation properties and the possibility to use meshes with hanging nodes make DG methods interesting also from a practical point of view. Recently, it has been shown that DG methods can be extended to computational grids characterized by polytopic elements, cf. Ref. [1, 2, 3, 4, 6, 8, 14, 15, 16, 27, 33, 38, 46]. In particular, the efficient approach presented in [27] is based on defining a local polynomial discrete space by making use of the bounding box of each element [32]: this technique together with a careful choice of the discontinuity penalization parameter permits the use of polytopal elements which can be characterized by faces of arbitrarily small measure and as shown in [25], see also [6], possibly by an unbounded number of faces.

On the other hand, the development of fast solvers and preconditioners for the linear system of equations arising from high-order DG discretization is been developed. A recent strand of the literature has focused on multilevel techniques, including Schwarz domain decomposition methods, cf. Ref. [4, 7], and two-level and multigrid techniques, cf. Ref. [6, 9]. The efficiency of those methods is more evident in the case of polygonal grids, because the flexibility of the element shape couples very well with the possibility to easily define agglomerated meshes, which is the key ingredient for the developing of multigrid algorithms. In [6] a two-level scheme and W-cycle multigrid method is developed to solve the linear system of equations arising from high-order discretization introduced in [27]. One iteration of the proposed methods consists of an iterative application of the smoothing Richardson operator and the subspace correction step. In particular, the latter is based on a nested sequence of discrete polynomial spaces where the underlying polytopal grid of each subspace is defined by agglomeration. While being faster than other classical iterative methods, the agglomeration approach presents itself some limitations. When the finest grid is unstructured and characterized by polytopic elements, there is the possibility that its very small edges could be inherited by the coarser levels until the one where the linear system is solved with a direct method. In this case the presence of small faces negatively affects the condition number of the associated matrix: indeed, according to [27], the discontinuity penalization parameter is defined locally in each face as the inverse of its measure.

In this paper we aim to overcome this issue by solving the same linear system through a multilevel method characterized by a sequence of non-nested agglomerated meshes in order to make sure that the number of faces of the agglomerates does not blows up as the number of levels of our multigrid method increases. This can be achieved for example based on employing edge-coarsening techniques in the agglomeration procedures. The flexibility in the choice of the computational sub-grids leads to the definition of a non-nested multigrid method characterized by a sequence of non-nested multilevel discrete spaces, cf. Ref. [17, 49, 50], and where the discrete bilinear forms are chosen differently on each level, cf. Ref. [34, 35, 39]. The first non-nested multilevel method was introduced by Bank and Dupont in

[13]; a generalized framework was developed by Bramble, Pasciak and Xu in [23], and then widely used in the analysis of non-nested multigrid iterations, cf. Ref. [18, 19, 20, 21, 22, 24, 36, 43, 47, 48]. The method of [23], to whom we will refer as the BPX multigrid framework, is able to generalize also the multigrid framework that we will develop in this paper, but the convergence analysis relies on the assumption that $\mathcal{A}_j(I^j_{j-1}u, I^j_{j-1}u) \leq \mathcal{A}_{j-1}(u, u)$, which might not be guaranteed in the DG setting, as we will see in Sect. 4.2. Here $\mathcal{A}_j(\cdot, \cdot)$ and $\mathcal{A}_{j-1}(\cdot, \cdot)$ are two bilinear forms suitably defined on two consecutive levels, and I^j_{j-1} is the prolongation operator whose definition is not trivial, differently from the nested case. For this reason the convergence analysis will be presented based on employing the abstract setting proposed by Duan, Gao, Tan and Zhang in [30], which permits to develop a full analysis of V-cycle multigrid methods in a non-nested framework relaxing the hypothesis $\mathcal{A}_j(I^j_{j-1}u, I^j_{j-1}u) \leq \mathcal{A}_{j-1}(u, u)$. We will prove that our V-cycle scheme with non-nested spaces converges uniformly with respect to the discretization parameters provided that the number of smoothing steps, which depends on the polynomial approximation degree p, is chosen sufficiently large. This result extends the theory of [6] where W-cycle multigrid methods for high-order DG methods with nested spaces where proposed and analyzed.

The paper is organized as follows. In Sect. 2 we introduce the interior penalty DG scheme for the discretization of second-order elliptic problems on general meshes consisting of polygonal/polyhedral elements. In Sect. 3, we recall some preliminary analytical results concerning this class of schemes. In Sect. 4 we define the multilevel BPX framework for the V-cycle multigrid solver based on non-nested grids, and present the convergence analysis of the algorithm. The main theoretical results are validated through a series of numerical experiments in Sect. 5. In Sect. 6 we propose an improved version of the algorithm, obtained by choosing a smoothing operator based on a domain decomposition preconditioner.

2 Model problem and its DG discretization

We consider the weak formulation of the Poisson problem, subject to a homogeneous Dirichlet boundary condition: find $u \in V = H^2(\Omega) \cap H^1_0(\Omega)$ such that

$$\mathcal{A}(u, v) = \int_\Omega \nabla u \cdot \nabla v \; dx = \int_\Omega fv \; dx \qquad \forall v \in V, \tag{1}$$

with $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, a convex polygonal/polyhedral domain with Lipschitz boundary and $f \in L^2(\Omega)$. The unique solution $u \in V$ of problem (1) satisfies

$$\|u\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}. \tag{2}$$

In view of the forthcoming multigrid analysis, let $\{\mathcal{T}_j\}_{j=1}^J$ be a sequence of tessellation of the domain Ω, each of which is characterized by disjoint open polytopal elements κ of diameter h_κ, such that $\overline{\Omega} = \bigcup_{\kappa \in \mathcal{T}_j} \bar{\kappa}$, $j = 1, \ldots, J$. The mesh size of $\mathcal{T}_j$ is denoted by $h_j = \max_{\kappa \in \mathcal{T}_j} h_\kappa$. To each $\mathcal{T}_j$ we associate the corresponding discontinuous finite element space V_j, defined as

$$V_j = \{v \in L^2(\Omega) : v|_\kappa \in \mathcal{P}_{p_j}(\kappa), \kappa \in \mathcal{T}_j\},$$

where $\mathcal{P}_{p_j}(\kappa)$ denotes the local space of polynomials of total degree at most $p_j \geq 1$ on $\kappa \in \mathcal{T}_j$.

Remark 1 For the sake of brevity we use the notation $x \lesssim y$ to mean $x \leq Cy$, where $C > 0$ is a constant independent from the discretization parameters. Similarly we write $x \gtrsim y$ in lieu of $x \geq Cy$, while $x \approx y$ is used if both $x \lesssim y$ and $x \gtrsim y$ hold.

A suitable choice of $\{\mathcal{T}_j\}_{j=1}^J$ and $\{V_j\}_{j=1}^J$ leads to the *hp*-multigrid non-nested schemes. This method is based on employing, from one side, a set of non-nested partitions $\{\mathcal{T}_j\}_{j=1}^J$, such that the coarse level $\mathcal{T}_{j-1}$ is independent from $\mathcal{T}_j$, with the only constrain

$$h_{j-1} \lesssim h_j \leq h_{j-1} \qquad \forall \, j = 2, \ldots, J, \tag{3}$$

from the other side we assume that the polynomial degree vary from one level to another such that

$$p_{j-1} \leq p_j \lesssim p_{j-1} \qquad \forall \, j = 2, \ldots, J. \tag{4}$$

Additional assumptions on the grids $\{\mathcal{T}_j\}_{j=1}^J$ are outlined in the following paragraph.

2.1 Grid assumptions

For any $\mathcal{T}_j$, we define the *faces* of the mesh $\mathcal{T}_j$, $j = 1, \ldots, J$, as the intersection of the $(d-1)$-dimensional facets of neighbouring elements. This implies that, for $d = 2$, a *face* always consists of a line segment, however for $d = 3$, the *faces* of $\mathcal{T}_j$ are general shaped polygons. Thereby, we assume that each facets of an element $\kappa \in \mathcal{T}_j$ may be subdivided into a set of co-planar $(d-1)$-dimensional simplices and we refer to them as *faces*. In order to introduce the DG formulation, it is helpful to distinguish between boundary and interior element faces, denoted as $\mathcal{F}_j^B$ and $\mathcal{F}_j^I$, respectively. In particular, we observe that $F \subset \partial\Omega$ for $F \in \mathcal{F}_j^B$, while for any $F \in \mathcal{F}_j^I$ we assume that $F \subset \partial\kappa^\pm$, where $\kappa^\pm$ are two adjacent elements in $\mathcal{T}_j$. Furthermore, we denoted as $\mathcal{F}_j = \mathcal{F}_j^I \cup \mathcal{F}_j^B$ the set of all mesh faces of $\mathcal{T}_j$. With this notation, we assume that the sub-tessellation of element interfaces into $(d-1)$-dimensional simplices is given. Moreover, assume that the following assumptions hold, cf. [25, 26].

Assumption 1 *For any $j = 1, \ldots, J$, given $\kappa \in \mathcal{T}_j$ there exists a set of non-overlapping d-dimensional simplices $T_l \subset \kappa$, $l = 1, \ldots, n_\kappa$, such that for any face $F \subset \partial\kappa$ it holds that $\overline{F} = \partial\overline{\kappa} \cap \partial\overline{T_l}$ for some l, it holds $\cup_{l=1}^{n_\kappa} \overline{T_l} \subset \overline{\kappa}$, and the diameter h_κ of κ can be bounded by*

$$h_\kappa \lesssim \frac{d|T_l|}{|F|} \quad \forall \, l = 1, \ldots, n_\kappa.$$

Assumption 2 *For any $\kappa \in \mathcal{T}_j$, $j = 1, \ldots, J$, we assume that $h_\kappa^d \geq |\kappa| \gtrsim h_\kappa^d$, where $d = 2, 3$ is the dimension of Ω.*

Assumption 3 *Every polytopic element $\kappa \in \mathcal{T}_j$, $j = 1, \ldots, J$, admits a sub-triangulation into at most m_κ shape-regular simplices $\{\mathfrak{s}_i\}_{i=1}^{m_\kappa}$, for some $m_\kappa \in \mathbb{N}$, such that $\overline{\kappa} = \cup_{i=1}^{m_\kappa} \overline{\mathfrak{s}_i}$ and*

$$|\mathfrak{s_i}| \gtrsim |\kappa| \quad \forall i = 1, \ldots, m_\kappa,$$

Assumption 4 *Let $\mathcal{T}_j^{\#} = \{\mathcal{K}\}$, denote a covering of Ω consisting of shape-regular d dimensional simplices $\mathcal{K}$. We assume that, for any $\kappa \in \mathcal{T}_j$, there exists $\mathcal{K} \in \mathcal{T}_j^{\#}$ such that $\kappa \subset \mathcal{K}$ and*

$$\max_{\kappa \in \mathcal{T}_j} card\{\kappa' \in \mathcal{T}_j : \kappa' \cap \mathcal{K} \neq \emptyset, \mathcal{K} \in \mathcal{T}_j^{\#} \ such \ that \ \kappa \subset \mathcal{K}\} \lesssim 1.$$

Remark 2 Assumption 1 is needed in order to obtain the trace inequalities of Lemma 1 and Lemma 2. Assumption 2 and 3 are required for the inverse estimates of Lemma 5 and Theorem 6. Assumption 4 guarantees the validity of the approximation result and error estimetes of Lemma 4 and Theorem 5, respectively.

Remark 3 Assumptions 1 allows to employ polygonal and polyhedral elements possibly characterized by face of degenerating Hausdorff measure as well as unbounded number of faces, cf. [25], see also [6].

2.2 DG formulation

In order to introduce the DG discretization of (1), we firstly need to define suitable jump and average operators across the faces $F \in \mathcal{F}_j$, $j = 1, \ldots, J$. Let $\boldsymbol{\tau}$ and v be sufficiently smooth functions. For each internal face $F \in \mathcal{F}_j^I$, such that $F \subset \partial\kappa^{\pm}$, let $\mathbf{n}^{\pm}$ be the outward unit normal vector to $\partial\kappa^{\pm}$, and let $\boldsymbol{\tau}^{\pm}$ and $v^{\pm}$ be the traces of the functions $\boldsymbol{\tau}$ and v on F from $\kappa^{\pm}$, respectively. The jump and average operators across F are then defined as follows:

$$[\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}^+ \cdot \mathbf{n}^+ + \boldsymbol{\tau}^- \cdot \mathbf{n}^-, \qquad \{\!\{\boldsymbol{\tau}\}\!\} = \frac{\boldsymbol{\tau}^+ + \boldsymbol{\tau}^-}{2}, \qquad F \in \mathcal{F}_j^I,$$

$$[\![v]\!] = v^+ \mathbf{n}^+ + v^- \mathbf{n}^-, \qquad \{\!\{v\}\!\} = \frac{v^+ + v^-}{2}, \qquad F \in \mathcal{F}_j^I,$$

$$\{\!\{\boldsymbol{\tau}\}\!\} = \boldsymbol{\tau}, \qquad\qquad [\![v]\!] = v \ \mathbf{n}, \qquad\qquad F \in \mathcal{F}_j^B,$$

cf. [11]. With this notation, the bilinear form $\mathcal{A}_j(\cdot, \cdot) : V_j \times V_j \to \mathbb{R}$ corresponding to the symmetric interior penalty DG method on the j-th level is defined by

$$\mathcal{A}_j(u, v) = \sum_{\kappa \in \mathcal{T}_j} \int_\kappa (\nabla u + \mathcal{R}_j([\![u]\!])) \cdot (\nabla v + \mathcal{R}_j([\![v]\!])) dx + \sum_{F \in \mathcal{F}_j} \int_F \sigma_j [\![u]\!] \cdot [\![v]\!] \ ds,$$

where $\sigma_j \in L^\infty(\mathcal{F}_j)$ denotes the interior penalty stabilization function, which is defined by

$$\sigma_j(x) = \begin{cases} C_\sigma^j \max\limits_{\kappa \in \{\kappa^+, \kappa^-\}} \left\{\dfrac{p_j^2}{h_\kappa}\right\}, & x \in F, \ F \in \mathcal{F}_j^I, \ F \subset \partial\kappa^+ \cap \partial\kappa^-, \\ C_\sigma^j \dfrac{p^2}{h_\kappa}, & x \in F, \ F \in \mathcal{F}_j^B, \ F \subset \partial\kappa^+ \cap \partial\Omega, \end{cases}$$

with $C_\sigma^j > 0$ independent of p, $|F|$ and $|\kappa|$, and $\mathcal{R}_j : [L^1(\mathcal{F}_j)]^d \to [V_j]^d$ is the lifting operator on the space V_j, defined as

$$\int_\Omega \mathcal{R}_j(\mathbf{q}) \cdot \boldsymbol{\eta} = -\int_{\mathcal{F}_j} \mathbf{q} \cdot \{\!\{\boldsymbol{\eta}\}\!\} \ ds \quad \forall \ \boldsymbol{\eta} \in [V_j]^d.$$

We refer to [11] for more details.

Remark 4 Here, the formulation with the lifting operators $\mathcal{R}_j$ allows to introduce the discrete gradient operator $\mathcal{G}_j : V_j \to [V_j]^d$, defined as

$$\mathcal{G}_j(v) = \nabla_j v + \mathcal{R}_j([\![v]\!]) \quad \forall \, j = 1, \ldots, J, \tag{5}$$

where ∇_j is the piecewise gradient operator on the space V_j. The role of $\mathcal{G}_j$ will be clarified in Sect. 4.2.

The goal of this paper is to develop non-nested V-cycle multigrid schemes to solve the following problem posed on the finest level V_J: find $u_J \in V_J$ such that

$$\mathcal{A}_J(u_J, v_J) = \int_\Omega f v_J \, dx \quad \forall v_J \in V_J. \tag{6}$$

By fixing a basis for V_J, i.e. $V_J = span\{\phi_J^k\}_k$, formulation (6) results in the following linear system of equations

$$\mathbf{A}_J \mathbf{u}_J = \mathbf{f}_J,$$

where $\mathbf{u}_J$ is the vector of unknowns.

3 Preliminary results

In this section we recall some preliminary results which form the basis of the convergence analysis presented in the next section.

Lemma 1 *Assume that the sequence of meshes $\{\mathcal{T}_j\}_{j=1}^J$, satisfies Assumption 1 and let $\kappa \in \mathcal{T}_j$, then the following bound holds*

$$\|v\|_{L^2(\partial\kappa)}^2 \lesssim \frac{\epsilon}{h_\kappa} \|v\|_{L^2(\kappa)}^2 + \frac{h_\kappa}{\epsilon} |v|_{H^1(\kappa)}^2 \quad \forall v \in H^1(\kappa),$$

where h_κ is the diameter of κ and $\epsilon > 0$ is a positive number.

The proof of Lemma 1 is given in Appendix A.

Lemma 2 *Assume that the sequence of meshes $\{\mathcal{T}_j\}_{j=1}^J$ satisfies Assumption 1 and let $\kappa \in \mathcal{T}_j$. Then, the following bound holds*

$$\|v\|_{L^2(\partial\kappa)}^2 \lesssim \frac{p_j^2}{h_\kappa} \|v\|_{L^2(\kappa)}^2 \quad \forall v \in \mathcal{P}_{p_j}(\kappa).$$

We refer to [25] for the proof.

On each discrete space $\{V_j\}$, $j = 1, \ldots, J$, we consider the following DG norm:

$$\|w\|_{DG,j}^2 = \sum_{\kappa \in \mathcal{T}_j} \int_\kappa |\nabla w|^2 \, dx + \sum_{F \in \mathcal{F}_j} \int_F \sigma_j |[\![w]\!]|^2 \, ds. \tag{7}$$

The well-posed of the DG formulation is established in the following lemma.

Lemma 3 *The following continuity and coercivity bounds, respectively, hold*

$$\mathcal{A}_j(u, v) \lesssim \|u\|_{DG,j} \|v\|_{DG,j} \quad \forall u, v \in V_j,$$
$$\mathcal{A}_j(u, u) \gtrsim \|u\|_{DG,j}^2 \qquad\qquad \forall u \in V_j,$$

Next, we recall the following approximation result, which is an analogous bound presented in [27, Theorem 5.2].

Lemma 4 *Let Assumption 4 be satisfied, and let $v \in L^2(\Omega)$ such that, for some $k \geq 0$, $v|_\kappa \in H^k(\kappa)$ for each $\kappa \in \mathcal{T}_j$. Then there exists a projection operator $\tilde{\Pi}_j : L^2(\Omega) \to V_j$ such that*

$$\|v - \tilde{\Pi}_j v\|_{H^q(\Omega)} \lesssim \frac{h_j^{s-q}}{p_j^{k-q}} \|v\|_{H^k(\Omega)}, \quad for \quad 0 \leq q \leq k,$$

where $s = \min\{p_j + 1, k\}$ and $p_j \geq 1$.

The result presented in Lemma 4 leads to the following error bounds for the underlying interior penalty DG scheme. The error in the energy norm has been proved in [27], see also [25]. L^2-estimates can be found in [6].

Theorem 5 *Assume that Assumptions 1 and 4 hold. We denote by $u_j \in V_j$, $j = 1, \ldots, J$, the DG solution of problem (6) posed on level j, i.e.,*

$$\mathcal{A}_j(u_j, v_j) = \int_\Omega f v_j \; dx \quad \forall v_j \in V_j.$$

If the solution u of (1) satisfies $u|_\kappa \in H^k(\kappa)$, $k \geq 2$, then

$$\|u - u_j\|_{DG,j} \lesssim \frac{h_j^{(s-1)}}{p_j^{(k-\frac{3}{2})}} \|u\|_{H^k(\Omega)}, \quad \|u - u_j\|_{L^2(\Omega)} \lesssim \frac{h_j^s}{p_j^{k-1}} \|u\|_{H^k(\Omega)},$$

where $s = \min\{p_j + 1, k\}$ and $p_j \geq 1$.

Remark 5 We point out that the bounds in Theorem 5 are optimal in h and sub-optimal in p of a factor $p^{\frac{1}{2}}$ and p for the DG-norm and the L^2-norm, respectively. Optimal error estimates with respect to p can be shown, for example, by using the projector of [31] for quadrilateral meshes providing the solution belongs to a suitable augmented Sobolev space. The issue of proving optimal estimates as the ones in [31] on polytopic meshes is an open problem and it is under investigation. In the following, we will write:

$$\|u - u_j\|_{DG,j} \lesssim \frac{h_j^{(s-1)}}{p_j^{(k-1-\frac{\mu}{2})}} \|u\|_{H^k(\Omega)}, \quad \|u - u_j\|_{L^2(\Omega)} \lesssim \frac{h_j^s}{p_j^{k-\mu}} \|u\|_{H^k(\Omega)},$$

where $s = \min\{p_j + 1, k\}$, $p_j \geq 1$, and $\mu \in \{0, 1\}$ for optimal and suboptimal estimates, respectively.

We also need to introduce an appropriate inverse inequality, cf. [26].

Lemma 5 *Assume that Assumptions 2 and 3 hold. Then, for any $v \in V_j$, $j = 1, \ldots, J$, the following inverse estimate holds*

$$\|\nabla u\|_{L^2(\kappa)}^2 \lesssim p_j^4 h_\kappa^{-2} \|u\|_{L^2(\kappa)}^2 \qquad \forall \kappa \in \mathcal{T}_j.$$

Thanks to the inverse estimate of Lemma 5, it is possible to obtain the following upper bound on the maximum eigenvalue of $\mathcal{A}_j(\cdot, \cdot)$. We refer to [5] for a similar result on standard grids, and to [6] for its extension to polygonal grids.

Theorem 6 *Let Assumptions 1, 2 and 3 be satisfied. Moreover, we assume that* $h_j = \max_{\kappa \in \mathcal{T}_j} h_\kappa \approx h_\kappa \; \forall \kappa \in \mathcal{T}_j$, *for* $j = 1, \ldots, J$. *Then*

$$\mathcal{A}_j(u, u) \lesssim \frac{p_j^4}{h_j^2} \|u\|_{L^2(\Omega)}^2 \quad \forall u \in V_j, \quad j = 1, \ldots, J.$$

4 The BPX-framework for the V-cycle algorithms

The analysis presented in this section is based on the general multigrid theoretical framework already employed and developed in [23] for non-nested spaces and non-inherited bilinear forms. In order to develop a geometric multigrid, the discretization at each level V_j follows the one already presented in [9], where a W-cycle multigrid method based on nested subspaces is considered. The key ingredient in the construction of our proposed multigrid schemes is the inter-grid transfer operators.

Firstly, we introduce the operators $A_j : V_j \to V_j$, defined as

$$(A_j u, v) = \mathcal{A}_j(u, v) \quad \forall u, v \in V_j, \quad j = 1, \ldots, J, \tag{8}$$

and we denote as $\Lambda_j \in \mathbb{R}$ the maximum eigenvalue of $A_j \; \forall \; j = 2, \ldots, J$. Moreover, let Id_j be the identity operator on level V_j. The smoothing scheme, which is chosen to be the Richardson iteration, is then characterized by the following operators:

$$B_j = \Lambda_j \mathrm{Id}_j \quad j = 2, \ldots, J.$$

The prolongation operator connecting the coarser space V_{j-1} to the finer space V_j is denoted by I_{j-1}^j. Since the two spaces are non-nested, i.e. $V_{j-1} \not\subset V_j$, it cannot be chosen as the "natural injection operator". The most natural way to define the prolongation operator is the L^2-projection, i.e. $I_{j-1}^j : V_{j-1} \to V_j$

$$(I_{j-1}^j v_H, w_h)_{L^2(\Omega)} = (v_H, w_h)_{L^2(\Omega)} \; \forall w_h \in V_j, \tag{9}$$

The restriction operator $I_j^{j-1} : V_j \to V_{j-1}$ is defined as the adjoint of I_{j-1}^j with respect to the $L^2(\Omega)$-inner product, i.e.,

$$(I_j^{j-1} w_h, v_H)_{L^2(\Omega)} = (w_h, I_{j-1}^j v_H)_{L^2(\Omega)} \; \forall v_H \in V_{j-1}.$$

For our analysis, we also need to introduce the operator $P_j^{j-1} : V_j \to V_{j-1}$ such that:

$$\mathcal{A}_{j-1}(P_j^{j-1} w_h, v_H) = \mathcal{A}_j(w_h, I_{j-1}^j v_H) \quad \forall v_H \in V_{j-1}, w_h \in V_j.$$

According with (8), problem (6) can be written in the following equivalent form: find $u_J \in V_J$ such that

$$A_J u_J = f_J, \tag{10}$$

where $f_J \in V_J$ is defined as $(f_J, v)_{L^2(\Omega)} = \int_\Omega f v \, dx \; \forall v \in V_J$. Given an initial guess $u_0 \in V_J$, and choosing parameters $m_1, m_2 \in \mathbb{N}$, the multigrid V-cycle iteration algorithm for the approximation of u_J is outlined in Algorithm 1. In particular, $\mathsf{MG}_{\mathcal{V}}(J, f_J, u_k, m_1, m_2)$ represents the approximate solution obtained after one

---

Algorithm 1 Multigrid V-cycle iteration for the solution of problem (10)

Initialize $u_0 \in V_J$;
for $k = 0, 1, \ldots$ **do**
    $u_{k+1} = \mathsf{MG}_{\mathcal{V}}(J, f_J, u_k, m_1, m_2)$;
    $u_k = u_{k+1}$;
end for

Algorithm 2 One iteration of the Multigrid V-cycle scheme on the level $j \geq 2$

---

if j=1 **then**
    $\mathsf{MG}_{\mathcal{V}}(1, g, z_0, m_1, m_2) = A_1^{-1}g$.
else
    Pre-smoothing:
    for $i = 1, \ldots, m_1$ **do**
        $z^{(i)} = z^{(i-1)} + B_j^{-1}(g - A_j z^{(i-1)})$;
    end for

    Coarse grid correction:
    $r_{j-1} = I_j^{j-1}(g - A_j z^{(m_1)})$;
    $e_{j-1} = \mathsf{MG}_{\mathcal{V}}(j - 1, r_{j-1}, 0, m_1, m_2)$;
    $z^{(m_1+1)} = z^{(m_1)} + I_{j-1}^j e_{j-1}$;

    Post-smoothing:
    for $i = m_1 + 2, \ldots, m_1 + m_2 + 1$ **do**
        $z^{(i)} = z^{(i-1)} + B_j^{-1}(g - A_j z^{(i-1)})$;
    end for

    $\mathsf{MG}_{\mathcal{V}}(j, g, z_0, m_1, m_2) = z^{(m_1+m_2+1)}$.
end if

---

iteration of our non-nested V-cycle scheme, which is defined by induction: if we consider the general problem of finding $z \in V_j$ such that

$$A_j z = g, \tag{11}$$

with $j \in \{2, \ldots, J\}$ and $g \in L^2(\Omega)$, then $\mathsf{MG}_{\mathcal{V}}(j, g, z_0, m_1, m_2)$ represents the approximate solution of (11) obtained after one iteration of the non-nested V-cycle scheme with initial guess $z_0 \in V_j$ and m_1, m_2 number of pre-smoothing and post-smoothing steps, respectively. The recursive procedure is outlined in Algorithm 2, where we also observe that on the level $j = 1$ the problem is solved by using a direct method.

4.1 Convergence analysis

We first define the following norms on each discrete space V_j

$$|||v|||_{s,j} = \sqrt{(A_j^s v, v)_{L^2(\Omega)}} \qquad \forall \, s \in \mathbb{R}, \; v \in V_j, \quad j = 1, \ldots, J.$$

To analyze the convergence of the algorithm, for any $j = 2, \ldots, J$ we set $G_j = \mathrm{Id}_j - B_j^{-1} A_j$ and let G_j^* be its adjoint respect to $\mathcal{A}_j(\cdot, \cdot)$. Following [30], we make three standard assumptions in order to prove the convergence of Algorithm 1:

A.1 Stability estimate: $\exists\, C_Q > 0$ such that

$$\|(\mathrm{Id}_j - I_{j-1}^j P_j^{j-1})u_h\|_{1,j} \le C_Q \|u_h\|_{1,j} \qquad \forall u_h \in V_j, \quad j = 2, \dots, J;$$

A.2 Regularity-approximation property: $\exists\, C_1 > 0$ such that

$$\left|\mathcal{A}_j\big((\mathrm{Id}_j - I_{j-1}^j P_j^{j-1})u_h, u_h\big)\right| \le C_1 \frac{\|u_h\|_{2,j}^2}{\Lambda_j} \qquad \forall u_h \in V_j, \quad j = 2, \dots, J,$$

where $\Lambda_j = \max \lambda_i(A_j)$;

A.3 Smoothing property: $\exists\, C_R > 0$ such that

$$\frac{\|u_h\|_{L^2(\Omega)}}{\Lambda_j} \le C_R\big(\mathcal{R}u_h, u_h\big) \qquad \forall u_h \in V_j, \quad j = 2, \dots, J,$$

where $\mathcal{R} = \big(\mathrm{Id}_j - G_j^* G_j\big)A_j^{-1}$.

The convergence analysis of the V-cycle method is described by the following theorem that gives an estimate for the error propagation operator, which is defined as

$$\begin{cases} \mathbb{E}_{1,m_1,m_2}^{\mathsf{V}} v &= 0, & j = 1, \\ \mathbb{E}_{j,m_1,m_2}^{\mathsf{V}} v &= (G_J^*)^{m_2}(\mathrm{Id}_j - I_{j-1}^j P_j^{j-1} + I_{j-1}^j \mathbb{E}_{j-1,m_1,m_2}^{\mathsf{V}} P_j^{j-1})G_j^{m_1}v, & j > 1. \end{cases}$$

Theorem 7 *If Assumptions **A.1**, **A.2** and **A.3** hold, then*

$$\left|\mathcal{A}_j(\mathbb{E}_{j,m,m}^{\mathsf{V}} u, u)\right| \le \delta_j \mathcal{A}_j(u, u) \qquad \forall u \in V_j, \quad j = 2, \dots, J$$

where $\delta_j = \frac{C_1 C_R}{m - C_1 C_R} < 1$, provided that $m > 2 C_1 C_R$.

We refer to [30] for the proof of Theorem 7 in an abstract setting. In the following, we prove the validity of Assumptions **A.1**, **A.2** and **A.3** for the algorithm presented in this section. We start with a two-level approach, i.e. $J = 2$, so we will consider the two-level method for the solution of (6), based on two spaces $V_{J-1} \not\subset V_J$. The generalization to the V-cycle method will be given at the end of this section.

4.2 Verification of Assumption **A.1**

In order to verify Assumption **A.1** for the two-level method we first show a stability result of the prolongation operator I_{J-1}^J. In the following, we also consider the L^2-projection operator on the space V_J defined as

$$Q_J : L^2(\Omega) \to V_J, \text{ such that } (Q_J u, v_J)_{L^2(\Omega)} = (u, v_J)_{L^2(\Omega)} \quad \forall v_J \in V_J.$$

Remark 6 From the definition of I_{J-1}^J given in (9), it holds $I_{J-1}^J = Q_J|_{V_{J-1}}$.

Moreover, we need the following approximation result which shows that any $v_j \in V_j$, $j = J - 1, J$, can be approximated by an H^1-function, see [7]. Let $\mathcal{G}_j$ be the discrete gradient operator (5) introduced in Remark 4, and consider the following problem: $\forall v_j \in V_j$, find $\mathcal{H}(v_j) \in H_0^1(\Omega)$ such that

$$\int_\Omega \nabla \mathcal{H}(v_j) \cdot \nabla w \; dx = \int_\Omega \mathcal{G}_j(v_j) \cdot \nabla w \; dx \quad \forall w \in H_0^1(\Omega). \tag{12}$$

It is shown in [7] that $\mathcal{H}(v_j)$ possesses good approximation properties in terms of providing an H^1-conforming approximant of the discontinuous function v_j:

Theorem 8 *Let Ω be a bounded convex polygonal/polyhedral domain in $\mathbb{R}^d$, $d = 2, 3$. Given $v_j \in V_j$, we write $\mathcal{H}(v_j) \in H_0^1(\Omega)$ to be the approximation defined in (12). Then, the following approximation and stability results hold:*

$$\|v_j - \mathcal{H}(v_j)\|_{L^2(\Omega)} \lesssim \frac{h_j}{p_j} \|\sigma_j^{\frac{1}{2}} [\![v_j]\!]\|_{L^2(\mathcal{F}_h)}, \qquad |\mathcal{H}(v_j)|_{H^1(\Omega)} \lesssim \|v_j\|_{DG,j}. \tag{13}$$

We make use of the previous result in order to show the following stability result of the prolongation operator:

Lemma 6 *There exists a positive constant $\mathsf{C}_{\mathsf{stab}}$, independent of the mesh size such that*

$$\|I_{J-1}^J v_H\|_{DG,J} \leq \mathsf{C}_{\mathsf{stab}}(p_J) \; \|v_H\|_{DG,J-1} \quad \forall v_H \in V_{J-1},$$

here $\mathsf{C}_{\mathsf{stab}}(p_J) \approx p_J$.

Proof Let $v_H \in V_{J-1}$, by the definition of the DG-norm (7), we need to estimate:

$$\|I_{J-1}^J v_H\|_{DG,J}^2 = \|\nabla_J (I_{J-1}^J v_H)\|_{L^2(\mathcal{T}_J)}^2 + \|\sigma_J^{\frac{1}{2}} [\![I_{J-1}^J v_H]\!]\|_{L^2(\mathcal{F}_J)}^2. \tag{14}$$

We next bound each of the two terms on the right hand side. For the first one let be $\mathcal{H}_H = \mathcal{H}(v_H)$ defined as in (12). Then:

$$\begin{aligned}
\|\nabla_J (I_{J-1}^J v_H)\|_{L^2(\mathcal{T}_J)}^2 \leq\; &\|\nabla_J (I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H))\|_{L^2(\mathcal{T}_J)}^2 \\
&+ \|\nabla_J (\mathcal{H}_H - \widetilde{\Pi}_J(\mathcal{H}_H))\|_{L^2(\mathcal{T}_J)}^2 + |\mathcal{H}_H|_{H^1(\Omega)}^2,
\end{aligned} \tag{15}$$

where we have added and subtracted the terms $\nabla_J \widetilde{\Pi}_J(\mathcal{H}_H))$ and $\nabla \mathcal{H}_H$. The second term of the right hand above side can be estimated using the interpolation bounds of Lemma 4, the Poincaré inequality for $\mathcal{H}_H \in H_0^1(\Omega)$ and the second bound of (13):

$$\|\nabla_J (\mathcal{H}_H - \widetilde{\Pi}_J(\mathcal{H}_H))\|_{L^2(\mathcal{T}_J)}^2 \lesssim |\mathcal{H}_H|_{H^1(\Omega)}^2 \lesssim \|v_H\|_{DG,J-1}^2.$$

In order to estimate the first term on the right hand side in (15) we observe that, since $I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H) \in V_J$, it is possible to make use of the inverse inequality of Lemma 5, that leads to the following bound:

$$\|\nabla_J (I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H))\|_{L^2(\mathcal{T}_J)}^2 \lesssim p_J^4 h_J^{-2} \|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|_{L^2(\mathcal{T}_J)}^2. \tag{16}$$

By adding and subtracting $\mathcal{H}_H$ to $\|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)}$ we obtain

$$\|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)} \lesssim \|I_{J-1}^J v_H - \mathcal{H}_H\|^2_{L^2(\mathcal{T}_J)} + \|\mathcal{H}_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)}. \tag{17}$$

Using Lemma 4 and the Poincaré inequality we have

$$\|\mathcal{H}_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)} \lesssim \frac{h_J^2}{p_J^2} \|\mathcal{H}_H\|^2_{H^1(\Omega)} \lesssim \frac{h_J^2}{p_J^2} \|v_H\|^2_{DG,J-1},$$

whereas the term $\|I_{J-1}^J v_H - \mathcal{H}_H\|^2_{L^2(\mathcal{T}_J)}$ can be estimate as follow:

$$\|I_{J-1}^J v_H - \mathcal{H}_H\|^2_{L^2(\mathcal{T}_J)} \lesssim \|I_{J-1}^J v_H - Q_J(\mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)} + \|\mathcal{H}_H - Q_J(\mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)}$$

Using Remark 6, the continuity of Q_J with respect to the L^2-norm, Lemma 4 and (13) we have

$$\begin{aligned}
\|I_{J-1}^J v_H - \mathcal{H}_H\|^2_{L^2(\mathcal{T}_J)} &\lesssim \|Q_J(v_H - \mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)} + \|\mathcal{H}_H - Q_J(\mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)} \\
&\lesssim \|v_H - \mathcal{H}_H\|^2_{L^2(\mathcal{T}_J)} + \|\mathcal{H}_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)} \\
&\lesssim \frac{h_J^2}{p_J^2} \|\sigma_J^{\frac{1}{2}} [\![v_H]\!]\|^2_{L^2(\mathcal{F}_J)} + \frac{h_J^2}{p_J^2} \|\mathcal{H}_H\|^2_{H^1(\Omega)} \lesssim \frac{h_J^2}{p_J^2} \|v_H\|^2_{DG,J-1}.
\end{aligned}$$

Thanks to the previous estimates and inequalities (17), it holds

$$\|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|^2_{L^2(\mathcal{T}_J)} \lesssim \frac{h_J^2}{p_J^2} \|v_H\|^2_{DG,J-1}, \tag{18}$$

the previous estimate, together with (16), (15) and the bound $|\mathcal{H}_H|^2_\Omega \lesssim \|v_H\|^2_{DG,J-1}$ leads to

$$\|\nabla_J(I_{J-1}^J v_H)\|^2_{L^2(\mathcal{T}_J)} \lesssim p_J^2 \|v_H\|^2_{DG,J-1}. \tag{19}$$

Next we bound the second term on the right hand side in (14). By the definition of the jump term and remembering that $[\![\mathcal{H}_H]\!] = 0 \; \forall F \in \mathcal{F}_J$ since $\mathcal{H}_H \in H_0^1(\Omega)$, it holds

$$\|\sigma_J^{\frac{1}{2}} [\![I_{J-1}^J v_H]\!]\|^2_{L^2(\mathcal{F}_J)} \lesssim \tag{20}$$

$$\sum_{\kappa \in \mathcal{T}_J} \frac{p_J^2}{h_\kappa} \left(\|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|^2_{L^2(\partial\kappa)} + \|\widetilde{\Pi}_J(\mathcal{H}_H) - \mathcal{H}_H\|^2_{L^2(\partial\kappa)} \right),$$

where we also used the definition of σ_J. Now, we first observe that we could use the trace inequality of Lemma 2 in order to obtain

$$\|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|^2_{L^2(\partial\kappa)} \lesssim \frac{p_J^2}{h_J} \|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|^2_{L^2(\kappa)}. \tag{21}$$

To bound the second term on the right hand side in (20) we make use of the continuous trace inequality on polygons of Lemma 1 with $\epsilon = p_J$, the approximation property of Lemma 4 and the Poincaré inequality:

$$\begin{aligned}
\|\widetilde{\Pi}_J(\mathcal{H}_H) - \mathcal{H}_H\|^2_{L^2(\partial\kappa)} &\lesssim \frac{p_J}{h_J} \|\widetilde{\Pi}_J(\mathcal{H}_H) - \mathcal{H}_H\|^2_{L^2(\kappa)} + \frac{h_J}{p_J} |\widetilde{\Pi}_J(\mathcal{H}_H) - \mathcal{H}_H|^2_{H^1(\kappa)} \\
&\lesssim \frac{p_J}{h_J} \frac{h_J^2}{p_J^2} \|\mathcal{H}_H\|^2_{L^2(\kappa)} + \frac{h_J}{p_J} \|\mathcal{H}_H\|^2_{H^1(\kappa)} \lesssim \frac{h_J}{p_J} |\mathcal{H}_H|^2_{H^1(\kappa)}.
\end{aligned}$$

From the previous inequality and the bound (21), (20) becomes:

$$\|\sigma_J^{\frac{1}{2}}[\![I_{J-1}^J v_H]\!]\|_{L^2(\mathcal{F}_J)}^2 \lesssim \frac{p_J^4}{h_J^2} \|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|_{L^2(\mathcal{T}_J)}^2 + p_J |\mathcal{H}_H|_{H^1(\Omega)}^2$$

$$\lesssim p_J^2 \|v_H\|_{DG,J-1}^2,$$

where we also used inequality (18). This estimate together with (19) lead to

$$\|I_{J-1}^J v_H\|_{DG,J} \le \mathsf{C}_{\mathsf{stab}}(p_J) \, \|v_H\|_{DG,J-1} \quad \forall v_H \in V_{J-1}.$$

where $\mathsf{C}_{\mathsf{stab}}(p_J) \approx p_J$.

We can use the previous result in order to prove that Assumption **A.1** holds. We first observe that also the operator P_J^{J-1} satisfies a similar stability estimate as the one of I_{J-1}^J, that is

$$\|P_J^{J-1} v_h\|_{DG,J-1}^2 \lesssim \mathcal{A}_{J-1}(P_J^{J-1} v_h, P_J^{J-1} v_h) = \mathcal{A}_J(v_h, I_{J-1}^J P_J^{J-1} v_h)$$

$$\lesssim \|v_h\|_{DG,J} \|I_{J-1}^J P_J^{J-1} v_h\|_{DG,J} \lesssim \mathsf{C}_{\mathsf{stab}}(p_J) \, \|v_h\|_{DG,J} \|P_J^{J-1} v_h\|_{DG,J},$$

from which it follows

$$\|P_J^{J-1} v_h\|_{DG,J-1} \lesssim \mathsf{C}_{\mathsf{stab}}(p_J) \, \|v_h\|_{DG,J}.$$

Proposition 1 *Assumption **A.1** holds with $C_Q \approx p_J^2$.*

Proof Let $v_H \in V_{J-1}$, making use of Lemma 3 we have

$$\mathcal{A}_J(I_{J-1}^J v_H, I_{J-1}^J v_H) \lesssim \|I_{J-1}^J v_H\|_{DG,J}^2 \lesssim p_J^2 \, \|v_H\|_{DG,J-1}^2 \lesssim p_J^2 \, \mathcal{A}_{J-1}(v_H, v_H),$$

and similarly it holds

$$\mathcal{A}_{J-1}(P_J^{J-1} u_h, P_J^{J-1} u_h) \lesssim p_J^2 \, \mathcal{A}_J(u_h, u_h) \quad \forall u_h \in V_J. \tag{22}$$

Let $u_h \in V_J$ and fix $v_H = P_J^{J-1} u_h$, then the following inequality holds:

$$\mathcal{A}_J(I_{J-1}^J P_J^{J-1} u_h, I_{J-1}^J P_J^{J-1} u_h) \lesssim p_J^2 \, \mathcal{A}_{J-1}(P_J^{J-1} u_h, P_J^{J-1} u_h). \tag{23}$$

By adding and subtracting u_h to both arguments of $\mathcal{A}_J(\cdot, \cdot)$ on the left hand side of (23), and using (22) we obtain

$$\underbrace{\mathcal{A}_J((\mathrm{Id}_J - I_{J-1}^J P_J^{J-1}) u_h, (\mathrm{Id}_J - I_{J-1}^J P_J^{J-1}) u_h)}_{=\|(\mathrm{Id}_J - I_{J-1}^J P_J^{J-1}) u_h\|_{1,J}^2} \lesssim \underbrace{\left(p_J^2 (p_J^2 - 2) + 1 \right)}_{\le p_J^4} \mathcal{A}_J(u_h, u_h),$$

that concludes the proof.

4.3 Verification of Assumption **A.2**

In order to show the validity of Assumption **A.2** we need the following standard approximation result, which is proved in Appendix B.

Lemma 7 *Let Assumptions 1 - 4 hold. Then*

$$\|(\mathrm{Id}_J - I_{J-1}^J P_J^{J-1})v_J\|_{L^2(\Omega)} \lesssim \frac{h_J^2}{p_J^{2-\mu}}\|v_J\|_{2,J} \quad \forall v_J \in V_J. \tag{24}$$

Thanks to Lemma 7, it is possible to show the following theorem:

Theorem 9 *The regularity-approximation property* **A.2** *holds with* $C_1 \approx p_J^{2+\mu}$.

Proof Theorem 6 gives the following bound of the maximum eigenvalue of A_J: $\Lambda_J \lesssim \frac{p_J^4}{h_J^2}$. Using Lemma 7, the above bound on Λ_J, and the symmetry of $\mathcal{A}_J(\cdot,\cdot)$ we have, for all $v \in V_J$:

$$\mathcal{A}_J((\mathrm{Id}_J - I_{J-1}^J P_J^{J-1})v, v) \leq \|v\|_{2,J}\|(\mathrm{Id}_J - I_{J-1}^J P_J^{J-1})v\|_{0,J} \lesssim \frac{h_J^2}{p_J^{2-\mu}}\|v\|_{2,J}^2$$

$$\lesssim p_J^{2+\mu}\frac{\|v\|_{2,J}^2}{\Lambda_J}.$$

that concludes the proof.

4.4 Verification of Assumption **A.3**

Proposition 2 *Assumption* **A.3** *holds with* $C_R = 1$.

Proof We have:

$$\mathcal{R} = (\mathrm{Id}_J - G_J^* G_J)A_J^{-1} = \left(\frac{2}{\Lambda_J}A_J - \frac{1}{\Lambda_J^2}A_J A_J\right)A_J^{-1} = \frac{1}{\Lambda_J}\left(\mathrm{Id}_J + \left(\mathrm{Id}_J - \frac{1}{\Lambda_J}A_J\right)\right),$$

and so

$$(\mathcal{R}u, u)_{L^2(\Omega)} = \frac{\|u_h\|_{L^2(\Omega)}}{\Lambda_J} + \left(\left(\mathrm{Id}_J - \frac{1}{\Lambda_J}A_J\right)u, u\right)_{L^2(\Omega)}.$$

We now prove that $\left(\mathrm{Id}_J - \frac{1}{\Lambda_J}A_J\right)$ is a positive definite operator. By contradiction, let us suppose that there exists a function $\overline{u} \in V_J$, $\overline{u} \neq 0$, such that $\left(\left(\mathrm{Id}_J - \frac{1}{\Lambda_J}A_J\right)\overline{u}, \overline{u}\right)_{L^2(\Omega)} < 0$, then

$$\Lambda_J(\overline{u}, \overline{u})_{L^2(\Omega)} < \mathcal{A}_J(\overline{u}, \overline{u}), \tag{25}$$

by Lemma 3 and the symmetry of the bilinear form $\mathcal{A}_J(\cdot,\cdot)$, the eigenfunctions $\{\phi_k^J\}_{k=1}^{N_J}$ satisfy

$$\mathcal{A}_J(\phi_k^J, v) = \lambda_k^J(\phi_k^J, v)_{L^2(\Omega)} \qquad \forall v \in V_J,$$

where $0 < \lambda_1^J \leq \lambda_2^J \leq \cdots \leq \lambda_{N_J}^J = \Lambda_J$. The set of eigenfunctions is an orthonormal basis for the space V_J, i.e. $(\phi_i^J, \phi_j^J)_{L^2(\Omega)} = \delta_{ij}$, and they satisfy $\mathcal{A}_J(\phi_i^J, \phi_j^J) =$

$\lambda_i^J \delta_{ij}$, where δ_{ij} is the Kronecker symbol. Since $\{\phi_k^J\}_{k=1}^{N_J}$ is a basis of the space V_J, we can write $\overline{u} = \sum_{k=1}^{N_J} c_k \phi_k^J$, so that (25) becomes

$$\Lambda_J \sum_{i,j=1}^{N_J} c_j (\phi_j^J, \phi_i^J)_{L^2(\Omega)} c_i < \sum_{i,j=1}^{N_J} c_j \mathcal{A}_J(\phi_j^J, \phi_i^J) c_i = \sum_{i,j=1}^{N_J} c_j \lambda_i^J (\phi_i^J, \phi_j^J)_{L^2(\Omega)} c_i,$$

$$\Rightarrow \quad \Lambda_J \sum_{i=1}^{N_J} c_i^2 < \sum_{i,j=1}^{N_J} c_i^2 \lambda_i^J,$$

which is a contradiction. We then deduce that $\left(\mathrm{Id}_J - \frac{1}{\Lambda_J} A_J\right)$ is a positive definite operator.

Remark 7 We observe that, as we need to satisfy the condition $m > 2C_1 C_R$ of Theorem 7, we can guarantee the convergence of the method choosing the number m of smoothing steps such that $m \gtrsim p_J^{2+\mu}$, which is in agreement to what proved for W-cycle algorithms in [9] and [6] on nested grids.

Remark 8 The analysis of this section can be generalized to the full V-cycle algorithm with $J > 2$ as follows: Assumption **A.3** is verified with $C_R = 1$ also on the arbitrary levels $j, j - 1$, because each level j satisfies Assumption **A.3** with constant $C_R^j = 1$. Assumptions **A.2** and **A.1** are satisfied with $C_1 = \max_j \{C_1^j\}$ and $C_Q = \max_j \{C_Q^j\}$, respectively, where C_1^j and C_Q^j are the same as the ones defined in the previous analysis but on the level j.

5 Numerical results

In this section we present several numerical results to test the theoretical convergence estimates provided in Theorem 7. We focus on a two dimensional problem on the unit square $\Omega = (0,1)^2$. For the simulations, we consider the sets of polygonal grids shown in Figure 1. Each polygonal element mesh is generated through the Voronoi Diagram algorithm by using the software package `PolyMesher` [44]. In particular the finest grids (Level 4) of Figure 1 consist of 512 (Set 1), 1024 (Set 2), 2048 (Set 3) and 4096 (Set 4) elements. Starting from the number of elements of each initial mesh, a sequence of coarse, non-nested partitions is generated: each coarse mesh is built independently from the finer one, with the only constrain that the number of element is approximately $1/4$ of the finer one. An example of sequence of non-nested partitions is shown in Figure 2.

First of all, we verify the estimate of Lemma 6, numerically evaluating $\mathsf{C}_{\mathrm{stab}}(p) \approx p$, where p is the polynomial approximation degree. To this aim we consider three pairs of non-nested grids, where the number of elements of the coarser grid is the number of the finer divided by 4: for each pair, we compute the value of $\mathsf{C}_{\mathrm{stab}}(p)$ as a function of p. Figure 3 show that, as expected, $\mathsf{C}_{\mathrm{stab}}(p)$ depends linearly on p and is independent of the mesh-size h.

We now consider the grids shown in Set 1 and in Set 2 of Figure 1, and numerically evaluate the constant δ_j in Theorem 7 based on selecting the Richardson smoother with $m_1 = m_2 = m = 3p^2$, cf. Figure 4. Here, we observe that δ_2 and

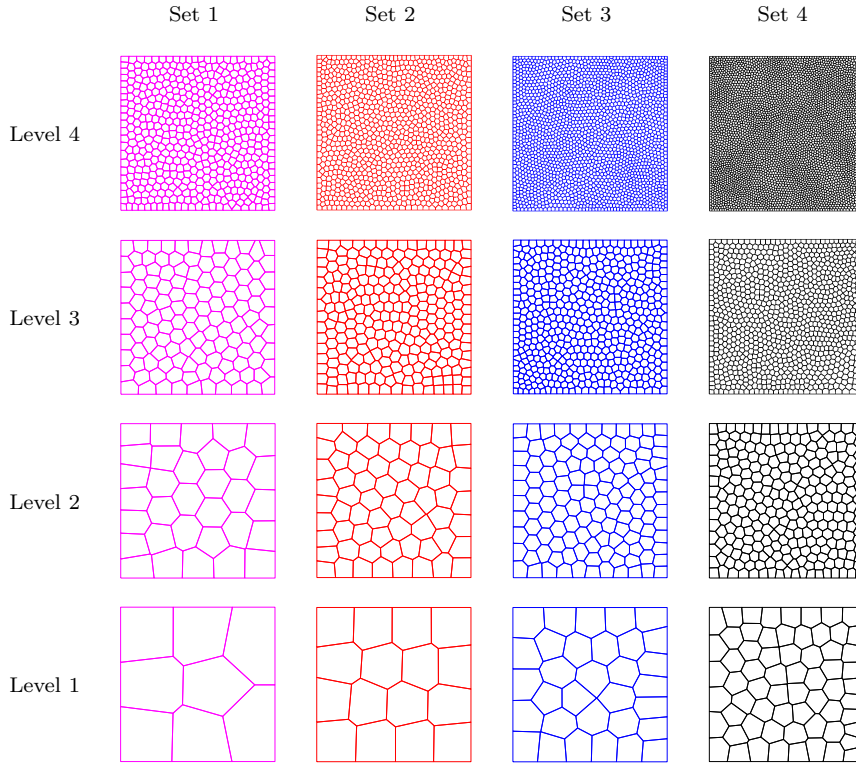|	Set 1	Set 2	Set 3	Set 4

Level 4

Level 3

Level 2

Level 1

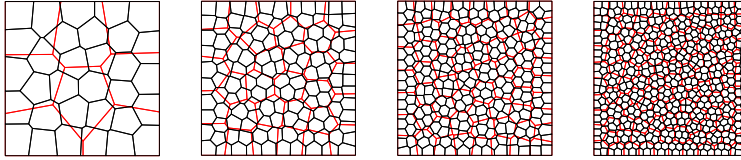Fig. 1: Sets of non-nested grids employed for numerical simulations.

Fig. 2: Example of non-nested partition.

δ_3 are asymptotically constant, as the polynomial degree p increases showing that our two-level and V-cycle algorithms are uniformly convergent also with respect to p provided that $m \approx p^2$.

Next, we investigate the performance of the iterative Multigrid non-nested V-cycle algorithm presented in Sect. 4. In order to do that, we solve the Poisson problem with homogeneous Dirichlet boundary conditions on the unit square $\Omega = (0,1)^2$, and we compute the number of iterations needed by our V-cycle algorithm to reduce the relative residual error below a given tolerance of 10^{-8}, by varying the polynomial degree of approximation and the granularity of the finest grid. In
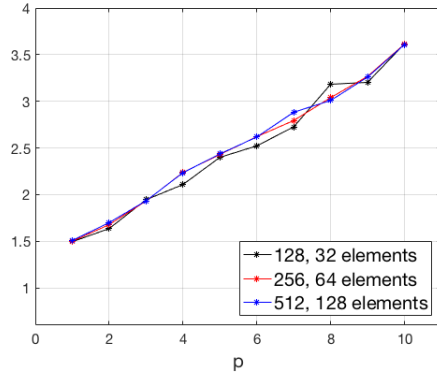
Fig. 3: Estimates of $C_{stab}(p)$ in Lemma 6 as a function of p. Non-nested Voronoi meshes as shown in Figure 1.
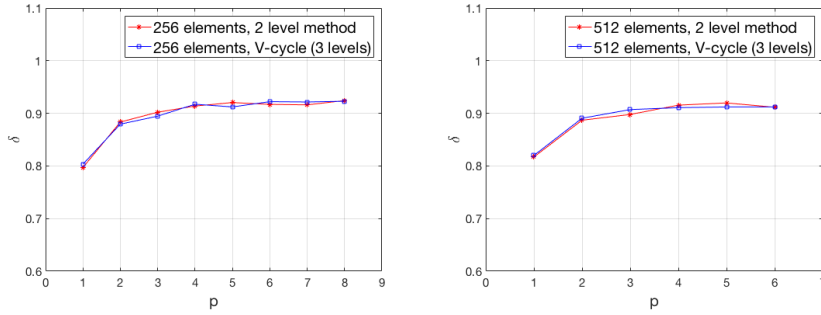

Fig. 4: Estimates of δ_2 and δ_3 in Theorem 7 as a function of p, with $m_1 = m_2 = m = 3p^2$ and two polyhedral grids of 256 (left) and 512 (right) elements.

Table 1 we report the convergence factors

$$\rho_J = \exp\left(\frac{1}{N_{it,J}} \ln \frac{\|\mathbf{r}_{N_{it,J}}\|}{\|\mathbf{r}_0\|} \right),$$

where $N_{it,J}$ is the iteration counts needed to attain convergence of the h-version of the V-cycle scheme with J levels, where $J = 2, 3, 4$, while $\mathbf{r}_{N_{it,J}}$ and $\mathbf{r}_0$ are the final and initial residual vectors, respectively. Here the polynomial approximation degree on each level is chosen as $p_j = 1$, $j = 1, \ldots, J$, while we vary the number of elements of the finest grid and the number of smoothing steps ($m_1 = m_2 = m$). According to Theorem 7, the convergence factor is independent from the spatial discretization step h, indeed, for a fixed $J \in \{2, 3, 4\}$ ad a fixed number of pre-smoothing steps m, the convergence factor is roughly constant between the 4 sets of grids. In particular, this means that the number of iterations needed from the proposed V-cycle method to attain the convergence is not influenced by the mesh refinement, contrarily of what we observe for the Conjugate Gradient (CG) method. As expected, the convergence factor is reduced by increasing the number of smoothing step.

Table 1: Converge factors ρ_J of V-cycle multigrid method as a function of m and a comparison with the iteration counts of the CG method ($C_\sigma^j \equiv C_\sigma = 10$, $p = 1$).

	Set 1, $N_{iter}^{CG} = 410$			Set 2, $N_{iter}^{CG} = 600$		
	2 levels	3 levels	4 levels	2 levels	3 levels	4 levels
$m = 3$	0.77	0.83	0.83	0.82	0.84	0.85
$m = 5$	0.69	0.76	0.78	0.74	0.77	0.79
$m = 8$	0.63	0.69	0.72	0.66	0.70	0.73
	Set 3, $N_{iter}^{CG} = 867$			Set 4, $N_{iter}^{CG} = 1228$		
	2 levels	3 levels	4 levels	2 levels	3 levels	4 levels
$m = 3$	0.79	0.85	0.93	0.78	0.84	0.87
$m = 5$	0.72	0.79	0.82	0.71	0.78	0.81
$m = 8$	0.65	0.72	0.76	0.64	0.72	0.74

Table 2: Converge factors ρ_J (and iterations count) of the V-cycle methods as a function of the number m of pre-smoothing steps and comparison with the iteration counts of the CG method ($C_\sigma^j \equiv C_\sigma = 10$, $p = 3$).

	Set 1, $N_{iter}^{CG} = 2212$			Set 2, $N_{iter}^{CG} = 3223$		
	2 levels	3 levels	4 levels	2 levels	3 levels	4 levels
$m = 3$	0.99 (3306)	0.98 (992)	0.98 (955)	0.97 (616)	0.98 (852)	0.98 (1024)
$m = 5$	0.96 (429)	0.97 (566)	0.97 (591)	0.95 (396)	0.96 (523)	0.97 (626)
$m = 8$	0.94 (296)	0.95 (367)	0.95 (388)	0.94 (277)	0.95 (339)	0.95 (403)
	Set 3, $N_{iter}^{CG} = 4174$			Set 4, $N_{iter}^{CG} = 6689$		
	2 levels	3 levels	4 levels	2 levels	3 levels	4 levels
$m = 3$	-	0.98 (1061)	0.98 (860)	-	0.97 (699)	0.98 (823)
$m = 5$	0.96 (428)	0.97 (648)	0.97 (527)	0.95 (392)	0.96 (435)	0.96 (508)
$m = 8$	0.94 (288)	0.96 (418)	0.95 (341)	0.93 (273)	0.94 (290)	0.95 (335)

We have repeated the same set of experiments employing $p_j = 3$, $\forall j = 1, \ldots, J$; the results are reported in Table 2, where we also have reported the iterations count (between parenthesis). Firstly, a comparison between Table 1 and Table 2 confirms that the convergence factor increases as p grows up if the number of smoothing steps is kept fixed. Secondly, we observe that if the number of smoothing step is kept too small then the convergence of the method could not be guaranteed: indeed, according to Theorem 7, a uniformly convergent (also with respect to p) solver require a number of smoothing steps $m > 2C_1 C_Q \gtrsim p^{2+\mu}$ as shown in Figure 4. If m is big enough, we observe that also in this case the number of iterations does not depend from the granularity of the underlying mesh, while the iterations count of the Conjugate Gradient method is growing if h decrease.

6 Additive Schwarz smoother

In order to improve the convergence properties of the V-cycle algorithm studied above, we define in this section a domain decomposition preconditioner that we will use as a smoothing operator instead of the Richardson iteration. To this end, let $\mathcal{T}_j$ and $\mathcal{T}_{j-1}$ be respectively the finer and the coarser non-nested meshes, satisfying the grid assumptions given in Sect. 2.1. We then introduce the *local* and *coarse*

solvers, that are the key ingredients in the definition of the smoother on the space V_j, $j = 2, \ldots, J$.

Local Solvers. Let us consider the finest mesh $\mathcal{T}_j$ with cardinality N_j, then for each element $\kappa_i \in \mathcal{T}_j$, we define a local space V_j^i as the restriction of the DG finite element space V_j to the element $\kappa_i \in \mathcal{T}_j$:

$$V_j^i = V_j|_{\kappa_i} \equiv \mathcal{P}_{p_j}(\kappa_i) \qquad \forall i = 1, \ldots, N_j,$$

and for each local space, the associated local bilinear form is defined by

$$\mathcal{A}_j^i : V_j^i \times V_j^i \to \mathbb{R}, \quad \mathcal{A}_j^i(u_i, v_i) = \mathcal{A}_j(R_i^T u_i, R_i^T v_i) \quad \forall u_i, v_i \in V^i,$$

where $R_i^T : V_j^i \to V_j$ denotes the classical extension by-zero operator from the local space V_j^i to the global V_j.

Coarse Solver. The natural choice in our contest is to define the coarse space V_j^0 to be exactly the same used for the *Coarse grid correction* step of the V-cycle algorithm introduced in Sect. 4, that is

$$V_j^0 = V_{j-1} \equiv \{v \in L^2(\Omega) : v|_\kappa \in \mathcal{P}_{p_{j-1}}(\kappa), \kappa \in \mathcal{T}_{j-1}\},$$

the bilinear form on V_j^0 is then given by

$$\mathcal{A}_j^0 : V_j^0 \times V_j^0 \to \mathbb{R}, \quad \mathcal{A}_j^0(u_0, v_0) = \mathcal{A}_{j-1}(u_0, v_0) \quad \forall u_0, v_0 \in V_j^0.$$

We also define the injection operator from V_j^0 to V_j: conversely with respect to the case where the coarser mesh is obtained by agglomeration, here the injection operator is not trivial, and it is defined as the prolongation operator introduced in Sect. 4, that is $R_0^T : V_j^0 \to V_j$, $R_0^T = I_{j-1}^j$. By introducing the projection operators $P_i = R_i^T \tilde{P}_i : V_j \to V_j$, $i = 0, 1, \ldots, N_j$, where

$$\tilde{P}_i : V_j \to V_j^i, \quad \mathcal{A}_j^i(\tilde{P}_i v_h, w_i) = \mathcal{A}_j(v_h, R_i^T w_i) \quad \forall w_i \in V_j^i, \quad i = 1, \ldots, N_j,$$
$$\tilde{P}_0 : V_j \to V_j^0, \quad \mathcal{A}_j^0(\tilde{P}_0 v_h, w_0) = \mathcal{A}_j(v_h, R_0^T w_0) \quad \forall w_0 \in V_j^0,$$

the additive Schwarz operator is defined by $P_{ad} = \sum_{i=0}^{N_j} (R_i^T (A_j^i)^{-1} R_i) A_j \equiv B_{ad}^{-1} A_j$, where $B_{ad}^{-1} = \sum_{i=0}^{N_j} (R_i^T (A_j^i)^{-1} R_i)$ is the preconditioner. Then, the *Additive Schwarz* smoothing operator with m steps consists in performing m iterations of the *Preconditioned Conjugate Gradient* method using B_{ad} as preconditioner. In Algorithm 3 we outline the V-cycle multigrid method using P_{ad} as a smoother. Here, $\mathsf{MG}_{\mathcal{AS}}(j, g, z_0, m_1, m_2)$ denotes the approximate solution of $A_j z = g$ obtained after one iteration, with initial guess z_0 and m_1, m_2 pre- and post-smoothing steps, respectively. Here, the smoothing step is performed by the algorithm $ASPCG$, i.e., $z = ASPCG(A, z_0, g, m)$ represents the output of m steps of *Preconditioned Conjugate Gradient* method applied to the linear system of equations $Ax = g$, by using B_{as} as preconditioner and starting with the initial guess z_0.

The numerical performance of Algorithm 3 are reported in Tables 3, 4 and 5, for the corresponding V-cycle algorithm with $J = 2, 3, 4$ levels. The simulations are similar to the ones described in the previous section: here we used the grids of Set 2, 3 and 4 of Figure 1, and we varied the polynomial degree $p \in \{1, 3, 5\}$. Firstly,

Algorithm 3 One iteration of Multigrid V-cycle scheme with AS-smoother

---

Pre-smoothing:
if j=1 **then**
  $\mathsf{MG}_{\mathcal{AS}}(1, g, z_0, m_1, m_2) = A_1^{-1} g.$
else
  Pre-smoothing:
  $z^{(m_1)} = ASPCG(A_j, z_0, g, m_1);$

  Coarse grid correction:
  $r_{j-1} = I_j^{j-1}(g - A_j z^{(m_1)});$
  $e_{j-1} = \mathsf{MG}_{\mathcal{AS}}(j-1, r_{j-1}, 0, m_1, m_2);$
  $z^{(m_1+1)} = z^{(m_1)} + I_{j-1}^j e_{j-1};$

  Post-smoothing:
  $z^{(m_1+m_1+1)} = ASPCG(A_j, z^{(m_1+1)}, g, m_2);$

  $\mathsf{MG}_{\mathcal{AS}}(j, g, z_0, m_1, m_2) = z^{(m_1+m_2+1)}.$
end if

Table 3: Iteration counts of the V-cycle solvers with the Additive Schwarz smoother as a function of m ($C_\sigma^j \equiv C_\sigma = 10$, $p = 1$).

	Set 2			Set 3			Set 4		
	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.
$m=3$	18	18	18	18	18	18	20	20	20
$m=5$	9	9	9	9	9	9	10	10	10
$m=8$	5	5	5	5	5	5	5	5	5
	$N_{iter}^{CG} = 600$			$N_{iter}^{CG} = 867$			$N_{iter}^{CG} = 1228$		

Table 4: Iteration counts of the V-cycle solvers with the Additive Schwarz smoother as a function of m ($C_\sigma^j \equiv C_\sigma = 10$, $p = 3$).

	Set 2			Set 3			Set 4		
	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.
$m=3$	63	64	64	57	59	59	59	60	60
$m=5$	27	27	27	25	25	25	26	26	26
$m=8$	13	13	13	13	14	14	14	14	14
	$N_{iter}^{CG} = 3223$			$N_{iter}^{CG} = 4174$			$N_{iter}^{CG} = 6689$		

we observe that, also in this case, the number of iteration does not increase with the number of elements in the underlying mesh for a fixed number of smoothing steps m; moreover, we does not observe the constrain from below required to the number of smoothing steps with respect to the degree of approximation: the method converges also with high degree of approximation and a small number of smoothing steps. Finally, Table 6 shows the numerical results relatives to an example of hp-multigrid, characterized by a choice of different polynomial degrees of approximation between non-nested space: also in this case we observe that the number of iterations is independent of the granularity of the finest mesh, and we have convergence for any choice of smoothing steps m.

Table 5: Iteration counts of the V-cycle solvers with the Additive Schwarz smoother as a function of m ($C_\sigma^j \equiv C_\sigma = 10$, $p = 5$).

	Set 2			Set 3			Set 4		
	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.
$m = 3$	148	156	156	125	132	132	149	158	157
$m = 5$	59	59	58	51	51	51	59	60	60
$m = 8$	26	26	26	24	24	24	27	27	27
	$N_{iter}^{CG} = 7676$			$N_{iter}^{CG} = 11525$			$N_{iter}^{CG} = 15814$		

Table 6: Iteration counts of the hp-version of the V-cycle solvers with the Additive Schwarz smoother as a function of m. Here the polynomial degree on each space is $p_j = j$ for $j = 1, 2, 3, 4$.

	Set 2			Set 3			Set 4		
	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.
$m = 3$	85	86	86	79	80	80	83	85	84
$m = 5$	35	35	35	32	32	32	33	33	33
$m = 8$	17	17	17	17	17	17	17	18	17
	$N_{iter}^{CG} = 5108$			$N_{iter}^{CG} = 7697$			$N_{iter}^{CG} = 10572$		

A Proof of Lemma 1

Proof (of Lemma 1) We follow the idea of [29, Proof of Lemma 1.49]. First of all, we observe that

$$\|v\|_{L^2(\partial \kappa)}^2 = \sum_{F \subset \partial \kappa} \|v\|_{L^2(F)}^2. \tag{26}$$

For each face $F \subset \partial \kappa$ let $T_F \subset \kappa$ be a d-dimensional simplex sharing the face F with κ and satisfying the Assumption 1: in T_F we define a function σ_F as follow:

$$\sigma_F : \mathbf{x} \in \overline{T_F} \mapsto \sigma_F(\mathbf{x}) = \frac{|F|}{d|T_F|}(\mathbf{x} - \mathbf{v}_F),$$

where $\mathbf{v}_F$ is the vertex of the simplex T_F opposite to the face F. We observe that:

- $\sigma_F(\mathbf{x}) \cdot \mathbf{n}_F = \frac{|F|}{d|T_F|}\tilde{h}$ $\forall \mathbf{x} \in F$, where $\tilde{h}$ is the height of the simplex respect to the face F, that is also $\tilde{h} = \frac{d|T_F|}{|F|}$, then $\sigma_F|_F \cdot \mathbf{n}_F = 1$;
- $\sigma_F|_{F'} \cdot \mathbf{n}_{F'} = 0$ $\forall$ faces $F' \subset \partial T_F, F' \neq F$;

then we have:

$$\|v\|_{L^2(F)}^2 = \int_F |v|^2 d\sigma = \int_{\partial T_F} |v|^2 \sigma_F \cdot \mathbf{n}_F d\sigma = \int_T \nabla \cdot (|v|^2 \sigma_F) d\mathbf{x}$$

$$= \int_T 2v \nabla v \cdot \sigma_F d\mathbf{x} + \int_T |v|^2 \nabla \cdot \sigma_F d\mathbf{x};$$

now the following properties hold for σ_F:

- $\nabla \cdot \sigma_F = \nabla \cdot \frac{|F|}{d|T_F|}(\mathbf{x} - \mathbf{v}_F) = \frac{|F|}{d|T_F|}\nabla \cdot \mathbf{x} = \frac{|F|}{|T_F|}$;
- $\|\sigma_F\|_{[L^\infty(T_F)]^d} = \frac{|F|}{d|T_F|}h_T \leq \frac{|F|}{d|T_F|}h_\kappa$,

which implies:

$$\|v\|_{L^2(F)}^2 \leq 2\|\sigma_F\|_{[L^\infty(T_F)]^d}\|v\nabla v\|_{[L^1(T_F)]^d} + \frac{|F|}{|T_F|}\|v\|_{L^2(T_F)}^2,$$

$$\leq 2\frac{|F|}{d|T_F|}h_\kappa\|v\|_{L^2(T_F)}|v|_{H^1(T_F)} + \frac{|F|}{|T_F|}\|v\|_{L^2(T_F)}^2,$$

using the Assumption 1 we have

$$\|v\|^2_{L^2(F)} \le 2C \|v\|_{L^2(T_F)} |v|_{H^1(T_F)} + \frac{Cd}{h_\kappa} \|v\|^2_{L^2(T_F)}.$$

By using Young Inequality we could bound

$$\|v\|_{L^2(T_F)} |v|_{H^1(T_F)} \le \frac{1}{2} \Big(\frac{\epsilon}{h_\kappa} \|v\|^2_{L^2(T_F)} + \frac{h_\kappa}{\epsilon} |v|^2_{H^1(T_F)} \Big),$$

where we have chosen $\epsilon \ge 1$. Using the previous inequality we have

$$\|v\|^2_{L^2(F)} \le 2Cd \Big(\frac{\epsilon}{h_\kappa} \|v\|^2_{L^2(T_F)} + \frac{h_\kappa}{\epsilon} |v|^2_{H^1(T_F)} \Big). \tag{27}$$

we observe that (27) holds $\forall F \subset \partial\kappa$. Then, thanks to (26), we have

$$
\begin{aligned}
\|v\|^2_{L^2(\partial\kappa)} = \sum_{F \subset \partial\kappa} \|v\|^2_{L^2(F)} &\le \sum_{F \subset \partial\kappa} 2Cd \Big(\frac{\epsilon}{h_\kappa} \|v\|^2_{L^2(T_F)} + \frac{h_\kappa}{\epsilon} |v|^2_{H^1(T_F)} \Big) \\
&= 2Cd \Big(\frac{\epsilon}{h_\kappa} \sum_{F \subset \partial\kappa} \|v\|^2_{L^2(T_F)} + \frac{h_\kappa}{\epsilon} \sum_{F \subset \partial\kappa} |v|^2_{H^1(T_F)} \Big) \\
&\le 2Cd \Big(\frac{\epsilon}{h_\kappa} \|v\|^2_{L^2(\kappa)} + \frac{h_\kappa}{\epsilon} |v|^2_{H^1(\kappa)} \Big),
\end{aligned}
$$

where in the last inequality we have used the fact that the simplices of the set $\{T_F : F \subset \partial\kappa\}$ satisfy Assumption 1, in the sense that they are disjoints and $\cup_{F \subset \partial\kappa} \overline{T_F} \subset \overline{\kappa}$.

B Proof of Lemma 7

In order to show Lemma 7 we follow the analysis presented in [30], by firstly showing two preliminary results making use of the properties presented in Sect. 3.

Lemma 8 *Let Assumptions 1 - 4 hold, and let $\widetilde{\Pi}_j$ be the projection operator on V_j as defined in Lemma 4, for $j = J, J - 1$. Then*

$$\|\widetilde{\Pi}_J w - I^J_{J-1} \widetilde{\Pi}_{J-1} w\|_{L^2(\Omega)} \lesssim \frac{h_J^2}{p_J^2} \|w\|_{H^2(\Omega)} \quad \forall w \in H^2(\Omega).$$

Proof Using the triangular inequality, Remark 6 and the approximation estimates of Lemma 4 we have:

$$
\begin{aligned}
\|\widetilde{\Pi}_J w - I^J_{J-1} \widetilde{\Pi}_{J-1} w\|_{L^2(\Omega)} &\le \\
&\le \|\widetilde{\Pi}_J w - w\|_{L^2(\Omega)} + \|w - Q_J w\|_{L^2(\Omega)} + \|Q_J w - I^J_{J-1} \widetilde{\Pi}_{J-1} w\|_{L^2(\Omega)} \\
&= \|\widetilde{\Pi}_J w - w\|_{L^2(\Omega)} + \min_{z_h \in V_J} \|w - z_h\|_{L^2(\Omega)} + \|Q_J(w - \widetilde{\Pi}_{J-1} w)\|_{L^2(\Omega)} \\
&\le \|\widetilde{\Pi}_J w - w\|_{L^2(\Omega)} + \| w - \widetilde{\Pi}_J w\|_{L^2(\Omega)} + \|w - \widetilde{\Pi}_{J-1} w\|_{L^2(\Omega)} \\
&\lesssim \frac{h_J^2}{p_J^2} \|w\|_{H^2(\Omega)} + \frac{h_{J-1}^2}{p_{J-1}^2} \|w\|_{H^2(\Omega)} \lesssim \frac{h_J^2}{p_J^2} \|w\|_{H^2(\Omega)},
\end{aligned}
$$

where in the last inequality we also used hypothesis (3) and (4).

Lemma 9 *Let Assumptions 1 - 4 hold. Let be $g \in L^2(\Omega)$ and denote by $w_j \in V_j$ the solution of $\mathcal{A}_j(w_j, v) = (g, v)_{L^2(\Omega)}$ $\forall v \in V_j$ with $j = J - 1, J$. Then the following inequality holds:*

$$\|w_J - I^J_{J-1} w_{J-1}\|_{L^2(\Omega)} + \|w_{J-1} - P^{J-1}_J w_J\|_{L^2(\Omega)} \lesssim \frac{h_J^2}{p_J^{2-\mu}} \|g\|_{L^2(\Omega)}. \tag{28}$$

Proof Consider the unique solution $w \in V$ of the problem

$$\mathcal{A}(w, v) = (g, v)_{L^2(\Omega)} \quad \forall v \in V.$$

Using Theorem 5 we have

$$\|w - w_j\|_{L^2(\Omega)} \lesssim \frac{h_j^2}{p_j^{2-\mu}} \|w\|_{L^2(\Omega)}, \quad j = J - 1, J. \tag{29}$$

Using the triangular inequality and Remark 6 we have:

$$
\begin{aligned}
\|w_J - I_{J-1}^J w_{J-1}\|_{L^2(\Omega)} &\leq \|w_J - w\|_{L^2(\Omega)} + \|w - \widetilde{\Pi}_J w\|_{L^2(\Omega)} \\
&\quad + \|\widetilde{\Pi}_J w - I_{J-1}^J \widetilde{\Pi}_{J-1} w\|_{L^2(\Omega)} + \|I_{J-1}^J \widetilde{\Pi}_{J-1} w - Q_J w\|_{L^2(\Omega)} \\
&\quad + \|Q_J w - I_{J-1}^J w_{J-1}\|_{L^2(\Omega)} \\
&= \|w_J - w\|_{L^2(\Omega)} + \|w - \widetilde{\Pi}_J w\|_{L^2(\Omega)} + \|\widetilde{\Pi}_J w - I_{J-1}^J \widetilde{\Pi}_{J-1} w\|_{L^2(\Omega)} \\
&\quad + \|Q_J(\widetilde{\Pi}_{J-1} w - w)\|_{L^2(\Omega)} + \|Q_J(w - w_{J-1})\|_{L^2(\Omega)} \\
&\leq \|w_J - w\|_{L^2(\Omega)} + \|w - \widetilde{\Pi}_J w\|_{L^2(\Omega)} + \|\widetilde{\Pi}_J w - I_{J-1}^J \widetilde{\Pi}_{J-1} w\|_{L^2(\Omega)} \\
&\quad + \|\widetilde{\Pi}_{J-1} w - w\|_{L^2(\Omega)} + \|w - w_{J-1}\|_{L^2(\Omega)}.
\end{aligned}
$$

Using (29), Lemma 4 and Lemma 8, we have

$$
\begin{aligned}
\|w_J - I_{J-1}^J w_{J-1}\|_{L^2(\Omega)} &\lesssim \frac{h_J^2}{p_J^{2-\mu}} \|w\|_{H^2(\Omega)} + \frac{h_J^2}{p_J^2} \|w\|_{H^2(\Omega)} + \frac{h_J^2}{p_J^2} \|w\|_{H^2(\Omega)} \\
&\quad + \frac{h_{J-1}^2}{p_{J-1}^2} \|w\|_{H^2(\Omega)} + \frac{h_{J-1}^2}{p_{J-1}^{2-\mu}} \|w\|_{H^2(\Omega)}.
\end{aligned}
$$

From the elliptic regularity assumption (2) and hypothesis (3) and (4), we can write

$$\|w_J - I_{J-1}^J w_{J-1}\|_{L^2(\Omega)} \lesssim \frac{h_J^2}{p_J^{2-\mu}} \|g\|_{L^2(\Omega)}. \tag{30}$$

Now, let $z_j \in V_j$ be the solution of:

$$\mathcal{A}_j(z_j, q) = (w_{J-1} - P_J^{J-1} w_J, q_j)_{L^2(\Omega)} \quad \forall q_j \in V_j, \quad j = J - 1, J;$$

Using (30) we get the following estimate:

$$\|z_{J-1} - I_{J-1}^J z_{J-1}\|_{L^2(\Omega)} \lesssim \frac{h_J^2}{p_J^{2-\mu}} \|w_{J-1} - P_J^{J-1} w_J\|_{L^2(\Omega)}.$$

Then, we have:

$$
\begin{aligned}
\|w_{J-1} - P_J^{J-1} w_J\|_{L^2(\Omega)}^2 &= \mathcal{A}_{J-1}(z_{J-1}, w_{J-1} - P_J^{J-1} w_J) \\
&= \mathcal{A}_{J-1}(z_{J-1}, w_{J-1}) - \mathcal{A}_J(I_{J-1}^J z_{J-1}, w_J) \\
&= (z_{J-1}, g) - (I_{J-1}^J z_{J-1}, g) = (g, z_{J-1} - I_{J-1}^J z_{J-1}) \\
&\lesssim \|g\|_{L^2(\Omega)} \frac{h_J^2}{p_J^{2-\mu}} \|w_{J-1} - P_J^{J-1} w_J\|_{L^2(\Omega)},
\end{aligned}
$$

from which, together with (30), inequality (28) follows.

Proof (of Lemma 7) For any $v_J \in V_J$ we consider the following equality:

$$\|(\mathrm{Id}_J - I_{J-1}^J P_J^{J-1})v_J\|_{L^2(\Omega)} = \sup_{0 \neq \phi \in L^2(\Omega)} \frac{\left(\phi, (\mathrm{Id}_J - I_{J-1}^J P_J^{J-1})v_J\right)_{L^2(\Omega)}}{\|\phi\|_{L^2(\Omega)}}. \quad (31)$$

Next, consider the solution z_j of the following problems

$$\mathcal{A}_j(z_j, v_j) = (\phi, v_j) \quad \forall v_j \in V_j, \text{ for } j = J, J-1.$$

By using the definition of P_J^{J-1} and Lemma 9, we have:

$$
\begin{aligned}
\left(\phi, (\mathrm{Id}_J - I_{J-1}^J P_J^{J-1})v_J)\right)_{L^2(\Omega)} &= \mathcal{A}_J(z_J, v_J) - \mathcal{A}_{J-1}(P_J^{J-1}z_J, P_J^{J-1}v_J) \\
&= \mathcal{A}_J(z_J - I_{J-1}^J z_{J-1}, v_J) + \mathcal{A}_J(I_{J-1}^J(z_{J-1} - P_J^{J-1}z_J), v_J) \\
&\leq \|v_J\|_{2,J}\left(\|z_J - I_{J-1}^J z_{J-1}\|_{L^2(\Omega)} + \|z_{J-1} - P_J^{J-1}z_J\|_{L^2(\Omega)}\right) \\
&\lesssim \|v_J\|_{2,J} \frac{h_J^2}{p_J^{2-\mu}} \|\phi\|_{L^2(\Omega)}.
\end{aligned}
$$

Using the last inequality together with (31) we get (24).

References

1. Antonietti, P.F., Brezzi, F., Marini, L.D.: Stabilizations of the Baumann-Oden DG formulation: the 3D case. Boll. Unione Mat. Ital. **1**(3), 629 – 643 (2008)
2. Antonietti, P.F., Facciolà, C., Russo, A., Verani, M.: Discontinuous Galerkin approximation of flows in fractured porous media. Technical Report 22/2016, MOX Report (2016)
3. Antonietti, P.F., Giani, S., Houston, P.: *hp*-version composite Discontinuous Galerkin methods for elliptic problems on complicated domains. SIAM J. Sci. Comput. **35**(3), A1417 – A1439 (2013)
4. 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 (2014)
5. 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 (2011)
6. Antonietti, P.F., Houston, P., Hu, X., Sarti, M., Verani, M.: Multigrid algorithms for *hp*-version Interior Penalty Discontinuous Galerkin methods on polygonal and polyhedral meshes. Calcolo pp. 1 – 30 (2017)
7. Antonietti, P.F., Houston, P., Smears, I.: A note on optimal spectral bounds for nonoverlapping domain decomposition preconditioners for *hp*-version Discontinuous Galerkin methods. Internat. J. Numer. Methods Engrg. **13**(4), 513–524 (2016)
8. Antonietti, P.F., Mazzieri, I.: DG methods for the elastodynamics equations on polygonal/polyhedral grids. in preparation (2017)
9. Antonietti, P.F., Sarti, M., Verani, M.: Multigrid algorithms for *hp*-Discontinuous Galerkin discretizations of elliptic problems. SIAM J. Numer. Anal. **53**(1), 598 – 618 (2015)
10. Arnold, D.N.: An Interior Penalty Finite Element method with discontinuous elements. SIAM J. Numer. Anal. **19**(4), 742 – 760 (1982)
11. Arnold, N.D., Brezzi, F., Cockburn, B., Marini, L.D.: Unified analysis of Discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal. **39**(5), 1749 – 1779 (2002)
12. Baker, G.A.: Finite element methods for elliptic equations using nonconforming elements. Math. Comp. **31**(137), 45 – 59 (1977)
13. Bank, R.E., Dupont, T.: An optimal order process for solving finite element equations. Math. Comp. **36**(153), 35 – 51 (1981)
14. Bassi, F., Botti, L., Colombo, A., Brezzi, F., Manzini, G.: Agglomeration-based physical frame dG discretizations: an attempt to be mesh free. Math. Models Methods Appl. Sci. **24**(8), 1495–1539 (2014)

15. Bassi, F., Botti, L., Colombo, A., Di Pietro, D.A., Tesini, P.: On the flexibility of agglomeration based physical space discontinuous Galerkin discretizations. J. Comput. Phys. **231**(1), 45–65 (2012)
16. Bassi, F., Botti, L., Colombo, A., Rebay, S.: Agglomeration based Discontinuous Galerkin discretization of the Euler and Navier-Stokes equations. Comput. & Fluids **61**, 77–85 (2012)
17. Braess, D., Verfürth, R.: Multigrid methods for nonconforming finite element methods. SIAM J. Numer. Anal. **22**(4), 979 − 986 (1990)
18. Bramble, J.H.: Multigrid Methods (Pitman Research Notes in Mathematics Series). Longman Scientific and Technical (1993)
19. Bramble, J.H., Kwak, D.Y., Pasciak, J.E.: Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems. SIAM J. Numer. Anal. **31**(6), 1746 − 1763 (1994)
20. Bramble, J.H., Pasciak, J.E.: The analysis of smoothers for multigrid algorithms. Math. Comp. **58**(198), 467 − 488 (1992)
21. Bramble, J.H., Pasciak, J.E.: New estimates for multilevel algorithms including the V-cycle. Math. Comp. **60**(202), 447 − 471 (1993)
22. Bramble, J.H., Pasciak, J.E.: Uniform convergence estimates for multigrid V-cycle algorithms with less than full elliptic regularity. SIAM J. Numer. Anal. **31**(6), 1746 − 1763 (1994)
23. Bramble, J.H., Pasciak, J.E., Xu, J.: The analysis of multigrid algorithms with nonnested space or noninherited quadratic forms. Math. Comp. **56**(193), 1–34 (1991)
24. Bramble, J.H., Zhang, X.: Uniform convergence of the multigrid V-cycle for an anisotropic problem. Math. Comp. **70**(234), 453 − 470 (2001)
25. Cangiani, A., Dong, Z., Georgoulis, E.H.: hp-Version space-time Discontinuous Galerkin methods for parabolic problems on prismatic meshes. SIAM J. Sci. Comput. **39**(4), A1251 − A1279 (2017)
26. Cangiani, A., Dong, Z., Georgoulis, E.H., Houston, P.: hp-Version Discontinuous Galerkin methods for advection-diffusion-reaction problems on polytopic meshes. ESAIM Math. Model. Numer. Anal. **50**(3), 699 − 725 (2016)
27. Cangiani, A., Georgoulis, E.H., Houston, P.: hp-Version discontinuous Galerkin methods on polygonal and polyhedral meshes. Math. Models Methods Appl. Sci. **24**(10), 2009 − 2041 (2014)
28. Cockburn, B., Karniadakis, G.E., Shu, C.W.: Discontinuous Galerkin Methods. Theory, computation and applications. Springer Berlin (2000)
29. Di Pietro, D.A., Ern, A.: Mathematical aspects of discontinuous Galerkin methods. Springer Science & Business Media (2011)
30. Duan, H.Y., Gao, S.Q., Tan, R.C.E., Zhang, S.: A generalized BPX multigrid framework covering nonnested V-cycle methods. Math. Comp. **76**(257), 137–152 (2007)
31. Georgoulis, E.H., Suli, E.: Optimal error estimates for the hp-version Interior Penalty Discontinuous Galerkin finite element method. IMA J. Numer. Anal. **25**(1), 205–220 (2005)
32. Giani, S., Houston, P.: Domain decomposition preconditioners for Discontinuous Galerkin discretizations of compressible fluid flows. Numer. Math. Theory Methods Appl. **7**(2), 123 − 148 (2014)
33. Giani, S., Houston, P.: hp-Adaptive composite Discontinuous Galerkin methods for elliptic problems on complicated domains. Numer. Methods Partial Differential Equations **30**(4), 1342 − 1367 (2014)
34. Gopalakrishnan, J., Kanschat, G.: Application of unified DG analysis to preconditioning DG methods. In: K.J. Bathe (ed.) Computational Fluid and Solid Mechanics 2003, pp. 1943 − 1945. Elsevier Science (2003)
35. Gopalakrishnan, J., Kanschat, G.: A multilevel Discontinuous Galerkin method. Numer. Math. **95**(3), 527 − 550 (2003)
36. Gopalakrishnan, J., Pasciak, J.E.: Multigrid for the Mortar finite element method. SIAM J. Numer. Anal. **37**(3), 1029 − 1052 (2000)
37. Hesthaven, J.S., Warburton, T.: Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications, 1st edn. Springer Berlin (2008)
38. Lipnikov, K., Vassilev, D., Yotov, I.: Discontinuous Galerkin and mimetic finite difference methods for coupled Stokes-Darcy flows on polygonal and polyhedral grids. Numer. Math. **126**(2), 321 − 360 (2014)

39. Monk, P., Zhang, S.: Multigrid computation of vector potentials. J. Comput. Appl. Math. **62**(3), 301 – 320 (1995)
40. Nitsche, J.: Über ein Variationsprinzip zur Lösung von Dirichlet Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. Abh. Math. Semin. Univ. Hambg. **36**, 9 – 15 (1971)
41. Reed, W., Hill, T.: Triangular mesh methods for the neutron transport equation. Tech. Rep. LA-UR-73-479, Los Alamos Scientific Laboratory (1973)
42. Rivière, B.: Discontinuous Galerkin methods for solving elliptic and parabolic equations. SIAM (2008)
43. Scott, L.R., Zhang, S.: Higher-dimensional nonnested multigrid methods. Math. Comp. **58**(198), 457 – 466 (1992)
44. Talischi, C., Paulino, G., Pereira, A., Menezes, I.: Polymesher: A general-purpose mesh generator for polygonal elements written in matlab. Struct. Multidiscip. Optim. **45**(3), 309–328 (2012)
45. Wheeler, M.F.: An elliptic collocation Finite Element method with Interior Penalties. SIAM J. Numer. Anal. **15**(1), 152 – 161 (1978)
46. Wiresaet, D., Kubatko, E.J., Michoski, C.E., Tanaka, S., Westerink, J.J., Dawson, C.: Discontinuous Galerkin methods with nodal and hybrid modal/nodal triangular, quadrilateral, and polygonal elements for nonlinear shallow water flow. Comput. Methods Appl. Mech. Engrg. **270**, 113 – 149 (2014)
47. Xu, X., Chen, J.: Multigrid for the mortar element method for P1 nonconforming element. Numer. Math. **88**(2), 381 – 389 (2001)
48. Xu, X., Li, L., Chen, W.: A multigrid method for the Mortar-type Morley element approximation of a plate bending problem. SIAM J. Numer. Anal. **39**(5), 1712 – 1731 (2002)
49. Zhang, S.: Optimal-order nonnested multigrid methods for solving Finite Element equations I: on quasi-uniform meshes. Math. Comp. **55**(191), 23 – 36 (1990)
50. Zhang, S., Zhang, Z.: Treatments of discontinuity and bubble functions in the multigrid method. Math. Comp. **66**(219), 1055 – 1072 (1997)